

WICC 2017

XIX Workshop
de Investigadores
en Ciencias de la
Computación

27 y 28 de Abril
Ciudad de Buenos Aires



ISBN 978-987-42-5143-5



Índice

WICC 2017	9
AUTORIDADES	10
COMITÉ ORGANIZADOR LOCAL (ITBA)	10
COMITÉ ACADÉMICO WICC 2017	11
COORDINADORES DE ÁREA WICC 2017	13
AGENTES Y SISTEMAS INTELIGENTES	15
AGENTES INTELIGENTES Y WEB SEMÁNTICA: PREPROCESAMIENTO DE TEXTO DE REDES SOCIALES	16
AGENTES INTELIGENTES. MODELOS FORMALES Y APLICACIONES	21
APLICACIÓN DE TÉCNICAS APRENDIZAJE AUTOMÁTICO PARA ESTIMAR LA CALIDAD DE LA VOZ EN ESCALA GRBAS	26
APLICACIONES DE ANÁLISIS DE INFORMACIÓN TEXTUAL: CORPUS LINGÜÍSTICOS	31
APRENDIZAJE AUTÓNOMO EN SISTEMAS INTELIGENTES	35
ARTIFICIAL BEE COLONY APLICADO A LA OPTIMIZACIÓN NUTRICIONAL DE MENÚS ESCOLARES	40
BÚSQUEDA DE ENTORNO VARIABLE (VNS) PARA EL PROBLEMA DE PLANIFICACIÓN DE MÁQUINAS PARALELAS IDÉNTICAS	45
CONFIANZA Y REPUTACIÓN DE AGENTES EN SISTEMAS MULTI-AGENTE PARA ENTORNOS DINÁMICOS	48
CONOCIMIENTO COMPARTIDO Y RAZONAMIENTO ARGUMENTATIVO COLABORATIVO PARA ENTORNOS DE MÚLTIPLES AGENTES EN AMBIENTES DISTRIBUIDOS	53
DESARROLLO DE SISTEMAS DE ANÁLISIS DE TEXTO	58
DISEÑO DE ALGORITMOS EVOLUTIVOS HÍBRIDOS OPTIMIZADOS PARA BICLUSTERING	63
EVALUACIÓN DE SENSORES Y PLANTA MOTRIZ UTILIZADOS EN LA CONSTRUCCIÓN DE DRONES	68
FUSIÓN DE ALGORITMOS BAYESIANOS Y ÁRBOLES DE CLASIFICACIÓN COMO PROPUESTA PARA LA CLASIFICACIÓN SUPERVISADA DE FALLOS DE EQUIPOS EN UN LABORATORIO DE CÓMPUTOS.	72
HACIA LA DEFINICIÓN DE UN AGENTE GENERADOR DE CONOCIMIENTO DE VALOR SOCIAL PARA POBLACIONES EN RIESGO	77
IMPLEMENTACIÓN DE OPERADORES DE CONSOLIDACIÓN DE ONTOLOGÍAS EN DATALOG +/-	81
INTEGRACIÓN DE ONTOLOGÍAS DATALOG ± MEDIANTE CONSOLIDACIÓN A TRAVÉS DE DEBILITAMIENTO	86
INTELIGENCIA COMPUTACIONAL APLICADA A LA OPTIMIZACIÓN MULTIOJETIVO DE PROBLEMAS DE SCHEDULING CON RESTRICCIONES	91
LIBRERÍA DE METAHEURÍSTICAS PARA PROBLEMAS DE OPTIMIZACIÓN CONTINUOS Y DISCRETOS	96
MODELO PARA ANALIZAR MENSAJES Y DETECTAR ACTITUDES PELIGROSAS A TRAVÉS DE ANÁLISIS DE SENTIMIENTOS CON ALGORITMOS DE APRENDIZAJES	100
PLANIFICACIÓN ESTOCÁSTICA BASADA EN MARKOV PARA LA COMPOSICIÓN Y RECOMENDACIÓN DE SERVICIOS WEB	105
PROBLEMAS COMPLEJOS RESUELTOS CON METAHEURÍSTICAS	110
PROCESO DE ELICITACIÓN DE REQUERIMIENTOS PARA LA CONSTRUCCIÓN DE MODELOS PREDICTIVOS BASADOS EN SISTEMAS INTELIGENTES	115
REVISIÓN DE CREENCIAS EN BASES DE CONOCIMIENTO HORN Y SU COMPLEJIDAD COMPUTACIONAL	120
SER HUMANO VIRTUAL	125
TÉCNICAS DE INTELIGENCIA ARTIFICIAL APLICADAS A PROBLEMAS CIENTÍFICO-TECNOLÓGICOS	130
TOMA DE DECISIONES INDIVIDUALES Y COLECTIVAS PARA SISTEMAS MULTI-AGENTE EN ENTORNOS DISTRIBUIDOS	135
USO DE METAHEURÍSTICAS PARA EL DISEÑO DE COBERTURAS CON RADIO FRECUENCIA APLICADAS A LA TELESUPERVISIÓN DE YACIMIENTOS PETROLEROS.	140
ARQUITECTURA, REDES Y SISTEMAS OPERATIVOS	146
ADAPTACIÓN DE UN MIDDLEWARE DE INTERNET DE LAS COSAS PARA GESTIONAR EL PROCESO DE FERMENTADO EN LA INDUSTRIA DE MANUFACTURA DEL TÉ	147
AMPLIANDO LA VIDA ÚTIL DE LAS WSN POR MEDIO DE LOS PROTOCOLOS DE RUTEO, MODIFICACION DE AODV	152
ANÁLISIS COMPARATIVO DE PRESTACIONES DE TRÁFICO DE VIDEO MULTICAST EN REDES IPV6	158
ANÁLISIS DE EFICIENCIA EN ARQUITECTURAS MULTIPROCESADOR PARA APLICACIONES DE TRANSMISIÓN Y PROCESAMIENTO DE DATOS	163
AUTOMATIZACIÓN DE SISTEMAS DE CULTIVOS HIDROPÓNICOS	168
DIAGNÓSTICO Y METODOLOGÍA PARA LA IMPLEMENTACIÓN DE "INTERNET OF THINGS" EN EL PLANEAMIENTO Y DESARROLLO DE CIUDADES INTELIGENTES	171
ESTUDIO DE HERRAMIENTAS DE SIMULACIÓN EN REDES DEFINIDAS POR SOFTWARE	176
IMPLEMENTACIÓN DE MIDDLEWARE PUBLICADOR/SUBSCRIPTOR PARA APLICACIONES WEB DE MONITOREO	181
IMPLEMENTANDO IOT EN LA UNIVERSIDAD NACIONAL DE MISIONES	186
LAS COMUNICACIONES EN LAS ZONAS RURALES: POSIBLES SOLUCIONES PARA SU DESARROLLO	192
SISTEMA DE ACCESO Y AUTENTICACIÓN EN REDES DEFINIDAS POR SOFTWARE	197
UN PROTOTIPO DE PLATAFORMA DE CIUDADES INTELIGENTES	202
UN SISTEMA DE VIRTUALIZACIÓN DISTRIBUIDA	207

BASES DE DATOS Y MINERÍA DE DATOS.....	212
ANÁLISIS DE INFORMACIÓN DE REDES SOCIALES (TWITTER).....	213
ANÁLISIS Y ELABORACIÓN DE DATOS PARA EL DESARROLLO DE UN SISTEMA DE INDICADORES DE AYUDA SOCIAL.....	218
APACHE PIG EN HADOOP SOBRE CASSANDRA	223
APLICACIONES DE ONTOLOGÍAS A PROBLEMAS LINGÜÍSTICOS: BASES DE CONOCIMIENTO BASADAS EN TEXTO NO ESTRUCTURADO	227
CARACTERIZACIÓN DEL ASPIRANTE A INGRESAR A LA LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN.....	231
CONTRIBUCIONES A LAS BASES DE DATOS NO CONVENCIONALES	236
DATA VIZ EN BIBLIOTECAS UNIVERSITARIAS.....	241
DESARROLLO DE CAPACIDADES CIENTÍFICO-TECNOLÓGICAS PARA LA GESTIÓN DE DATOS MASIVOS	246
DETECCIÓN DE ATAQUES DOS CON HERRAMIENTAS DE MINERÍA DE DATOS	251
DETERMINACIÓN DEL RENDIMIENTO ACADÉMICO UNIVERSITARIO	256
EL PROCESO DE EXTRACCIÓN DE CONOCIMIENTO EN LA DETERMINACIÓN DEL PERFIL DEL AUTOR Y LA ATRIBUCIÓN DE AUTORÍA	261
EVALUACIÓN DE TÉCNICAS DE CLASIFICACIÓN PARA PREDECIR EL RENDIMIENTO ACADÉMICO DE INGRESANTES A LA UNIVERSIDAD EN TEMAS DE MATEMÁTICA.....	266
EXTENSIONES BIOMÉTRICAS PARA BASES DE DATOS OBJETO- RELACIONALES.....	271
FORMULACIÓN DE ESQUEMAS DE ALMACENAMIENTO DE DATOS MÉDICOS PARA APLICAR MINERÍA DE DATOS EN EL DIAGNÓSTICO DE ENFERMEDADES	275
GEOMETRÍA COMPUTACIONAL Y BASES DE DATOS.....	280
GRANDES DATOS Y ALGORITMOS EFICIENTES PARA APLICACIONES DE ESCALA WEB	285
INDEXACIÓN Y BÚSQUEDA SOBRE DATOS NO ESTRUCTURADOS.....	291
LA GESTIÓN DE LA INFORMACIÓN CIENTÍFICA EN ABIERTO: ¿ESTAMOS MAXIMIZANDO LA VISIBILIDAD WEB?	296
MINERÍA DE DATOS APLICADA A LA EDUCACIÓN: MODELO DE DESERCIÓN UNIVERSITARIA EN LA UNIVERSIDAD TECNOLÓGICA NACIONAL FACULTAD REGIONAL ROSARIO	301
MINERÍA DE DATOS EN RESPECTO DE LA INCIDENCIA DEL CLIMA SOBRE EL DESPERFECTO EN EL ALUMBRADO PÚBLICO.....	305
MINERÍA DE DATOS PARA ANÁLISIS DEL MICROBIOMA HUMANO	310
MINERÍA DE DATOS Y BIG DATA. APLICACIONES EN SEÑALES Y TEXTOS.....	315
MODELO DE ANÁLISIS DE INFORMACIÓN DESESTRUCTURADA UTILIZANDO TÉCNICAS DE RECOPIACIÓN Y MINERÍA WEB	320
MODELO PARA PREDECIR LA CANTIDAD DE GRADUADOS DE INGENIERÍA DE UTM APLICANDO TÉCNICAS DE MINERÍA DE DATOS	325
PERSONALIZACIÓN DE LA EDUCACIÓN A TRAVÉS DE SISTEMAS RECOMENDADORES DINÁMICOS.....	330
PROCESAMIENTO Y RECUPERACIÓN EN BASES DE DATOS MASIVAS.....	334
PROPUESTA DE PROCESOS COMPLEMENTARIOS PARA UN SISTEMA DE RECUPERACIÓN DE INFORMACIÓN	339
SELECCIÓN DE BIOMATERIALES UTILIZADOS EN IMPLANTES DENTALES APLICANDO TÉCNICAS DE MINERÍA DE DATOS.....	344
SISTEMA DE APOYO A LAS DECISIONES (DDS) PARA LA PRODUCTIVIDAD DE LAS UNIVERSIDADES: IMPLEMENTACIÓN DE TABLEROS DE CONTROL	349
TÉCNICAS DE ANÁLISIS DE INFORMACIÓN EN NEUROCIENCIAS APLICADAS	354
TECNOLOGÍAS DE PROCESAMIENTO DE DATOS MASIVOS	359
TRANSFORMANDO DATOS DE BIODIVERSIDAD EN LINKED DATA	364
UN MÉTODO PARA LA CONSTRUCCIÓN Y MANTENIMIENTO DE UNA CMDB PARA ÁREAS DE INFORMÁTICA INTERNAS DE EMPRESAS DE SERVICIOS: GESTIÓN DE FUENTES DE INFORMACIÓN	369
COMPUTACIÓN GRÁFICA, IMÁGENES Y VISUALIZACIÓN.....	374
ANÁLISIS DE CAPTURAS DE MOVIMIENTOS PARA LA ANIMACIÓN DE HUMANOS VIRTUALES	375
EL USO DE LA REALIDAD VIRTUAL INMERSIVA EN TERAPIAS MOTRICES	379
EVALUACIÓN DE ALGORITMOS DE REFOCALIZACIÓN PARA RADIANCIAS DE ARQUITECTURA PLENOPTICS 1.0.....	384
EVALUACIÓN DEL ERROR EN LA DETECCIÓN DE PUNTOS DE BORDE EN IMÁGENES SAR POLARIMÉTRICAS	389
FRAMEWORK PARA LA GENERACIÓN DE TEMPLATES EN SISTEMAS DE CATÁLOGOS DE REALIDAD AUMENTADA	393
IDENTIFICACIÓN NO SUPERVISADA DE FAUNA MARINA MEDIANTE PROCESAMIENTO DE IMÁGENES FOTOGRAFICAS	398
INTERFAZ NATURAL DE USUARIO PARA EL CONTROL DE ROBOT MÓVIL CON GESTOS FACIALES Y MOVIMIENTOS DEL ROSTRO USANDO CÁMARA RGB	403
MEDICIÓN DEL TAMAÑO DE PARTÍCULAS DE MINERALES MEDIANTE PROCESAMIENTO DIGITAL DE IMÁGENES.....	408
MODELOS DE INTERACCIÓN Y APLICACIONES EN REALIDAD VIRTUAL MEDIANTE DISPOSITIVOS MÓVILES.....	414
VISUALIZACIÓN BASADA EN SEMÁNTICA	418
VISUALIZACIÓN Y REALIDAD AUMENTADA EN EL CAMPO DE LAS CIENCIAS GEOLÓGICAS	422
INGENIERÍA DE SOFTWARE	427
ANÁLISIS DE ENFOQUES DE APLICACIONES PARA DISPOSITIVOS MÓVILES	428
ANÁLISIS DE LA USABILIDAD DE LOS SITIOS WEB DE UNIVERSIDADES: ELABORACIÓN DE UN CONJUNTO DE BUENAS PRÁCTICAS APLICABLES EN SU DISEÑO Y CONSTRUCCIÓN.....	432
ANÁLISIS DEL MODELO WQM PARA MÉTRICAS DE USABILIDAD WEB	437

APLICACIÓN DE LA LÓGICA DIFUSA EN LA EVALUACIÓN DE MODELOS CONCEPTUALES DE PROCESOS DE NEGOCIO	441
APROXIMACIONES PARA EL DESARROLLO MULTIPLATAFORMA Y MANTENIMIENTO DE APLICACIONES MÓVILES	446
ASISTENCIA DIRIGIDA POR ONTOLOGÍAS AL DISEÑO ARQUITECTÓNICO DE VIDEOJUEGOS	451
BÚSQUEDA DE ESTRATEGIAS PARA LA CLASIFICACIÓN DEL CONTENIDO EN FOROS TÉCNICOS DE DISCUSIÓN.....	456
CLASIFICACIÓN Y EVALUACIÓN DE MÉTRICAS DE MANTENIBILIDAD APLICABLES A PRODUCTOS DE SOFTWARE LIBRE	460
COORDINACIÓN DE DISPOSITIVOS EN AMBIENTES UBICUOS MEDIANTE COREOGRAFÍAS	465
DEFINICIÓN DE UNA ARQUITECTURA DE PROCESOS UTILIZANDO LA METODOLOGÍA BPTRENDS PARA LA APLICACIÓN DEL CICLO DE VIDA BPM	470
DERIVACIÓN DE CASOS DE PRUEBA A PARTIR DE ESCENARIOS	475
DESARROLLO DE APLICACIONES MÓVILES 3D	480
DESARROLLO E IMPLEMENTACIÓN DE UN MODELO BASADO EN ESTÁNDARES PARA EL LOGRO Y EVALUACIÓN DE LA USABILIDAD EN APLICACIONES INFORMÁTICAS	484
ELICITACIÓN DEL VOCABULARIO DEL CONTEXTO AYUDADA POR MAPAS CONCEPTUALES	489
ESPECIFICACIONES FORMALES TEMPRANAS DEL COMPORTAMIENTO DE SISTEMAS DE SOFTWARE	494
EVALUACIÓN DE LA ACCESIBILIDAD WEB. CASO DE ESTUDIO: SITIOS WEB DE LA UNCA.....	499
EVALUACIÓN DE LA USABILIDAD EN SITIOS WEB MÓVILES ADAPTATIVOS	504
EVALUACIÓN DE TRANSFORMACIONES DE MODELOS.....	509
EXTENDIENDO TRANSFORMACIONES MDA CON METAMODELO DE PATRONES DE DISEÑO	514
FORTEALECIMIENTO DE LA CALIDAD EN PROCESOS DE SOFTWARE Y PROCESOS DE GESTIÓN PARA LA MEJORA DE LAS SOCIEDADES DEL CONOCIMIENTO	519
GESTIÓN CUANTITATIVA DE PROYECTOS Y ENTREGA CONTINÚA EN ENTORNOS ÁGILES	525
GESTIÓN DE PROYECTO DE SOFTWARE: UN MÉTODO BASADO EN GAMIFICACIÓN PARA MEJORAR LA CALIDAD DEL PRODUCTO Y DESEMPEÑO DE EQUIPOS DE DESARROLLO.....	531
HACIA UN MODELO DE EVALUACIÓN DE CALIDAD PARA DATOS BASADO EN LA NORMA ISO/IEC 25012	536
HERRAMIENTA DE CALENDARIZACIÓN PARA PROYECTOS DESARROLLADOS UTILIZANDO XP.....	541
IMPACTO DEL FACTOR PEOPLEWARE EN LA IMPLANTACIÓN DE SISTEMAS INFORMÁTICOS	546
INGENIERÍA DE REQUISITOS PARA PROYECTOS CRM	550
INGENIERÍA DE SOFTWARE DIRIGIDA POR MODELOS APLICADA A SISTEMAS ROBÓTICOS USANDO LOS ESTÁNDARES DE LA OMG	555
INGENIERÍA DE SOFTWARE PARA SISTEMAS DISTRIBUIDOS	560
INTEGRACIÓN DE ARQUITECTURA DE SOFTWARE EN EL CICLO DE VIDA DE LAS METODOLOGÍAS ÁGILES. UNA PERSPECTIVA BASADA EN REQUISITOS	566
INTERFAZ GRÁFICA DE USUARIO: EL USUARIO COMO PROTAGONISTA DEL DISEÑO	570
LA GESTIÓN DEL CONOCIMIENTO EN PEQUEÑAS Y MEDIANAS FÁBRICAS DE SOFTWARE EN EL ÁREA METROPOLITANA DE BUENOS AIRES.....	575
¿LAS REDES SOCIALES CUMPLEN CON LOS CRITERIOS DE ACCESIBILIDAD?	580
LÍNEAS DE PRODUCTOS DE SOFTWARE: HACIA LA CREACIÓN DE TÉCNICAS Y MÉTODOS PARA MEJORAR EL REUSO	585
MODELADO DE NEGOCIOS ORIENTADO A ASPECTOS CON AOP4ST	591
MODELO DE GESTIÓN DE LA CALIDAD ORIENTADO A EMPRESAS DE SOFTWARE Y SERVICIOS INFORMÁTICOS (SSI) DE LA REPÚBLICA ARGENTINA.....	596
MODELO DE SISTEMA DE GESTIÓN DE CALIDAD PARA UN INSTITUTO DE INVESTIGACIÓN UNIVERSITARIO	601
PRÁCTICAS DE LA INGENIERÍA DE REQUERIMIENTOS EN LAS ORGANIZACIONES DE DESARROLLO	607
PROGRAMA DE I+D+I EN INGENIERÍA DE EXPLOTACIÓN DE INFORMACIÓN.....	612
PROPUESTA DE INVESTIGACIÓN: DIAGNÓSTICO E IMPACTO DE LAS TI/SI EN PYMES DE LA REGIÓN CENTRO	617
REUSO ORIENTADO A SERVICIOS: COMPATIBILIDAD Y COMPLEJIDAD DE SERVICIOS	622
SELECCIÓN DE ATRIBUTOS DE CALIDAD DE DATOS EN SISTEMAS DE GESTIÓN DE APRENDIZAJE BAJO LA FAMILIA DE NORMAS ISO/IEC 25000.....	627
SELECCIÓN DE METODOLOGÍAS ÁGILES E INTEGRACIÓN DE ARQUITECTURAS DE SOFTWARE EN EL DESARROLLO DE SISTEMAS DE INFORMACIÓN.....	632
TECNOLOGÍA MOBILE APLICADA A LAS INSTITUCIONES EDUCATIVAS	637
TECNOLOGÍA Y APLICACIONES DE SISTEMAS DE SOFTWARE: CALIDAD E INNOVACIÓN EN PROCESOS, PRODUCTOS Y SERVICIOS	642
TRATAMIENTO DE EVIDENCIAS DIGITALES FORENSES EN DISPOSITIVOS MÓVILES	648
UN ENFOQUE INTEGRADOR PARA DISEÑAR Y EVALUAR INTERFACES DE USUARIO WEB	653
UN MODELO DE CALIDAD MIXTO COMO SOPORTE A LA MEJORA DE LOS PRODUCTOS SOFTWARE CON IMPACTO EN LOS PROCESOS ORGANIZACIONALES	658
UNA PLATAFORMA TECNOLÓGICA PARA EL SEGUIMIENTO Y EVALUACIÓN DE CALIDAD DE PROYECTOS ÁGILES.....	663
INNOVACIÓN EN EDUCACIÓN EN INFORMÁTICA	667
AGENTE HORNERO. AMPLIANDO LAS POSIBILIDADES DE APRENDER A PROGRAMAR	668
APLICACIONES DE SOFTWARE CON TECNOLOGÍA NFC EN DISPOSITIVOS MÓVILES, PARA NIÑOS CON DIVERSIDAD FUNCIONAL	673
APROXIMANDO LAS CIENCIAS DE LA COMPUTACIÓN A LA ESCUELA SECUNDARIA.....	678
ARTICULACIÓN DE CONTENIDOS CURRICULARES DE LAS DISCIPLINAS SISTEMAS OPERATIVOS Y SISTEMAS DE TIEMPO REAL CON	

SISTEMAS EMBEBIDOS EN LA CARRERA INGENIERÍA EN INFORMÁTICA	683
AVANCES EN TORNO A LA FORMACIÓN EN ACCESIBILIDAD WEB	687
CATEDRA DE PROYECTO FIN DE CARRERA UNA GESTIÓN DE CALIDAD.....	691
EL PENSAMIENTO COMPUTACIONAL EN EL ÁMBITO UNIVERSITARIO	696
EXPERIMENTACIÓN EN LA ENSEÑANZA DE REQUERIMIENTOS	700
FRAMEWORK PARA LA CONSTRUCCIÓN DE ESTRATEGIAS DIDÁCTICAS PARA LA ENSEÑANZA INICIAL DE LA PROGRAMACIÓN DE COMPUTADORAS	704
GRUPO DE INVESTIGACIÓN PARA EL DESARROLLO DE VIDEOJUEGOS DENTRO DE CARRERAS DE INFORMÁTICA	709
IDENTIFICANDO FORTALEZAS Y DEBILIDADES EN LA FORMACIÓN DE GRADUADOS DE ANALISTA Y LICENCIATURA EN COMPUTACIÓN PARA MEJORAR LA ENSEÑANZA EN EL MARCO DE NUESTRAS ASIGNATURAS	713
INFLUENCIA DE LA EDUCACIÓN INVERSA EN EL APRENDIZAJE Y ADQUISICIÓN DE COMPETENCIAS TRANSVERSALES	718
LA TECNOLOGÍA COMO VEHÍCULO DE ARTICULACIÓN NIVEL MEDIO / UNIVERSIDAD	723
PAUTAS DE DISEÑO DE ACTIVIDADES FORMATIVAS PARA EL DESARROLLO DE COMPETENCIAS ESPECÍFICAS DE ORDEN SUPERIOR Y GENÉRICAS	728
PROGRAMAR EN LA ESCUELA: NUEVOS DESAFÍOS EN LAS AULAS	732
PROMOCIÓN DEL PENSAMIENTO COMPUTACIONAL PARA FAVORECER LA FORMACIÓN DE RECURSOS HUMANOS EN DISCIPLINAS STEM	737
ROBÓTICA APLICADA COMO ESTRATEGIA DE ENSEÑANZA DE PROGRAMACIÓN	743
TASA – TECNOLOGÍA ASISTIVA MEDIANTE SISTEMA ARDUINO	747
INNOVACIÓN EN SISTEMAS DE SOFTWARE.....	752
AGENTES INTELIGENTES Y WEB SEMÁNTICA: HACIA LA VERBALIZACIÓN DE UN SUBCONJUNTO DE UML EN UNA HERRAMIENTA GRÁFICA WEB.....	753
AGRICULTURA DE PRECISIÓN Y DATA-DRIVEN AGRICULTURE EN REGIONES DE CLIMAS ÁRIDOS	758
ANÁLISIS E IMPLEMENTACIÓN DE NUEVAS TECNOLOGÍAS PARA LA WEB MÓVIL.....	763
AUMENTO DE LA PRECISIÓN POSICIONAL EMPLEANDO TÉCNICAS Y ALGORITMOS PARA EL TRATAMIENTO DEL ERROR EN RECEPTORES GNSS DE BAJO COSTO	767
COMPLETITUD DE LOS MÉTODOS DE ACCESO A DATOS BASADO EN ONTOLOGÍAS: ENFOQUES, PROPIEDADES Y HERRAMIENTAS	772
CONSTRUCCIÓN DE SISTEMAS BASADOS EN REDES DE CONOCIMIENTO PARA LA GESTIÓN.....	777
DISEÑO DE SISTEMA IoT DE MONITOREO Y ALARMA PARA PERSONAS MAYORES	782
DISEÑO Y DESARROLLO DE INTERFACES CON INTERACCIÓN FÍSICA UTILIZANDO DISPOSITIVOS MÓVILES.....	785
EDUCIÓN EMOCIONAL DE UN INDIVIDUO EN CONTEXTOS MULTIMODALES EN COMPUTACIÓN AFECTIVA	790
ELABORACIÓN DE MODELOS CUANTITATIVOS PARA LA EVALUACIÓN INSTITUCIONAL	795
ESTRATEGIAS DE IoT PARA LOGRAR CIUDADES DIGITALES SEGURAS, MÁS INCLUSIVAS Y SUSTENTABLES	800
EVALUACIÓN DE SOFTWARE PARA EL DESARROLLO INDUSTRIAL	805
GENERACIÓN DE METAMODELOS ONTOLÓGICOS QUE REPRESENTEN AL ARTEFACTO SISTEMA DE INFORMACIÓN (SI)	810
HERRAMIENTA DE MODELADO Y ANÁLISIS ESTOCÁSTICO DE SISTEMAS BIOLÓGICOS	815
INFORMÁTICA Y TECNOLOGÍAS EMERGENTES.....	819
INGENIERÍA DE SOFTWARE PARA DESARROLLAR UNA PLATAFORMA ACADÉMICA PROTOTÍPICA DE eHEALTH	824
INTEROPERABILIDAD ENTRE LENGUAJES DE MODELADO CONCEPTUAL EN CROWD	828
LINKED DATA Y ONTOLOGÍAS EN UNA HERRAMIENTA GRÁFICA WEB	833
MÉTODOS Y TÉCNICAS PARA DESARROLLOS DE APLICACIONES UBICUAS.....	838
MISIÓN CUBE SAT FS2017: DESARROLLO DE SOFTWARE PARA UNA MISIÓN SATELITAL UNIVERSITARIA.....	843
PROCESAMIENTO Y ANÁLISIS DE DATOS ESPACIALES Y TEMPORALES RELATIVOS A ESPACIOS URBANOS	848
RASPBERRY PI COMO SERVIDOR PORTÁTIL DE CONTENIDOS PARA SER CONSUMIDOS DESDE DISPOSITIVOS MÓVILES	853
REALIDAD VIRTUAL Y ADIESTRAMIENTO EN SISTEMAS CRÍTICOS.....	857
REALIDAD VIRTUAL Y AUMENTADA, BIG DATA Y DISPOSITIVOS MÓVILES: APLICACIONES EN TURISMO	862
SERVICIOS DE CLOUD COMPUTING E INFORMÁTICA CONTEXTUALIZADA APLICADA A UNA APP TURÍSTICA PARA LA CIUDAD DE PUERTO IGUAZÚ MISIONES.....	868
SISTEMA AUTOMÁTICO PARA EL MONITOREO OPERACIONAL DE LOS DESMONTES USANDO IMÁGENES SATELITALES.....	873
SUBSISTEMA DE RECOLECCIÓN DE ELEMENTOS DE HARDWARE INTEGRADO A UN SISTEMA DE GESTIÓN, INVENTARIO Y MONITOREO DE HARDWARE CON ALERTAS AUTOMÁTICAS (S.R.E.H.)	878
PROCESAMIENTO DE SEÑALES Y SISTEMAS DE TIEMPO REAL	883
ALTERNATIVAS EFICIENTES PARA PROCESAMIENTO Y COMUNICACIONES EN SISTEMAS DE TIEMPO REAL.....	884
DETECCIÓN EN TIEMPO REAL DE MALEZAS A TRAVÉS DE TÉCNICAS DE VISIÓN ARTIFICIAL	889
IMPLEMENTACIONES DE GNSS RTK EN SISTEMAS EMBEBIDOS AUTÓNOMOS	894
MODELOS Y MÉTODOS COMPUTACIONALES EN INGENIERÍA (A17)	898
MONITOR DE PROFUNDIDAD ANESTÉSICA.....	902
PROCESAMIENTO DE SEÑALES VIBRO-ACÚSTICAS. ANÁLISIS DE CASOS DE ESTUDIO, MODELACIÓN, PROTOTIPADO Y EXPERIMENTACIÓN.....	905

RECONOCIMIENTO DE PATRONES Y MODELADO EN SEÑALES DE ELECTROCARDIOGRAMA: DETECCIÓN TEMPRANA DE ISQUEMIA E INFARTO DE MIOCARDIO	909
RED INALÁMBRICA DE SENSORES APLICADA A LA PREDICCIÓN DEL COMPORTAMIENTO DE INCENDIOS FORESTALES.....	914
REDES DE SENSORES INTELIGENTES PARA MONITOREO DE DATOS REMOTOS	919
SISTEMA DE CONTROL ELECTRÓNICO DE ESTABILIDAD PARA ESTUDIAR DATOS DE INTERÉS ESTRATÉGICO SOBRE CULTIVOS.....	924
SISTEMA DE MONITOREO CONTINUO DE NIVELES DE DENSIDAD DE POTENCIA ELECTROMAGNÉTICA PRESENTES EN EL MEDIO AMBIENTE	928
PROCESAMIENTO DISTRIBUIDO Y PARALELO	933
ALGORITMOS, ESTRATEGIAS Y ANÁLISIS DE ARQUITECTURAS ORIENTADOS AL MANEJO DE DATOS MASIVOS.....	934
ARQUITECTURAS MULTIPROCESADOR EN COMPUTACIÓN DE ALTO DESEMPEÑO: SOFTWARE, MÉTRICAS, MODELOS Y APLICACIONES	939
CÓMPUTO DE ALTAS PRESTACIONES. FUNDAMENTOS DE CÓMPUTO PARALELO Y DISTRIBUIDO. CONSTRUCCIÓN Y EVALUACIÓN DE APLICACIONES	945
ESTUDIO DE LAS MEJORAS DE APLICAR FOG COMPUTING EN LA DISTRIBUCIÓN DE SERVICIOS EN CLOUD COMPUTING	951
EVOLUCIÓN Y TENDENCIAS EN SISTEMAS PARALELOS Y DISTRIBUIDOS.....	956
HIVE: FRAMEWORK DE SINCRONIZACIÓN DE OBJETOS EN LA NUBE PARA SISTEMAS DISTRIBUIDOS MULTIPLATAFORMA	962
IMPLEMENTACIÓN DE PLATAFORMA HPC DINÁMICA PARA LA RESOLUCIÓN DE PROBLEMAS DE ALTA COMPLEJIDAD COMPUTACIONAL.....	965
MÉTODO DE REDUCCIÓN DE INCERTIDUMBRE BASADO EN HPC Y METAHEURÍSTICAS HÍBRIDAS APLICADO A LA PREDICCIÓN DE INCENDIOS FORESTALES.....	970
PARALELISMO HÍBRIDO APLICADO A SOLUCIONES DE PROBLEMAS DE DATOS MASIVOS	974
PERFORMANCE DE ARQUITECTURAS MULTIPROCESADOR: TÉCNICAS DE MODELADO Y SIMULACIÓN EN HPC Y CLOUD COMPUTING	979
PERFORMANCE DE CLOUD COMPUTING PARA HPC: DESPLIEGUE Y SEGURIDAD.....	984
SISTEMAS DE PROCESAMIENTO DE ALTAS PRESTACIONES RECONFIGURABLES	988
SEGURIDAD INFORMÁTICA	992
ALGORITMO DE CIFRADO PARA SISTEMAS MÓVILES	993
ANÁLISIS COMPARATIVO DE DISTINTAS TOOLKITS PARA EL RECONOCIMIENTO BIOMÉTRICO DE PERSONAS MEDIANTE VOZ.....	996
ANÁLISIS DE METODOLOGÍAS DE RECOLECCIÓN DE DATOS DIGITALES.....	1000
ANÁLISIS DE SEGURIDAD EN REDES WIRELESS UTILIZANDO DISPOSITIVOS MÓVILES	1005
ANÁLISIS DEL ESTADO DEL ARTE DE SISTEMAS DE SOPORTE PARA EL ESTUDIO DE VULNERABILIDADES EN SISTEMAS WEB	1009
ANONIMATO INCONDICIONAL EN SISTEMAS DE VOTO ELECTRÓNICO PRESENCIAL.....	1014
APLICABILIDAD DE NORMAS DE GESTIÓN DE LA CONFIGURACIÓN A MEDIANAS Y PEQUEÑAS INSTALACIONES TIC DE ORGANISMOS PÚBLICOS	1019
APROXIMACIÓN A LA SEGURIDAD DE LAS COMUNICACIONES EN INTERNET DE LAS COSAS.....	1024
ARQUITECTURA DE SEGURIDAD POR CAPAS EN SISTEMAS CRÍTICOS	1028
CONTROLES Y MÉTRICAS ASOCIADAS EN EL CONTEXTO DE LA CIBERDEFENSA	1033
CRIOGRAFÍA POST CUÁNTICA	1038
DISEÑO Y DESARROLLO DE UN PROTOTIPO DE APLICACIÓN PARA LA GESTIÓN DE LAS PERICIAS EN INFORMÁTICA FORENSE ADAPTADA AL SISTEMA JURÍDICO ARGENTINO (GEPiF).....	1043
ESTEGANOGRAFÍA SIMULADA PARA ANÁLISIS DE EFECTOS SOBRE PORTADORES IMAGEN	1048
ESTUDIO COMPARATIVO DE BUENAS PRÁCTICAS PARA LA RECOLECCIÓN DE LA EVIDENCIA DIGITAL	1053
HERRAMIENTAS DE CRIPTOANÁLISIS.....	1057
INTRODUCCIÓN AL ESTUDIO DE ALGORITMOS CRIPTOGRÁFICOS LIVIANOS	1061
LONGITUD DE LA CLAVE RSA VS PODER COMPUTACIONAL	1065
OTP-VOTE: AVANCES EN LA GENERACIÓN DE UN MODELO DE VOTO ELECTRÓNICO.....	1069
PROPUESTA DE UN MODELO DE PROCESO PARA RESOLVER VULNERABILIDADES DE SEGURIDAD EN INFRAESTRUCTURA UTILIZANDO HERRAMIENTAS DE COMPUTACIÓN COGNITIVA.....	1074
RED DE COMUNICACIÓN Y SEGURIDAD INALAMBRICA: IMPLEMENTACIÓN DEL PROGRAMA “REDCoFi, WIFI AL ALCANCE DE TODOS” EN LA FI UNLZ	1079
SEGURIDAD EN ENTORNOS BPM: FIRMA DIGITAL Y GESTIÓN DE CLAVE.....	1084
SEGURIDAD EN REDES LAS INDUSTRIALES: CLAVE PARA LA CIBERDEFENSA DE LAS INFRAESTRUCTURAS CRÍTICAS	1089
SEGURIDAD EN SERVICIOS WEB	1094
UTILIZACIÓN DE WATERMARKING PARA SEGURIDAD EN LA NUBE: EL CASO DE LAS IMÁGENES MÉDICAS.....	1099
VERIFICACIÓN DEL HABLANTE MEDIANTE DISPOSITIVOS MÓVILES EN ENTORNOS RUIDOSOS.....	1104
TECNOLOGÍA INFORMÁTICA APLICADA EN EDUCACIÓN	1108
ANÁLISIS Y DETECCIÓN TEMPRANA DE DESERCIÓN ESTUDIANTIL EN UTN FRLP	1109
ANÁLISIS Y EVALUACIÓN DE DESEMPEÑOS ESCRITOS EN ENTORNOS VIRTUALES DE APRENDIZAJE: CATEGORÍAS DISCURSIVAS Y RÚBRICA	1114

APORTES DE LAS TIC DESDE EL ENFOQUE ONTOSEMIÓTICO: ANÁLISIS DE LOS ERRORES EN LA MODELIZACIÓN DE PROBLEMAS DE PROBABILIDAD Y ESTADÍSTICA EN LA FI-UNLZ	1119
CONSTRUYENDO UN CURSO MASIVO EN LÍNEA SOBRE ACCESIBILIDAD WEB. PLANTEOS Y DESAFÍOS DE LOS NUEVOS ENTORNOS DE APRENDIZAJE	1124
DESARROLLO DE COMPETENCIAS EN CARRERAS DE INGENIERÍA CON APOYO EN TECNOLOGÍA DE SIMULACIÓN: IMPACTO EN LA SATISFACCIÓN Y LA CALIDAD PERCIBIDA POR LOS ALUMNOS	1129
DESARROLLO DE RECURSOS TIC Y MÉTODOS COMPUTACIONALES APLICADOS	1134
DESGRANAMIENTO TEMPRANO EN MATERIAS DE PRIMER AÑO EN LAS CARRERAS DE LA FACULTAD DE INGENIERÍA DE LA UNJU: EVALUACIÓN DE LA INFLUENCIA DE FACTORES COGNITIVOS	1139
DETECCIÓN AUTOMÁTICA DE CONFLICTOS GRUPALES EN ENTORNOS DE APRENDIZAJE COLABORATIVO	1144
DETECCIÓN DE PERFILES DE RENDIMIENTO ACADÉMICO EN LA UNIVERSIDAD NACIONAL DEL ESTE DE PARAGUAY.....	1149
DISEÑO DE PLATAFORMA REMOTA PARA PRÁCTICAS DE LABORATORIO.....	1154
DISEÑO Y APLICACIÓN DE ESTRATEGIAS PARA LA ENSEÑANZA INICIAL DE LA PROGRAMACIÓN.....	1159
DISPOSITIVOS MÓVILES COMO SOPORTE PARA EL APRENDIZAJE COLABORATIVO DE PROGRAMACIÓN EN EL NIVEL UNIVERSITARIO INICIAL (RESULTADOS).....	1164
EDUCACIÓN CON TECNOLOGÍAS: LA GAMIFICACIÓN APLICADA PARA EL APRENDIZAJE DE LA PROGRAMACIÓN.....	1169
EDUCACIÓN CON TECNOLOGÍAS: LA ROBÓTICA EDUCATIVA APLICADA PARA EL APRENDIZAJE DE LA PROGRAMACIÓN.....	1174
EDUCACIÓN MEDIADA POR DISPOSITIVOS MÓVILES.....	1179
EL MUSEO DE INFORMÁTICA DE LA UNPA-UARG: ORGANIZACIÓN, ACCIONES Y DIFUSIÓN.....	1184
EL USO DE LA TECNOLOGÍA PARA LA INCLUSIÓN DE LOS DISMINUIDOS VISUALES EN LAS AULAS	1189
ELICITACIÓN DE REQUERIMIENTOS, CENTRADA EN EL USUARIO, PARA EL DESARROLLO DE UN SERIOUS GAME	1194
ELICITACIÓN Y ESPECIFICACIÓN DE REQUERIMIENTOS EN PERSVASIVE SERIOUS GAMES.....	1199
ENTORNO VIRTUAL FLEXIBLE COMO APOYO AL APRENDIZAJE EN EL ÁREA DE CIENCIAS EXACTAS	1204
ESPACIOS DIALÓGICOS CON TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN (TIC) COMO PORTADORES DE CONOCIMIENTOS EN CONSTRUCCIÓN EN EL APRENDIZAJE EN CARRERAS UNIVERSITARIAS.....	1209
ESTILOS Y PARADIGMAS DE INTERACCIÓN PERSONA-ORDENADOR Y SUS POSIBILIDADES PARA EL ESCENARIO EDUCATIVO	1213
ESTUDIO DE LA INFLUENCIA DEL USO DE SIMULACIÓN EN LA ENSEÑANZA DE REDES DE COMPUTADORAS EN EL NIVEL UNIVERSITARIO (RESULTADOS FINALES)	1219
FORMACIÓN AUTOMÁTICA DE GRUPOS COLABORATIVOS CONSIDERANDO ESTILOS DE APRENDIZAJE Y RENDIMIENTO ACADÉMICO	1225
GAMIFICACIÓN EN EL ÁMBITO UNIVERSITARIO – ANÁLISIS E IMPLEMENTACIÓN DE ELEMENTOS DE JUEGOS	1230
HERRAMIENTA DE EVALUACIÓN DE HARDWARE PEDAGÓGICO PARA LA EDUCACIÓN SUPERIOR.....	1235
IMPLEMENTACIÓN DE UN EVEA INSTITUCIONAL PARA ENRIQUECER LA ENSEÑANZA DE PREGRADO, GRADO Y POSGRADO DE LA UNTDF	1240
INICIATIVA 3D EN LA ESCUELA	1246
INSTALACIÓN DE UMREMOTO3D.....	1251
JUEGOS INTERACTIVOS EN ARDUINO Y JAVA, PARA MOTIVAR Y DESPERTAR EL INTERÉS EN INFORMÁTICA.....	1256
LABORATORIO DE ESCRITORIO PARA ENSEÑANZA DE ELECTRICIDAD Y ELECTRÓNICA.....	1261
LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN EN LOS PROCESOS DE ENSEÑAR Y APRENDER.....	1266
MEJORANDO ESCENARIOS DE APRENDIZAJE COLABORATIVO SOPORTADO POR COMPUTADORA.....	1271
MINERÍA DE DATOS Y UNA APLICACIÓN EN LA EDUCACIÓN SUPERIOR.....	1276
M-LEARNING CON REALIDAD AUMENTADA	1280
MODELO PARA EL DISEÑO DE PRUEBAS PERSONALIZADAS	1285
MODELO PARA LA EVALUACIÓN DEL REPOSITORIO DE OBJETOS DE APRENDIZAJE	1290
MOOC: PROPUESTA Y MEDICIÓN DE LA CALIDAD PERCIBIDA	1295
PROCESOS DE DISEÑO DE COMPONENTES PARA ESPACIOS VIRTUALES DE TRABAJO ORIENTADOS A LA EDUCACIÓN.....	1301
PRODUCCIÓN CREATIVA DE MATERIAL EDUCATIVO DIGITAL PARA UN ESPACIO INNOVADOR DE AULA INVERTIDA.....	1306
REALIDAD AUMENTADA, REALIDAD VIRTUAL E INTERACCIÓN TANGIBLE PARA LA EDUCACIÓN	1312
REPOSITORIOS DIGITALES CON CONTENIDOS ORIENTADOS A LAS NECESIDADES DE INSTITUCIONES EDUCATIVAS DE LA REGIÓN	1317
REPOSITORIOS DIGITALES DE ACCESO ABIERTO EN LA FACULTAD DE TECNOLOGÍA Y CIENCIAS APLICADAS (UNCA)	1323
RESOLVIENDO NUMÉRICAMENTE ECUACIONES DIFERENCIALES ORDINARIAS: DISEÑO DE UN SOFTWARE EDUCATIVO	1328
SUB-REDES DE GESTIÓN Y SUS HERRAMIENTAS COMO APOYO EN UN DEPARTAMENTO PEDAGÓGICO UNIVERSITARIO	1333
TECNOLOGÍA E INNOVACIÓN EN AMBIENTES DE APRENDIZAJE: DESARROLLO Y GESTIÓN.....	1337
TECNOLOGÍAS DE CÓDIGO ABIERTO PARA LA ENSEÑANZA DE LA INGENIERÍA: HERRAMIENTAS DE VIGILANCIA TECNOLÓGICA .	1342
TECNOLOGÍAS INNOVADORAS COMO MEDIADORAS DE PROCESOS DE ENSEÑANZA Y APRENDIZAJE. EXPLORACIÓN DE HERRAMIENTAS DE REALIDAD AUMENTADA	1347
TIC MODELOS Y OBJETOS DE ENSEÑANZA SU APLICACIÓN EN CARRERAS TECNOLÓGICAS DE NIVEL SUPERIOR.....	1352
UZI: MÁQUINA VIRTUAL SOBRE PLATAFORMA ARDUINO PARA ROBÓTICA EDUCATIVA.....	1357
TESIS DOCTORALES	1361
EVALUACIÓN DE RENDIMIENTO Y EFICIENCIA ENERGÉTICA DE SISTEMAS HETEROGÉNEOS PARA BIOINFORMÁTICA	1362

<i>MADE-MLEARN: MARCO PARA EL ANÁLISIS, DISEÑO Y EVALUACIÓN DE M-LEARNING</i>	1373
PLATAFORMA DE DESARROLLO DE LABORATORIOS REMOTOS DE REDES DE SENSORES INALÁMBRICOS BASADOS EN CLOUD COMPUTING	1383

WICC 2017

Autoridades

Instituto Tecnológico de Buenos Aires

Rector

Ing. José Luis Rocés

Red de Universidades Nacionales con Carreras de Informática (RedUNCI)

Coordinador Titular

Lic. Patricia Pesado (UNLP)

Coordinador Alterno

Ing. Marcelo Estayno (UNLZ)

Junta Directiva:

Lic. Horacio Kuna (UN Misiones)

Lic. Jorge Finocchietto (U CAECE)

Mg. Osvaldo Sposito (UN LaMatanza)

Lic. Claudia Russo (UNNOBA)

Lic. Guillermo Feierherd (UNTDF)

Lic. Fernanda Carmona (UN Chilecito)

Dra. Fabiana Piccoli (UNSL)

Lic. Hugo Padovani (U Morón)

Comité Organizador Local (ITBA)

Dra. Alicia Mon

Dra. Silvia Gómez

Ing. Franco Meola

Dra. Leticia Gómez

Mg. Alejandra Buquete

Lic. Valeria Soliani

Ing. Marcela Guerrero

Srta. Marta López

Sr. Adrian Etchevarne

Comité Académico WICC 2017

Fernández Slezak Diego	(UBA – Cs. Exactas)
Echeverria, Adriana	(UBA – Ingeniería)
Pesado, Patricia	(UN La Plata)
Rueda Sonia	(UN Sur)
Piccoli, Fabiana	(UN San Luis)
Aciti Claudio	(UNCPBA)
Balladini, Javier	(UN Comahue)
Spositto, Osvaldo	(UN La Matanza)
Alfonso, Hugo	(UN La Pampa)
Estayno, Marcelo	(UN Lomas de Zamora)
Feierherd, Guillermo	(UNTierra del Fuego)
Gil, Gustavo	(UN Salta)
Márquez, María Eugenia	(UN Patagonia Austral)
Otazú, Alejandra	(UN SanJuan)
Aranguren, Silvia	(UADER)
Buckle, Carlos	(UN Patagonia SJB)
Tugnarelli, Mónica	(UN Entre Ríos)
Dapozo, Gladys	(UN Nordeste)
Kantor Raul	(UN Rosario)
Kuna, Horacio	(UN Misiones)
Russo, Claudia	(UNNOBA)
Carmona, Fernanda	(UN Chilecito)
García Martínez, Ramón	(UN Lanús)
Duran Elena	(UN Santiago del Estero)
Castro Lechstaler Antonio	(Esc. Sup. Ejército)
Loyarte, Horacio	(UN Litoral)
Arroyo, Marcelo	(UN RioIV)
Fridlender, Daniel	(UN Córdoba)
Herrera Cognetta, Analía	(UN Jujuy)
Vivas, Luis	(UN Rio Negro)
Prato, Laura	(UN Villa María)
Scucimarri, Jorge	(UN Lujan)
Barrera, María Alejandra	(UN Catamarca)
Campazzo Eduardo	(UN La Rioja)
Oliveros, Alejandro	(UNTres de Febrero)
Luccioni, Griselda María	(UN Tucumán)
Morales, Martín	(UNAJ)
Zachman Patricia	(UN Chaco Austral)
Foti, Antonio	(UN del Oeste)
Forradelas, Raymundo	(UN de Cuyo)

Doumecq, Julio Cesar	(UNdeMardelPlata)
Díaz-Caro, Alejandro	(UNdeQuilmes)
Padovani Hugo	(UMorón)
De Vincenzi, Marcelo	(UAI)
Guerci, Alberto	(UBelgrano)
Panizzi, Marisa	(Ukennedy)
Bournissen Juan	(UAdventistadelPlata)
Finocchietto, Jorge	(UCAECE)
Alvarez Adriana	(UPalermo)
Grieco, Sebastián	(UCARosario)
Zanitti, Marcelo	(USalvador)
Giménez, Rosa	(UAconcagua)
Beyersdorf, Carlos	(UGastónDachary)
Guglianone, Ariadna	(UCEMA)
Cassol, Ignacio	(UAustral)
Rathmann, Liliana	(UAtlántida Argentina)
Bertone, Rodolfo	(UCALaPlata)
Gomez, Silvia	(ITBA)
Pincirolí, Fernando	(UChampagnat)

Coordinadores de Área WICC 2017

Agentes y Sistemas Inteligentes

Marcelo Falappa (UNS)
Marcelo Errecalde (UNSL)
Daniel Pandolfi (UNPA)

Arquitectura, Redes y Sistemas Operativos

Luis Marrone (UNLP)
Hugo Padovani (UMorón)
Daniel Arias Figueroa (UN Salta)

Bases de Datos y Minería de Datos

Laura Lanzarini (UNLP)
Claudia Deco (UNR)
Norma Herrera (UNSL)

Innovación en Educación Informática

Jorge Finocchietto (UCAECE)
Claudia Russo (UNNOBA)
Elena Durán (UNSE)

Computación Gráfica, Imágenes y Visualización

Martín Larrea (UNS)
María J. Abásolo (UNLP-UNCPBA)
Roberto Guerrero (UNSL)

Ingeniería de Software

Pablo Fillotrani (UNSur)
Pablo Thomas (UNLP)
Fernanda Carmona (UNChilecito)

Innovación en Sistemas de Software

Marcelo Estayno (UNLZ)
Guillermo Feierherd (UNTDF)
Osvaldo Sposito (UNLaM)
Gladys Dapozo (UNNE)

Procesamiento de Señales y Sistemas de Tiempo Real

Oscar Bría (INVAP)
Fernando Tinetti (UNLP)
Nelson Rodriguez (UNSJ)

Procesamiento Distribuido y Paralelo

Marcelo Naiouf (UNLP)
Fabiana Piccoli (UNSL)
Javier Ballardini (UNCOMA)

Tecnología Informática aplicada en Educación

Zulma Cataldi (UBA-UTN)
Alejandra Zangara(UNLP)
Mónica Tugnarelli (UNER)
Gustavo Gil (UNSalta)

Seguridad Informática

Paula Venosa (UNLP)
Javier Echaiz (UNSur)
Antonio Castro Lechtaller (IESE)

Jurado Tesis de Doctorado

Marcela Printista (UNSL)
Laura De Giusti (UNLP)
Silvia Castro (UNS)
Alejandra Cechich (UNCOMA)
Horacio Kuna (UNaM)
Orlando Micolini (UNC)

**Agentes y
Sistemas Inteligentes**

Agentes Inteligentes y Web Semántica: Preprocesamiento de Texto de Redes Sociales

Angel Addati

Sandra Roger

email: angeladdati@gmail.com , roger@fi.uncoma.edu.ar

Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

RESUMEN

El proyecto de investigación Agentes Inteligentes y Web Semántica, financiado por la Universidad Nacional del Comahue, tiene como objetivo general la generación de conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la Web, es decir, lo que se ha llamado la Web Semántica.

En particular, esta línea de investigación tiene como objetivo el desarrollo de una herramienta genérica y expansible para el preprocesamiento de lenguaje natural, para la normalización de texto. Para ello se ha definido su arquitectura básica, brindando la posibilidad de agregar nuevas funcionalidades de forma sencilla.

Palabras Clave: *Normalización de texto, Minería de opinión, Procesamiento de lenguaje natural, Preprocesamiento de texto, Tokenización de texto,*

CONTEXTO

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el contexto del proyecto de investigación Agentes Inteligentes y Web Semántica, en el contexto de una beca interna doctoral. El proyecto de investigación tiene prevista una duración de cuatro años,

desde enero del 2017 hasta diciembre del 2020.

1. INTRODUCCIÓN

Los cambios tecnológicos y el desarrollo de una cultura de las redes sociales, ha llevado a contar con bases de conocimientos globales de gran tamaño conformadas por las expresiones escritas de usuarios: sentimientos, gustos, personalidades, opiniones, tristezas, rencores, etc. Es decir, una base de datos enorme de textos de opinión.

El tema de estudio aparece naturalmente al investigar la literatura sobre minería de opinión. En estos desarrollos se utiliza generalmente como recursos los textos en redes sociales. Estos textos poseen sus propias particularidades de escritura. El trabajo estará orientado en esta dirección, abordando el preprocesamiento de texto por ser la primera etapa necesaria en la minería de opinión que se enfrenta a las dificultades del lenguaje en la redes sociales. Este preprocesamiento tiene como entrada el texto a analizar en lenguaje natural y genera como salida el texto ajustado para que sea más fácil de procesar.

En muchos de los trabajos de investigación, se invierte tiempo al estudio y desarrollo de esta primera etapa de preprocesamiento. Esta etapa es de vital importancia, ya que el texto proveniente de las redes sociales es de muy baja calidad: “*El crecimiento de los medios sociales ha ocasionado que la lingüística computacional se*

encuentre en contacto muy cercano con el lenguaje malo: texto que desafía nuestras expectativas sobre el vocabulario, el habla y la sintaxis” [1].

La ausencia de una herramienta de facto para el preprocesamiento y la carencia de una estandarización en esto, conlleva a que en cada desarrollo de minería de opinión se vuelva a programar un preprocesador que ataque los mismos temas, con distintas aproximaciones y eficiencias. Luego, tenemos distintas soluciones que resuelven con distintos grados de éxitos la misma problemática.

En la literatura existente se pueden apreciar algunos de los diferentes enfoques de cómo se ha abordado esta problemática

En [6,7] se realiza el preproceso de los tuits para tratar el uso de particular del lenguaje en la red como:

- ☐ Tratamiento de emoticones, cada emoticon se sustituye por una de estas cinco etiquetas: muy positivo, positivo, neutro, negativo y muy negativo.
- ☐ Normalización de URL's: Las direcciones son sustituidas por URL.
- ☐ Corrección de abreviaturas más frecuentes, sustituyéndolas por su forma reconocida.
- ☐ Normalización de risas.
- ☐ Tratamiento de elementos específicos de Twitter (#, @).

Por otro lado, en [12,8,9] se realizan un sistema para la resolución de la tarea de normalización de tweets basado en la concatenación de varios traductores. El trabajo describe un sistema de normalización de tweets en español. Buscando simplicidad y flexibilidad, su arquitectura es de pipeline ya que permite integrar, eliminar e intercambiar módulos de forma sencilla. Dichos módulos se comunican empleando un formato de representación intermedia codificado como texto de naturaleza estructurada y jerarquizada. Un tweet está

formado por términos y para cada término existe una serie de candidatos para su normalización.

En [11] presenta una aproximación lingüística basada en traductores de estados finitos para la normalización léxica de mensajes de Twitter en español. El desarrollo de esta propuesta consiste de traductores que son aplicados a token fuera del vocabulario. Los traductores implementan modelos lingüísticos que genera un conjunto de candidatos. Mediante un modelo estadístico se obtiene la secuencia de palabra más probable.

En [13] se utiliza un clasificador para detectar palabras mal formadas y generar posteriormente, candidatos de corrección basados en su similitud morfológica. Para la selección del candidato, se realiza lo siguiente: sobre la corrección más probable de la palabra se utiliza tanto la similitud de palabras como el contexto.

Apoyándonos en las investigaciones existentes, dentro del campo del procesamiento de lenguaje natural (PLN), vamos a definir el preprocesamiento como la primera transformación del texto de entrada. Su finalidad es generar un texto de mejor calidad para las etapas posteriores de trabajo. Por un texto de mejor calidad, entendemos a un texto segmentado, corregido (o con la menor cantidad posible de errores de entrada) y optimizado.

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación Agentes Inteligentes y Web Semántica tiene como objetivo general generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica.

Específicamente, esta línea se centra en el estudio de un sistema multiagente para la mejora en la calidad de los textos de opinión pública sobre un determinado tema de interés,

más precisamente sobre textos escritos en tweets.



Figura 1: Arquitectura básica del preprocesador

El trabajo se centrará en dos ejes: uno conceptual y otro práctico. Primero, se intentará definir cuál es la funcionalidad que debe tener el preprocesador. Aquí el foco estará en delinear la funcionalidad sin importar el uso que se le dará con posterioridad. Es decir, hallar una estandarización de lo que debe hacer el preprocesador. Segundo, se desarrollará una herramienta que pueda ser utilizada por

cualquier investigador para preprocesar sus datos. Para ello debe cubrir con lo especificado y debe tener la flexibilidad para ser expandida con nuevas características.

Basándonos en las distintas soluciones implementadas en otras investigaciones, el preprocesador se dividió en la ejecución de 6 etapas progresivas. En cada una de las etapas, se abordará una agrupación de problemas comunes que este debe resolver. La etapa contendrá módulos en donde estará la lógica de resolución de un problema. Una vez que el texto de entrada pasa por cada una de las etapas, se lo dará por preprocesado. La Figura 1 muestra la arquitectura básica para la normalización de texto planteada.

Cada etapa contiene sus propios desafíos, estrategias, definiciones e investigaciones asociadas. A continuación, se enumeran cada una de las etapas y se dará una breve descripción de las mismas:

ETAPA 1: Tokenización.

Aquí entrará el texto informal en formato plano y saldrá segmentado en tokens. El token será una palabra o un signo de puntuación. Para ello, es importante una correcta definición de los signos de puntuación y espacios para una apropiada separación en palabras

ETAPA 2: Clasificación en IV y OOV.

A esta fase entrará el texto segmentado en palabras y saldrán las palabras clasificadas en IV y OOV. Por IV se significa “*In Vocabulary*” cuya traducción es dentro del lenguaje. Las palabras dentro del lenguaje serán las palabras consideradas válidas, correctas y sin errores. Es decir, las palabras pertenecientes a nuestro lenguaje, bien escritas y reconocidas por los diccionarios. Por otro lado, OOV significa “*Out Of Vocabulary*”. La traducción es fuera del lenguaje. Son las palabras desconocidas o no reconocidas como co-

rrectas. Estas palabras son las candidatas a corregir si es posible.

ETAPA 3: Normalización.

Aquí entrará el texto segmentado en palabras clasificadas en IV y OOV. Por cada palabra OOV, se intentará corregirla y encontrar su palabra IV semánticamente equivalente. Aquí se aplicará diferentes heurísticas con el fin de buscar palabras IV candidatas para cada OOV. No necesariamente toda OOV tendrá una IV candidata, y los candidatos para una OOV pueden ser más de uno. Al finalizar, al texto segmentado en palabras clasificadas en IV y OOV, se le agregarán las palabras IV candidatas.

ETAPA 4: Selección.

En esta fase se deberá seleccionar por cada OOV una de las palabras IV candidatas. El algoritmo de selección podrá tener varios criterios, por ejemplo: el orden, la probabilidad, algún peso asociado al método aplicado para encontrar la palabra candidata, etc. La salida del proceso será la entrada agregando la palabra IV candidata elegida.

ETAPA 5: Capitalización.

Esta fase generará como salida la apropiada capitalización de todas las palabras, tanto las corregidas en la etapa anterior como las no corregidas. En los mensajes escritos en los medios sociales, la correcta capitalización es un problema, ya que los usuarios cometen errores, no respetan ortografías y utilizan la mayúscula para expresar sentimientos o énfasis. La finalidad de esta etapa es aplicar un criterio estandarizado al uso de mayúsculas y minúsculas.

ETAPA 6: Optimización.

Esta es probablemente la etapa más difusa de las 6 y la más compleja de acotar. Por optimización se va a entender a toda transformación de la entrada, no incluida en etapas anteriores, que la simplifique. Luego, aquí el foco estará en definir módulos con optimizaciones básicas o comunes y brindarle al desarrollador un marco de trabajo sencillo para expandir la funcionalidad. Algunos ejemplos de optimización podrán ser: la eliminación de palabras con poco contenido semántico, la lematización de palabras, la eliminación de signos de puntuación, la eliminación de palabras repetidas, etc.

3. RESULTADOS OBTENIDOS Y TRABAJOS FUTUROS

Inicialmente, se hizo un relevamiento de las diferentes estrategias de normalización de texto para poder crear un marco comparativo y poder evaluarlas. Del mismo surgió la arquitectura planteada anteriormente.

Se está trabajando en terminar de desarrollar el diseño profundizando el detalle de cada etapa para estudiar más a fondo cada problemática en particular. El objetivo es crear un preprocesador más potente y configurable para el usuario.

Además, se plantea generar un software funcional dedicado exclusivamente a la tarea de preprocesamiento de lenguaje natural orientado a textos escritos en redes sociales.

Finalmente, teniendo como partida el preprocesador desarrollado como primer componente, continuar en el desarrollo de los componentes subsiguientes para tener una iteración completa de un software de análisis de opinión.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante la realización de este sistema se espera lograr, como mínimo, la culminación de 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto.

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

5. BIBLIOGRAFÍA

- [1] Eisenstein, J. (2013). *What to do about bad language on the internet*. Atlanta, Georgia: Association for Computational Linguistics.
- [2] Graña Gil, J., Rodríguez, B., Mario, F., & Vilares Ferro, J. (2001). *Etiquetación robusta del lenguaje natural: preprocesamiento y segmentación*. Sociedad Española para el Procesamiento del Lenguaje Natural.
- [3] Padilla, A. P. (2004). *Técnicas lingüísticas aplicadas a la búsqueda textual multilingüe: ambigüedad, variación terminológica y multilingüismo*. Sociedad Española para el Procesamiento del Lenguaje Natural.
- [4] Ramírez Bustamante, F., Sánchez León, F., & Declerck, T. (1997). *Corrección gramatical y preprocesamiento*. Sociedad Española para el Procesamiento del Lenguaje Natural.
- [5] Dubiau, L., & Ale, J. M. (2013). *Análisis de Sentimientos sobre un Corpus en Español: Experimentación con un Caso de Estudio*. 14th Argentine Symposium on Artificial Intelligence, ASAI2013.
- [6] Vilares, D., Alonso, M. A., Gómez-Rodríguez, Carlos. *Una aproximación supervisada para la minería de opiniones sobre tuits en español en base a conocimiento lingüístico*. Procesamiento del Lenguaje Natural, Vol. 51. 2013, pp 127-134.
- [7] Vilares, J., Alonso, M. A., & Vilares, D. (2013). Prototipado Rápido de un Sistema de Normalización de Tuits: Una Aproximación Léxica.
- [8] Alegria, I., Aranberri, N., Fresno, V., Gamallo, P., Padró, L., Vicente, I. S., . . . Zubiaga, A. (2013). Introducción a la Tarea Compartida Tweet-Norm 2013: Normalización Léxica de Tuits en Español.
- [9] Alegria, I., Etxeberria, I., & Labaka, G. (2013). Prototipado Rápido de un Sistema de Normalización de Tuits: Una Aproximación Léxica. CEUR-WS.org.
- [10] Orquín, A. F., Rodríguez, K. V., Amable, A. C., Martín, R. P., Echarte, Á. L., & Morera, D. C. (2009). *Sistema para el pre-procesamiento de textos para el Procesamiento del Lenguaje Natural*.
- [11] Porta, J., & Sancho, J. L. (2013). *Word normalization in Twitter using finite-state transducers*.
- [12] Alegria, I., Etxeberria, I., Labaka, G. *Una Cascada de Transductores Simples para Normalizar Tweets*. CEUR Workshop Proceedings. Vol. 1086, 2013. pp. 15-19.
- [13] Han, B, Baldwin, T. *Lexical Normalisation of Short Text Messages: Makn Sens a #Twitter*. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies – Vol. 1. 2011 pp. 368-378. Association for Computational Linguistics

Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación

Marcelo Amaolo¹ Daniel Dolz¹ Guillermo Grosso¹ Pablo Kogan¹
 María Eda Cornejo² Sonia Sommer² Jorge Rodríguez¹ Gerardo Parra¹

email: {marcelo.amaolo, ddolz, guillermo.grosso, pablo.kogan}@fi.uncoma.edu.ar
 medacornejo@gmail.com, soniasommer@yahoo.com
 {j.rodri, gparra}@fi.uncoma.edu.ar

¹ *Grupo de Investigación en Lenguajes e Inteligencia Artificial*
 Departamento de Teoría de la Computación - Facultad de Informática
 UNIVERSIDAD NACIONAL DEL COMAHUE

² *Consejo Provincial de Educación*
 MINISTERIO DE EDUCACIÓN DE LA PROVINCIA DEL NEUQUÉN

Resumen

En este trabajo se presenta el proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación*. El proyecto tiene como objetivo general el estudio y desarrollo de técnicas de Inteligencia Artificial para dotar de inteligencia, conocimiento y capacidades cognitivas a agentes inmersos en ambientes complejos.

Asimismo, el proyecto se abocará al estudio de modelos formales buscando identificar posibles aplicaciones en el contexto educativo. En este sentido, se procurará desarrollar modelos y categorías que contribuyan a la producción del marco teórico que se ocupa de estudiar la inclusión de la computación en la educación, asumiendo este campo de conocimiento como disciplina teórica en construcción.

En el proyecto convergen diferentes líneas de investigación en el contexto de los Agentes Inteligentes, que articuladas describen el objeto de estudio y definen las

actividades de investigación: modelos formales e identificación de posibles aplicaciones al campo educativo.

Palabras Clave: Inteligencia Artificial, Agentes Inteligentes, Lenguajes Formales, Aplicaciones en Educación, Soporte al Medio Académico, Enseñanza de las Ciencias de la Computación.

Contexto

Esta propuesta se presenta en el ámbito de las iniciativas desarrolladas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática y en el contexto del Convenio Marco de Colaboración firmado entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén, durante 2016. El convenio persigue el objetivo de contribuir recíprocamente al desarrollo de actividades de investigación, formación de recursos humanos y

promoción de la enseñanza de las Ciencias de la Computación durante la escolaridad obligatoria.

Este trabajo se desarrolla en el marco del proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica y por el Consejo Provincial de Educación en el contexto del Convenio Marco de Colaboración. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

Las actividades concretadas en el ámbito de la investigación se plantean como articuladas a un conjunto de proyectos de extensión [3, 5, 10], en ejecución durante 2017, con intención de construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el campo de la praxis.

1. Introducción

Los agentes inteligentes son piezas de software que pueden ejecutarse sin control humano para cumplir metas provistas por los usuarios [11, 9, 4, 7]. Un agente gestiona, procesa y accede a información de su ambiente e interactúa, posiblemente, con otros agentes [8, 13]. En este contexto, los modelos y lenguajes formales, proveen mecanismos de abstracción y de representación de dicha información [2, 14]

Asimismo, los lenguajes formales subyacentes y las técnicas de aprendizaje y razonamiento proveen formas avanzadas para manipular la información y extraer nuevo conocimiento, definiendo así un ambiente adecuado para resolver problemas reales y, por ende, complejos.

En el proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación* se estudiarán cuestiones fundacionales de la teoría y la construcción de aplicaciones basadas en Agentes Inteligentes inmersos en ambientes dinámicos.

Además, el proyecto tiene como objetivo identificar posibles campos de aplicación en el ámbito de la educación y avanzar en el desarrollo de modelos y agentes inteligentes que apoyen los procesos de enseñanza de la computación en diferentes contextos educativos.

En este sentido, se procurará desarrollar modelos y categorías que contribuyan a la producción del marco teórico que se ocupa de estudiar la inclusión de la computación en la educación, asumiendo este campo de conocimiento como disciplina teórica en construcción.

En este ámbito se presentan escenarios que describen nuevas instancias educativas que requieren del desarrollo de modelos conceptuales y herramientas tecnológicas. Antecedentes próximos en esta línea de investigación y desarrollo son el Agente Hornero construido, en el marco del proyecto de investigación anterior, para apoyar los procesos de enseñanza y de aprendizaje de la programación a partir de los torneos de programación [5].

Además confluyen en este sentido los enfoques metodológicos diseñados y puestos en ejecución en el marco de los proyectos de extensión *Agentes Robots*, que proponen el aprendizaje de la programación a partir de la construcción colaborativa de Agentes Autónomos [3, 10].

La formación en conceptos fundamentales de las Ciencias de la Computación en el ámbito de la educación obligatoria es reconocida como prioritaria en forma creciente[1, 12]. Sin embargo, la presencia en forma sostenida y rigurosa es aún un proceso en desarrollo en la mayoría de los sistemas educativos[6].

El contexto descrito pone en evidencia la necesidad de desarrollar líneas de investigación y desarrollo en el campo de los agentes inteligentes y modelos formales aplicados a la educación. De esta manera contribuir a la elaboración de construcciones teóricas y tecnológicas que apoyen los procesos de inserción de las Ciencias de la Computación a la educación obligatoria.

2. Líneas de Investigación

Este proyecto plantea la convergencia de al menos dos líneas de investigación en relación al estudio de modelos formales y la identificación de posibles aplicaciones en el campo educativo. En principio, se proyecta abordar las siguientes líneas, las que articuladas describen el objeto de estudio y definen las actividades de investigación:

- Agentes Inteligentes. Modelos Formales.
- Agentes Inteligentes. Identificación de posibles aplicaciones al campo educativo.

Los modelos y lenguajes formales proveen mecanismos de abstracción y de representación de la información necesaria para que un agente pueda interactuar de manera adecuada con el ambiente que lo circunda. La primer línea del proyecto se abocará al estudio de los fundamentos y la teoría de agentes inteligentes y de la construcción de aplicaciones basadas en

tales agentes. Asimismo, se examinarán y propondrán modelos formales relevantes y adecuados para esta tarea.

La segunda línea se ocupará de la identificación de dominios específicos en el ámbito de la educación susceptibles a ser tratados desde el cuerpo de conocimiento desarrollado en el campo de los agentes inteligentes y los modelos formales. Por otro lado se buscará avanzar en el desarrollo de agentes que contribuyan al mejoramiento de los procesos de enseñanza y aprendizaje de las Ciencias de la Computación en el contexto de la educación obligatoria.

3. Resultados Esperados

En el contexto de las actividades proyectadas se espera:

- Desarrollar agentes inteligentes con capacidades cognitivas, de manera que puedan actuar en forma autónoma y sin intervención humana.
- Formalizar el enfoque como un agente de información inteligente, cuya meta es gestionar, procesar y acceder información.
- Analizar las propiedades computacionales de los modelos formales, a fin de evaluar su utilización en la implementación de agentes inteligentes que se desenvuelven en tiempo real.
- Utilizar las formalizaciones teóricas desarrolladas en la resolución de problemas estándares conocidos en la literatura de Inteligencia Artificial y de problemas simples del mundo real.
- Avanzar en la identificación y descripción de los elementos que permitan demarcar el campo de la computación escolar y su relación con

otros campos de conocimiento.

- Diseñar modelos conceptuales para la instanciación del aprendizaje colaborativo a campos problemáticos específicos.
- Desarrollar e implementar prototipos de agentes inteligentes que soporten los modelos conceptuales para el aprendizaje colaborativo.

4. Actividades de Extensión relacionadas al Proyecto

El proyecto busca construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el ámbito de la Extensión Universitaria en relación a la aplicación de agentes al contexto educativo. Las iniciativas de referencia se expresan en los siguientes Proyectos de Extensión:

- Torneos de Programación para promover el aprendizaje en la Escuela Media. Año 2017, avalado por Resolución FaI 086/16 [5].
- Vamos a la Escuela. Acercando las Ciencias de la Computación a la Escuela Media. Año 2017, avalado por Resolución FaI 087/16 [10].
- Agentes Robots: Divulgando Computación en la Escuela Media". Año 2017, avalado por Resolución FaI 088/16 [3].

5. Formación de Recursos Humanos

La ejecución del proyecto de investigación brindará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. Asimismo, será un ámbito adecuado para la realización de tesis de grado de la

carrera de Licenciatura en Ciencias de la Computación.

Se espera, además, la consolidación como investigadores de los miembros más recientes del grupo de investigación. El área científico-tecnológica principalmente impactada por la formación de recursos humanos en el marco de este proyecto será el área de Tecnologías de la Información y de las Comunicaciones.

Referencias

- [1] ACM and CSTA. *K-12 Computer Science Framework..* 2016.
- [2] G. Bel-Enguix, M. D. Jiménez- López, and C. Martín-Vide, editors. *New Developments in Formal Languages and Applications.* 2008.
- [3] L. Cecchi and G. Grosso. *Proyecto de Extensión Agentes Robots: Divulgando Computación en la Escuela Media.* 2017, avalado por Resolución FaI 088/16.
- [4] R. Feldmann, G. Brewka, and S. Wenzel. Planning with Prioritized Goals. In *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning*, 2006.
- [5] I. Godoy and P. Kogan. *Torneos de Programación para promover el aprendizaje en la Escuela Media.* 2017, avalado por Resolución FaI 086/16.
- [6] Google and Gallup. Trends in the state of computer science in u.s. k-12 schools. 2016.
- [7] Hanheide, Marc and Hawes, Nick and Wyatt, Jeremy and Gibelbecker,

- Moritz and Brenner, Michael and Sjöo, Kristoffer and Aydemir, Alper and Jensfelt, Patric and Zender, Hendrik and Kruijff, Geert-Jan. A Framework for Goal Generation and Management. 2010.
- [8] M. Huhns and L. Stephens. Multiagent Systems and Societies of Agents. In G. Weiss, editor, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, pages 79-120. The MIT Press, Cambridge, Massachusetts, 1999.
- [9] D. Poole, A. Mackworth, and R. Goebel. *Computational Intelligence: A Logical Approach*. Oxford University Press, 1998.
- [10] J. Rodríguez and R. Zurita. *Proyecto de Extensión Vamos a la Escuela: Acercando las Ciencias de la Computación a la Escuela Media*. 2017, avalado por Resolución Fal 087/16.
- [11] S. Russell and P. Norvig. *Artificial Intelligence: A modern approach*. Prentice Hall, New Jersey, third edition, 2009.
- [12] M. Smith. Computer science for all. *The White House*, 2016.
- [13] G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, Cambridge, Massachusetts, 1999.
- [14] Z. Ésik, C. Martin-Vide, and V. Mitran, editors. *Recent Advances in Formal Languages and Applications*. 2006.

Aplicación de Técnicas Aprendizaje Automático para Estimar la Calidad de la Voz en Escala GRBAS

García Mario Alejandro¹, Rosset Ana Lorena², Moyano Miguel Alejandro¹, Ramírez Héctor Emilio¹, Melgralejo Samara Ofelia¹, Carrillo Florencia Noel¹

¹Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN FRC)

²Universidad Nacional de Córdoba (UNC)

RESUMEN

La valoración de la calidad vocal mediante el análisis audio-perceptual es parte de la rutina clínica de evaluación de pacientes con trastornos de la voz. La debilidad de este método reside en la subjetividad y en la necesidad de que sea realizada por oyentes experimentados. Este proyecto tiene como objetivo particular la realización de una clasificación automática de la calidad vocal, valuada en la escala GRBAS, a través de características extraídas del análisis acústico de la señal y técnicas de aprendizaje automático.

Palabras clave: *machine learning, deep learning, voice quality, GRBAS*

CONTEXTO

Este trabajo de investigación se desarrolla en el marco del proyecto “Análisis acústico de la voz con técnicas de aprendizaje automático” (UTN3947) de la Universidad Tecnológica Nacional, Facultad Regional Córdoba y cuenta con la colaboración del Departamento de Investigación Científica, Extensión y Capacitación "Raquel Maurette", Escuela de Fonoaudiología, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba.

1. INTRODUCCIÓN

Se intenta reconocer, de forma automática, características del análisis acústico de la voz que permitan clasificar muestras de audio. El estudio se enfoca en la medición de la calidad vocal según la escala GRBAS. La clasificación se realizará aplicando principalmente modelos de deep learning, un subgrupo de técnicas del campo de aprendizaje automático (machine learning). Las grabaciones de la voz, la clasificación de los ejemplos y la validación de los resultados se realizarán por especialistas en análisis de la voz de la Escuela de Fonoaudiología de la Universidad Nacional de Córdoba. El análisis acústico se realizará en conjunto (especialistas vocales e integrantes de UTN) y el modelado y desarrollo de los clasificadores por los integrantes de UTN.

GRBAS: La escala GRBAS es un método de valoración perceptivo- auditivo de la voz. Surge de la necesidad de estandarizar la valoración subjetiva y de interrelacionar los aspectos auditivos y fisiológicos de la producción vocal. Está basada en estudios del año 1966 de la *Japan Society of Logopedics and Phoniatrics* [1] y posteriormente divulgada y descripta por Minoru Hirano en el año 1981 [2]. Consiste en la valoración de la fuente glótica a través de 5 parámetros que forman el acrónimo GRBAS:

G: (Grade) Grado general de disfonía R: (*Roughness*) Rugosidad, irregularidad de la onda glótica.

B: (*Breathiness*) Soplosidad, sensación de escape de aire en la voz.

A: (*Astheny*) Astenia, pérdida de potencia.

S: (*Strain*) Tensión, sensación de hiperfunción vocal.

Puede valorarse de dos maneras: a través de 4 grados, desde el 0 al 3 o mediante un valor en un rango continuo de 0 a 100. En ambas el 0 es ausencia de disfonía y el 3 o 100 implican disfonía severa. La escala fue mundialmente adoptada y validada en numerosos países [3], [4], [5], [6]. Actualmente se utiliza en la investigación y de manera rutinaria en los consultorios de los profesionales que hacen clínica vocal. Sirve como metodología simple y al alcance de la mano para valorar la evolución pre- post tratamiento. La debilidad de este método reside en la subjetividad de la valoración de la voz y en la necesidad de que sea realizada por oyentes experimentados en la escucha y la disociación de los parámetros [7], [8].

Análisis acústico: Existen otras formas de analizar la voz de manera más objetiva a través del análisis acústico. Éste consiste en la digitalización de la señal vocal y su análisis mediante gráficos como el Espectrograma, el espectro FFT (*Fast Fourier Transform*) o LPC (*Linear Predictive Coding*) y medidas numéricas de perturbación de la señal, como Jitter, Shimmer y HNR (*Harmonics to Noise Ratio*). El análisis acústico, a pesar de ser más objetivo, siempre necesita de la intervención del evaluador y esto es lo que también le impone una cierta subjetividad. Por este motivo surge la necesidad de la estandarización de todos los pasos en los que el sujeto evaluador va a intervenir: elección

del material de habla que se graba, grabación, elección de los análisis a realizar (no se puede analizar todas las voces, más o menos disfónicas, con las mismas medidas) y finalmente, el análisis físico- acústico y fisiológico de los datos obtenidos. Para lograr una integración de la valoración subjetiva (GRBAS u otras escalas) con el análisis acústico, se han realizado numerosos trabajos de correlación [9], [10], algunos relacionados a la voz normal y otros a diferentes patologías. Por ejemplo, el trabajo de Nuñez Batalla, F. et al [11] es un referente y establece una relación entre el parámetro de Astenia del GRBAS y el Espectrograma de banda angosta.

Aprendizaje automático: El aprendizaje automático o *machine learning* es un campo de las ciencias de la computación que abarca el estudio y la construcción de algoritmos capaces de aprender y hacer predicciones. Estas predicciones se pueden tomar como una clasificación de los datos de entrada a partir del reconocimiento de patrones existentes en los mismos. Existen varios enfoques de *machine learning*. Estos difieren en el objetivo, tipo de entrenamiento, inspiración (por ejemplo matemática, estadística, biológica, etc.), eficiencia y complejidad entre otras características. Algunos de estos enfoques son redes neuronales artificiales, reglas de asociación, máquinas de vectores de soporte (*support vector machines*), árboles de decisión, redes bayesianas y análisis de *clusters*. *Deep Learning* es una rama del aprendizaje automático. Está compuesto por un grupo de algoritmos que intentan clasificar los datos en abstracciones de alto nivel mediante el uso de estructuras jerárquicas complejas. Algunas de las técnicas son *deep neural networks*, *convolutional neural networks* y *deep belief networks*.

Aprendizaje automático y análisis acústico:

Los trabajos más importantes de los últimos años sobre *machine learning* y análisis acústico tienen como objetivo reconocer lo que se dice (*speech recognition*) y quién lo dice (*speaker recognition*). Los datos más frecuentemente utilizados como entrada a los modelos de *machine learning* son los Coeficientes Cepstrales en las Frecuencias de Mel (MFCC) o Coeficientes de Predicción Lineal Perceptual (PLPs) calculados directamente sobre la señal acústica y sobre sus primer y segunda derivada [12].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para el desarrollo de esta línea de investigación se toman tanto los últimos avances del campo de reconocimiento del habla, como la totalidad de los datos que brinda el análisis acústico clásico. Los valores devueltos por las técnicas de análisis acústico y los datos de originales (los valores muestreados que permanecen en los archivos de audio) son entradas potenciales para entrenar modelos de redes neuronales diseñados especialmente. El diseño implica la determinación del tipo de patrones que debe ser reconocido en cada una de las capas de la red, donde los primeros niveles reaccionan a estímulos de estructura simple y los más profundos lo hacen a las relaciones más complejas. El enfoque principal del proyecto es obtener en las primeras etapas de un modelo de *deep learning* valores equivalentes a los resultados del análisis acústico y buscar una correlación entre estos valores y los niveles de calidad vocal (GRBAS) en las últimas etapas del modelo.

Se programaron las siguientes tareas:

Grabación y entrevista:

Se realizará el protocolo de grabación de voz sugerido por Dejonckere *et al* [13] y se agregarán 2 aspectos extra:

- 1- una vocal /A/ sostenida en un tono e intensidad cómodos,
- 2- una vocal /A/ en una intensidad levemente ascendida respecto a la normal,
- 3- una frase simple estandarizada,
- 4- un fragmento de habla encadenada (serie automática) y
- 5- un glissando con la vocal /A/.

Lugar:

Departamento de Investigación Científica, Extensión y Capacitación "Raquel Maurette"

Entrevista:

Sobre cada persona se registrarán datos como edad, profesión/ocupación, género, antecedentes de problemas vocales, etc. Estos datos se coleccionarán en formularios predefinidos para luego cargarse en una base de datos que guarde la relación con los archivos de audio y los análisis acústicos posteriores.

Clasificación de las grabaciones

Los especialistas clasificarán según la escala GRBAS cada una de las grabaciones y se cargarán los datos en la base de datos.

Análisis acústico

Para cada grabación se realizará un análisis acústico (Espectrograma de banda angosta, Espectro FFT, Jitter y HNR) y se cargarán los resultados en la base de datos.

Software: Praat

Creación de modelos de aprendizaje automático (con datos de análisis acústico como entrada)

Se desarrollarán, entrenarán y evaluarán los modelos de aprendizaje en el lenguaje Python (principalmente con la librería scikit-learn).

Creación de modelos de deep learning (con audio y/o espectrograma como entrada)

Se desarrollarán, entrenarán y evaluarán los modelos de aprendizaje en el lenguaje Python (principalmente con el framework Keras sobre Theano).

3. RESULTADOS ESPERADOS

La estimación de la calidad vocal lograda de forma automática se debe comparar con la estimación de los expertos. Es importante notar que la clasificación puede diferir entre especialistas.

El nivel de acuerdo entre los especialistas de la voz se calculará con el coeficiente kappa de Cohen [72X], al igual que en el trabajo de De Bodt *et al* [73X].

El acuerdo entre la clasificación automática y los especialistas humanos también se calculará con el coeficiente kappa de Cohen, como Villa-Cañas *et al* en [74X].

Además, se calculará la desviación del modelo obtenido respecto a la media de las clasificaciones y se comparará con las desviaciones de los especialistas. Se espera obtener valores de variación significativamente menores.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo del proyecto está formado por un docente/investigador de la UTN FRC, dos docentes/investigadores de la UNC y cuatro

alumnos de la carrera de grado de la UTN FRC.

Además de formación de los alumnos participantes, el conocimiento generado por el proyecto se incorporará a las cátedras de los docentes de la UTN y UNC.

5. REFERENCIAS

- [1] Isshiki, N., Yanagihara, N., & Morimoto, M. (1966). *Approach to the objective diagnosis of hoarseness*. Folia Phoniatica et Logopaedica, 18(6), 393- 400.
- [2] Hirano, M. (1981). *Clinical examination of voice* (Vol. 5). Springer.
- [3] Yun, Y. S., Lee, E. K., Baek, C. H., & Son, Y. I. (2005). *The correlation of GRBAS scales and laryngeal stroboscopic findings for the assessment of voice therapy outcome in the patients with vocal nodules*. Korean Journal of Otolaryngology- Head and Neck Surgery, 48(12), 1501- 1505.
- [4] Hui, H., Weijia, K., & Shusheng, G. (2007). *The Validation of Acoustic Analysis and Subjective Judgment Scales of Several Voice Disorders* [J]. Journal of Audiology and Speech Pathology, 3, 010.
- [5] Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., & Hoffman, H. T. (2007). *Reliability of clinician- based (GRBAS and CAPE- V) and patient- based (V-RQOL and IPVI) documentation of voice disorders*. Journal of Voice, 21(5), 576- 590.
- [6] Jesus, L. M., Barney, A., Couto, P. S., Vilarinho, H., & Correia, A. (2009, December). *Voice quality evaluation using cape- v and GRBAS in european Portuguese*. In MAVEBA (pp. 61- 64).

- [7] Kreiman, J., & Gerratt, B. R. (2010). *Perceptual assessment of voice quality: past, present, and future*. SIG 3 Perspectives on Voice and Voice Disorders, 20(2), 62- 67.
- [8] Núñez- Batalla, F., Díaz- Molina, J. P., García- López, I., Moreno- Méndez, A., Costales- Marcos, M., Moreno- Galindo, C., & Martínez- Cambor, P. (2012). El espectrograma de banda estrecha como ayuda para el aprendizaje del método GRABS de análisis perceptual de la disfonía. *Acta Otorrinolaringológica Española*, 63(3), 173- 179.
- [9] Freitas, S. V., Pestana, P. M., Almeida, V., & Ferreira, A. (2015). *Integrating Voice Evaluation: Correlation Between Acoustic and Audio- Perceptual Measures*. *Journal of Voice*, 29(3), 390- e1.
- [10] ELISEI, N. G. (2013, May). Percepción auditiva de voces patológicas. In XIV Reunión Nacional y III Encuentro Internacional De La Asociación Argentina de Ciencias del Comportamiento.
- [11] Nuñez Batalla, F., Corte Santos, P., Señaris Gonzalez, B., Rodriguez Prado, N., Suárez Nieto, C. (2004) Evaluación espectral de la hipofunción vocal. *Acta Otorrinolaringol. Esp.* 55:327- 333.
- [12] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... & Kingsbury, B. (2012). *Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups*. *Signal Processing Magazine, IEEE*, 29(6), 82- 97.
- [13] Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier- Buchman, L., Friedrich, G., ... & Woisard, V. (2001). *A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques*. *European Archives of Oto- rhinolaryngology*, 258(2), 77- 82.
- [14] Cohen J. *A coefficient of agreement for nominal scales*. *Educ Psych Measurement* 1960;20:37—46.
- [15] De Bodt, M. S., Wuyts, F. L., Van de Heyning, P. H., & Croux, C. (1997). *Test-retest study of the GRBAS scale: influence of experience and professional background on perceptual rating of voice quality*. *Journal of Voice*, 11(1), 74-80.
- [16] Villa-Cañas, T., Orozco-Arroyave, J. R., Arias-Londono, J. D., Vargas-Bonilla, J. F., & Godino-Llorente, J. I. (2013, September). *Automatic assessment of voice signals according to the GRBAS scale using modulation spectra, Mel frequency Cepstral Coefficients and Noise Parameters*. In *Image, Signal Processing, and Artificial Vision (STSIVA), 2013 XVIII Symposium of* (pp. 1- 5). IEEE.

Aplicaciones de Análisis de Información Textual: Corpus Lingüísticos

Julio Castillo, Marina Cardenas

Laboratorio de Investigación de Software LIS, Dpto. Ingeniería en Sistemas de Información
Facultad Regional Córdoba, Universidad Tecnológica Nacional

{ jotacastillo, ing.marinacardenas }@gmail.com

Resumen

En este artículo se describe un proyecto de investigación relacionado al análisis y procesamiento de información textual, tal como el reconocimiento de paráfrasis o la implicación de textos.

En ese contexto se describe la creación de una herramienta para construir corpus lingüísticos que pueden ser utilizados como material de entrenamiento para sistemas de minería de datos y de extracción de información, en especial sobre texto no estructurado.

Palabras clave: análisis de texto, extracción de información, corpus.

Contexto

El proyecto denominado análisis de texto no estructurado (ADT) se encuentra consolidado dentro de la línea de investigación relacionada con lingüística computacional y es llevado a cabo en el Laboratorio de Investigación de Software LIS1 del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba.

Por otra parte, este proyecto se encuentra dentro del grupo de investigación denominado Grupo de Inteligencia Artificial (o GIA) de la UTN-FRC.

El grupo de investigación nuclea diferentes proyectos de investigación que se

hayen todos en temáticas concernientes a la inteligencia artificial entre las que podemos destacar análisis de imágenes, algoritmos evolutivos, y su aplicabilidad en problemas de la ingeniería, de las ciencias naturales, y de las ciencias sociales.

Este grupo está compuesto de becarios, pasantes, docentes investigadores y doctores.

1. Introducción

Mediante este proyecto se propone abordar la detección de similitudes entre diferentes textos. Si bien se trata de un área de investigación compleja, se han llevado a cabo varias aproximaciones tendientes a clasificar si dos fragmentos de textos se relacionan entre sí, y en ese caso, qué grado de relación presentan entre sí [1][2].

A los efectos de poder abordar este tipo de problemas, es necesario contar con conjunto de datos (material lingüístico) apropiado para poder construir algoritmos de aprendizaje sobre los mismos [3][4][5].

Por ello, la creación de una herramienta que pueda ser utilizada por anotadores humanos es de vital importancia en la construcción de un corpus. En el proyecto se ha desarrollado una herramienta que permite procesar y etiquetar diferentes conjuntos de datos.

Entre los objetivos generales de construcción de este asistente podemos enumerar tres principales:

- Proveer un mecanismo de clasificación de pares de texto con paráfrasis y

¹ www.investigacion.frc.utn.edu.ar/mslabs/

facilitar la lectura y estudio de otros corpus.

- Contar con una herramienta que permita automatizar la clasificación de fragmentos de textos en base a la presencia o ausencia de ciertos fenómenos lingüísticos que se mantengan en el mismo texto o que se mantenga entre dos fragmentos de textos.
- Generación de corpus etiquetado. El corpus permite incrementar la capacidad de reconocimiento y de clasificación de los algoritmos utilizados.

La construcción de material de entrenamiento involucra a anotadores humanos por lo cual el proceso de construcción es costoso, lento y a menudo es fuente de errores. La herramienta desarrollada permite automatizar la creación y acelerar el proceso de etiquetado, a la vez que posibilita realizar una trazabilidad entre los anotadores humanos y conocer su aporte en la construcción del corpus. Como consecuencia de agilizar el proceso de etiquetado es posible bajar los costos asociados en su construcción.

La información lingüística asociada a cada corpus creado dependerá del problema que se necesite abordar, pero el corpus debe ser consistente y cada elemento que lo constituya (texto o pares de fragmentos de textos) debe poseer el mismo tipo de anotaciones lingüísticas.

El tamaño del corpus que se puede crear puede ser tan grande como se requiera. La versión actual de la herramienta desarrollada soporta un máximo de archivos individuales de cuatro gigabytes de tamaño.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de investigación en las que se enmarca el proyecto de análisis de texto es el área de inteligencia artificial, más

concretamente una sub-especialidad que se denomina computación lingüística.

Los desarrollos de esta línea de investigación, lo constituyen, por un lado, las herramientas elaboradas para facilitar el análisis y procesamiento de archivos de texto, y por el otro, los sistemas de reconocimiento de implicación o de reconocimiento de paráfrasis entre dos fragmentos de textos.

La innovación del proyecto concierne a los nuevos métodos propuestos para el análisis y procesamiento de textos, como así también a los algoritmos creados para abordar las problemáticas anteriormente mencionadas. Los algoritmos diseñados aprovechan las diferentes características que se pueden aprender de los textos y que son recolectados y creados a partir de las herramientas de procesamiento de textos.

Son múltiples las posibles sub-disciplinas que podrían valerse de los resultados de este proyecto, entre las que podemos destacar a las tareas de recuperación de información, evaluación de las traducciones automáticas [6], evaluación de la calidad de las traducciones, reconocimiento de paráfrasis [7] e implicación de textos[8][9][10] . Adicionalmente, la creación de corpus es una actividad de relevancia y de impacto en las tareas relacionadas al procesamiento del lenguaje.

3. Resultados

Para poder construir una herramienta que permita construir corpus lingüísticos se investigaron y definieron el conjunto de fenómenos lingüísticos de interés. Se definieron cuatro tipos de fenómenos que son los que permite registrar el software asistente de creación de corpus.

La clasificación se definió en base al tipo de fenómeno presente en un fragmento de texto. Se identificaron y clasificaron en Fenómenos a nivel Léxico, Morfológico, Semántico, y Sintáctico.

Algunos de los fenómenos Léxicos que se registran son anglicismos, arcaísmos, barbarismos, cultismos, eufemismos, neologismos, entre otros fenómenos.

En cuanto a los fenómenos morfológicos se pueden registrar lexemas, morfemas y gramemas.

Los fenómenos sintácticos que se registran son anáforas, flexiones, pronombres, y concondancias entre otros fenómenos. Como fenómenos semánticos se registran la antonimia, homonimia, polisemia y sinonimia.

La caracterización de estos fenómenos facilita el proceso de etiquetado a los anotadores humanos, al tiempo que provee información lingüística de textos que luego pueden ser utilizados por diversos algoritmos de clasificación.

En cuanto al proceso de desarrollo, se comenzó con una especificación de los requerimientos funcionales y no funcionales, y posteriormente, se realizó el análisis, diseño e implementación. En todas las etapas se utilizó UML como lenguaje de modelado. La creación de esta herramienta siguió el proceso de desarrollo unificado, pero esta es la excepción más que la regla, debido a que normalmente en el proyecto se utilizan metodologías ágiles.

El software construido posee las siguientes funcionalidades:

- Lectura de corpus de implicación de textos y de paráfrasis.
- Carga de nuevos pares del corpus.
- Búsqueda y posicionamiento de un par dentro del corpus.
- Selección de substrings de fragmentos de texto en base a una clasificación lingüística elegida por el anotador humano.
- Clasificación de los fenómenos en categorías y subcategorías definidas previamente.
- Generación de un nuevo corpus de datos. Estos corpus generados conforman un corpora lingüístico que constituyen un recurso necesario para muchas aplicaciones

de análisis y procesamiento del lenguaje [11][12].

Uno de los corpus generados consiste en corpus de implicación de textos ampliados.

Los corpus ampliados han demostrado ser de utilidad en diversas tareas y ya que permiten incrementar la efectividad en la clasificación de textos [13].

Actualmente se está trabajando en extender las funcionalidades de esta herramienta para que sea capaz de informar:

- Aquellos pares de textos en los cuales hay disidencia en cuanto a su clasificación.
- Porcentaje de fragmentos de textos en los cuales hay coincidencias y disidencias.
- Consistencia del material de entrenamiento.
- Sesiones de usuario, para que los anotadores puedan suspender y continuar el proceso de etiquetado en el momento que lo deseen.
- Extensión de los tipos de archivos y corpus que es capaz de reconocer la herramienta.

4. Formación de Recursos Humanos

El equipo de investigación está formado por docentes investigadores del Laboratorio de Investigación de Software LIS² del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba, se detallan a continuación los responsables del proyecto:

- El Dr. Julio Castillo desarrolló su tesis de doctorado en Ciencias de la Computación en la temática de implicación de textos y paráfrasis, y coordina las actividades del proyecto y dirige a los integrantes miembros del equipo.
- La Mg. Ing. Marina Cardenas está evaluando la posibilidad de desarrollar

² www.investigacion.frc.utn.edu.ar/mmlabs/

su tema de tesis de doctorado en la misma temática con una variación del enfoque desde el punto de vista de los sistemas de Generación del Lenguaje Natural (NLG), y dirige a los integrantes miembros de equipo.

- Participan del proyecto alumnos que necesitan realizar su práctica supervisada que es uno de los requisitos para la obtención del grado de Ingeniero. Los alumnos que intervienen aprenden a realizar actividades de investigación, y cómo integrarse en un equipo existente. También generalmente participan por año uno o dos becarios alumnos a los que se les enseña como trabajar en un proyecto de investigación. Adicionalmente, se prevé que becarios de investigación de posgrado puedan realizar actividades en el marco del presente proyecto.

5. Bibliografía

- [1] Castillo J.; Cardenas M. Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia 2010, LNCS, vol. 6433, pp. 366-375, 2010.
- [2] Castillo J. Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task. TAC, 2009.
- [3] Judith K lavans and Philip Resnik. The Balancing Act. Combining Symbolic and Statistical Approaches to Language. MIT Press. 1996.
- [4] C. Manning and H. Schütze. Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge, MA, 1999.
- [5] D. Lin and P. Pantel. DIRT - Discovery of Inference Rules from Text. In Proceedings of ACM SIGKDD Conference
- on Knowledge Discovery and Data Mining, pages 323–328, 2001.
- [6] Castillo, Julio and Estrella, Paula. Semantic textual similarity for MT evaluation. Proceedings of the Seventh Workshop on Statistical Machine Translation. WMT '12. 2012.
- [7] Dolan, W., Brockett, C., Castillo, J. and Vanderwende, L. (2010). Mining phrase pairs from an structured resource. WO/2010/135204. 2010.
- [8] C. Monz and M. de Rijke. Light-Weight Entailment Checking for Computational Semantics. In P. Blackburn and M. Kohlhase, editors, Inference in Computational Semantics (ICoS-3), pages 59–72, 2001.
- [9] Appelt, Douglas E., Jerry R. Hobbs, John Bear, David Israel, Megumi Kameyama, and Mabry Tyson. "The SRI MUC-5 JV-FASTUS Information Extraction System", Proceedings, Fifth Message Understanding Conference (MUC- 5), Baltimore, Maryland, August 1993.
- [10] FeldmanR., and Hirsh H. Exploiting Background Information in Knowledge Discovery from Text. Journal of Intelligent Information Systems. 1996.
- [11] Lewis, D. D. Evaluating and optimizing autonomous text classification systems. In Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval (Seattle, US, 1995), pp. 246-254. 1995.
- [12] M. Craven and J. Shavlik. Using Neural Networks for Data Mining. *Future Generation Computer Systems*, 13, pp. 211-229. 1997.
- [13] Stefan. T; Stefanowitsch A. (2006). Corpora in Cognitive Linguistics. Corpus - Based Approaches to Syntax and Lexis, Berlin: Mouton, pág. 117. 2006.

Aprendizaje Autónomo en Sistemas Inteligentes

González Daniel y Postay Juan Domingo

Unidad del Programa de Transferencia de Tecnologías Informáticas (PROTTI)
 Departamento de Sistemas - Área Tecnología Robótica
 Universidad John F. Kennedy, Argentina
 Bartolomé Mitre 1441, C1037ABA, CABA, Argentina.
 dgonzalez11@kennedy.edu.ar - jpostay@kennedy.edu.ar

RESUMEN

El objetivo de este trabajo es desarrollar nuevos algoritmos de aprendizaje para agentes inteligentes. En una etapa inicial, investigamos la manera que el ser humano adquiere el conocimiento basándonos en las teorías de aprendizaje desarrolladas en otras disciplinas científicas, es decir, el trabajo tiene características transdisciplinarias ya que posee un enfoque desde la ciencia cognitiva en su conjunto. Debemos destacar que esta investigación no tiene intención de realizar un sistema o programa para lograr un objetivo particular sino demostrar como un sistema puede adquirir el conocimiento a partir de la experiencia, y cómo puede emplearse con fines prácticos. Fue necesario producir un hardware y software básicos donde implementar los algoritmos de aprendizaje. Luego, se desarrollaron los módulos correspondientes al tipo de aprendizaje particular. Existen varias etapas previstas, donde se modelaron y modelan los distintos tipos, basados en las diferentes teorías. La primera de ellas es la que mostramos en este trabajo, donde se han desarrollado algoritmos que emulan el condicionamiento clásico. El resto se desarrollará en la etapa que comienza. Consideramos que se trata de un proyecto original e innovador, sirviendo de punto de partida de muchas investigaciones posteriores.

Palabras clave: reglas de producción, algoritmos de aprendizaje, condicionamiento clásico.

CONTEXTO

El proyecto se inserta en el grupo de investigaciones en inteligencia artificial y robótica impulsadas por el Departamento de Sistemas y el Programa de Transferencia de Tecnología Informática (PROTTI). Este es

un programa de transferencia tecnológica en el área de Informática y Sistemas. A través de él, se ponen a disposición de toda la comunidad los desarrollos realizados por alumnos y docentes de la carrera de sistemas. El objetivo del PROTTI es participar en el desarrollo social y económico de la comunidad aportando conocimiento en el ámbito de sistemas a empresas, particulares u organizaciones estatales. Esta investigación fue financiada en su totalidad con recursos provenientes de la universidad.

1. INTRODUCCIÓN

Gran parte de las capacidades del cerebro humano son utilizadas para predecir el futuro, tanto el comportamiento de la naturaleza como el de otras personas. Para lograr los objetivos de la forma más adecuada, tanto en el medio natural como en el social, los seres humanos tratan de realizar predicciones basadas en conocimientos adquiridos con anterioridad. La inteligencia artificial produce sistemas que intentan emular el conocimiento humano en las máquinas, cuando el aprendizaje se realiza a través de la

experiencia, nos encontramos frente a un agente inteligente [González, 2013]. Si este conocimiento es almacenado mediante reglas de producción, denominamos a este sistema, agente inteligente basado en reglas.

Estas reglas no son estáticas sino dinámicas, no permanecen inmutables desde su origen sino que pueden ser absorbidas por otras más generales, desdoblarse a ámbitos más reducidos o simplemente entrar en conflicto y ser eliminadas.

El modelado de agentes o la adquisición de un determinado comportamiento, implica el desarrollo de dos módulos o subsistemas fundamentales. El primero de ellos es el sistema que genera las reglas, encargado del aprendizaje, mientras que el segundo, el planificador evalúa las metas u objetivos y genera las tareas que se requieren para alcanzar los mismos [García Martínez, Servente y Pasquini, 2003]. De acuerdo a estas últimas, las interfaces (actuadores) actúan con el medio, ejecutando movimientos (acciones). Adicionalmente, este conocimiento puede ser transferido a otros agentes, compartiendo los conocimientos adquiridos y, de ese modo, potenciar el sistema.

Desde hace ya una década, los métodos y técnicas utilizados para modelar el comportamiento de los agentes son muy debatidas y presentan configuraciones muy variadas [Giráldez Bretón, 1999, P. 18-24]. Tanto los sistemas expertos como las redes neuronales poseen ámbitos de aplicación donde producen mejores resultados. Las redes parecen tener muy buen desempeño en medios difusos, cercanos al medio físico pero aún no han dado muestra de resultados evidenciables al tratar la abstracción simbólica propia del ser humano. En cambio, los sistemas inteligentes, evoluciones de los sistemas expertos, se desempeñan eficazmente en este medio. Es posible complementar ambas técnicas para lograr resultados más satisfactorios. Este es uno de los objetivos que pretendemos desarrollar en una segunda etapa que está comenzando. Durante el ciclo anterior nos centramos en el aprendizaje del agente.

Debemos destacar que este trabajo no tiene intención de realizar un sistema o programa para lograr un objetivo particular, sino mostrar como el sistema adquiere el conocimiento. Se

podría decir que el trabajo tiene una finalidad epistemológica y otra pragmática, ya que, por un lado, se pretende conocer cómo es posible adquirir conocimiento en máquinas, y por otro aplicarlo en sistemas artificiales.

La denominada Inteligencia Artificial Fuerte, pretende producir máquinas inteligentes que simulen o no la inteligencia humana. Este desarrollo se encuentra enmarcado dentro de los llamados agentes inteligentes autónomos, estos son sistemas que adquieren el conocimiento interactuando con el entorno, limitados por su diseño, es decir por la teoría de aprendizaje que se propone modelar [Moriello, 2004, p. 55-67]. Para lograr que el agente situado adquiriera los conocimientos, en su interacción con el medio, que le permiten alcanzar sus objetivos de un modo más eficiente, se requiere del desarrollo de un conjunto de módulos de software y una serie de sensores y actuadores. Estos interactúan con el medio, mientras que un sistema que hemos denominado Generador de Reglas, resulta ser el encargado de interpretar los datos de los sensores a través de alguna de las teorías de aprendizaje propuestas, y producir las reglas donde se almacena el conocimiento adquirido. En este último módulo reside el componente que diferencia nuestro trabajo, de otros trabajos sobre agentes, y donde se encuentra el núcleo del objetivo mismo que nos hemos propuesto. A continuación mencionamos algunas de las teorías seleccionadas.

Las teorías del aprendizaje conforman un variado conjunto de marcos teóricos que algunas veces se complementan y otras se contradicen. A fin de organizar las diferentes corrientes, mostramos la siguiente clasificación [Perez Gomez y Sacristán, 1992]:

Teorías asociacionistas:

Estas teorías se basan en el conductismo que postulan el aprendizaje por condicionamiento estímulo-respuesta.

- Condicionamiento básico o clásico (Pavlov, Watson, Guthrie.) y condicionamiento

instrumental u operante (Hull, Thorndike y Skinner.)

Teorías mediacionales:

- Se tiene en cuenta la importancia de las variables y reorganización interna interna.
- Aprendizaje social, condicionamiento por imitación de modelos (Bandura, Lorenz, Tinbergen, Rosenthal)

Teorías cognitivas

- Teoría de la Gestalt (Kofka, Köhler, Vhertheimer, Maslow y Rogers.)
- Psicología genético-cognitiva (Piaget, Bruner, Ausubel, Inhelder.)
- Psicología genético-dialéctica (Vigotsky, Luria, Leontiev, Rubinstein, Wallon.)
- La teoría del procesamiento de información (Gagné, Newell, Simon, Mayer, Pascual, Leone.)

Nuestra intención es tomar los principios teóricos de algunas de ellas para que el sistema elabore su propio conocimiento. El grado de dificultad para hacerlo varía enormemente según la teoría escogida. Para nuestro primer modelo escogimos el condicionamiento clásico debido a sus particulares características [Perez y Cruz, 2003]. Previamente hemos realizado un proceso de análisis y selección que no está cerrado definitivamente. Por el momento hemos incluido las siguientes etapas:

- Condicionamiento básico o clásico.
- Condicionamiento instrumental u operante.
- Psicología genético-cognitiva.
- La teoría del procesamiento de información.

En este informe presentamos los resultados obtenidos hasta el momento sobre el desarrollo de la primera etapa: El condicionamiento básico o clásico.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Se desarrolló una plataforma base que permite probar los algoritmos de aprendizaje. El conocimiento adquirido por estos módulos se almacena mediante reglas de producción, se medirá la utilidad del mismo de acuerdo al cambio de comportamiento de un sistema móvil para alcanzar un determinado objetivo. Consta de una parte móvil, el robot, y una fija, una PC donde reside la

parte central e inteligente del sistema, comunicados vía bluetooth. En la PC, se probarán los distintos algoritmos de aprendizaje, basados en diferentes teorías ya mencionadas. Se detallan las fases del desarrollo:

- Se diseñó un robot móvil basado en un microcontrolador ATmega 328 (Arduino Uno).
- Se diseñó y construyó un agente inteligente basado en reglas de producción y se instaló en una computadora (PC).
- Se elaboró una base de datos donde se almacenan reglas de producción, datos históricos y actuales de los sensores, y planificación de tareas.
- Se instaló en la PC y el robot un módulo de comunicaciones bluetooth que permite el intercambio de información entre ambos.
- Se desarrolló un protocolo de comunicaciones bidireccional para enviar información desde los sensores del robot y órdenes desde la PC.
- Se desarrolló la programación del Arduino.
- Se desarrolló un módulo de aprendizaje de reglas basado en aprendizaje por reforzamiento que emula el condicionamiento clásico.
- Se realizaron pruebas con el sistema, actualmente se están revisando los resultados de dichas pruebas y modificando parte del programa.

La metodología de desarrollo de software empleada, es decir el conjunto de métodos, principios y reglas que posibilitan realizar el desarrollo del software, están planteadas como una serie de tareas, que pueden repetirse de manera recursiva, orientadas a la construcción y prueba de dicho programa, o de una serie de módulos que conforman el mismo.

Para este desarrollo, hemos escogido el ciclo de vida de prototipos por considerarlo el más adecuado para el modo en que se plantea este desarrollo. Nos permite incorporar nuevas funcionalidades a partir de la finalización de las anteriores.

El desarrollo por prototipos ayuda a comprender las funcionalidades requeridas por el usuario, sobre todo si este no tiene una idea clara o acabada de lo que desea. Se produce una versión preliminar de lo que será el producto final, con una funcionalidad acotada que se irá aumentando gradualmente a partir de ciclos sucesivos de las especificaciones

del sistema, hasta lograr el sistema final o solución buscada [Fernández, 2001].

3. RESULTADOS OBTENIDOS/ESPERADOS

Se realizaron pruebas para diferentes objetivos del agente. Se emplearon tres casos diferentes. Se lograron mayormente resultados satisfactorios, sin embargo se encontró que una gran cantidad de reglas generadas que no se ajustaron a lo esperado. Luego de un análisis se decidió incorporar un módulo de validación entre el objetivo y la regla formada que categorice el nivel de la misma. Cuanto más reforzamiento, o sea cuantas más veces se produce la situación esperada se incrementa el nivel de la regla, diferenciándose de este modo las reglas casuales. Después de cierto tiempo las reglas generadas con bajo nivel de reforzamiento son descartadas. Se probaron algunas reglas frente a situaciones nuevas obteniendo un resultado satisfactorio; se logró el comportamiento esperado. Quedan pendientes algunas modificaciones para mejorar los resultados. Una vez corregidos los pendientes se evaluará el sistema de acuerdo a los indicadores de desempeño (KPI) desarrollados para registrar las mejoras respecto al alcance de una meta u objetivo.

Como evolución del sistema proponemos un mayor grado de autonomía del equipo móvil incorporando al mismo una placa raspberry pi, eliminando la PC y la dependencia de la misma, la incorporación de otros algoritmos de aprendizaje y se completará el Planificador. También proponemos la incorporación de redes neuronales como interface entre el sistema central y el medio, ya que estas logran un buen desempeño en entornos con ruido. A grandes rasgos, el sistema podría verse como un sistema central determinístico y otro periférico que interconecta el núcleo con el mundo real de características analógicas, semejante al propuesto por Jerry Fodor (La modularidad de la mente, 1983).

Tratándose de un sistema de propósito general las aplicaciones son innumerables. Mediante estos sistemas se puede ofrecer una solución a muchos problemas aún no resueltos, o bien resueltos por caminos menos eficientes.

4. FORMACIÓN DE RECURSOS HUMANOS

EL grupo de investigación está integrado actualmente por dos docentes y cinco alumnos. La participación es abierta aunque existe un cupo máximo de colaboradores.

En cuanto a las TFI, se han finalizado y aprobado dos trabajos sobre sistemas expertos. En curso de desarrollo se han relevado: una tesis sobre redes neuronales y otra sobre algoritmos genéticos. Podemos agregar varias propuestas para TFI aún no evaluadas.

5. BIBLIOGRAFIA

- Arduino. (2016, 30 de agosto). Arduino Genuino.: <https://www.arduino.cc>
- Barrientos A, y otros (1997); Fundamentos de Robótica. Madrid, España: McGraw-Hill.
- Fernández, C. (2001). Modelo de Prototipo. Universidad Estatal de Milagro, Ecuador.
Recuperado de: <http://es.slideshare.net/yanezcabrera/modelo-de-prototipo>
- García Martínez, R., Servente, M. y Pasquini, D. (2003), Sistemas Inteligentes (pp. 149-280), Buenos Aires, Editorial Nueva Librería.
- Giráldez Bretón J. (1999). Modelo de Toma de Decisiones y Aprendizaje en Sistemas Multi-Agente. Universidad Politécnica de Madrid, Departamento de Inteligencia Artificial, España.
Recuperado de: <http://oa.upm.es/1255/1/10199908.pdf>
- González, E. (2013), Aprendizaje y Planificación en Sistemas Inteligentes Autónomos. Tesis de Maestría. Universidad Tecnológica Nacional, Buenos Aires, Argentina.
Recuperado de: <http://sistemas.unla.edu.ar/sistemas/gisi/tesis/gonzalez-trabajofinaldeespecialidad.pdf>
- Igoe T. (2011). Making Things Talk, Sebastopol, O'Reilly Media Inc, [Versión electrónica]: http://www.proalias.com/books/Making_Things_Talk_Second_Edition.pdf
- Making Things Talk Second Edition.pdf Mataric M. J. (2007). The Robotics Primer. EUU: Massachusetts Institute of Technology.
- Moriello, S. (2004). Agentes Físicos Autónomos. Tesis de Maestría, Universidad Tecnológica Nacional, Buenos Aires, Argentina.
Recuperado de: <http://posgrado.frba.utn.edu.ar/investigacion/tesis/MIS-2006-Moriello.pdf>
- Pérez, Andrés M.; Cruz, Julio Eduardo
Conceptos de condicionamiento clásico en los

campos básicos y aplicados
Interdisciplinaria, vol. 20, núm. 2, 2003, pp. 205-
227
-Centro Interamericano de Investigaciones
Psicológicas y Ciencias Afines
Buenos Aires, Argentina Recuperado de:

<http://www.redalyc.org/pdf/180/18020204.pdf>
-Perez Gomez y Sacristán (1992). Comprender y
Transformar la Enseñanza. Madrid, España, Ed.
Morata. Recuperado de:
[http://dragodsm.com.ar/pdf/dragodsm-educacion-
comprender-y-transformar-ensenanza-06-2012.pdf](http://dragodsm.com.ar/pdf/dragodsm-educacion-comprender-y-transformar-ensenanza-06-2012.pdf)

Artificial Bee Colony Aplicado a la Optimización Nutricional de Menús Escolares

Gerónimo, Verónica¹; Navarro, Carina²; Paz, Fabiola³;
Herrera Cogna, Analía⁴

Facultad Ingeniería

Universidad Nacional de Jujuy

geronimo_vero@hotmail.com¹; navarro.carinav@gmail.com²;
fabypaz@gmail.com³; anihco@yahoo.com.ar⁴

RESUMEN

El presente trabajo describe el desarrollo de un prototipo, diseñado para optimizar los nutrientes de diversos menús escolares destinados a estudiantes del nivel secundario, de manera tal que alcancen una alimentación saludable. La optimización se centra en minimizar el contenido de grasas totales del menú seleccionado, teniendo presente no solo las cantidades de los nutrientes que deben ser consumidos mínimamente, sino también las cantidades máximas de los alimentos seleccionados para los menús. Dada la complejidad del caso de estudio, se utilizó como técnica de resolución, una variante del algoritmo de Colonia de Abejas Artificiales (ABC) la cual es lo suficientemente flexible, eficiente y robusta para ser aplicada a problemas de optimización.

Palabras Clave: Optimización, Metaheurísticas, Colonia de Abejas Artificiales, Programación Lineal, Restricciones.

CONTEXTO

El presente trabajo es un resumen de la investigación realizada para elaborar el trabajo final de carrera de grado. Utiliza un algoritmo de metaheurísticas, llamado Colonias de Abejas Artificiales (ABC), perteneciente a la inteligencia de enjambre, para la optimización de diversos nutrientes de un menú determinado. Los ajustes son proporcionados por la cátedra de Trabajo Final de Sistemas, de la carrera Licenciatura en Sistemas dictada en la Facultad de

Ingeniería de la Universidad Nacional de Jujuy.

1. INTRODUCCIÓN

Las personas en edad escolar necesitan alimentarse bien a fin de desarrollarse adecuadamente, evitar las enfermedades y tener energía para estudiar y estar físicamente activos. Las ventajas derivadas de un buen nivel de nutrición no solo se refieren a un adecuado crecimiento físico de las personas, sino que también implican, en los niños y adolescentes, un óptimo desarrollo del nivel madurativo, permitiéndoles que desplieguen las habilidades cognitivas necesarias para comprometerse en experiencias educativas y así poder alcanzar satisfactoriamente sus objetivos [1].

La optimización de menús se realizará considerando los datos de la Escuela Provincial Agrotécnica N° 14, de la Localidad de Cieneguillas del Departamento de Santa Catalina, Provincia de Jujuy.

El algoritmo de Colonia de Abejas Artificiales propuesto por primera vez en 2005 por Karaboga et al. [2], representa una de las tantas alternativas de inteligencia artificial para optimizar problemas del mundo real por medio de la programación lineal en los que el número de variables en juego puede ser tan amplio como se desee.

Si bien esta metaheurística se ha probado exitosamente en diferentes funciones benchmark [3] cada una de ellas representa un caso particular teórico que no garantiza una

eficiencia absoluta del algoritmo frente a un caso de estudio real [2].

1.1 Formulación matemática de los menús escolares

Para describir y formular el problema se utiliza un modelo matemático establecido en la programación lineal. Se busca por lo tanto determinar la cantidad de cada alimento a incluir en cada menú, asegurando los requerimientos nutricionales adecuados para una dieta saludable.

El objetivo principal de este modelo es minimizar el contenido de grasas totales para cada comida. Otro aspecto importante a destacar en la problemática de estudio son las restricciones a considerar. Se definen dos tipos de restricciones:

- *Restricciones de Nutrientes:* para asegurar que el desayuno/almuerzo proporcione la cantidad mínima de cada nutriente.
- *Restricciones de Alimentos:* para asegurar que no se incluya demasiada cantidad de un tipo de alimento.

Los menús tanto del desayuno como del almuerzo deben confeccionarse de manera tal que satisfaga los requerimientos nutricionales y se proporcionen las cantidades adecuadas de cada macronutriente (nutrientes necesarios que constituyen la parte principal de una dieta, de los cuales se encuentran los hidratos de carbono, las proteínas y las grasas). Se determinó un valor energético total de 2300 calorías por alumno/día, las cuales se distribuyen en porcentajes para cada macronutriente obteniendo un 60 % de hidratos de carbono, 15 % de proteínas y 25 % de grasas. A su vez, se dividen para cada comida a lo largo del día de la siguiente manera: 345 cal. para el desayuno, 920 cal. para el almuerzo, 345 cal. para la merienda y 690 cal. para la cena.

Tanto el desayuno, el almuerzo, la merienda como la cena, la función a minimizar queda determinada por las cantidades (gramos) de grasa que aporta cada 100 gr. del ingrediente CRUDO seleccionado para un determinado menú.

Las restricciones nutricionales del desayuno/merienda mínimas a consumir son:

12,94 gr. de proteínas, 51,75 gr. de hidratos de carbono y 9,58 gr. de grasas. Para evitar una excesiva cantidad de un tipo de alimento, no debe incluirse más de 50 gr. de lácteos, 5 gr. de infusiones, 25 gr. de azúcares, 20 gr. de chocolates, 60 gr. de panificados y 45 gr. de agregados. A continuación en la **Figura 1** queda representado el modelo matemático.

$$\text{MIN } Z = 27,5 * LAC + 0,0 * INF + 0,0 * AZU + 2,3 * CHOC + 1,2 * PAN + 9,0 * DL + 99,5 * MAN + 0,3 * MER + 22,1 * QSO$$

Sujeto a:

Restricciones Nutricionales

$$26,4 * LAC + 0,1 * INF + 0,0 * AZU + 4,1 * CHOC + 8,2 * PAN + 8,1 * DL + 0,25 * MAN + 0,9 * MER + 18,9 * QSO \geq 12,94 \text{ (Proteínas)}$$

$$38,2 * LAC + 0,4 * INF + 99,5 * AZU + 78,5 * CHOC + 50,1 * PAN + 61,9 * DL + 1,0 * MAN + 23,2 * MER + 0,0 * QSO \geq 51,75 \text{ (H.de C.)}$$

Restricciones de Límites

$$LAC \leq 50 \text{ grs.}$$

$$INF \leq 5 \text{ grs.}$$

$$AZU \leq 25 \text{ grs.}$$

$$CHOC \leq 20 \text{ grs.}$$

$$PAN \leq 60 \text{ grs.}$$

$$DL + MAN + MER + QSO \leq 45 \text{ grs.}$$

Restricciones de No negatividad

$$LAC, INF, AZU, CHOC, PAN, DL, MAN, MER, QSO \geq 0 \text{ grs.}$$

Figura 1. Modelo matemático Desayuno/Merienda. Se identifican las siguientes cantidades de ingredientes: LAC: gr de lácteos, INF: gr de infusiones (té, mate y café), AZU: gr de azúcares, CHOC: gr de chocolates, PAN: gr de panificados, DL: gr de dulce de leche, MAN: gr de manteca, MER: gr de mermelada y QSO: gr de queso.

Para el caso del Almuerzo/Cena las restricciones nutricionales deben contener 34,5 gr. de proteínas, 138 gr. de hidratos de carbono y 25,5 gr. de grasas. En el caso de la cena se debe proporcionar 25,87 gr. de proteínas, 103,5 gr. de hidratos de carbono y 19,16 gr. de grasas. Para evitar demasiada cantidad de un tipo de alimento, no debe incluirse más de 110 gr. de carnes, 280 gr. de vegetales tipo "A" y "B", 180 gr. de vegetales tipo "C", 130 gr. de frutas, 35 gr. de panificados, 80 gr. de cereales y legumbres, 30 gr. de azúcares, 20 gr. de lácteos, 70 gr. de queso y 15 gr. de aceite. El modelo matemático para este menú se observa en la **Figura 2**.

$$\begin{aligned} \text{MIN } Z = & 17,09 * \text{CNE} + 0,4 * \text{VEG_A} + 0,2 * \text{VEG_B} + 0,0 * \text{VEG_C} + \\ & 0,4 * \text{FRU} + 0,0 * \text{PAN} + 0,9 * \text{CER_LEG} + 0,0 * \text{AZU} + \\ & 27,5 * \text{LAC} + 22,1 * \text{QSO} + 99,8 * \text{ACEI} \end{aligned}$$

Sujeto a:

Restricciones Nutricionales

$$18,5 * \text{CNE} + 1 * \text{VEG_A} + 1,3 * \text{VEG_B} + 0,0 * \text{VEG_C} + 0,3 * \text{FRU} + 0,0 * \text{PAN} + 23,9 * \text{CER_LEG} + 0,0 * \text{AZU} + 26,4 * \text{LAC} + 18,9 * \text{QSO} + 0,0 * \text{ACEI} \geq 34,5 \text{ (Proteínas)}$$

$$0,0 * \text{CNE} + 3,5 * \text{VEG_A} + 5,8 * \text{VEG_B} + 0,0 * \text{VEG_C} + 12,4 * \text{FRU} + 0,0 * \text{PAN} + 56,5 * \text{CER_LEG} + 99,5 * \text{AZU} + 38,2 * \text{LAC} + 0,0 * \text{QSO} + 0,0 * \text{ACEI} \geq 138,0 \text{ (H.de C.)}$$

Restricciones de Limites

$$\begin{aligned} \text{CNE} & \leq 110 \text{ grs.} \\ \text{VEG_A_B} & \leq 280 \text{ grs.} \\ \text{VEG_C} & \leq 180 \text{ grs.} \\ \text{FRU} & \leq 130 \text{ grs.} \\ \text{PAN} & \leq 35 \text{ grs.} \\ \text{CER_LEG} & \leq 80 \text{ grs.} \\ \text{AZU} & \leq 30 \text{ grs.} \\ \text{LAC} & \leq 20 \text{ grs.} \\ \text{QSO} & \leq 70 \text{ grs.} \\ \text{ACEI} & \leq 15 \text{ grs.} \end{aligned}$$

Restricciones de No negatividad

$$\text{CNE, VEG_A, VEG_B, VEG_C, FRU, PAN, CER_LEG, AZU, LAC, QSO, ACEI} \geq 0 \text{ grs.}$$

Figura 2. Modelo matemático Almuerzo/Cena. Se identifican las siguientes variables de decisión: CNE: gr de carnes, VEG_A: gr de vegetales tipo A, VEG_B: gr de vegetales tipo B, VEG_C: gr de vegetales tipo C, FRU: gr de frutas, PAN: gr de panificados, CER_LEG: gr de cereales y/o legumbres, AZU: gr de azúcares, LAC: gr de lácteos, QSO: gr de queso y ACEI: gr de aceite.

2. METAHEURÍSTICA COLONIA DE ABEJAS ARTIFICIALES

El algoritmo de colonia de abejas artificiales se inspiró en el comportamiento inteligente de forrajeo de las abejas, y es utilizado exitosamente para resolver diferentes problemas de optimización.

ABC es un algoritmo de optimización en el que las soluciones del problema se denominan fuentes de alimentos, las que son modificadas por las abejas artificiales con el objeto de encontrar aquellas con mayor cantidad de néctar. En el algoritmo existen abejas empleadas y observadoras, que se mueven en el espacio en busca de fuentes de alimentos, teniendo en cuenta su experiencia y de la de sus compañeros de colmena para reposicionarse. Mientras que las llamadas abejas exploradoras vuelan y eligen aleatoriamente las fuentes sin considerar su experiencia, y si la localización de la fuente es mayor, memorizan la posición y olvidan la anterior. De esta manera, ABC combina métodos de búsqueda local, mediante las abejas empleadas y observadoras, con métodos de búsqueda global mediante las abejas exploradoras, intentando equilibrar el

proceso de la exploración y de la explotación del espacio de búsqueda [4].

El algoritmo ABC se ha propuesto originalmente para problemas de optimización sin restricciones [5]. Sin embargo para poder implementar esta técnica como método de resolución, fue necesario diseñar una versión del algoritmo que tenga incorporado un mecanismo para el control de las restricciones en el algoritmo base, ya que el caso de estudio cuenta con ciertos tipos de limitaciones observadas en las **Figuras 1 y 2**.

2.1 Mecanismo para el control de las restricciones.

Este proceso para el manejo de restricciones fue incorporado en varias técnicas de optimización como PSO, ACO, ABC, entre otros [6]. El proceso más utilizado debido a su sencillez del método fue el propuesto por Deb [7]. Este trabajo se utiliza el método de Deb con algunas modificaciones propuesto en [8], la cual se optó también por utilizar este proyecto ya que la problemática son muy similares. Este método considera las siguientes reglas:

- 1) una solución factible siempre es preferible a una solución no factible
- 2) entre dos soluciones factibles, es preferible aquella con mayor valor (o menor valor según sea el caso) en su función objetivo
- 3) entre dos soluciones no factibles, es preferible aquella que viole el menor número de restricciones (número de violaciones de restricciones o NVR)
- 4) en el caso que ambas violen el mismo número de restricciones, se elige aquella con menor valor en la suma de porcentaje de violaciones (suma de violaciones de restricciones o SVR) con respecto a los recursos disponibles [9].

3. LINEAS DE INVESTIGACION Y DESARROLLO

En el trabajo, respecto a la planificación de menús escolares, se deja abierta la posibilidad de considerar la incorporación del aporte nutricional de los micronutrientes y el factor de corrección de alimentos cocidos para el cálculo de las cantidades de los

macronutrientes. Por otra parte, es de gran interés considerar el costo de los alimentos, que permitan respetar el presupuesto asignado a cada menú.

En cuanto al algoritmo, una de las extensiones futuras se centra en disminuir el costo computacional para problemas con varias restricciones de diversas naturalezas, como sería este caso.

4. RESULTADOS OBTENIDOS/ESPERADOS

Se toma como base un ejemplo de menú diario lo detallado en la **Tabla 1** para las pruebas experimentales.

Tabla 1. Menús

comida	Nombre
Desayuno	Leche con Chocolate y Pan
Almuerzo	Guiso de Lenteja y Compota de pera

4.1 Parámetros del algoritmo

Los parámetros especificados en la **Tabla 2**, son los que intervienen en la ejecución del algoritmo ABC para el caso de estudio. Es importante destacar que los valores asignados a los parámetros expuestos fueron elegidos experimentalmente.

Tabla 2. Parámetros del algoritmo ABC

NOMBRE	SÍMBOLO	VALOR
tamaño de la colonia	tam_colonia	30
Número de soluciones (fuentes de alimento)	Num_fuentes	15
Número máximo de iteraciones	max_iteraciones	150
Límite máximo	limite_max	tam_colonia*5%
Rango de modificación	MR	0.8
Periodo de producción explorador	SPP	0.1 * tam_colonia * Dim
Factor de tolerancia dinámico	E	Se inicializa en 0.1 y varía en el rango [0.01, 0.001, 0.0001, 0].

4.2 Resultados del menú del Desayuno

Para el caso del “Desayuno” se observa que el algoritmo propuesto alcanza una solución aproximada a la solución óptima (9,58 gr de grasa) presentando un error de 0,03 siendo altamente aceptable. Se observa en la **Figura 3** las cantidades de cada alimento, señalado en la **Tabla 1**, que se deben incluir en el desayuno, para que los macronutrientes aportados sean los adecuados. Se aproxima satisfactoriamente a los valores recomendados de la información calórica en **Figura 4** y con una distribución

del 61% de Hidratos de carbono, 14% de proteínas y 24% de grasas representados en la **Figura 5**.

Resultados de la Optimización		
(valor óptimo) Z = 9.61125 (grs. en Grasas)		
Variable	Alimento	Cantidad (grs.)
X1	Leche en polvo	31.9322
X2	Chocolate	10
X3	Azucar	9.7586
X4	Pan	50

Figura 3. Resultados del algoritmo ABC para el desayuno. Cantidades en gr de cada ingrediente a utilizar en el menú.

Información calórica		
Calorías totales =	357.494	(cal.)
Hidratos de Carbono =	219.232	(cal.)
Proteínas =	51.7604	(cal.)
Grasas =	86.5021	(cal.)

Figura 4. Resultados de la información calórica del algoritmo ABC para el desayuno, para el cual se consideraba un aporte calórico de aprox. 345 cal.

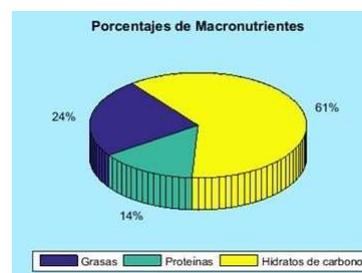


Figura 5. Resultados de la distribución de macronutrientes para el desayuno.

4.3 Resultados del menú del Almuerzo

Para el caso del “Almuerzo” el algoritmo se aproxima a la solución óptima (25,5 gr de grasa) con una diferencia de 1,05 gr., que al igual que en el desayuno no influye significativamente en los nutrientes del menú. En la **Figura 6** calcula las cantidades de cada alimento seleccionada en la **Tabla 3** para el almuerzo. La información calórica alcanzada se observa en **Figura 7**, con una distribución del 60% de Hidratos de carbono, 15% de proteínas y 25% de grasas visualizadas en **Figura 8**.

Resultados de la Optimización

(valor óptimo) Z = (grs. en Grasas)

Variable	Alimento	Cantidad (grs.)
X1	Acete	10
X2	Arroz	59.9966
X3	Carne vacuna	80
X4	Cebolla	30
X5	Lenteja	56.5654
X6	Tomate	40.0266
X7	Azúcar	29.9515
X8	Pera	130
X9	Pan	20.0009

Figura 6. Resultados del algoritmo ABC para el Almuerzo. Cantidades en gr de cada ingrediente a utilizar en el menú.

Información calórica

Calorías totales = (cal.)

Hidratos de Carbono = (cal.)

Proteínas = (cal.)

Grasas = (cal.)

Figura 7. Resultados de la información calórica del algoritmo ABC para el Almuerzo, para el cual se consideraba un aporte calórico de aprox. 920 cal.

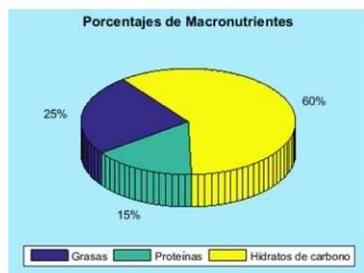


Figura 8. Resultados de la distribución de macronutrientes para el Almuerzo

5. FORMACION DE RECURSOS HUMANOS

El desarrollo forma parte de un conjunto de trabajos finales de características similares. En este caso se encuentra conformado por: 2 estudiantes de Licenciatura en Sistemas y 2 Licenciadas en Sistemas.

Dirección de tesis

Tesis de Grado. Título: “Sistema de optimización aplicado a la elaboración de productos alimenticios mediante la técnica de enjambres de partículas (PSO)”. *Tesista:* Fabiola Patricia Paz. Carrera de Licenciatura en Sistemas, Facultad de Ingeniería – UNJu.

Tutores: M.A. Azar y A. Herrera Cognitiona. Finalizada 2014.

Tesis de Grado. Título: “Prototipo de una app para optimizar el recorrido de la recolección de residuos mediante la metaheurística colonias de hormigas (ACO)”. *Tesista:* Franco Gastón Figueroa. Carrera de Licenciatura en Sistemas, Facultad de Ingeniería – UNJu. *Tutores:* F.P. Paz, A. Herrera Cognitiona y M. A. Azar. En curso desde 2016.

6. BIBLIOGRAFIA

- [1] Espínola V. y Brünner M., «Alimentación saludable en la escuela. Hacia una adecuada alimentación de los niños en edad escolar,» Ciudad Autónoma de Buenos Aires, Argentina, 2014.
- [2] Karaboga, D. An idea based on honey bee swarm for numerical optimization (Vol. 200). Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, 2005.
- [3] Liang, J. J., Runarsson, T. P., Mezura-Montes, E., Clerc, M., Suganthan, P. N., Coello, C. C., & Deb, K. (2006). Problem definitions and evaluation criteria for the CEC 2006 special session on constrained real-parameter optimization. *Journal of Applied Mechanics*, 2006.
- [4] Karaboga D., Akay N., A comparative study of Artificial Bee Colony algorithm, *Applied Mathematics and Computation*, no 214, pp. 108-132, 2009.
- [5] Justo A., «Un algoritmo basado en la colonia artificial de abejas con búsqueda local para resolver problemas de optimización con restricciones,» Veracruz, México, 2014.
- [6] Araoz M., «Uso de tolerancias dinámicas para el Manejo de Restricciones en problemas de optimización usando la Colonia Artificial de Abejas,» Veracruz, México, 2013.
- [7] Deb K., «An efficient constraint handling method for genetic algorithms,» *Computer Methods in Applied Mechanics and Engineering*, vol. 186, pp. 311-338, 2000.
- [8] Paz F., Azar M., Herrera Cognitiona A., Optimización por enjambre de partículas con restricciones aplicada a la elaboración de productos alimenticios, Segundo Congreso Argentino de Ingeniería – CADI 2014, Tucumán, Argentina, ISBN 978-987-1662-51-7, Septiembre 2014.
- [9] Paz F., Azar M., Herrera Cognitiona A., Pérez Otero N., Una alternativa para el mecanismo de manejo de restricciones en algoritmos PSO, Segundo Congreso Argentino IPCTIIC 2013, Córdoba, Argentina, ISBN 978.88.96.471.25.8, Noviembre 21-22, 2013.

Búsqueda de Entorno Variable (VNS) para el Problema de Planificación de Máquinas Paralelas Idénticas

Gatica Claudia Ruth, Esquivel Susana Cecilia

LIDIC - Departamento de Informática- FCFMyN Universidad Nacional de San Luis-
Ejército de los Andes 950 – Local 106, Tel: (266) 4420823, crgatica,esquivel@unsl.edu.ar

Resumen

En el presente trabajo se estudio un algoritmo de Búsqueda de Entornos Variables (VNS) básico en el cual el ordenamiento o secuencia de las estructuras de entornos fueron provistas por el método de cuadrados latinos conocido como LHS [3] y seleccionadas mediante un análisis estadístico basado en el *ranking* del *test* de Friedman [2]. Dos variantes de VNS básico: VNS-R1 y VNS-R2 en los cuales las estructuras de entornos fueron basadas en un ordenamiento aleatorio para el problema de planificación (*scheduling*) de máquinas paralelas idénticas sin restricciones fueron contrastadas.

Hemos estudiado tal problema a través de cuatro funciones objetivo: la máxima tardanza (*Maximum Tardiness: T_{max}*), la tardanza media (*Average Tardiness: T_{avg}*), la tardanza pondera total (*Total Weighted Tardiness : T_{wt}*) y el número ponderado de tareas tardías (*Weighted Number of Tardy Jobs: N_{wt}*).

El conjunto de *benchmarks* usado se divide en ocho escenarios de 125 instancias cada uno. Tales instancias del problema se construyeron en base a datos seleccionados de la *OR-Library* [6] correspondientes a problemas de tardanza ponderada. Se obtuvieron los valores de los óptimos conocidos o *benchmarks* mediante la aplicación de reglas de despacho y heurísticas conocidas en la literatura [7, 8].

Palabras clave: *Scheduling*, Máquinas Paralelas Idénticas, algoritmo de búsqueda en Entornos Variables (VNS).

Contexto

La línea de investigación se enmarca en el paradigma de “Metaheurísticas de Trayectoria”, en el proyecto de investigación: “Diseño de distribución de turbinas en parques eólicos usando métodos de optimización aproximados”. Tal proyecto esta en ejecución desde el 1 de enero de 2014 e inserto dentro del Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC).

1. Introducción

El problema en estudio es la planificación (*scheduling*) de máquinas paralelas idénticas con respecto a cuatro funciones objetivo: T_{max} , T_{avg} , T_{wt} y N_{wt} . La notación usada en la literatura [7, 8] para describir el problema es una tripla: $(\alpha | \beta | \delta)$. El primer campo α describe el ambiente de máquinas, el segundo campo β indica las restricciones entre las tareas que son asignadas a las máquinas. Por último el tercer campo δ provee la o las funciones objetivo a ser optimizadas. Acorde a esta notación, nuestro problema se describe mediante $(P_m | | \delta)$, el campo δ representa las funciones objetivo mencionadas. Tal problema se ha considerado en [8] de complejidad NP-duro cuando $2 \leq m \leq n$ (m es el número de máquinas y n el número de tareas).

2. Líneas de Investigación y Desarrollo

La motivación de nuestro trabajo es la observación de suponer que el ordenamiento de las estructuras de entorno es importante para el desempeño del algoritmo VNS. En implementaciones estándares, el orden en que se exploran los entornos se fija comúnmente de ante mano y es un asunto bastante crítico, pero una política usada comúnmente es seleccionar los entornos de acuerdo a su complejidad.

En la revisión de la literatura, el enfoque de búsqueda de entornos variables VNS y sus variantes han sido ampliamente estudiados y aplicados para resolver varios problemas de planificación y en especial para el caso de máquina única. En este contexto, un método VNS fue propuesto por Liao y Chen [4]. En otros trabajos relacionados [5], los autores proponen una nueva estrategia para seleccionar de forma iterativa el entorno en el enfoque de Entorno Variable Descendente (VND). La idea básica en su estrategia fue, dado un conjunto de estructuras de k entornos N_1, N_2, \dots, N_k , se asigna un peso w_i a cada uno. El peso w_i refleja el rendimiento relativo de cada entorno durante un proceso previo a la ejecución del algoritmo. Luego, en un procedimiento iterativo del algoritmo, la selección del entorno N_i que debe aplicarse en cada iteración j está basado en su peso w_{ij} . En [9] y [10] los autores presentan cuestiones, entre otras, como las siguientes:

- Es importante el orden en el cual los entornos pueden ser designados?
- Cuál es el orden en el cual los entornos deberían ser aplicados?

En nuestro trabajo se propuso una nueva estrategia para fijar el orden de exploración de los entornos, que consiste en aplicar el método de cuadrados latinos conocido como LHS [3]. Para ello tomamos una muestra en el espacio de las permutaciones de los entornos y seleccionamos 30 combinaciones de ellos, entonces ejecutamos el VNS básico para estas 30 configuraciones diferentes. Para el análisis

estadístico de los resultados se utilizó un paquete de software denominado CONTROLTEST que provee varios *test* estadísticos de los cuales elegimos el *test* de Friedman [2] para clasificar los resultados, esto es, elegir la mejor configuración c_i y luego realizar comparaciones de a pares y analizar diferencias significativas.

De estos experimentos se eligieron las mejores configuraciones de VNS básico, luego se realizaron 30 ejecuciones de los algoritmos VNS-R1 y VNS-R2 los cuales son una variante del VNS básico y la secuencia de ordenamiento de los entornos está basada en la aleatoriedad, el primero permite entornos repetidos en la secuencia y el segundo no se permiten entornos repetidos en la secuencia. Como último paso se realizó un estudio comparativo mediante el análisis estadístico nuevamente utilizando el *test* de Friedman.

3. Resultados y Objetivos

Los resultados alcanzados fueron estudiados con varias métricas de evaluación propuestas. Los escenarios del conjunto de instancias son los siguientes: **I**: $n = 40$ y $m = 2$; **II**: $n = 40$ y $m = 5$; **III**: $n = 100$ y $m = 2$; y **IV**: $n = 100$ y $m = 5$, **V**: $n=40$ y $m=15$, **VI**: $n=40$ y $m=30$, **VII**: $n=100$ y $m=15$; **VIII**: $n=100$ y $m=30$, donde n es el número de tareas y m es el número de máquinas.

En cada uno de los escenarios se realizaron las comparaciones usando el *ranking* de Friedman los resultados muestran que el entorno VNS basado en la configuración seleccionada mediante LHS es de mejor performance con respecto a los entornos VNS-R1 y VNS-R2. Cuando se observan en comparaciones de a pares los valores de *p-values* ajustados como los *p-values* no ajustados todos son menores a $\alpha \leq 0.05$, lo

que implica que todos presentan diferencia significativa estadísticamente. Salvo en el escenario VI: $n=40$ y $m=30$ con respecto a las funciones T_{max} y N_{wt} no hay diferencias significativa en los resultados de VNS basado en la configuración LHS con respecto a VNS-R1 y de VNS-R2.

A través de los resultados obtenidos podríamos responder a las cuestiones planteadas por distintos autores, afirmando que, salvo en algunas excepciones, el orden usado en las estructuras de entornos es importante en la performance del algoritmo VNS.

4. Formación de Recursos Humanos

El presente trabajo corresponde al plan de trabajo de la tesis de Maestría: “Desarrollo y Aplicación de Metaheurísticas para resolver problemas de Planificación de Máquinas Paralelas Idénticas”, de la Carrera de Posgrado: Maestría en Ciencias de la Computación. El trabajo se divide en dos etapas. La Etapa I involucra la profundización en el desarrollo y aplicación de Metaheurísticas basadas en poblaciones y trayectorias para resolver problemas de planificación de máquina paralela idénticas sin restricciones. La Etapa II comprende la obtención de un conjunto de *benchmarks* para instancias del problema y colocarlos a disposición de interesados en la investigación del problema de planificación (*scheduling*) de máquinas idénticas paralelas vía web.

5. Bibliografía

- [1] E.G. Talbi, "Metaheuristics from design to implementation", by John Wiley & Sons, Canada, 2009.
- [2] Joaquín Derrac and Salvador García and Daniel Molina and Francisco Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms", Swarm and Evolutionary Computation, 2011.
- [3] T. Bartz-Beielstein, “Experimental Research in Evolutionary Computation”, The New Experimentalism, Springer, 2006.
- [4] Liao C. J., Cheng C. C., “A variable neighborhood search for minimizing single machine weighted earliness and tardiness with common due date”, Computers and Industrial Engineering, 52(4), 404-413, 2007.
- [5] Hiba Yahyaoui, Saoussen Krichen, Bilel Derbel, El-Ghazali Talbi, “A Variable Neighborhood Descent for solving the Single Machine Total Weighted Tardiness Problem”, IEEE, 2013.
- [6] J.E. Beasley, ORLIB, <http://people.brunel.ac.uk/mastjjb/jeb/orlib/wtinfo.html>.
- [7] Pinedo M., “Scheduling: Theory, Algorithms and System”, Prentice Hall, 1995.
- [8] Morton T. and Pentico D., “Heuristic Scheduling Systems, John Wiley and Sons”, 1993, New York
- [9] M. J. Geiger, M. Sevaux, Stefan Voß, “Neighborhood Selection in Variable Neighborhood Search”, MIC 2011: The IX Metaheuristics International Conference, 2011.
- [10] Bin Hu and Günther R. Raidl, “Variable Neighborhood Descent with Self-Adaptive Neighborhood-Ordering”, Institute of Computer Graphics and Algorithms Vienna University of Technology Favoritenstraße 9–11/1861, 1040 Vienna, Austria, 2012.

Confianza y Reputación de Agentes en Sistemas Multi-agente para Entornos Dinámicos

Federico Joaquín Luciano H. Tamargo Alejandro J. García

Instituto de Ciencias e Ingeniería de la Computación (UNS–CONICET),
Departamento de Ciencias e Ingeniería de la Computación,
Universidad Nacional del Sur, Bahía Blanca, Argentina.

San Andrés 800, (B8000CPB).

Tel: (0291) 459-5135 / Fax: (0291) 459-5136

e-mail: { federico.joaquin, lt, ajg }@cs.uns.edu.ar

RESUMEN

Esta línea de investigación se enfoca en el área de confianza y reputación de agentes en sistemas multi-agente. Su objetivo general es el análisis, desarrollo y formalización de la dinámica de la confianza y reputación de los agentes, a partir de la interacción con sus pares en el marco de un sistema multi-agente (SMA). Esto involucra el desarrollo y formalización de técnicas de representación y actualización del grado de confianza y de reputación de un agente, y también, la integración de estas técnicas con mecanismos de razonamiento automático y dinámica de creencias. En particular, se espera avanzar en el estudio y desarrollo de formalismos para aplicaciones de naturaleza dinámica y distribuida, que combinen mecanismos de confianza, dinámica de creencias y argumentación.

Palabras claves: *Confianza y reputación, Sistemas Multi-Agente, Dinámica del conocimiento.*

CONTEXTO

Esta línea de investigación se realizará dentro del ámbito del Instituto de Ciencias e Ingeniería de la Computación (ICIC) de doble dependencia entre CONICET y la Universidad Nacional del Sur, y dentro del Departamento de Ciencias e Ingeniería de la Computación de dicha

Universidad. La investigación está asociada con los proyectos de investigación: “Formalismos para el tratamiento de confianza y reputación en sistemas multi-agente” (PGI 24/ZN32) y “Argumentación y dinámica de creencias para mejorar las capacidades de razonamiento y representación de conocimiento en sistemas multi-agente” (PGI 24/N035) ambos financiados por la Universidad Nacional del Sur. Además, el primer autor posee una beca doctoral financiada por CONICET.

INTRODUCCIÓN

En [SS05], Sabater y Sierra sostienen que la importancia de la confianza y reputación en sociedades humanas está fuera de discusión, por lo cual no es sorprendente que varias disciplinas, cada una desde una perspectiva diferente, haya estudiado y utilizado ambos conceptos. La investigación científica en el área de mecanismos computacionales de confianza y reputación en sociedades virtuales, es una disciplina reciente orientada a incrementar la fiabilidad y performance de comunidades electrónicas. En artículos recientes como [Chan10, Bhu10, Jos08, Jos07, SS05, Del03, McKChe02, AM02, eBAY02, Mont02, Mui02, ARH00, GrSlo00, BK01] podemos notar que en ciencias de la computación hay dos elementos que han contribuido sustancialmente a incrementar el interés en confianza y reputación:

el paradigma de SMAs y la evolución creciente del comercio electrónico.

El estudio de confianza y reputación tiene muchas aplicaciones en tecnologías de Comunicación e Información. Estos sistemas han sido reconocidos como factores claves para el éxito de la adopción del comercio electrónico. Los mismos son usados por agentes de software inteligentes como un mecanismo para buscar compañeros confiables y como un incentivo en la toma de decisiones acerca de si se tiene en cuenta un contrato. La reputación es usada en el mercado electrónico como un mecanismo para evitar fraudes y estafas [AM02, eBAY02, Del03]. Los e-markets no son el único campo de aplicación; por ejemplo, en [TGFS14, TGFS12, BK01] usan la confianza para mejorar la performance de mecanismos de revisión de creencias. Otra importante área de aplicación en tecnología de agentes es el trabajo de equipo y cooperación [Mont02]. Además, entre los formalismos de argumentación que se proponen sobre sistemas multi-agente, también hay algunos que usan la información de confianza y reputación para razonar [PSM11, Tang12]. Las investigaciones en confianza y reputación se requieren tanto para ambientes colaborativos como competitivos [LLRJ12, C11], y considerando tanto valores de confianza como valores de desconfianza [HCJ10].

En la actualidad, existen pocos trabajos que estudien una visión general de confianza y reputación desde el punto de vista de ciencias de la computación. Por ejemplo, Dellarocas en [Del03] presenta una visión general de mecanismos de reputación online que son usados en sitios web comerciales. En el área de confianza, Grandison [GrSlo00] examina varias definiciones de confianza que existen en la literatura y proveen una definición de confianza para aplicaciones de Internet. Hay también algunas propuestas que establecen una tipología

de reputación [Mui02] y confianza [McKChe02].

El estudio y desarrollo de agentes inteligentes, y la investigación en el área de SMA, han demandado en los últimos años el estudio y desarrollo de formalismos de representación y mantenimiento de conocimiento. En particular esto ha dado un fuerte impulso al área de Revisión de Creencias [Han99, HFCF01, FKS02] donde se estudia formalmente parte de estos temas. Además, en los últimos años se han producido importantes avances en un área relativamente joven que se enfoca en la Revisión de Creencias en SMA, entre cuyos artículos se pueden destacar [DGB97, Can98, LW99, LW01, TGFS12, TGFS14, TGFS16].

En esta línea de investigación se pretende realizar un aporte en el estudio y desarrollo de nuevos formalismos, técnicas y métodos de representación y actualización del grado de confianza y de reputación de un agente, y su integración con mecanismos de razonamiento automático. Este aporte es de importancia para mejorar las capacidades de los agentes que sean utilizados en aplicaciones concretas que requieran el uso de agentes inteligentes y sistemas multi-agente. Este tipo de aplicaciones son requeridas en la actualidad para la resolución de problemas de mucha complejidad en ciencias de la computación y en otras disciplinas.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación, toma como punto de partida varias publicaciones vinculadas a la misma. Por ejemplo, en el libro [FC10], se realiza un análisis sistemático y acabado de los sistemas computacionales de confianza; en particular, sus autores distinguen lo desarrollado como la visión global del naciente campo de la modelización de confianza y los modelos computacionales de confianza. Se discute la

confianza y el concepto aliado de la reputación de una gama de diferentes orígenes. En particular, se plantean como objetivo poder sistematizar una teoría general de la 'confianza' para proporcionar un modelo orgánico de este fenómeno muy complejo y dinámico en el desarrollo cognitivo, afectivo, y de niveles sociales (interactivos y colectivos).

En el artículo [TGFS12], publicado en *Journal of The Knowledge Engineering Review*, se ha desarrollado una primera propuesta para modelar la dinámica de unabase de creencias de agentes basada en la credibilidad de la información que es recibida de sus pares informantes. En particular, se han definido operadores de revisión priorizados y no priorizados, donde la credibilidad de la información es usada para aceptar o rechazar la información recibida.

En [Tam12], se ha publicado una primera aproximación para utilizar la credibilidad de la información recibida en una extensión de un sistema de argumentación estructurada.

En [TGFS14], publicado en *Artificial Intelligence Journal*, se propuso operadores de expansión, contracción y revisión de órdenes de credibilidad de informantes. En esta propuesta un agente que tiene un orden de credibilidad sobre sus informantes, puede modificar dinámicamente ese orden ante la presencia de nueva información recibida de sus pares.

En [TGFS16], publicado recientemente en la revista *Knowledge and Information Systems*, se realiza una propuesta para compartir información entre agentes, donde cada agente puede tener su propio orden de credibilidad entre sus informantes.

RESULTADOS OBTENIDOS/ESPERADOS

La implementación de aplicaciones que integren adecuadamente mecanismos de confianza y

reputación en sistemas multi-agente, requiere del estudio y desarrollo de nuevos formalismos teóricos. En esta línea de investigación, se consideran varios objetivos específicos para tal fin:

- Proponer un formalismo de actualización de valores de confianza y reputación usando la experiencia generada a partir de la interacción con otros agentes del sistema. Se estudiará en particular cómo un cambio en las creencias de un agente puede afectar los valores de confianza que el mismo mantiene de sus informantes.
- Analizar y definir un formalismo de confianza y reputación multi-contexto. La confianza y la reputación son propiedades dependientes del contexto. Un modelo de confianza y reputación que administra un único contexto, es diseñado para asociar un único valor de confianza o reputación a cada miembro del SMA sin tomar en consideración el contexto. Sin embargo, un modelo de multi-contexto permite la posibilidad de tratar con varios contextos, manteniendo diferentes valores de confianza y reputación asociados a cada contexto particular de un agente.
- Formalizar un modelo de mantenimiento de confianza que, además de considerar los elementos indicados en los objetivos específicos anteriores, permita integrar a ellos el concepto de la opinión de la mayoría. Dado que muchas propuestas de la literatura consideran únicamente a la opinión de la mayoría como método de evaluación, se planea integrar esta propuesta a las propuestas basadas en un orden de desconfianza entre los agentes.
- Evaluar la relación entre mecanismos de mantenimiento de valores de confianza simbólicos y numéricos, y desarrollar una integración de ambas propuestas.

- Analizar los formalismos mencionados tanto en ambientes colaborativos como competitivos, y considerar tanto valores de confianza como valores de desconfianza

FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación forma parte de los trabajos que se realizarán para la concreción de un Doctorado en Ciencias de la Computación.

BIBLIOGRAFÍA

- [AM02] Amazon (2002). Amazon Auctions. <http://auctions.amazon.com>.
- [ARH00] A. Abdul-Rahman and S. Hailes. Supporting trust in virtual communities. In *Proceedings of the Hawaii's International Conference on Systems Sciences (Maui, Hawaii, 2000)*.
- [BK01] Barber, K. S. and Kim, J. (2001). Belief Revision Process based on Trust: Simulation Experiment. In: *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada*. pp.112.
- [Bhu10] Touhid Bhuiyan, Yue Xu, Audun Josang. A Review of Trust in Online Social Networks to Explore New Research Agenda. *International Conference on Internet Computing*, pp. 123-128, 2010.
- [C11] Cholvy L (2011) How strong can an agent believe reported information? 11th European conference, ECSQARU 2011, Belfast, UK, June 29–July 1, 2011, pp 386–397
- [Can98] Cantwell, J. 1998. Resolving conflicting information. *Journal of Logic, Language and Information* 7(2):191–220.
- [Chan10] Changyu Dong, Naranker Dulay. Shinren: Non-monotonic Trust Management for Distributed Systems. *IFIPTM, 2010*. pp. 125-140.
- [Del03] Dellarocas, C. 2003. The digitalization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms. *Management Science*.
- [DGB97] Dragoni, A.F.; Giorgini, P.; and Baffetti, M. 1997. Distributed belief revision vs. belief revision in a multi-agent environment: First results of a simulation experiment. In *MAAMAW*, 45–62.
- [eBAY02] eBay (2002). eBay. <http://www.eBay.com>.
- [FC10] R. Falcone. C. Castelfranchi. Trust theory: A socio-cognitive and computational model, 2010. Wiley Ed. ISBN-13: 978-0470028759
- [FKS02] M. A. Falappa, G. Kern-Isberner, G. R. Simari. Explanations, Belief Revision and Defeasible Reasoning. *Artificial Intelligence Journal*, 141: pp. 1-28, 2002.
- [GrSl00] Grandison, T. & Sloman, M. (2000). A survey of trust in Internet application, *IEEE, Communications Surveys, Fourth Quarter, 2000*.
- [Han99] Hansson, S. O. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer 1999.
- [HCJ10] Harwood WT, Clark JA, Jacob JL (2010) Networks of trust and distrust: towards logical reputation systems. In: Gabbay DM, van der Torre L (eds) *Logics in Security*, Copenhagen, Denmark.
- [HFCF01] S. O. Hansson, E. L. Fermé, J. Cantwell, M. A. Falappa. Credibility Limited Revision. *The Journal of Symbolic Logic*, Volume 66, Number 4: pp. 1581-1596, 2001.
- [Jos07] A. Josang, R. Ismail, C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, Volume 43, Number 2: pp. 618-644, 2007.

- [Jos08] Audun Josang, Touhid Bhuiyan, Yue Xu, Clive Cox. Combining Trust and Reputation Management for Web-Based Services. *TrustBus*. 2008. pp. 90-99.
- [McKChe02] McKnight, D. H. & Chervany, N. L. (2002). Notions of Reputation in Multi-Agent Systems: A Review. In: *Proceedings of the 34th Hawaii International Conference on System Sciences*.
- [Mont02] Montaner, M., Lopez, B. & de la Rosa, J. (2002). Developing Trust in Recommender Agents. In: *Proceedings of the First International Joint Conference on Autonomous Agents and Multi agent Systems (AAMAS-02)*, Bologna, Italy. pp.304-305.
- [LLRJ12] Luke Teacy WT, Luck M, Rogers A, Jennings NR (2012) An efficient and versatile approach to trust and reputation using hierarchical bayesian modelling. *Artif Intell* 193:149–185
- [LW99] Liu, W., and Williams, M. 1999. A framework for multi-agent belief revision, part i: The role of ontology. In *Australian Joint Conference on A. I.*, 168–179.
- [LW01] Liu, W., and Williams, M. 2001. A framework for multi-agent belief revision. *Studia Logica* 67(2):291–312.
- [Mui02] Mui, L., Halberstadt, A. & Mohtashemi, M. (2002). Notions of Reputation in Multi-Agent Systems: A Review. In: *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, Bologna, Italy. pp.280-287.
- [PSM11] Parsons S., Sklar E., McBurney P. (2011) Using argumentation to reason with and about trust. In: *Argumentation in multi-agent systems—8th international workshop, ArgMAS 2011, Taipei, May 3, 2011, Revised Selected Papers*. pp 194–212.
- [SS05] Sabater, J and Sierra C. 2005. Review on Computational Trust and Reputation Models. *Artificial Intelligence Review* 24:33-60.
- [Tang12] Tang Y, Cai K, McBurney P, Sklar E, Parsons S (2012) Using argumentation to reason about trust and belief. *J Log Comput* 22(5):979–1018.
- [TGFS12] Tamargo, L. H., Garcia, A. J., Falappa, M. A., and Simari, G. R. Modeling knowledge dynamics in multi-agent systems based on informants. *Journal of The Knowledge Engineering Review (KER)* 27 (1) (2012) 87-114.
- [Tam12] Luciano H. Tamargo, Sebastián Gottifredi, Alejandro J. García, Marcelo A. Falappa and Guillermo R. Simari. Deliberative DeLP agents with multiple informants. *Journal Inteligencia Artificial v 15 nro 49* pp 13-30. *Asociación Española de Inteligencia Artificial (AEPIA)*. ISSN 1137-3601 (2012).
- [TGFS14] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa and Guillermo R. Simari. On the revision of informant credibility orders. *Artificial Intelligence Journal (AIJ)*. Vol. 212 (2014), pp: 36-58. ISSN: 0004-3702. ELSEVIER SCIENCE BV. doi:10.1016/J.ARTINT.2014.03.006.
- [TGFS16] Luciano H. Tamargo, Sebastián Gottifredi, Alejandro J. García and Guillermo R. Simari. Sharing beliefs among agents with different degrees of credibility. *Knowledge and Information Systems (KAIS)*. Vol. 47. Num. 43. (2016), pp: 1-33. SPRINGER. ISSN: 0219-1377.

Conocimiento Compartido y Razonamiento Argumentativo Colaborativo para Entornos de Múltiples Agentes en Ambientes Distribuidos

Ramiro A. Agis

Sebastian Gottifredi

Alejandro J. García

Instituto de Ciencias e Ingeniería de la Computación (UNS–CONICET),
Departamento de Ciencias e Ingeniería de la Computación,
Universidad Nacional del Sur, Bahía Blanca, Argentina
{ramiro.agis, sg, ajg}@cs.uns.edu.ar

1. Resumen

Los Sistemas Multi-Agente constituyen un área en continuo crecimiento para el desarrollo de aplicaciones comerciales e industriales de gran escala ya que proveen de manera más natural soluciones a problemas complejos. En este tipo de sistemas, cada agente tiene capacidades limitadas e información incompleta sobre su entorno. Dicha información puede estar en contradicción con la información de otros agentes del sistema, y la resolución de este tipo de conflictos no es trivial. Esta línea de investigación se enfoca en mejorar las capacidades de razonamiento, representación de conocimiento, e interacción de agentes que participan en Sistemas Multi-Agente, los cuales colaboran y comparten su conocimiento en entornos dinámicos.

Palabras clave: Sistemas Multi-Agente, Representación de Conocimiento y Razonamiento, Argumentación.

2. Contexto

La presente línea de investigación se encuentra inserta en el marco de los proyectos PGI 24/ZN32 “*Formalismos para el tratamiento de confianza y reputación en Sistemas Multi-Agente*” y PGI 24/N035 “*Argumentación y Dinámica de Creencias para mejorar las capaci-*

dades de razonamiento y representación de conocimiento de Sistemas Multi-Agente”, financiados por la Universidad Nacional del Sur (UNS), ambos llevados a cabo dentro del Departamento de Ciencias e Ingeniería de la Computación (DCIC), UNS.

3. Introducción

Un agente es una entidad computacional autónoma, que percibe su entorno a través de sensores y actúa en ese entorno utilizando efectores. Decir que es autónomo significa que tiene algún tipo de control sobre su propio comportamiento y que puede actuar sin la intervención de otros agentes o humanos. Actualmente los agentes tienen un campo de aplicación muy amplio y existen muchos tipos de agentes diferentes (por ejemplo: reactivos, deliberativos, inteligentes, de interface, colaborativos, etc.) los cuales a su vez están orientados a distintos entornos de aplicación. Los agentes involucran aportes de varias áreas de Inteligencia Artificial como Resolución de Problemas Distribuidos, e Inteligencia Artificial Paralela. Es por esto, que los agentes heredan: modularidad, velocidad (gracias al paralelismo), confiabilidad (gracias a la redundancia), fácil mantenimiento, reusabilidad e independencia de la plataforma.

En un Sistema Multi-Agente (SMA), [DRM05], varios agentes interactúan para conseguir algún objetivo o realizar alguna tarea común. En este tipo de sistemas, cada agente tiene información incompleta y capacidades limitadas, el control del sistema es distribuido, los datos están descentralizados, y la computación es asincrónica. Además, los agentes se desenvuelven en un entorno dinámico y cambiante, el cual no puede predecirse y se ve afectado por las acciones que son llevadas a cabo por los agentes y también por humanos. Por lo tanto, todo lenguaje de especificación o implementación de agentes debe considerar primitivas para la interacción [GGS09]. Debido a las características enunciadas antes, los SMA constituyen un área en continuo crecimiento para el desarrollo de aplicaciones comerciales e industriales de gran escala. Esto se debe principalmente a que proveen de manera más natural soluciones a problemas complejos.

En un SMA los distintos agentes pueden percibir o inferir información diferente (y potencialmente contradictoria e incompleta) sobre el entorno en el que se encuentran. El razonamiento colaborativo consiste en que los agentes – aprovechando sus capacidades de representación de conocimiento y razonamiento – puedan combinar entre todos dicha información para realizar nuevas inferencias difíciles de realizar individualmente, con el objetivo de resolver colaborativamente problemas complejos.

La argumentación constituye un área de estudio de especial interés en el ámbito de la Inteligencia Artificial (ver por ejemplo [RS09]), principalmente, porque permite razonar con información incompleta e incierta, y permite manejar inconsistencias en los sistemas basados en conocimiento. Este tipo de razonamiento es particularmente atractivo para toma de decisiones, y dentro de la Inteligencia Artificial existe particular interés en abordar este tipo de problemas [HV06, FEGG8]. La argumentación ha evolucionado como un mecanismo atractivo

para formalizar el razonamiento de sentido común [PV02, FGKS11, GRS07]. En la literatura se evidencia un gran desarrollo tanto de la formalización de diferentes *frameworks* de argumentación abstracta [BGG05, NBD08]; como así también de sistemas de argumentación estructurados (o basados en reglas) [AK07, DKT06, GS04].

La programación lógica rebatible (*DeLP* por sus siglas en inglés) es un ejemplo de sistema de argumentación estructurado en el cual se puede representar conocimiento en un programa lógico rebatible, en forma de hechos, reglas estrictas y reglas rebatibles. El mecanismo de razonamiento realiza un análisis exhaustivo que considera argumentos a favor y en contra, con el objetivo de encontrar cuales conclusiones están garantizadas por el programa [GS04]. En el último tiempo, el campo de aplicación de la argumentación se ha expandido velozmente, en gran parte debido a los avances teóricos, pero también gracias a la demostración exitosa de su uso práctico en un gran número de dominios de aplicación, tales como el razonamiento legal [PS02], la ingeniería del conocimiento [TGS09], los sistemas multi-agente [PSJ98, AMP02], y el *e-government* [ABM05]. DeLP permite a los agentes representar información potencialmente contradictoria de forma declarativa, y además provee un mecanismo de inferencia para garantizar las conclusiones inferidas.

El razonamiento colaborativo en SMA es un área con muchas aristas por explorar. Los enfoques actuales todavía tienen falencias por resolver, particularmente en situaciones donde los agentes pueden tener conocimiento contradictorio y/o posiciones en conflicto con otros agentes. Por lo tanto, en esta línea de investigación se busca explotar los beneficios de la argumentación y DeLP para equipar a los agentes con mecanismos que les permitan razonar exitosamente de manera conjunta y colaborativa en diferentes contextos.

4. Líneas de Investigación y Desarrollo

La línea de investigación propuesta se enfoca en mejorar las capacidades de razonamiento, representación de conocimiento, e interacción de agentes que participan en SMA, los cuales colaboran y comparten su conocimiento en entornos dinámicos. En particular, se busca mejorar la capacidad de los agentes para razonar cuando se intenta resolver problemas de manera conjunta y colaborativa, y además, mejorar la capacidad de aprovechar la información que reciben de sus pares.

El objetivo general de esta línea de investigación es el desarrollo y formalización de técnicas de razonamiento avanzadas, combinando programación en lógica rebatible, argumentación, actualización de bases deductivas, y mecanismos computacionales de confianza y reputación. Su aporte está orientado a la problemática asociada con la especificación e implementación de agentes inteligentes deliberativos, en el contexto de su participación en un sistema multi-agente colaborativo.

Entre los objetivos particulares se encuentran:

- Desarrollar y formalizar métodos de razonamiento automático basados en argumentación, para un entorno de múltiples agentes, los cuales pueden colaborar en dicho razonamiento.
- Desarrollar métodos para equipar a los sistemas de argumentación con la posibilidad de tener en cuenta cambios dinámicos.
- Desarrollar nuevas técnicas de mantenimiento de bases de conocimiento donde los agentes puedan recibir información de múltiples informantes y que involucren mecanismos computacionales de confianza y reputación en SMA.
- Extender la programación lógica rebatible para entornos distribuidos donde puedan coexistir múltiples agentes que comparten información.

La importancia de esta investigación radica en el estudio y desarrollo de nuevos formalismos, técnicas y métodos para su aplicación en las áreas de agentes inteligentes y sistemas multi-agente. Estos sistemas se utilizan en la actualidad para la resolución de problemas complejos en ciencias de la computación y en otras disciplinas. Su campo de aplicación ha ido en aumento en los últimos años, fundamentalmente por los desarrollos teóricos y prácticos producidos en el área de inteligencia artificial distribuida. La aplicación de estos sistemas seguirá en aumento, lo cual requerirá el desarrollo de nuevos formalismos teóricos para su futura aplicación.

Un aspecto de la importancia de esta investigación radica en mejorar la capacidad de resolver problemas de manera colaborativa, utilizando agentes deliberativos autónomos, donde se pueda compartir la información y razonar aprovechando las capacidades individuales. Al intentar resolver un problema de manera conjunta, cada uno de los agentes puede brindar su aporte individual de varias maneras: completando la solución con información relevante que solamente este agente dispone; dar razones a favor de una conclusión; dar razones en contra de una conclusión; o también aportando un criterio de preferencia que permita resolver un conflicto entre conclusiones contradictorias.

Pensar en resolver esta clase de problemas simplemente uniendo las bases de conocimiento de los agentes participantes es impracticable por diferentes razones. Por un lado, los agentes deben poder mantener ciertos niveles de privacidad sobre la información que disponen; además, los agentes que tienen grandes volúmenes de información deberían poder restringirse a aportar solamente la información relevante a la cuestión a tratar, y por otro lado, los agentes no solo pueden colaborar con información, sino también con otros elementos como criterios que permitan resolver conflictos entre diferentes argumentos o información contradictoria.

El estudio de agentes en Ciencias de la Computación involucra tanto a agentes de software como a agentes físicos. Los agentes de software se han convertido en una herramienta fundamental para aplicaciones de naturaleza distribuida en entornos dinámicos, y tienen aplicación directa en áreas como robótica cognitiva, comercio electrónico, y asistentes para toma de decisiones. El desarrollo de tecnología de agentes inteligentes permitirá además lograr nuevos avances en áreas fundacionales de las Ciencias de la Computación como sistemas operativos distribuidos, bases de datos distribuidas, y lenguajes de programación en paralelo.

5. Resultados Esperados

Con el fin de cumplir los objetivos antes mencionados, en esta línea de investigación se espera lograr los siguientes resultados a corto plazo:

- Definir una estrategia de intercambio de argumentos que permita a los agentes compartir conocimiento entre sí, y que contemple y resuelva los potenciales conflictos que surjan de dicho intercambio de información.
- Definir una estrategia que permita a los agentes generar reglas a partir del razonamiento colaborativo sobre las diferentes percepciones parciales e individuales del entorno.
- Definir una estrategia de toma de decisiones entre múltiples agentes que utilice argumentación colectiva, basada en confianza y reputación de múltiples informantes.

6. Formación de Recursos Humanos

Dentro de esta línea de investigación se lleva a cabo la tesis de Doctor en Ciencias de la Computación de Ramiro A. Agis, bajo la direc-

ción de Alejandro J. García y Sebastian Gottifredi, en desarrollo dentro del Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) perteneciente al Instituto de Ciencias e Ingeniería de la Computación (ICIC), instituto de doble dependencia de la Universidad Nacional del Sur y CONICET. Actualmente, el LIDIA cuenta con investigadores, becarios y estudiantes de posgrado trabajando intensamente en las áreas de Razonamiento bajo Incertidumbre e Inconsistencia, Web Semántica, Razonamiento sobre Preferencias, Robótica Cognitiva, Argumentación Rebatible, Revisión de Creencias y Sistemas Multi-Agente.

7. Bibliografía

- [AK07] L. Amgoud, S. Kaci, *An argumentation framework for merging conflicting knowledge bases*, Int. J. Approx. Reasoning 45(2): 321-340, 2007.
- [AMP02] L. Amgoud, N. Maudet, S. Parsons, *An argumentation-based semantics for agent communication languages*, in: Proc. of the 15th. ECAI, Lyon, France, 2002, pp. 38-42.
- [ABM05] K. Atkinson, T. J. M. Bench-Capon, P. McBurney, *Multi-agent argumentation for e-democracy*, in: Proceedings of the Third European Workshop on Multi-Agent Systems, Brussels, Belgium, Koninklijke Vlaamse Academie, 2005, pp. 35-46.
- [BGG05] P. Baroni, M. Giacomin, G. Guida, *Scn-recursiveness: a general schema for argumentation semantics*, Artificial Intelligence, 168 (1-2), 2005, 162-210.
- [DRM05] Dastani, M.; van Riemsdijk, M. B.; and Meyer, J.-J. C. 2005. *Programming multi-agent systems in 3apl*. In Multi-Agent Programming. 39-67.
- [DKT06] P. M. Dung, R. A. Kowalski, F. Toni, *Dialectic proof procedures for assumption-based, admissible argumentation*, Artificial Intelligence, 170(2): 114-159, 2006.

- [FGKS11] M. Falappa, A. García, G. Kern-Isberner, G. Simari. *On the evolving relation between Belief Revision and Argumentation*. The Knowledge Engineering Review Journal v26:1, pp. 35-43, Cambridge University Press, 2011.
- [FEGG8] E. Ferretti, M. Errecalde, A. J. García, G. R. Simari, *Decision Rules and Arguments in Defeasible Decision Making*. In the 2nd International Conference on Computational Models of Argument (COMMA'08), Toulouse, France, 2008.
- [GRS07] A. J. García, N. D. Rotstein, G. R. Simari, *Dialectical explanations in defeasible argumentation*, In Khaled Mellouli, editor, ECSQARU, volume 4724 of Lecture Notes in Computer Science, pages 295–307, Springer, 2007.
- [GS04] A. J. García, G. R. Simari, *Defeasible logic programming: An argumentative approach*, Journal of Theory and Practice of Logic Programming, 4 (1), 2004, pp. 95-138.
- [GG09] S. Gottifredi, A. J. García, G. R. Simari, *Argumentation Systems and Agent Programming Languages*, In AAI Fall Symposium: The Uses of Computational Argument, Washington D.C., USA, 2009.
- [HV06] D. L. Hitchcock, B. Verheij (eds.), *Arguing on the Toulmin Model*. New Essays in Argument Analysis and Evaluation (Argumentation Library, Volume 10), Dordrecht: Springer-Verlag, 2006.
- [MH02] Mui, L. Halberstadt, A. and Mohtashemi, M. *Notions of Reputation in Multi-Agent Systems: A Review*. In: First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02). pp. 280-287, 2002.
- [NBD08] F. S. Nawwab, T. J. M. Bench-Capon, P. E. Dunne, *A Methodology for Action-Selection using Value-Based Argumentation*, COMMA 2008: 264-275.
- [PSJ98] S. Parsons, C. Sierra, N. Jennings, *Agents that Reason and Negotiate by Arguing*, Journal of Logic and Computation 8, 1998, pp. 261-292.
- [PS02] H. Prakken, G. Sartor, *The role of logic in computational models of legal argument - a critical survey*, in: A. Kakas, F. Sadri (eds.), Computational Logic: Logic Programming and Beyond, Springer, 2002, pp. 342-380.
- [PV02] H. Prakken, G. Vreeswijk, *Logical Systems for Defeasible Argumentation*, in: D. Gabbay, F. Guenther (eds.), Handbook of Philosophical Logic, Kluwer Academic Publishers, 2002, pp. 219-318.
- [RS09] I. Rahwan, G. Simari (eds.), *Argumentation in Artificial Intelligence*, Springer-Verlag, 2009, ISBN 978-0-387-98196-3.
- [TGKS8] M. Thimm, A. J. Garcia, G. Kern-Isberner, G. R. Simari. *Using Collaborations for Distributed Argumentation with Defeasible Logic Programming*. Proc. of the Twelfth Int. Workshop on Non-Monotonic Reasoning (NMR'08), pp 179-188. 2008.
- [TGS09] M. Tucat, A. J. García, G. R. Simari, *Using Defeasible Logic Programming with Contextual Queries for developing Recommender Servers*, In AAI Fall Symposium: The Uses of Computational Argument, Washington D.C., USA, 2009.

Desarrollo de Sistemas de Análisis de Texto

Julio Castillo¹, Marina Cardenas¹, Adrián Curti¹, Osvaldo Casco¹
 Martín Navarro¹, Nicolás Hernández¹, Melisa Velazco¹

¹Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
 { jotacastillo, ing.marinacardenas}@gmail.com

Resumen

En este proyecto se busca utilizar técnicas de aprendizaje automático para analizar y procesar textos que pueden estar en formato estructurado como no estructurado. Se han desarrollado un conjunto de herramientas que pueden ser utilizadas en el área de computación lingüística para diversos fines, entre los que se encuentran construcción de material de entrenamiento, procesamiento de datos estructurados y detección de similitudes entre fragmentos de textos.

Los problemas que se abordan en este proyecto son, entre otros, reconocimiento de implicación de textos e identificación de paráfrasis.

En este artículo se presenta la línea de investigación en la que se encuentra el proyecto, y se describen tres herramientas desarrolladas en el mismo.

Palabras clave: análisis de texto, extracción de información, corpus.

Contexto

El presente proyecto denominado Análisis de Texto (ADT) es un proyecto homologado por la SCyT de la UTN, y se enmarca dentro del área de computación lingüística. El mismo es desarrollado en el Laboratorio de Investigación de Software LIS¹ del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC).

¹ www.investigacion.frc.utn.edu.ar/mslabs/

A su vez, este proyecto se encuentra dentro del grupo de investigación denominado Grupo de Inteligencia Artificial (o GIA) de la UTN-FRC.

Este grupo nuclea proyectos de una línea de investigación relacionada al área de inteligencia artificial, redes neuronales artificiales, autómatas celulares, análisis y procesamiento de imágenes, minería de datos, y su aplicabilidad a la resolución de problemas de las ciencias naturales y de las ciencias sociales. El grupo se conforma por doctores, ingenieros, licenciados, becarios y pasantes.

En el mismo se investigan temas de ciencia básica, como puede ser el estudio de los momentos de aprendizaje de redes neuronales artificiales y su relación con autómatas celulares, como así también, se estudian aspectos de ciencia aplicada, como la estimación del cálculo del riesgo de la vivienda urbana para la salud, que se aplica concretamente en el campo de ciencias sociales.

1. Introducción

Mediante este proyecto se propone abordar el problema del análisis e interpretación de textos no estructurados, extracción de información [1] y minería de datos [2] basados en técnicas de aprendizaje por computadora [3][4][5], en especial las basadas en redes neuronales artificiales [6][7], máquinas kernel [8], y árboles de decisión entre otras [9]. Así, la línea de investigación de aprendizaje automático por computadora es otra de las líneas que intervienen y dirigen el proyecto de investigación.

En el marco de este proyecto se han desarrollado, y se están desarrollando actualmente, varios sistemas de análisis y procesamiento de texto, entre los que se destacan:

- Software de Asistente de Creación de Corpus: es un software que permite construir material de entrenamiento para aplicaciones de minería de datos sobre texto no estructurado.
- Sistema de Mapeo de Datos: Software que permite manipular orígenes de datos estructurados y centralizarlos para un posterior análisis con técnicas de recuperación de información o de minería de datos.
- Sistema de detección de similitudes en archivos de código fuente: Es un sistema que permite analizar archivos de código fuente escritos en diferentes lenguajes de programación e informar el grado de similitud entre los mismos.

El Software de Asistente de Creación de Corpus (ACC) se desarrolla con el objetivo de facilitar la construcción de material de entrenamiento que se necesita en los algoritmos de aprendizaje supervisado. La calidad y el tamaño del conjunto de entrenamiento impactan directamente en la efectividad del algoritmo de clasificación. Adicionalmente, el programa permite acelerar el tiempo necesario para la confección del material de entrenamiento, como así también brinda trazabilidad respecto de los expertos humanos que contribuyeron a cada parte del corpus. Esto permite establecer métricas y calcular la confianza del material de entrenamiento construido.

Entre las aplicaciones que podrían utilizar el material construido podemos señalar a traducción automática asistida por computador, creación de corpus de paráfrasis, creación de corpus para

implicación de textos, resumen automático, entre otras posibles aplicaciones.

El otro sistema es denominado Sistema de Mapeo de Datos (SMD) y se desarrolla con el objetivo de realizar la manipulación y procesamiento desde diferentes fuentes y orígenes de datos y almacenarlos en un repositorio común y centralizado mediante una base de datos en un motor SQL Server.

De esta manera, es posible obtener datos estructurados desde diferentes orígenes y registrarlos en un nuevo repositorio normalizado que permite facilitar el proceso de análisis y búsqueda sobre textos.

Un último sistema que se está desarrollando se denomina sistema de detección de similitudes en archivos de código fuente (SDS), y tiene como objetivo proporcionar métricas de similitud textual entre programas escritos en diversos lenguajes de programación.

Estas métricas principalmente son realizadas a nivel léxico y sintáctico.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de investigación en las que se enmarca el proyecto de análisis de texto es el área de inteligencia artificial, específicamente la sub-área de lingüística computacional en las que se desarrollan aproximaciones a las problemáticas de extracción de información, paráfrasis y minería de datos en textos.

Algunas de las técnicas utilizadas son basadas en redes neuronales artificiales, y otras en árboles de decisión, entre otros algoritmos de aprendizaje supervisado utilizados.

En este contexto, se han desarrollado varios productos software que abordan una de las principales problemáticas de esta línea de investigación, que incluyen la construcción de corpus para diversas aplicaciones de análisis de texto, análisis, clasificación de información almacenada en formato estructurado para su posterior consulta y procesamiento, y software de

detección de similitudes de fragmentos comunes en archivos detexto.

La innovación se ve reflejada en los nuevos algoritmos que se elaboran para realizar el procesamiento de textos, muchos de ellos son utilizados en las herramientas desarrolladas, mientras que otros podrán ser utilizados posteriormente gracias al uso de los sistemas de análisis de texto expuestos en este artículo. Actualmente, se están desarrollando algoritmos de reconocimiento de paráfrasis que hacen uso del etiquetado de la información lingüística que se registra mediante el software de asistente de creación de corpus.

En la siguiente sección se presentan los productos software desarrollados hasta el momento.

3. Resultados

En esta sección se presenta una descripción de las herramientas realizadas explicando su motivación y resultados alcanzados hasta el momento.

3.1 Sistema de Mapeo de Datos:

Consiste de una aplicación de escritorio junto con una aplicación web, que permite manipular múltiples orígenes de datos estructurados, tales como archivos correspondientes a información contenida en diversos motores de bases de datos, y concentrarlos en un repositorio común (una base de datos sql-server), para poder aplicar posteriormente técnicas de análisis de textos.

Su construcción está motivada por la necesidad de procesar y analizar información distribuida en repositorios de datos diferentes. Como ventaja destacamos la centralización de la información, y como desventaja observamos que este procedimiento debe realizarse en diferentes instantes de tiempo a los efectos de mantener actualizada la información en el repositorio.

La Figura 1 muestra la interfaz principal de la aplicación de mapeo de datos.



Figura 1. Pantalla de Mapeo de Datos

3.2 Asistente de Creación de Corpus:

Este software permite la lectura y edición de corpus en los formatos provistos por el NIST (National Institute of Standards and Technology)² y por el CLEF (Cross Evaluation Language Forum)³. Al mismo tiempo, es posible realizar un etiquetado adicional sobre estos corpus agregando información lingüística.

También permite trabajar con corpus multilingües, es decir con pares de texto en idiomas diferentes.

También es posible realizar la clasificación de un conjunto de fenómenos de origen léxico, sintáctico, semánticos y morfológicos, mediante la selección de subcadenas entre las que se sostiene un determinado fenómeno lingüístico. Esto es especialmente útil en tareas como la implicación de textos o en la detección de paráfrasis.

En la figura 2 puede observarse la interfaz principal de esta herramienta que permite realizar una construcción semiautomática de corpus para facilitar la tarea de los anotadores humanos.

² www.nist.gov

³ <http://www.clef-initiative.eu/>

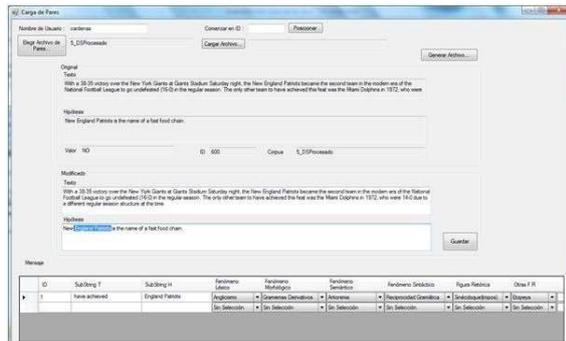


Figura 2. Asistente de Creación de Corpus

3.3. Sistema de detección de similitudes:

Se está desarrollando una herramienta software con el propósito de medir similitudes entre archivos de código fuente entre diferentes lenguajes de programación.

La misma consta de una interfaz gráfica que facilita la inspección de múltiples archivos, y ayuda a encontrar aquellos archivos con mayores similitudes. Por el momento las similitudes son evaluadas a nivel léxico-sintáctica.

Este software hace uso de diferentes herramientas de detección de similitudes y las resume en porcentajes. A su vez permite la comparación exhaustiva entre un conjunto de archivos y brinda un ranking de similitud que puede ser utilizado en la reutilización de códigos fuentes, y en la detección de uso de patrones en el código fuente, entre otras aplicaciones.

Se prevé continuar ampliando y mejorando este sistema, migrándolo a un entorno web y reduciendo el tiempo de procesamiento de los archivos utilizando técnicas de paralelismo.

La Figura 3 muestra la interfaz principal del sistema de detección de similitudes en archivos de código fuente.



Figura 3. Software de detección de similitud entre archivos fuente

4. Formación de Recursos Humanos

El equipo de investigación y desarrollo de software, está formado por docentes investigadores de la Universidad Tecnológica Nacional, Facultad Regional Córdoba, que a continuación se detallan:

- Actualmente el Dr. Julio Castillo está guiando a becarios de grado y de posgrado, como así también dirección de prácticas profesionales supervisadas y pasantías.
- Así mismo la Mg. Marina Cardenas está evaluando la posibilidad de desarrollar su tema de tesis de doctorado (en Ingeniería en Sistemas en la Universidad Tecnológica Nacional- FRC) en la misma temática con una variación del enfoque desde el punto de vista de los sistemas de Generación del Lenguaje Natural (NLG). Adicionalmente, realiza la guía de becarios de grado y de posgrado.
- También participan alumnos que realizan su práctica supervisada como parte de los requisitos para la obtención del grado de Ingeniero, haciendo aportes en el proyecto.
- Año tras año se capacita y forma a alumnos becarios que participan y aprenden desarrollando diversas tareas en el proyecto de investigación, lo que permite complementar su formación curricular desde el punto de vista científico.

Referencias

[1] Judith K lavans y Philip Resnik. The Balancing Act. Combining Symbolic and Statistical Approaches to Language. MIT Press, 1996.

[2] C. Manning y H. Schutze. Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge, MA, 1999.

[3] Castillo J. Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task. TAC, 2009.

[4] Castillo J., Cardenas M. Using Sentence Semantic Similarity Based on WordNet in Recognizing Textual Entailment. Iberamia 2010, LNCS, vol. 6433, pp. 366-375, 2010.

[5] Castillo J. Using Machine Translation Systems to Expand a Corpus in Textual Entailment. Proceedings of the Iccetal 2010, LNCS, vol. 6233, pp.97-102, 2010.

[6] Feldman R. y Hirsh H.. Exploiting Background Information in Knowledge Discovery from Text. Journal of Intelligent Information Systems, 1996.

[7] Lewis, D.. Evaluating and optimizing autonomous text classification systems. In Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval. Seattle, US, págs. 246-254, 1995.

[8] Castillo J. An approach to Recognizing Textual Entailment and TE Search Task using SVM. Procesamiento del Lenguaje Natural 44, 139-145, 2010. 4, 2010.

[9] M. Craven y J. Shavlik. Using Neural Networks for Data

Mining. Future Generation Computer Systems, 13, págs. 211-229, 1997.

[10] Stefan Th. Y Anatol Stefanowitsch. Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis, Berlin: Mouton, pág. 117, 2006.

Diseño de Algoritmos Evolutivos Híbridos Optimizados para Biclustering

LÍNEA DE INVESTIGACIÓN

Macarena Anahí Latini, Dra. Rocío Cecchini, Dra. Jessica Andrea Carballido

¹ Instituto de Ciencias e Ingeniería de la Computación Universidad Nacional del Sur - CONICET
Bahía Blanca - Argentina

E-mail: jac@cs.uns.edu.ar

RESUMEN

El objetivo general de esta línea de investigación consiste en diseñar nuevas técnicas computacionales que ayuden a descubrir potenciales conexiones entre datos presentados en forma de matriz pertenecientes a distintos campos de aplicación. Más específicamente, se planea desarrollar una estrategia evolutiva hibridada con búsqueda local especialmente diseñada para biclustering de datos. En tal sentido, se busca desarrollar una herramienta que pueda asistir a investigadores de distintas disciplinas en la inferencia de relaciones entre datos procedentes de grandes volúmenes de información.

CONTEXTO

Esta línea se enmarca en el siguiente proyecto de investigación subsidiado por la UNS:

PGI-UNS Tema: Modelado predictivo en Bioinformática basado en el desarrollo de técnicas de Computación Evolutiva y Aprendizaje Automático. Código 24/N042. Entidad financiadora: Secretaría de Ciencia y Técnica, UNS. Institución de ejecución: DCIC, UNS. Período de vigencia es 01/01/2016-31/12/2019.

El mismo está dirigido por la Dra. Carballido y se encuentra acreditado para el programa de Incentivos.

Además, la línea también es financiada por el siguiente proyecto:

PIP 2013-2015. Tema: Diseño de Modelos Predictivos en Bioinformática basados en técnicas de Minería de Datos y Aprendizaje

Automático. Código: 112-2012-0100471CO. Director: Dr. Ignacio Ponzoni. Entidad financiadora: CONICET. Institución de ejecución: Laboratorio de Investigación y Desarrollo en Computación Científica, UNS. Monto otorgado: \$360.000 (trescientos sesenta mil pesos). Período de ejecución: Iniciado el 25/8/2014, duración: 3 años.

1. Introducción

Más allá de esto, el biclustering es una metodología que tiene gran variedad de posibles aplicaciones, por lo que se espera abordar en esta línea de investigación un aporte también para alguna de ellas. Por nombrar algunas:

- Marketing: búsqueda de grupos de clientes con un comportamiento similar dado una gran base de datos de datos de clientes que contienen sus propiedades y los registros de compra pasados;
- Seguros: identificación de los grupos de usuarios de seguros con un costo promedio de alta demanda; identificación de fraudes;
- Planificación urbanística: identificación de grupos de viviendas de acuerdo al tipo de casa, el valor y la ubicación geográfica;
- Estudios de terremotos: observación de agrupaciones de epicentros sísmicos para identificar zonas peligrosas;
- WWW: clasificación de documentos; la agrupación de datos para descubrir grupos de patrones de acceso similares.

Por último, como objetivo a largo plazo se busca contribuir a la formación de RRHH en Bioinformática, la cual constituye un área científica de vacancia en la Argentina. Asimismo, cabe destacar que esta área está siendo fuertemente impulsada por el LIDeCC¹ a través de actividades de cooperación internacional y ejecución de proyectos bilaterales, y participando en la creación de la Asociación Argentina de Bioinformática y Biología Computacional.

2. Líneas de investigación y desarrollo

Biclustering

Las características comunes principales de las técnicas de clustering se resumen en la búsqueda de conjuntos disjuntos de datos, de tal manera que aquellos datos que se encuentren en un mismo clúster presenten un comportamiento similar frente a todas las columnas de la matriz. Además, cada uno de los datos debe pertenecer a un único clúster (y no a ninguno) al final del proceso. Las técnicas de biclustering [a, b] se presentan como una alternativa más flexible, ya que permiten que las agrupaciones se formen no solo en base a una dimensión, sino que sea posible formar biclusters que contengan datos que presenten un comportamiento similar frente a un subconjunto de las columnas de la matriz. Esta característica es muy importante, ya que aumenta la capacidad de extracción de información a partir de un mismo conjunto de datos, pudiendo ignorar determinadas columnas frente a las cuales un grupo de datos no presenten un comportamiento coherente.

Otro aspecto significativo que diferencia a las técnicas de biclustering frente a las declustering es la forma en que las agrupaciones son hechas, ya que ahora se permite el solapamiento (datos que pueden estar contenidos en varios biclusters a la vez), así como que existan datos que no se hayan incluido en ningún subconjunto. Esta característica aporta más flexibilidad a este tipo de técnicas, ya que no obliga a incluir cada dato en una agrupación determinada, sino que un determinado dato no pertenecerá a ningún

bicluster si su valor no se ajusta a ninguno de los patrones.

Un punto importante consiste en determinar la medida de evaluación que permite analizar la calidad de los biclusters. Existen medidas generales, y se pueden definir medidas ad-hoc de acuerdo al problema particular siendo resuelto. En este sentido, la Dra. Carballido realizó recientemente una revisión teórica de métodos evolutivos para biclustering y las distintas medidas de evaluación utilizadas, la cual fue presentada en [c]. Allí se sugiere como medida de evaluación más completa el Error Virtual [d], ya que puede detectar tanto biclusters que presenten patrones de desplazamiento como de escalado. Otro tipo interesante de biclusters son los biclusters basados en una evolución coherente de sus valores. Algunos algoritmos de biclustering intentan solucionar el problema de encontrar evoluciones coherentes a lo largo de las filas o columnas de una matriz sin tener en cuenta sus valores exactos. Esta forma totalmente distinta de considerar a un bicluster se tiene en cuenta en importantes algoritmos [e, f, g]. Estos serán analizados en una de las aplicaciones propuestas en este plan, donde se buscan perfiles compartidos entre filas de la matriz (que representan expresión de genes, como se verá más adelante).

Algoritmos genéticos: hibridación y optimización de operadores

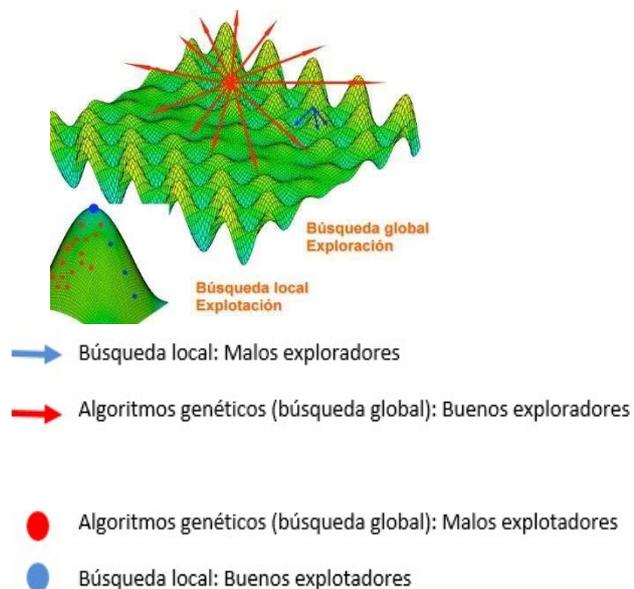
Un algoritmo memético [h] es la combinación entre:

- Una búsqueda global basada en poblaciones,
- Una heurística de búsqueda local realizada por cada individuo.

Es importante aclarar que la búsqueda global no implica necesariamente un algoritmo genético. Sin embargo, en este plan se proyecta utilizar esta técnica como base de la metodología desarrollada. La clave está en que los algoritmos genéticos son buenos “exploradores de soluciones” pero malos “explotadores”, en cambio los algoritmos de búsqueda local son

¹ Laboratorio de Investigación y Desarrollo en Computación Científica (LIDeCC), Depto. Ciencias e

buenos “explotadores” de soluciones prometedoras, mientras que son malos “exploradores” [i]. De esta manera surge la idea de combinarlos. Gráficamente:



Luego, como primer paso, pretendemos profundizar en tres puntos de interés que se deberían tener en cuenta al momento de desarrollar un algoritmo memético a partir de un algoritmo genético. Cómo hibridarlo es una cuestión que se puede responder parcialmente, debido a que es sabido que la hibridación consiste en combinarlo con una estrategia de búsqueda local. Cuál es el método de búsqueda local que en términos generales mejor desempeño produce, será analizado inicialmente. Proponemos también investigar dónde hibridarlo, lo cual será estudiado de acuerdo al mejor momento del proceso de evolución en el cual debería intervenir la búsqueda local. Por último en este sentido, planeamos también responder a la pregunta de cuándo hibridarlo, y con esto nos referimos a los tipos de problemas en los que es aconsejable utilizar este tipo de estrategia sinérgica.

Por otro lado, en lo que refiere a la mejora en la calidad de los operadores, se proyecta estudiar una forma de cruzamiento inteligente que aproveche el conocimiento de las características de las técnicas de biclustering en el proceso de reproducción. Para esto planteamos la hipótesis de realizar un proceso recursivo que permita combinar de manera eficaz dos o más biclusters produciendo uno de mejor calidad que sus

ancestros [j]. Una de las ideas surge de observar la efectividad del método de Ward implementado en el algoritmo recursivo de Lance-Williams [k]. La recursividad aplicada al biclustering se fundamenta en que esta técnica es ideal para objetos estáticos ya que se puede calcular previamente la medida de calidad para distintos grupos y utilizar estas medidas para instancias posteriores. Asimismo, se estudiará en detalle el criterio de terminación del algoritmo, para asegurar la convergencia a soluciones variadas en el contexto de cada problema [l].

Aplicación en Bioinformática: Datos de expresión de genes

La tecnología de microarray permite analizar niveles de expresión de miles de genes en distintas condiciones. Esta información se presenta en forma de matriz de números reales donde cada elemento representa el nivel de expresión del gen en la condición correspondiente [m]. Una de las aplicaciones más populares de las técnicas de biclustering está justamente relacionada a la extracción de información de genes que se expresan de manera coherente en datos de microarray [m]. Un ejemplo particular de esto es la inferencia de redes de regulación de genes (GRNs) usando biclustering, que representan relaciones entre genes que podrían constituir agrupamientos encargados de alguna función celular determinada [n]. En este sentido, la directora propuesta viene participando en el último período de investigaciones orientadas a dicho tema de investigación [o, p]. Luego, en el contexto de este plan, otra de las hipótesis planteada es la de encontrar pares de genes co-expresados; es decir, que tienen un perfil de expresión similar. De esta forma, para una aplicación de *eSalud* muy promisoría, se podría usar la metodología para inferir nuevos posibles marcadores tumorales a partir de marcadores conocidos [q, r].

3. Resultados esperados

Con esta línea de investigación se espera principalmente lograr una contribución teórica en algoritmos meméticos, a partir del diseño y desarrollo de una nueva estrategia híbrida que combine de manera eficaz las mejores

características de los algoritmos evolutivos y de un método de búsqueda local, mejorando así las capacidades de búsqueda global y local de cada una de estas estrategias respectivamente. Se diseñará la metodología con el fin de encontrar uno o varios biclusters en grandes cantidades de datos contenidos en una matriz. En principio, en la búsqueda local se analiza la posibilidad de incorporar un proceso recursivo para la reparación y/u optimización de los biclusters encontrados.

Cabe destacar la experiencia de la Dra. Carballido en el área de computación evolutiva, temática que ha desarrollado desde su propia tesis doctoral hasta la actualidad, aplicándola a distintos problemas de optimización. También ha tenido un primer acercamiento a la hibridación básica de algoritmos genéticos en una beca doctoral dirigida con anterioridad.

Asimismo, la Dra. Carballido ha dedicado los últimos años de su trabajo al estudio de técnicas de biclustering para datos de expresión obtenidos de experimentos de microarray. Por este motivo, una de las aplicaciones a abordar en una primera instancia pertenece al área de bioinformática. En particular, el problema que se espera atacar consiste en encontrar relaciones entre distintos genes a partir de información sobre su nivel de actividad. El desarrollo de metodologías con estas características podría resultar de suma utilidad en la investigación del área de *eSalud*. En este sentido ya está comprobada la eficacia de los enfoques de biclustering para seleccionar conjuntos de genes óptimos para la determinación de pronósticos de estratos específicos de pacientes en base a características moleculares de tumores. En particular, para este plan de trabajo se plantea la siguiente hipótesis: dado un gen conocido que se sabe que constituye un marcador tumoral comprobado, otro gen con un perfil de expresión similar podría ser también asociado a dicha enfermedad. El perfil de expresión similar será encontrado con la técnica de biclustering.

4. Formación de recursos humanos

En el contexto de esta línea intervienen el Dr. Ignacio Ponzoni (Investigador CONICET), la Dra. Jessica Carballido (Investigadora

CONICET), la Dra. Rocío Cecchin (Investigadora CONICET), la Dra. Julieta Dussaut (becaria POSDOC CONICET) y la Ing. Jimena Martínez (becaria Doctoral CONICET) principalmente. Todos sus planes de investigación están estrechamente vinculados a esta línea. Además constituye el eje principal del plan de doctorado de la Ing. Macarena Latini.

De este modo, esta línea se suma contribuirá a la formación de integrantes en distintos niveles de CONICET, en una disciplina de fuerte proyección en la actualidad, tal como es bioinformática. Con lo cual, nuestro laboratorio (LIDeCC) seguirá afianzándose en la formación de nuevos investigadores, tal como ha sido su tradición desde su conformación como grupo de investigación en el año 1996.

5. Referencias

- a. Madeira SC, Oliveira AL (2004). "Biclustering Algorithms for Biological Data Analysis: A Survey". *IEEE Transactions on Computational Biology and Bioinformatics* 1 (1): 24–45. doi:10.1109/TCBB.2004.2. PMID 17048406.
- b. Kriegel, H.-P.; Kröger, P.; Zimek, A. (March 2009). "Clustering High Dimensional Data: A Survey on Subspace Clustering, Pattern-based Clustering, and Correlation Clustering". *ACM Transactions on Knowledge Discovery from Data* 3 (1): 1–58. doi:10.1145/1497577.1497578.
- c. Carballido J.A., Gallo C.A., Dussaut J.S., Ponzoni I. "On Evolutionary Algorithms for Biclustering of Gene Expression Data", *Current Bioinformatics* (2015). 10 3 259-267(9).
- d. Beatriz Pontes, Federico Divina, Raúl Giráldez, Jesús S. Aguilar-Ruiz (2007), *Virtual Error: A New Measure for Evolutionary Biclustering*, *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics Vol 4447 of the series Lecture Notes in Computer Science* pp 217-226

- e Everitt, B. S., Landau, S. and Leese, M. (2001), *Cluster Analysis*, 4th Edition, Oxford University Press, Inc., New York; Arnold, London.
- f Hartigan, J. A. (1975), *Clustering Algorithms*, New York: Wiley.
- g Jain, A. K. and Dubes, R. C. (1988), *Algorithms for Clustering Data*, New Jersey: Prentice–Hall.
- h Moscato, P., Cotta, C.: A gentle introduction to memetic algorithms. In Glover, F., Kochenberger, G., eds.: *Handbook of Metaheuristics*. Kluwer Academic Publishers, Boston MA (2003) 105–144.
- i Ishibuchi, H., Yoshida, T., Murata, T.: Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling. *IEEE Transactions on Evolutionary Computation* 7 (2003) 204–223.
- j Eiben, A.E., Raue, P.E., Ruttkay, Z.: Genetic algorithms with multi-parent recombination. In Davidor, Y., Schwefel, H.P., Manner, R., eds.: *Parallel Problem Solving From Nature III*. Volume 866 of *Lecture Notes in Computer Science*. Springer-Verlag (1994) 78–87
- k Cormack, R. M. (1971), "A Review of Classification", *Journal of the Royal Statistical Society, Series A*, 134(3), 321-367.
- l Safe M.D., Carballido J.A., Ponzoni I, Brignole N.B. "On Stopping Criteria for Genetic Algorithms" *Lecture Notes in Artificial Intelligence*, Vol. 3171, 405–413 (2004). Springer-Verlag.
- m Kluger Y., Basri R., Chang J., and Gerstein M. (2003). Spectral biclustering of microarray data: co-clustering genes and conditions. *Genome Research*, 13:703–716.
- n Ma, S.; Kosorok, M.R. "Identification of differential gene pathways with principal component analysis", *Bioinformatics*, 25:882-889, 2009.
- o Gallo C.A., Cecchini R.L., Carballido J.A., Micheletto S., Ponzoni I. "Discretization of gene expression data revised". *Briefings in Bioinformatics* (2015). 1-13.
- p Gallo C.A., Carballido J.A., Ponzoni I. "Discovering Time-Lagged Rules from Microarray Data using Gene Profile Classifiers", *BMC Bioinformatics* (2011). 12:123.
- q Wang YK, Print CG, Crampin EJ (2013) Biclustering reveals breast cancer tumour subgroups with common clinical features and improves prediction of disease recurrence. *BMC genomics* 14: 102. Ali Oghabian, Sami Kilpinen, Sampsa Hautaniemi, Elena Czeizle, *Biclustering Methods: Biological Relevance and Application in Gene Expression Analysis* <http://dx.doi.org/10.1371/journal.pone.0090801>

Evaluación de Sensores y Planta Motriz Utilizados en la Construcción de Drones

Pedro D. Foresi¹, Paola I. Beltramini^{1,2}, Martín R. Herrera¹, José R. Gallardo³, Graciela I. Carranza⁴ & Diego M. Lobo¹

(1) *Departamento de Electrónica, Facultad de Tecnología y Ciencias Aplicadas, UNCa.*
pdforesi7@yahoo.com.ar, pbeltramini@tecno.unca.edu.ar, ing_martin_herrera@yahoo.com.ar

(2) *Laboratorio de Sistemas Embebidos, Facultad de Tecnología y Ciencias Aplicadas, UNCa, Argentina.*

(3) *Ing. Electrónico, Asesor Externo joserogallardo@gmail.com*

(4) *Facultad de Ciencias Económicas y Administración, UNCa. gicarranza@yahoo.com.ar*

AREA TEMATICA: Tecnología Informática Aplicada en Educación

RESUMEN

La utilización y desarrollo de vehículos aéreos no tripulados (UAVs o drones), han revolucionado a la sociedad, por ser dispositivos útiles, versátiles y de relativo bajo costo para un gran número de aplicaciones civiles, militares, sanitarias, educativas, etc.

La Facultad de Tecnología y Ciencias Aplicadas de la UNCa, como institución educativa, a través del Laboratorio de Sistema Embebidos (LaSE), el Centro de Investigación y Desarrollo en Modelado, Simulación y Optimización de Sistemas Electrónicos (CIDMOS) y el Laboratorio de Electrónica, dispone de espacios de investigación y desarrollo. En este marco, y motivados por el interés de iniciar una línea de investigación y contribuir a la enseñanza de la utilización de sistemas embebidos en aplicaciones de esta naturaleza, se propuso la evaluación y caracterización de sensores y planta motriz factibles de utilizar para el pilotaje y estabilización de un dron.

Se espera un impacto positivo sobre la Unidad Académica, por cuanto permitirá disponer de equipamiento tecnológico para la realización de los ensayos y verificaciones propuestas, reforzar la formación de los estudiantes de ingeniería a partir de la utilización y programación de diferentes sistemas embebidos en esta área, apoyar la conformación de grupos interdisciplinarios para la generación de ideas y proyectos, y propender al funcionamiento de los citados laboratorios.

Palabras Claves: UAVs, Sistema Embebido, Sensores, Planta motriz, Educación

CONTEXTO

Esta investigación se enmarca en el Proyecto iniciado en el año 2016, a través del cual se pretende introducir al grupo de trabajo en el estudio, evaluación y caracterización de los distintos tipos comerciales de sensores y planta motriz que se utilizan o pueden utilizar para el pilotaje y estabilización de un Drone o cuadricóptero.

El proyecto fue presentado en la convocatoria de Proyectos I+D 2016, siendo el organismo que lo financia la Secretaria de Ciencia y Tecnología de la Universidad Nacional de Catamarca (UNCa) a través del Programa de Desarrollo Científico y Tecnológico.

El ámbito de desarrollo es la Facultad de Tecnología y Ciencias Aplicadas de la UNCa, con la colaboración del Laboratorio de Sistemas Embebidos (LaSE) y el Laboratorio de Electrónica, mediante el aporte del espacio físico, materiales, equipamiento e instrumental necesarios para su realización.

1. INTRODUCCIÓN

En los últimos años, los importantes avances tecnológicos, han posibilitado el desarrollo de sensores inteligentes capaces de medir distintas magnitudes físicas con un solo chip, lográndose así, pequeños sensores microelectromecánicos denominados MEMS (Sistemas Micro Electro Mecánicos). La miniaturización lograda por estas nuevas técnicas de fabricación hace que estos dispositivos puedan ser implementados en sistemas de navegación conocidos como “Strapdown”, término que se refiere a que el sistema de medida se encuentra ajustado al marco de referencia del objeto en prueba, es decir, se mueve de la misma forma que lo hace dicho objeto (Ardila Pérez, 2012).

Estos sensores pueden adquirirse de manera independiente o integrados en un “módulo de navegación inercial” (IMU), el cual es un dispositivo que incorpora al menos un acelerómetro y un giroscopio orientados axialmente. De esta forma proporcionan los datos de aceleración lineal y velocidad angular sobre el eje en el que se encuentran ubicados. Habitualmente estos módulos incorporan un microcontrolador, encargado de recoger, acondicionar, y transmitir los datos entregados por los sensores, en forma de niveles de voltaje, a otro sistema que se encargará de procesarlos para su uso. (Preckler Clemente, 2012).

El desarrollo de estos MEMS, de microprocesadores y de sistemas de propulsión de tamaño reducido, menor peso y con más prestaciones que nunca, permitieron la construcción y permanente evolución de vehículos aéreos no tripulados (UAVs, VANTs o drones), tanto para fines militares, civiles e incluso educativos, existiendo una amplia variedad de formas, tamaños, configuraciones y características en su diseño.

Su auge en el mercado ha sido tal que cada vez hay un mayor número de empresas, Organismos públicos y privados, que han conducido sus desarrollos en esta dirección. Incluso a nivel educativo nacional, muchas Universidades públicas y privadas realizan investigaciones en desarrollos de VANTs.

El este sentido, se intenta a través del proyecto iniciar una línea de investigación en esta temática, a partir del estudio del comportamiento y caracterización de los sensores y sistemas de control de su planta motriz utilizados en la construcción de los referidos vehículos aéreos no tripulados. Sensores tales como los requeridos para el control de altitud (Sensor de presión); para monitorizar los movimientos de posición (Sensor de aceleración); para detectar los movimientos de cabeceo, giro y viraje (Giróscopo) y para definir su orientación (Magnetómetro). La Planta motriz, conformada por cuatro motores y sus respectivos controlador de velocidad, responsables, a nivel de hardware, para que el dron se mantenga en el aire, con la ayuda de los comandos a nivel de software emitidos por el sistema de estabilización.

Para la lectura de los valores captados por estos sensores digitales, es frecuente el uso de lenguajes de programación de alto nivel, como el C y/o assembler, lenguaje este último de bajo nivel utilizado en computadoras, microprocesadores, microcontroladores y otros circuitos integrados programables.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La principal línea de investigación de este proyecto es el uso de las nuevas tecnologías en la evaluación del comportamiento de los sensores y planta motriz y determinación de posibles factores que condicionan o limitan su utilización en los UAVs, dando solución al problema de disponibilidad de información técnica confiable de calidad.

Se pretende además contribuir con la educación de carreras técnicas que se dictan en las Instituciones educativas (públicas o privadas) de nuestro medio, a partir de la utilización y programación de sistemas embebidos (equipo electrónico basado en microprocesadores, que realiza el procesamiento de datos digitalmente, pero que a diferencia de una computadora personal, su hardware y software están específicamente diseñados y optimizados para ejecutar una función determinada o resolver un problema concreto de forma eficiente) para la

lectura y operación de los sensores y motores, control y comunicación entre módulos, etc.

Las tareas a desarrollar durante el proceso de investigación, se agrupan en:

1. Investigación documental de los distintos tipos de sensores y planta motriz que se requieren y utilizan en la construcción de cuadricópteros.
2. Análisis de las posibles soluciones y selección de los dispositivos específicos y equipamiento a adquirir en función de la disponibilidad presupuestaria.
3. Confirmación teórica - práctica de las soluciones propuestas y de los resultados, mediante la utilización de un sistema embebido.
4. Confección de tablas e indicadores coherentes, propuestos.
5. Evaluación de los resultados obtenidos mediante comunicados y reportes a conferencias y reuniones científicas.
6. Comunicación de los resultados mediante publicaciones a revistas.

3. RESULTADOS OBTENIDOS / ESPERADOS

En una etapa inicial, se encaró la búsqueda de información documental técnica sobre los diferentes sensores comerciales factibles de utilizar en aplicaciones de esta naturaleza, el comportamiento de los sensores y los factores que condicionan o limitan su utilización, como así también sus características técnicas, disponibilidad y precios en el mercado local y nacional.

Se compró el sensor MPU-6050 de la firma InvenSense, que es un sensor de costo accesible, especialmente teniendo en cuenta el hecho de que combina un acelerómetro y un giroscopio en el mismo dispositivo (figura 1).



Figura 1: Sensor MPU-6050 (15x21mm)

Se montó el sensor en una plataforma consistente en una combinación de planos inclinados regulables en dos ejes, a la cual se le agregó una escala graduada, lo que permitió variar los ángulos de inclinación y realizar las primeras pruebas de funcionamiento, tomando lecturas de aceleración y velocidad angular. Para la adquisición de dichos datos, se utilizó una herramienta de desarrollo que incorpora un DSC (controlador digital de señales), de Freescale Semiconductor, conformado por un microcontrolador y un DSP (procesador digital de señales). El DSC pertenece a la familia del core 56F800E el cual puede correr hasta 32 MHz de velocidad.

Los datos leídos se transmitieron a una PC para su procesamiento, cotejando las lecturas con los indicados en la plataforma (Figura2).



Figura 2. Prueba de funcionamiento sensor MPU-6050

Durante el transcurso del año se avanzará en la compra otros sensores y diferentes motores, a

efectos de realizar análisis comparativos, destacando ventajas y desventajas de cada uno. Asimismo, el proyecto propende además contribuir con nuevos conocimientos a la comunidad educativa en general, ya que podrán difundirse sus conclusiones en el ámbito académico tanto universitario como en las escuelas técnicas.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por cinco ingenieros, cuatro de ellos docentes de la carrera Ing. Electrónica de la Facultad de Tecnología y Ciencias Aplicadas de la U.N.Ca., en cátedras afines al tema de investigación, técnicamente capacitados para afrontar las diversas vertientes teóricas y prácticas del proyecto, los que se complementan con una Lic. En Matemática, especialista en Didáctica.

Si bien gran parte del grupo posee la misma formación de grado, han adquirido experiencia y recibido capacitación en diversas instancias sobre problemáticas afines a la temática del proyecto, en áreas de matemática, sistemas embebidos y educación, logrando una importante diversidad del equipo de investigación.

Cuatro de los investigadores están categorizados en el Programa de Incentivos de la Secretaría de Políticas Universitarias (SPU) perteneciente al Ministerio de Educación de la Nación.

El equipo se completa con un estudiante avanzado de Ingeniería Electrónica, quien se capacita en el uso y programación de diversas placas de sistemas embebidos disponibles en la facultad para la lectura de sensores digitales.

5. BIBLIOGRAFIA

Ardila Pérez, Luis Eduardo, Eslava Garzón, Johan Sebastián. “*Diseño, Calibración y Prueba de una Unidad de Medida Inercial*

(Imu) Inalámbrica para Cohetes” - Revista de la Facultad de Ingeniería, Año 13, n° 25, enero-junio de 2012.

Bruno Saravia, A.R. et al. “*Diseñando Sistemas Embebidos con Librerías de Microchip*”. Mcelectronics, Buenos Aires, 2012.

Escamilla Nuñez R., *Diseño, construcción, instrumentación de un Vehículo Aéreo no Tripulado (UAV)*, México D.F., Noviembre 2010.

Ferrer Mínguez, G. “*Integración Kalman de sensores inerciales INS con GPS en un UAV*”, Trabajo Final de Carrera, 27/04/2009. <http://upcommons.upc.edu/handle/2099.1/6930>

Landi, Diego, *La industria de los drones crece en la Argentina con desarrollos locales*, <http://www.infotechnology.com/comunidad/La-industria-de-los-drones-crece-en-la-Argentina-con-desarrollos-locales-20140416-0004.html#sthash.ERdifBP8.dpuf>, 2014. Fecha consulta 25 abril 2016.

Naranjo, C. C. M., “*Analysis and modeling of MEMS based inertial sensors*”. School of Electrical Engineering, Kungliga Tekniska Hgskolan, Stockholm, 2008.

Pérez García, M.A. et al., “*Instrumentación Electrónica*”, Thomsom, España, 2004.

Ramos Vicedo M., *Diseño e implementación de un quadcopter basado en microcontrolador arduino*, Valencia, España, 2014.

Serrano E., Sirne R. O. y La Mura G., *Rotaciones, secuencia aeroespacial y cuaterniones. Una revisión de las relaciones fundamentales*. Ciencia y Tecnología, ISSN 1850-0870, pp. 11-28, 2014.

<http://dronecenter.blogspot.com.ar/p/informacion-basica.html>, junio 2016.

<http://www.drones-argentina.com.ar/category/argentina/>, junio 2016.

<https://www.invensense.com/products/motion-tracking/6-axis/MPU6050>, Junio 2016.

Fusión de Algoritmos Bayesianos y Árboles de Clasificación como Propuesta para la Clasificación Supervisada de Fallos de Equipos en un laboratorio de Cómputos.

Ing. Corso Cynthia, Ing. Maldonado Calixto, Ing. Pereyra Florencia, Srta. Martínez Gimena, Sr. Donnet Matías.

Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información
 Departamento Ingeniería en Sistemas de Información
 Facultad Regional Córdoba/Universidad Tecnológica Nacional
 Maestro M. López esq. Cruz Roja-Ciudad Universitaria-Córdoba
cynthia@bbs.frc.utn.edu.ar/calixto_maldonado@hotmail.com/pereyraflorencia@gmail.com
gimena_martinez@bbs.frc.utn.edu.ar/donnetmatias@bbs.frc.utn.edu.ar

RESUMEN

Los algoritmos basados en redes bayesianas y árboles de decisión representan métodos que han resultado eficientes para la resolución de problemas de clasificación. Este trabajo pretende combinar estos algoritmos con el objetivo de obtener un modelo híbrido que permita aprovechar y combinar las ventajas de ambos. Con esta estrategia se pretende aumentar la precisión en los resultados de la clasificación supervisada. Este trabajo pretende detallar cual es el grado de precisión en la exactitud, cuando los algoritmos bayesianos son combinados con los árboles de decisión utilizando como recurso los métodos de fusión o ensamble *Grading* y *Vote*. Los modelos híbridos resultantes serán aplicados para la clasificación de eventos de fallos en equipos pertenecientes a un laboratorio de cómputos, con el propósito de aumentar su disponibilidad y mantenibilidad.

Palabras claves: *Métodos de fusión o ensamble, Algoritmos bayesianos, Árboles de decisión, fallos en equipos.*

CONTEXTO

Este trabajo pertenece al proyecto “Generación de Modelo Descriptivo para la caracterización de incidentes en equipos de un laboratorio de cómputos (Fase II)” PID-UTN3931. Correspondiente al periodo de ejecución 2016-2017 del Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (*GIDTSI*).

1. INTRODUCCIÓN

Un evento de fallo representa toda alteración o interrupción de un sistema (aparatos o equipos), en el cumplimiento de la función para la cual ha sido diseñado [1]. En el contexto de cualquier organización, la presentación de estos eventos resiente no solo el normal funcionamiento de las actividades programadas, sino que representa un impacto negativo en el aspecto económico.

La caracterización y detección de eventos de fallos ha sido estudiada por el área de computación como un problema de clasificación. Existen diversos algoritmos

que han sido diseñados para dar solución a esta problemática. Los modelos de algoritmos más utilizados son los árboles de decisión, redes neuronales y bayesianas. Dentro de estos modelos, los árboles de decisión son muy utilizados en este contexto [2], siendo el algoritmo *RandomTree* (árboles aleatorios) una alternativa viable como solución a problemas de clasificación.

Los árboles aleatorios han sido introducidos por Leo Brinman y Adele Cutler, y permiten abordar tanto problemas de regresión como clasificación. El algoritmo *RandomTree* permite la representación de un árbol diseñado al azar de un juego de posibles árboles denominados bosque. El mecanismo de clasificación en este algoritmo funciona de la siguiente forma

i) Los árboles de clasificación aleatorios consideran un vector que contiene las características de los datos de entrada ii) Cada árbol aleatorio realiza la clasificación y genera la etiqueta de clase que recibió la mayoría de los votos en la clasificación [3].

Obtener un clasificador eficiente no es una tarea simple. Cada clasificador se caracteriza porque emplea una representación diferente de los datos. Encontrar una representación de estos, que mejor se adapte con el problema a resolver, requiere de tiempo y de varios experimentos previos. El uso de distintos algoritmos clasificadores puede proporcionar información complementaria importante sobre la representación de los datos, como así también aumentar la precisión de los modelos obtenidos. Esto ha originado la necesidad de utilizar una fusión o ensamble de clasificadores como una alternativa apropiada para el tratamiento

de problemas de clasificación supervisada.

Un método de fusión o ensamble representa una agrupación de clasificadores, que combinan sus predicciones siguiendo un determinado esquema, con el objetivo de obtener una predicción más fiable que la que normalmente sería capaz de obtener de forma individual [4]. El estado del arte permite conocer la existencia de diversos métodos de fusión o ensamble, aunque no hay definiciones ni reglas que permitan dar cuenta sobre cual método de ensamble es más apropiado con respecto a otro.

Uno de los métodos de ensamble que ha sido referenciado y comparado en diversos trabajos con otros métodos híbridos, para demostrar su rendimiento y performance, fue el método conocido como *Grading* [5] [6] [7]. El mismo será considerado para la realización de experimentos junto con el algoritmo *Vote*.

Diversas investigaciones han profundizado sobre el uso de métodos bayesianos, que han demostrado ser tan competitivos como los árboles de decisión y las redes neuronales para el tratamiento de problemas de clasificación [8]. Las redes bayesianas representan las dependencias que existen entre los atributos a través de una distribución de probabilidad condicional en un grafo dirigido acíclico [9]. El clasificador Naive Bayes es un caso particular de red bayesiana, en el que se asume que los datos o características son condicionalmente independientes dado un atributo de clase [10]. En este trabajo se considera el algoritmo Naive BayesUpdateable que es una versión mejorada de Naive Bayes [11]. El algoritmo BayesNet al igual que Naive

BayesUpdateable permite la construcción de una red bayesiana utilizando diversos métodos de búsqueda (K2, HilClimber, TAN, BAN entre otros) y métricas que permiten medir la calidad del modelo resultante [12].

El objetivo de este trabajo es proponer modelos híbridos para la resolución de problemas de clasificación, mediante la aplicación de métodos de ensamble o fusión para la integración de algoritmos de redes bayesianas y árboles de clasificación. Por medio de experimentos en diferentes conjuntos de datos será posible determinar si la precisión de los algoritmos basados en la construcción de redes bayesianas (NaivesBayesUpdateable y BayesNet) mejora cuando se fusiona o combina con los métodos basado en la construcción con árboles de decisión aleatorios como lo es RandomTree.

Si la precisión de los modelos híbridos propuestos que arrojan los experimentos es aceptable, será factible se aplicación con el set de datos que almacena los eventos de fallos presentados en los componentes de hardware y software de los equipos (computadoras personales), que están en funcionamiento en un laboratorio de cómputos, que es la organización bajo estudio. La misma desempeña funciones de apoyo en el dictado de clases a cátedras pertenecientes a la carrera de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba.

La motivación para la realización de este trabajo experimental se basa en la necesidad de encontrar un método adecuado, que brinde soporte a un aspecto crítico en la organización bajo estudio, como lo es la necesidad de aumentar la

disponibilidad de los equipos, que en muchas situaciones se ve resentido por diagnósticos desacertados de eventos de fallos o por el tiempo excesivo que transcurre el tiempo para su resolución.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

La línea de investigación tiene como objetivo generar conocimiento no conocido y potencialmente útil para la toma de decisiones en el área de mantenimiento, específicamente en el tratamiento de fallos, tendientes a aumentar la disponibilidad y mantenibilidad en los equipos.

En el caso de esta línea de investigación se pretende el procesamiento de los datos que han sido obtenidos de los sistemas de gestión utilizados para el tratamiento de fallos en componentes de hardware y software de los equipos.

Este trabajo adopta las siguientes líneas de investigación y desarrollo:

- Aprendizaje automático.
- Arquitectura de modelos de fusión o ensamble.
- Aprendizaje automático.
- Modelos probabilísticos.
- Estadística.
- Gestión de eventos de fallos.
- ☐ Confiabilidad en equipos.

En esta segunda etapa del proyecto se focalizará en la ejecución de las siguientes tareas:

- Selección de diferentes set de datos para determinar la confiabilidad de los modelos híbridos resultantes.
- Reducción de la dimensionalidad de los datos del conjunto de entrenamiento,

evaluando y seleccionando los métodos de selección de atributos más apropiados.

- Selección del mecanismo de validación del modelo que resulte de aplicar los modelos de fusión o ensamble.
- Valoración de los algoritmos de redes bayesianas y árboles de decisión seleccionados, para determinar su comportamiento de manera individual.
- Realización de experimentos con los métodos de fusión o ensamble seleccionados, combinando los algoritmos de redes bayesianas y árboles de decisión considerados, con diferentes set de datos.
- Combinación de los algoritmos de clasificación que permitan describir de mejor forma y obtener mejor precisión del modelo de conocimiento obtenido.
- Analizar la precisión de los resultados obtenidos en cada experimento.
- Comparar los resultados para determinar la viabilidad de aplicación a la problemática estudiada.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los objetivos propuestos en el proyecto de investigación son:

- Proponer modelos híbridos con un nivel de precisión aceptable, para el tratamiento y clasificación de eventos de fallos en equipos.
- Fomentar, incentivar y difundir las actividades y resultados de investigación.

Uno de los resultados obtenidos en el primer año de ejecución del proyecto, fue una vista minable que permitió la unificación y el almacenamiento de los eventos de fallos. Ya que los fallos que afectan los componentes de hardware y software son manejados en diferentes

sistemas de gestión por parte de la organización bajo estudio.

La vista minable fue el resultado de adaptar el diseño del data mart propuesto en la primera fase del citado proyecto de investigación, que permitió solo el almacenamiento de los eventos fallos que afectaron a los componentes de hardware. El modelo dimensional resultante fue el esquema Estrella. La arquitectura que se adoptó para su explotación fue ROLAP (Relational On Line Analytical Processing), que permitió dar soporte a los requerimientos de información.

La herramienta seleccionada para realizar el proceso de explotación del data mart fue la herramienta de la suite Pentaho denominada Mondrian Schema Workbench. La adaptación del data mart en esta fase del proyecto, propició un contexto favorable para la realización de operaciones de integración, limpieza y estandarización de los datos considerados significativos para el tratamiento de eventos de fallos en equipos.

Los resultados esperados del proyecto se detallan a continuación:

- Diseño de diferentes enfoques de modelos híbridos que permitan abordar con una precisión aceptable para problemas de clasificación supervisada.
- Identificación de organizaciones e instituciones que puedan ser de utilidad la transferencia de los modelos resultantes.
- Formación de los recursos humanos que forman parte del proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto está conformado por docentes-investigadores pertenecientes a

la carrera de grado de Ingeniería en Sistemas de Información.

Todos los integrantes docentes del PID han participado del proceso de categorizaciones en investigación dentro del Programa de Incentivos del MECyT; así como en la categorización interna que posee la U.T.N.

Además colaboran en este proyecto tres becarios, dos alumnos y un graduado.

5. REFERENCIAS

- [1] Sols Alberto, Fiabilidad, mantenibilidad, efectividad: Un enfoque sistémico. Editorial Universidad Pontificia Comillas, Madrid, 2000.
- [2] Ding, Q. and Perrizo W., Decision Tree Classification of Spatial Data Streams Using Peano Count Trees. Proc. of the ACM 124 Symposium on Applied Computing, Madrid, España, 2002.
- [3] Pfahringer Bernhard, Random model trees: an effective and scalable regression method. University of Waikato, New Zealand, 2010.
- [4] Quintana Ramírez María José, Orallo José Hernández, Extracción Automática de conocimiento en Base de Datos e Ingeniería de Software, España, 2005.
- [5] Seewald Alexander, Furnkranz Johannes. Gradding Classifiers. <http://www.ofai.at/cgi-bin/get-tr?paper=ocfai-tr-2001-01.pdf>, 2001.
- [6] Ledesma Espino Ismael. Aprendizaje Automático en Conjunto de Clasificadores heterogéneo y Modelado de Agentes (tesis de doctorado). Universidad Carlos III de Madrid. 2004.
- [7] Seewald Alexander, Furnkranz Johannes. An Evaluation of Grading Classifiers. International Symposium on Intelligent Data Analysis. págs. 115-124, 2001.
- [8] Sahami, M. Learning Limited Dependence Bayesian Classifiers. 2th International Conference on Knowledge Discovery in Databases (KDD96), Menlo Park, CA, AAAI Press, 1996.
- [9] Singh, M. and G. M. Provan. Efficient Learning of Selective Bayesian Network Classifier. International Conference on Machine Learning. Philadelphia, PA., Computer and Information Science Department, University of Pennsylvania, 1995.
- [10] Witten, I. E. Data Mining. Practical Machine Learning Tools and Techniques with Java Implementations. San Francisco. California, 2000.
- [11] John H. George, Langley Pat. Estimating Continuous Distributions in Bayesian Classifiers. 11ava Conferencia de Incertidumbre en Inteligencia Artificial, págs. 338-345, 1995.
- [12] Pascual Mauricio Beltrán, Muñoz Martínez Azahara y Alamillos Muñoz Ángel, Redes bayesianas aplicadas a problemas de crédito scoring. Una aplicación práctica. Cuadernos de Economía. Elsevier. España. 2014.
- [13] Serra Araujo Basilio, Aprendizaje Automático: conceptos básicos y avanzados. Aspectos prácticos usando software Weka, Pearson Educación, Madrid, 2006.
- [14] Woycik, Mantenimiento y Reparación de equipos. Editorial: Cesarini, 1987.
- [15] Creus Solé Antonio., Fiabilidad y Seguridad: su aplicación en procesos industriales. 2da editorial, 2005.
- [16] Sushilkumar Rameshpant Kalmegh, Comparative Analysis of WEKA Data Mining Algorithm RandomForest, RandomTree and LADTree for Classification of Indigenous News Data, International Journal of Emerging Technology and Advanced Engineering, Volume 5, Issue 1, 2015.

Hacia la Definición de un Agente Generador de Conocimiento de Valor Social para Poblaciones en Riesgo

Federico Amigone¹

Jorge Rodríguez¹

Gerardo Parra¹

email: {fed.amigone, j.rodrig, gparra}@fi.uncoma.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

En este trabajo se presenta una línea de Investigación que busca analizar las características iniciales y describir el campo de acción de un agente autónomo que vehiculice la apertura y el descubrimiento de patrones, relaciones y clasificaciones en el contexto acotado a la base de datos de la aplicación Protocola[2].

Protocola es intensamente utilizada por la agrupación nacional Socorristas en Red para recolectar datos y analizar la realidad social en materia de abortos medicamentosos. En su rubro, la base de datos es una de las más completas de Sudamérica, constituyendo a la fecha un compendio multivariante de casos individuales tomados sobre la realidad social argentina.

De tal configuración emergen problemas que se presentan sugerentes y prometedores para la aplicación de técnicas de aprendizaje automático y minería de datos, como la búsqueda de los inesperados por descripción de la realidad multivariante, detección de asociaciones, tipificación, clasificación, regresión y predicción[1].

La taxonomía de problemáticas susceptibles a ser resueltas por las técnicas mencionadas constituyen un campo que deberá ser estudiado y acotado en función de las perspectivas de utilidad percibidas por parte de quienes se ocupan del cuidado de la salud pública en general y de las personas involucradas al uso

de la aplicación Protocola.

Para ello, resultará necesario clarificar cuáles de los numerosos modelos de aproximación[6] en el contexto de la Inteligencia Artificial resultan apropiados para el abordaje de la problemática y el diseño del agente.

Además, se deberá generar un marco de objetividad teórico integrando conclusiones científicas capitalizadas por el grupo con aportes producidos en la línea.

Palabras Clave: INTELIGENCIA ARTIFICIAL, AGENTES INTELIGENTES, LENGUAJES FORMALES, MINERÍA DE DATOS, APRENDIZAJE AUTOMÁTICO, COMPUTING FOR SOCIAL GOOD.

Contexto

Esta propuesta de trabajo se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática y en el marco de las actividades de Extensión Universitaria desarrolladas por la misma Facultad.

La Actividad de Extensión Universitaria[3] relacionada a esta Línea de Investigación tuvo por objetivo relevar e implementar los requerimientos comprendidos en el funcionamiento de la Organización No Gubernamental (ONG), sin fines de lucro y de alcance nacional, llamada Socorristas en Red. La Red utiliza un protocolo para sistematizar

información de utilidad para el sistema de salud y público en general en el marco de la atención integral a personas que interrumpen su embarazo por medios medicamentosos.

Este trabajo se desarrolla en el ámbito del Proyecto de Investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

Introducción

Los agentes inteligentes pueden ser definidos como entidades que se ejecutan sin interacción humana a efectos de lograr objetivos definidos por los usuarios[9,8]. Un agente, en tanto entidad autónoma, articula con un contexto sistémico específico accediendo y procesando datos a la vez que puede interactuar con otros agentes[7,12]. En dicho contexto, los modelos y lenguajes formales conforman la base teórica conceptual sobre la que diversas técnicas de razonamiento y aprendizaje se apoyan para lograr inferencias, relaciones, descubrimientos inductivos y nuevos conocimientos[4]. En ambientes específicos, en los que la percepción humana y las técnicas estadísticas tradicionales se ven superadas por la complejidad, los agentes se presentan como una alternativa de relevancia para la resolución de problemas complejos del mundo real.

Para el Proyecto de Investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación* resulta de importancia capital el desarrollo de la Línea que introduce el presente trabajo, ya que plantea el estudio sobre aspectos fundacionales de la teoría y la construcción de aplicaciones basadas en agentes inteligentes inmersos en ambientes dinámicos.

Un agente inteligente se presupone capaz de percibir su entorno, procesar tales percepciones y, en función de ciertas capacidades racionales, tomar la mejor decisión en términos conductuales o de procesamiento respecto del entorno circundante. La Línea de Investigación que

aquí se presenta persigue poner éstas características propias de un Agente Racional, al servicio de una problemática de salud concreta.

El contexto particular parte de la Actividad de Extensión Universitaria con la que articula el presente trabajo y que se propuso, con éxito, sistematizar e informatizar un protocolo colector de datos. A tal efecto, fueron relevados los requerimientos y estos señalaron la necesidad de brindar un soporte digital a formularios cuyo flujo de carga es controlado por las respuestas a aproximadamente 60 preguntas. La carga de cada formulario genera información estructurada de casos específicos vinculados a abortos medicamentosos, cuyo volumen de información hace necesario disponer de una instancia de análisis sobre determinadas variables, con un formato analítico específico. Las instancias colectoras de datos se encuentran distribuidas en 50 puntos del país, logrando así una base geográfica distribuida. A la fecha se registran aproximadamente 8 mil formularios cargados desde diversos puntos del país. Este conjunto de datos, es decir la base de datos de la aplicación *Protocola*, constituye el contexto sobre el que el Agente deberá desplegar sus características constitutivas.

Línea de Investigación

La Línea de Investigación presentada en este trabajo busca la convergencia entre una problemática concreta de la realidad social Argentina, como lo es el estado de situación y salud de la población constituida por mujeres que interrumpen embarazos, y las capacidades que ofrecen los agentes orientados a la inducción de información.

Tradicionalmente, los segmentos sociales relegados o que gravitan por fuera de los polos tecnológicos o corporativos no se han visto significativamente beneficiados por las capacidades tecnológicas que ofrece el campo de la Inteligencia Artificial.

Es por ello que poner al alcance de una organización social sin fines de lucro los resultados que el desempeño de un agente con capacidades cognitivas específicamente

diseñadas pudiera tener, constituye, desde el punto de vista financiero, una dificultad que se encuentra fuera del alcance de sus recursos económicos pero que éste proyecto se propone trascender.

El proyecto se abocará al estudio de los fundamentos y la teoría de agentes inteligentes así como del diseño y desarrollo de posibles alternativas de implementación[5]. Asimismo, se examinarán y propondrán modelos formales relevantes y adecuados para esta tarea.

En particular, se espera que el agente pueda realizar descubrimientos complejos y específicos[1] para su universo de discurso, entre los cuales se cuentan los siguientes:

- Descubrimiento de lo inesperado[10]: la mente humana subyuga la complejidad a través del poder de abstracción. Esto es, la detección de lo esencial a la resolución del problema y la eliminación de lo superfluo. De este marco emerge el Principio de la Parsimonia, muy aceptado en los años tempranos de la estadística, según el cual la solución más simple suele ser la mejor. Sin embargo, la irrupción de los agentes con capacidades cognitivas invita a abordar problemas complejos con nuevos enfoques que convoquen cada vez más variables a la descripción de un problema. Los agentes pueden viabilizar la detección de lo inesperado o de una nueva significación cuya existencia se ignoraba pero de la que el observador puede servirse para entender mejor la realidad. ¿Existe un patrón conductual común en la población de mujeres que abortan? ¿hay características etáreas que señalan patrones determinados? Se buscará conocer lo que emerge del conjunto de datos que subyace al dominio y de lo que hoy no se tiene confirmación ni noción absoluta de su existencia.
- Descubrimiento de asociaciones[13]: no resulta para nada trivial afirmar que la ocurrencia de un evento se da en términos de simultaneidad respecto de

otro, y que dicha ocurrencia sería mayor a la que tendría en soledad. Descubrir vínculos en términos de asociatividad de ocurrencias es una tarea candidata para un Agente Inteligente y que éste proyecto abordará por considerarla muy importante para quienes se ocupan de la salud poblacional. Actualmente, dentro del contexto de la base de datos Protocola, no se tienen certezas sobre cómo se correlacionan factores etáreas, familiares, religiosos, económicos o médicos en la población de estudio.

- Descubrir tipologías[5]: ¿Existirá un subconjunto de mujeres dentro de un rango etéreo que presentan características repetibles y constatables? Será una meta de diseño del Agente la capacidad de taxonomizar un conjunto potencialmente ingente de individuos a partir de un contexto multivariable.
- Predicciones[11]: bajo la forma de regresión o clasificación, la proyección del comportamiento de variables estudiadas resulta de particular importancia para asistir a la población de más alto riesgo.

Resultados Esperados

En el contexto de las actividades proyectadas se espera:

- Analizar los modelos formales a partir de su adecuación al abordaje de un agente de información inteligente, cuya meta será descubrir patrones, relaciones, categorías y desarrollar habilidades predictivas.
- Desarrollar un agente inteligente con capacidades cognitivas, de manera que pueda actuar en forma autónoma y sin intervención humana.
- Validar con los actores calificados y

relacionados al uso de la aplicación
 Protocola la utilidad del conocimiento
 producido por el agente.

Formación de Recursos Humanos

La ejecución del proyecto de investigación
 brindará un marco propicio para la iniciación
 y/o finalización de estudios de posgrado de los
 integrantes docentes. El área científico-
 tecnológica principalmente impactada por la
 formación de recursos humanos en el marco de
 este proyecto será el área de Tecnologías de la
 Información y de las Comunicaciones.

Referencias

- [1] T. Aluja. *La Minería de Datos: Entre la Estadística y la Inteligencia Artificial*. 2001.
- [2] F. Amigone. Aplicación Protocola. <http://protocola.fi.uncoma.edu.ar>, 2015.
- [3] F. Amigone. *Desarrollo de una aplicación web que dé soporte a las tareas que realiza la red nacional Socorristas en Red*. 2015, avalado por Resolución FaI 082/15.
- [4] G. Bel-Enguix, M. D. Jiménez-López, and C. Martín-Vide, editors. *New Developments in Formal Languages and Applications*. 2008
- [5] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. *From Data Mining to Knowledge Discovery in Data Bases*. 1996.
- [6] J. Friedman. *Data Minings and Statistics: What's the connection?* 2010.
- [7] Huhns and L. Stephens. Multiagent Systems and Societies of Agents. In G. Weiss, editor, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, pages 79-120. The MIT Press, Cambridge, Massachusetts, 1999.
- [8] D. Poole, A. Mackworth, and R. Goebel. *Computational Intelligence: A Logical Approach*. Oxford University Press, 1998.
- [9] S. Russell and P. Norvig. *Artificial Intelligence: A modern approach*. Prentice Hall, New Jersey, third edition, 2009.
- [10] P. M. Sankar K. Pal. *Patter Recognition Algorithms for Data Mining*. 2001.
- [11] T. K. Sung, N. Chang, and G. Lee. Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16(1):63-85, 1999.
- [12] G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, Cambridge, Massachusetts, 1999.
- [13] J. S. Yujing Zeng. *Statistical Approach to Clustering in Pattern Recognition*. 2001.

Implementación de Operadores de Consolidación de Ontologías en Datalog +/-

Ignacio Novello Pelayo

Cristian D. Pacifico

Cristhian A. D. Deagustini

Marcelo A. Falappa

Guillermo R. Simari

Area de Agentes y Sistemas Inteligentes
Facultad de Ciencias de la Administración
Universidad Nacional de Entre Ríos
Tavella 1424, (3200) Concordia, Entre Ríos

Laboratorio de Investigación y Desarrollo en Inteligencia Artificial
Instituto de Ciencias e Ingeniería de la Computación,
Universidad Nacional del Sur/CONICET
Alem 1253 - Bahía Blanca - Buenos Aires - Argentina
(0291) 459-5135

(inovello,cripac)@fcad.uner.edu.ar, (cadd,mfalappa,grs)@cs.uns.edu.ar

Resumen

En los últimos tiempos, la colaboración y el intercambio de información se han vuelto aspectos cruciales de muchos sistemas. En estos entornos es de vital importancia definir métodos automáticos para resolver conflictos entre el conocimiento compartido por distintos sistemas. Este conocimiento es frecuentemente expresado a través de ontologías que pueden ser compartidas por los sistemas que utilizan el mismo. Dentro de las diferentes posibilidades para expresar conocimiento en los últimos tiempos un lenguaje que ha recibido cada vez más atención es Datalog+/-; debido a ser un lenguaje que ofrece un alto nivel de expresividad por construir reglas en fragmentos de Lógica de Primer Orden, permitiendo su comprensión natural en la forma de esquemas de lógica clásica. Por otro lado, Datalog+/- como máquina de inferencia, tiene la propiedad de ser decidible, y (en la mayoría de los casos) tratable permitiendo manejar volúmenes masivos de datos de entornos reales.

Sin embargo, el uso del conocimiento, especialmente compartido, suele traer aparejados conflictos en el mismo que dificulta su explotación por procesos automatizados. Es decir, aceptar nuevas observaciones y datos suele traer aparejados violaciones a la integridad y consistencia del cuerpo de conocimiento. En este sentido, el principal desafío es eliminar los conflictos las inconsistencias e incoherencias que puedan surgir en el conocimiento expresado.

En la presente investigación se busca comprobar empíricamente la correctitud, computabilidad y eficien-

cia de operadores de contracción de kernel y de cluster para manejo de inconsistencias e incoherencias en ontologías Datalog+/- a través de la implementación de los mismos. Para esto, se analizará el diseño teórico de los operadores de contracción de kernel y cluster y su aplicación práctica para eliminar conflictos. A partir de un proyecto de software que implementa un intérprete Datalog+/-, se codificará las funcionalidades requeridas e implementarán los citados operadores.

Palabras Clave: Integración de Bases de Conocimiento, Revisión de Creencias, Representación de Conocimiento, Razonamiento.

1. Contexto

Esta línea de investigación se lleva a cabo en el marco de los siguientes proyectos de investigación:

- **“Representación de conocimiento y razonamiento argumentativo: Herramientas inteligentes para la web y las bases de conocimiento”**. Director: Guillermo R. Simari. 01/01/15 – 31/12/2018. Unidad coordinadora: Universidad Nacional del Sur.
- **“Combinación de Revisión de Creencias y Argumentación para mejorar las capacidades de Razonamiento y modelado de la Dinámica de Conocimiento en Sistemas Multi-agente”**. Director: Marcelo A. Falappa. PIP 112-20110101000. Unidad coordinadora:

Consejo Nacional de Investigaciones Científicas y Técnicas.

La línea de investigación se enmarca dentro del ámbito de colaboración entre el Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) del Instituto de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur/CONICET y el Área de Agentes y Sistemas Inteligentes de la Facultad de Ciencias de la Administración, Universidad Nacional de Entre Ríos.

2. Introducción

El trabajo colaborativo y el intercambio de información son aspectos cruciales de todo sistema de información, esto hace necesario contar con métodos automáticos y adecuados para el manejo de conflictos, esto es, inconsistencias e incoherencias que aparecen naturalmente en la medida que el conocimiento evoluciona. En entornos de este tipo, las bases de conocimiento son representadas a través de ontologías.

En particular y citando el análisis en [Dea15, DMFS16], el manejo de información conflictiva en bases de conocimientos es un importante problema que debe ser atacado [GCS10, HvHtT05, BQL07], especialmente cuando se integra conocimiento proveniente de diferentes fuentes [BHP09, AK05], o cuando tal conocimiento será explotado por procesos de razonamiento automáticos. El más conocido dentro de los conflictos es la inconsistencia. Inconsistencia es un concepto clásico dentro de la Representación de Conocimiento y el Razonamiento, y se refiere a teorías tales que no es posible encontrar para ellas un modelo; es decir, al menos una interpretación que haga verdaderas a todas las fórmulas de la teoría. En términos sintácticos, un conjunto de fórmulas es consistente si y sólo si para toda fórmula A , no es posible deducir tanto A como $\neg A$. Adicionalmente, en entornos ontológicos se encuentra otro tipo de conflicto, relacionados a un fenómeno conectado en cierta forma con inconsistencia pero a su vez con sus características propias: la incoherencia [FHP+06, BB97, KPSH05, SHCvH07, QH07]. Tal fenómeno surge cuando una porción de conocimiento expresado a través de un conjunto de reglas no puede ser aplicado sin generar problemas de consistencia; es decir, el conjunto de reglas no puede ser aplicado sin violar inevitablemente alguna de las restricciones impuestas al conocimiento, haciéndolas por lo tanto insatisfacibles. Intuitivamente podemos ver a la incoherencia como una inconsistencia latente o potencial. Consecuentemente, si este conjunto de reglas incoherente es considerado junto con hechos relevantes (hechos que activen las reglas en cuestión) entonces la violación es inevitable, generando inconsistencia. Esto marca la diferencia más grande con el concepto de inconsistencia: la incoherencia no necesariamente implica una contradicción; al punto de que incoherencia no

puede provocar una violación explícita sin ser considerada junto con instanciaciones de las reglas. Es decir, mientras que la inconsistencia es un conflicto puesto de manifiesto (explícito, una violación que ya está ocurriendo) la incoherencia se refiere a conflictos latentes (implícitos, pero que no necesariamente estén ocurriendo en el momento); se podría pensar por lo tanto en la incoherencia como un preludio de la inconsistencia. En estos entornos, para tratar con los conflictos planteados y recuperar la coherencia y la consistencia, existen dos enfoques:

- El primero de ellos implica modificar la información contenida en la ontología inconsistente de forma tal que la consistencia de la misma sea recuperada, para luego aplicar relaciones de consecuencia clásicas sobre el conocimiento consistente obtenido. De esta forma, se modifica la base de conocimiento original (e inconsistente), de manera que la base de conocimiento modificada, pueda aplicar relaciones de inferencia clásicas de manera segura. Este es uno de los objetivos perseguidos por la teoría de Revisión de Creencias [Gäro3, AGM85, Dal88, QLBo6, Han94].
- El segundo enfoque, propone modificar la forma en cómo el conocimiento es inferido mediante la definición de relaciones de inferencia alternativas, estableciendo de esta forma métodos de razonamiento tolerantes a inconsistencias, que son aquellos que pueden obtener conclusiones consistentes a partir de conocimiento inconsistente. Entre los ejemplos más importantes de tales métodos podemos citar a la Argumentación y el Razonamiento Rebatible [SL92, RSvBo9, GSo4, Pra10].

Siguiendo la línea del primer enfoque, trabajos previos [Dea15, DMFS16] han presentado de manera teórica, una alternativa para conseguir la consolidación de ontologías Datalog+/- . Este es un framework para realizar consultas de ontologías tratables y otras múltiples aplicaciones, basado en el lenguaje Datalog [CGL12]. Extiende el lenguaje incorporando características tales como cuantificadores existenciales para la cabeza de reglas y al mismo tiempo, restricciones en reglas de sintaxis para alcanzar decidibilidad y tratabilidad. La familia de lenguajes de ontologías Datalog permite un estilo modular de representación de conocimiento mediante el uso de reglas de forma similar a la usada en Programación Lógica, y su decidibilidad le permite manejar los volúmenes masivos de datos que podemos encontrar en aplicaciones hoy en día, haciéndola útil en diferentes campos como la consulta de ontologías, extracción de datos en web o intercambio de datos [LMS12]. En particular, la representación de conocimiento en ontologías Datalog+/- se lleva a cabo mediante el uso de (a) una Base de Datos: un conjunto de átomos que representan hechos acerca del mundo, e.g., alumno(pedro) (b) Tuple-generating Dependencies - TGDs: reglas que nos permiten obtener nuevos átomos mediante la activación

de las mismas como ser: $\text{alumno}(X) \rightarrow \text{persona}(X)$, (c) Equality-generating Dependencies: reglas que restringen la generación de átomos, por ejemplo: $\text{doctor}(D, P) \wedge \text{doctor}(D', P) \rightarrow D = D'$; y (d) Negative Constraints –NCs: reglas que expresan relaciones que no pueden existir entre átomos, e. g.: $\text{alto}(X) \wedge \text{bajo}(X) \rightarrow \perp$. Los métodos a implementar se enfocarán, por lo tanto, en controlar la relación entre los átomos y las TGDs presentes en la ontología, de forma que se respeten las EGDs y NCs en la KB.

Para ello se tomará como marco teórico los operadores de consolidación de ontologías Datalog+/- presentados en [Dea15, DMFS16] que, basándose en técnicas de Revisión de Creencias, atacan los conflictos citados anteriormente (inconsistencia e incoherencia). El primer enfoque (basado en las ideas de Hansson [Han94, Han99]) consiste atacar los conflictos, identificando conflictos mínimos conocidos como kernels y removiendo de ellos fórmulas para lograr resolverlos. En el segundo enfoque se utiliza una mirada más global, considerando la relación existente entre los diferentes kernels mediante la agrupación de los mismos en una superestructura denominada cluster, la cual es obtenida explotando una relación de solapamiento entre conflictos.

3. Líneas de Investigación y Desarrollo

Esta línea de investigación se enfoca en la implementación de procesos de resolución de conflictos en ontologías Datalog+/- a través del uso de formalismos enfocados en la resolución de incoherencias e inconsistencias provenientes de las áreas de Revisión de Creencias y Argumentación. Para ello distintos ejes deben ser investigados, que van desde la definición de incoherencias e inconsistencia en el entorno de ontologías Datalog+/- hasta las posibles aplicaciones que un método automático de integración de estas ontologías podría tener.

3.1. Definición de métodos de identificación de Incoherencias e Inconsistencias en Datalog+/-

Datalog+/- se ha vuelto un lenguaje muy popular en los últimos años, y numerosos estudios se han realizado acerca de sus propiedades de decidibilidad y la complejidad asociada a la respuesta de consultas en estas ontologías. Sin embargo, no ha habido mucho estudio acerca de los aspectos de Representación de Conocimiento en Datalog+/- . Muy poco trabajo se ha hecho acerca de inconsistencias en ontologías Datalog+/- . Peor aún es la situación respecto del concepto de incoherencia (*i. e.*, la imposibilidad de satisfacer cierto conjunto de reglas sin violar una restricción impuesta al conocimiento expresado en la ontología).

Uno de los ejes de esta línea de investigación es la definición formal del concepto de incoherencia en Datalog+/-, tomando como partida esfuerzos similares que han sido realizados para otros formalismos de representación de conocimiento, principalmente Description Logics. Adicionalmente, se procederá a identificar las propiedades que llevan a que un conjunto de TGDs sea incoherente, y las que hacen que una ontología Datalog+/- se vuelva inconsistente. De esta manera se podrán identificar tales casos, lo que será el primer paso para la posterior resolución de tales problemas.

Esto es sobremanera importante a fines de los objetivos de la presente línea de investigación, ya que es necesario tener identificadas tanto las características de cada tipo de conflicto como la relación entre las mismas para poder realizar una detección y tratamiento eficientes de conflictos. Para ver este más claramente considere el caso de la incoherencia: la detección de la misma involucra una aplicación iterativa de las TGDs, lo que debido a la suposición de mundo abierto es potencialmente infinita. Por lo tanto se deben definir métodos algorítmicamente realizables de detección de conjuntos insatisfacibles de TGDs minimales.

3.2. Resolución de Incoherencias e Inconsistencias en Datalog+/-

Una vez que se tiene definidos e identificados los conjuntos incoherentes de TGDs y aquellos conjuntos de átomos que provocan inconsistencias en la unión de varias ontologías Datalog+/-, se debe proceder a la resolución de estos conflictos. En esta línea de investigación esto se hará mediante la aplicación de técnicas derivadas de la Revisión de Creencias denominadas Kernel Contraction.

Este tipo de técnicas resuelve conflictos de incoherencia/inconsistencia tomando los conjuntos conflictivos mínimos y eligiendo de alguna forma que elemento remover de los mismos para solucionar el problema. En el caso de integración de ontologías Datalog+/- esto es la remoción de ciertos átomos y ciertas TGDs de la unión de todas las ontologías que se está integrando. Adicionalmente, se puede pensar en la definición de métodos de debilitamiento de reglas, en lugar de la remoción de las mismas. Esto no es una tarea trivial, ya que desde un punto de vista implementacional esto implica diversas decisiones de diseño que pueden tener un alto impacto en la eficiencia de los operadores implementados. Por ejemplo, es importante definir de manera adecuada como elegir el mejor candidato entre los átomos o TGDs que pueden eliminarse, lo que a su vez lleva a definir formas (automáticas) de obtener órdenes entre los candidatos. Como tales órdenes serán utilizados permanentemente por los operadores (cada vez que un conflicto deba ser resuelto) los algoritmos a desarrollar deberán utilizar estructuras de datos lo suficientemente ágiles en su dinamicidad para

permitir un reordenamiento eficiente de sus elementos. En particular, es necesario examinar distintas opciones de estructuras de datos y estructuras de control para encontrar aquellas más adecuadas para los fines planteados, tanto en lo referido a la detección de conflictos como a su posterior resolución.

3.3. Análisis computacional del comportamiento de los operadores de consolidación

Como se mencionó previamente, una línea importante de investigación involucra un estudio del comportamiento de los operadores desarrollados. Tal análisis involucra diferentes aspectos que deben ser considerados. Por un lado, es importante el realizar un estudio de la complejidad computacional de las diferentes tareas que involucra la resolución de incoherencias e inconsistencias, para establecer la tratabilidad o no de la tarea. Por otro lado, es interesante el saber como afecta el crecimiento de las ontologías (y de esta manera posiblemente la cantidad de conflictos a resolver) en el desempeño de los algoritmos desarrollados. Para esto se procederá a realizar diferentes análisis prácticos de la eficiencia de los algoritmos, aplicándolo a diferentes ontologías en distintos entornos de aplicación.

4. Resultados y Objetivos

El objetivo general de este trabajo es comprobar empíricamente la correctitud, computabilidad y eficiencia de los operadores de consolidación de ontologías en Datalog+/- . Para esto se analizará a fondo las diferentes características tanto de las ontologías como de los conflictos que en las mismas pueden aparecer, para de esta forma detectar aquello que pueda afectar, tanto positiva como negativamente, la eficiencia en la resolución de conflictos.

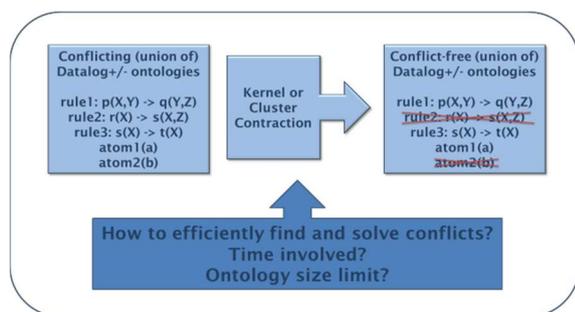


Figura 1: Aspectos a considerar.

Para conseguir este objetivo general hay diferentes objetivos específicos a perseguir

4.1. Estudiar los componentes intervinientes en la ejecución de un programa Datalog+/-

Se analizarán los diferentes componentes intervinientes en la resolución de consultas en Datalog+/- , y cómo los mismos colaboran a tal fin. De ser necesario se complementará la documentación existente en la literatura con diagramas explicativos tanto de funcionamiento de componentes aislados como de interacciones ocurrientes.

4.2. Diseñar e implementar diferentes soluciones para las operaciones requeridas

Se diseñarán de forma abstracta diferentes algoritmos que implementen las operaciones de consolidación en ontologías Datalog+/- . Esta tarea no es trivial ya que los diferentes tipos de conflictos posibles en una ontología Datalog+/- (esto es, incoherencias e inconsistencias) poseen características que los distinguen y pueden dificultar la tarea de encontrar los conflictos minimales (*kernels* en la terminología de Hansson [Han99]). Esto es particularmente cierto para el fenómeno de incoherencia, ya que la aplicación de reglas en Datalog+/- puede provocar una secuencia infinita, por lo que mecanismos adecuados deben ser ideados al momento de buscar los conjuntos insatisfacibles de reglas [DMFS16] para evitar estos inconvenientes derivados del uso de reglas existencialmente cuantificadas en su cabeza. Para ellos se estudiarán el uso de diferentes estructuras de control y de datos según el caso lo requiera.

4.3. Realizar análisis de complejidad y eficiencia de las soluciones propuestas

Las soluciones propuestas en el apartado anterior serán sometidas a análisis de complejidad y eficiencia. Se analizarán los recursos computacionales de tiempo y espacio que cada solución exija, para determinar su aplicabilidad en dominios de aplicación de la vida real.

4.4. Realizar pruebas unitarias, de integración y funcionales

Se utilizarán frameworks de pruebas con la finalidad de asegurar el correcto funcionamiento de los componentes del software de manera aislada y en grupo. Se pondrán a prueba los operadores desarrollados utilizando diferentes programas Datalog+/- .

5. Formación de Recursos Humanos

En la presente línea de investigación se enmarca el desarrollo de una tesis para la Maestría en Sistemas de la Información de la Universidad Nacional de Entre Ríos.

Referencias

- [AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. Symb. Log.*, 50(2):510–530, 1985.
- [AK05] L. Amgoud and S. Kaci. An argumentation framework for merging conflicting knowledge bases: The prioritized case. In *ECSQARU 2005*, pages 527–538, 2005.
- [BB97] D. Beneventano and S. Bergamaschi. Incoherence and subsumption for recursive views and queries in object-oriented data models. *Data Knowl. Eng.*, 21(3):217–252, 1997.
- [BHP09] E. Black, A. Hunter, and J. Z. Pan. An argument-based approach to using multiple ontologies. In *SUM*, pages 68–79, 2009.
- [BQL07] D. A. Bell, G. Qi, and W. Liu. Approaches to inconsistency handling in description logic based ontologies. In *OTM Workshops (2)*, pages 1303–1311, 2007.
- [CGL12] A. Cali, G. Gottlob, and T. Lukasiewicz. A general datalog-based framework for tractable query answering over ontologies. *J. Web Sem.*, 14:57–83, 2012.
- [Dal88] Mukesh Dalal. Investigations into a theory of knowledge base revision. In *AAAI*, pages 475–479, 1988.
- [Dea15] Cristhian Ariel David Deagustini. *Consolidación de Ontologías Datalog ±*. PhD thesis, Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, 2015.
- [DMFS16] Cristhian Ariel David Deagustini, Maria Vanina Martinez, Marcelo A. Falappa, and Guillermo Ricardo Simari. Datalog+ ontology consolidation. *J. Artif. Intell. Res. (JAIR)*, 56:613–656, 2016.
- [FHP+06] G. Flouris, Z. Huang, J. Z. Pan, D. Plexousakis, and H. Wache. Inconsistencies, negations and changes in ontologies. In *AAAI*, pages 1295–1300. AAAI Press, 2006.
- [Gär03] Peter Gärdenfors. *Belief revision*, volume 29. Cambridge University Press, 2003.
- [GCS10] S. A. Gómez, C. I. Chesñevar, and G. R. Simari. Reasoning with inconsistent ontologies through argumentation. *Appl. Artif. Intell.*, 24(1&2):102–148, 2010.
- [GSo4] A. J. García and G. R. Simari. Defeasible logic programming: An argumentative approach. *TPLP*, 4(1-2):95–138, 2004.
- [Han94] S. O. Hansson. Kernel contraction. *J. Symb. Log.*, 59(3):845–859, 1994.
- [Han99] S. O. Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Publishers, 1999.
- [HvHtTo5] Zhisheng Huang, Frank van Harmelen, and Annette ten Teije. Reasoning with inconsistent ontologies. In *Proc. of IJCAI 2005*, pages 454–459, 2005.
- [KPSH05] A. Kalyanpur, B. Parsia, E. Sirin, and J. A. Hendler. Debugging unsatisfiable classes in owl ontologies. *J. Web Sem.*, 3(4):268–293, 2005.
- [LMS12] T. Lukasiewicz, M. V. Martinez, and G. I. Simari. Inconsistency handling in datalog+/- ontologies. In *Proc. of ECAI*, pages 558–563, 2012.
- [Pra10] Henry Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1:93–124, 2010.
- [QH07] Guilin Qi and Anthony Hunter. Measuring incoherence in description logic-based ontologies. In *ISWC/ASWC*, pages 381–394, 2007.
- [QLBo6] Guilin Qi, Weiru Liu, and David A. Bell. Knowledge base revision in description logics. In *JELIA*, pages 386–398, 2006.
- [RSvBo9] I. Rahwan, G. R. Simari, and J. van Benthem. *Argumentation in artificial intelligence*, volume 47. Springer, 2009.
- [SHCvHo7] Stefan Schlobach, Zhisheng Huang, Ronald Cornet, and Frank van Harmelen. Debugging incoherent terminologies. *J. Autom. Reasoning*, 39(3):317–349, 2007.
- [SL92] G. R. Simari and R. P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artif. Intell.*, 53(2-3):125–157, 1992.

Integración de Ontologías Datalog \pm Mediante Consolidación a Través de Debilitamiento

Cristian D. Pacifico

**Cristhian A. D. Deagustini
Guillermo R. Simari**

Marcelo A. Falappa

Area de Agentes y Sistemas Inteligentes

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Tavella 1424, (3200) Concordia, Entre Ríos - (0345) 423-1400

Laboratorio de Investigación y Desarrollo en Inteligencia Artificial

Instituto de Ciencias e Ingeniería de la Computación - Universidad Nacional del Sur - CONICET

Alem 1253 - Bahía Blanca - Buenos Aires - Argentina - (0291) 459-5135

cripac@fcad.uner.edu.ar, (cadd,mfalappa,grs)@cs.uns.edu.ar

Resumen

En sistemas que manejan e intercambian grandes volúmenes de información es necesario contar con métodos adecuados para manejar conflictos que aparecen en la medida que sus bases de conocimiento evolucionan y se integra con otras fuentes de información. Este conocimiento es frecuentemente expresado por ontologías; no solo por el valor expresivo, sino también por las diferentes restricciones semánticas que se pueden aplicar al conocimiento, originadas por el dominio de aplicación o por los propios usuarios.

Esta línea de I+D profundiza en la resolución automatizada de conflictos en ontologías expresadas en Datalog \pm ; ya sea estos generados por inconsistencia o incoherencia. Se pretende definir métodos que resuelven los conflictos que surgen al integrar dos o más ontologías Datalog \pm . Una operación de integración sobre varias ontologías, debe restaurar la consistencia y coherencia en el todo, no solo considerando las características intrínseca de la información integrada, sino también meta-información relacionada con la confiabilidad de cada fuentes, la ocurrencia de información, etc. En particular, se definirán mecanismos resolución de incoherencias e inconsistencias que produzcan alguna forma de debilitamiento o modificación de las reglas; para lograr la mínima pérdida de información en el proceso de integración. Finalmente, se analizará los mecanismos descriptos para extenderlos a la integración con otras fuentes de datos, en particular con Bases de Datos Relacionales.

Palabras Clave: Integración de Bases de Conocimiento, Revisión de Creencias, Representación de Conocimiento, Razonamiento.

1. Contexto

Esta línea de investigación se lleva a cabo en el marco de los siguientes proyectos de investigación:

- **“Representación de conocimiento y razonamiento argumentativo: Herramientas inteligentes para la web y las bases de conocimiento”**. Director: Guillermo R. Simari. 01/01/15 – 31/12/2018. Unidad coordinadora: Universidad Nacional del Sur.
- **“Combinación de Revisión de Creencias y Argumentación para mejorar las capacidades de Razonamiento y modelado de la Dinámica de Conocimiento en Sistemas Multi-agente”**. Director: Marcelo A. Falappa. PIP 112-20110101000. Unidad coordinadora: Consejo Nacional de Investigaciones Científicas y Técnicas.

La línea de I+D se enmarca dentro del ámbito de colaboración entre el Laboratorio de Investigación y Desarrollo en Inteligencia Artificial (LIDIA) del Instituto de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur/CONICET y el Área de Agentes y Sistemas Inteligentes de la Facultad de Ciencias de la Administración, Universidad Nacional de Entre Ríos.

2. Introducción

2.1. Conflictos en Ontologías

En sistemas que manejan e intercambian grandes volúmenes de información es necesario contar con métodos automáticos y adecuados para el manejo de conflictos que aparecen naturalmente en la medida que el conocimiento evoluciona.

En los últimos tiempos, las Bases de Conocimiento expresadas mediante ontologías ha sido un temática de interés para áreas de investigación y desarrollo. En particular las ontologías son importantes no sólo por el conocimiento que pueden expresar, sino también por las diferentes restricciones semánticas que se pueden aplicar a dicho conocimiento.

El manejo de información conflictiva en bases de conocimiento es un importante problema que debe ser atacado [BQLO7], especialmente cuando se integra conocimiento proveniente de diferentes fuentes [AK05], o cuando tal conocimiento será explotado por procesos de razonamiento automáticos.

El más conocido de conflictos en información es la *inconsistencia*. El concepto se refiere a teorías tales que es imposible encontrarles un *modelo*; es decir, al menos una *interpretación* que haga verdaderas a todas sus fórmulas. En términos sintácticos, un conjunto de fórmulas es consistente si y sólo si para toda fórmula A , no es posible deducir tanto A como $\neg A$.

Adicionalmente, en entornos ontológicos se encuentra otro tipo de conflicto, relacionado a un fenómeno conectado con la inconsistencia pero a su vez con sus características propias: la *incoherencia* [FHP⁺06, QH07]. Este fenómeno surge cuando una porción de conocimiento expresado por un conjunto de reglas no pueda ser aplicado sin generar problemas de consistencia; es decir, el conjunto de reglas no puede ser aplicadas sin violar inevitablemente alguna de las restricciones impuestas al conocimiento, haciéndolas por lo tanto insatisfacibles.

En estos entornos, una posibilidad para tratar estos conflictos y recuperar la coherencia y la consistencia implica modificar la información contenida en la ontología inconsistente de forma tal que la consistencia sea recuperada. De esta forma, se modifica la base de conocimiento original (e inconsistente), de manera que la base de conocimiento modificada, pueda aplicar relaciones de inferencia clásicas de manera segura. Este es uno de los objetivos perseguidos por la teoría de *Revisión de Creencias* [Gäro3, AGM85, Han94].

Siguiendo este enfoque, trabajos previos [Dea16, DMFS16] han presentado de manera teórica, una alternativa para conseguir la consolidación de ontologías *Datalog*[±]. *Datalog*[±] es un framework para realizar consultas de ontologías tratables y otras múltiples aplicaciones, basado en el lenguaje *Datalog* [CGL12], incorporando características tales como cuantificadores existenciales para la cabeza de reglas y restricciones en sintaxis para alcanzar decidibilidad y, en una gran parte de los casos, incluso tratabilidad. En particular, la representación de conocimiento en ontologías *Datalog*[±] se lleva a cabo mediante el uso de (a) una *Base de Datos*: un conjunto de átomos que representan hechos acerca del mundo, e. g., *alumno(pedro)* (b) *Tuple-generating Dependencies (TGDs)*: reglas que permiten obtener nuevos átomos mediante la activación de las mismas como ser: *alumno(X) → persona(X)*,

(c) *Equality-generating Dependencies (EGDs)*: reglas que restringen la generación de átomos, por ejemplo: *doctor(D, P) ∧ doctor(D', P) → D = D'*; y (d) *Negative Constraints (NCs)*: reglas que expresan relaciones que no pueden existir entre átomos, e. g.: *alto(X) ∧ bajo(X) → ⊥*.

Un importante beneficio por utilizar bases de conocimientos es la posibilidad de manejar dinámica de conocimiento. Es decir, aceptar nueva información manteniendo la integridad y consistencia de la base. En este sentido, el desafío es eliminar los conflictos que se producen en el proceso de actualización de la base de conocimiento ante las novedades. Para esto, se tomará como marco teórico los operadores de consolidación de ontologías *Datalog*[±] presentados en [Dea16, DMFS16] que, basándose en técnicas de Revisión de Creencias, atacan los conflictos citados anteriormente (inconsistencia e incoherencia). El primer enfoque (basado en las ideas de Hansson [Han94, Han01]) consiste atacar los conflictos, identificando conflictos mínimos conocidos como *kernels* y removiendo de ellos fórmulas para lograr resolverlos. En el segundo enfoque se utiliza una mirada más global, considerando la relación existente entre los diferentes kernels mediante la agrupación de los mismos en una superestructura denominada *cluster*, la cual es obtenida explotando una relación de solapamiento entre conflictos.

2.2. Revisión de Creencias

El principal objetivo de la *Revisión de Creencias* o la *Teoría de Cambio*, es el modelado de la dinámica del conocimiento. Es decir, la conceptualización del modo en que cambia el conocimiento y se adapta a influencias externas, luego de recibir información nueva.

En los 80', Carlos Alchourrón, David Makinson y Peter Gärdenfors convergieron sus trabajos previos hacia la formalización del *modelo AGM* [AGM85]. Este trabajo es fundacional en la teoría de cambio y proveyó el marco de trabajo (framework) formal más general hasta ese momento para estudios de cambio de creencias.

El fundacional *modelo AGM* especificó un *modelo epistémico* de dinámica de conocimiento; es decir, se definió el formalismo en el cual las *creencias* (o el conocimiento) son representadas y en el cual diferentes clases de operaciones de cambio (*operadores*) pueden ser definidos. A partir de este trabajo, diferentes frameworks para la dinámica de conocimiento han sido propuestos, basándose en este "estilo" de presentación. Las operaciones básicas de cambio definidas en AGM, y tomadas en trabajos subsiguientes, sobre una base de conocimiento KB y una sentencia α , son:

- **Expansión:** expandir la KB mediante α resulta en una nueva KB^f que infiere a α .
- **Contracción:** contraer a KB , por α resulta en una nueva KB^f que no infiere a α .

- **Revisión:** revisar a KB por α resulta en una KB^r , en la cual α es *consistentemente* inferida.

El trabajo [AGM85] se basó en la utilización de teoremas de representación (también llamados caracterizaciones axiomáticas), los cuales sirven para brindar una clara conexión entre las construcciones de los operadores y las propiedades esperadas de los mismos. Tales teoremas se basan en mostrar la relación entre dos componentes fundamentales: los postulados y las construcciones de operadores.

Los postulados de racionalidad [AGM85, Gär88] determinan restricciones que los operadores deben satisfacer. Ellos tratan a los operadores como cajas negras, describiendo sus comportamientos con respecto a la entrada en casos básicos, pero no los mecanismos internos usados. Por otro lado las construcciones se refieren a la contraparte de los postulados, es decir, a definir como conseguir el objetivo esperado por parte del operador. Existen modelos equivalentes posteriores al AGM, tales como: *safe contractions* [AM85], *epistemic entrenchment* [GM88], *sphere systems* [Gro88], y *kernel contractions* [Han94].

Supongamos que KB representa el conocimiento actual y α es una nueva pieza de información. Supongamos que una contracción de KB por α es realizada. Siguiendo el modelo AGM, una *partial meet contraction* está basada en una selección entre los subconjuntos maximales de KB que no implican α . Alternativamente, otra posibilidad es realizar una selección entre los subconjuntos minimales de KB que implican α tal como en *safe contraction* [AM85]. Una variante más general del mismo enfoque fue introducido más tarde y es conocido como *kernel contraction* [Han94]. De esta forma, el proceso de resolución de conflictos es definido como la identificación de tales conflictos mínimos y la posterior remoción de alguna fórmula cada uno de ellos. Logrando así, que la KB se altere lo menos posible. Este estrategia es muy natural para un entorno de aplicación de bases de conocimiento finitas. Ha sido mostrado que las *safe contractions* y las *kernel contractions* son equivalentes a las *partial meet contractions*.

En otros trabajos en el área, otras operaciones adicionales fueron desarrolladas.

La operación de **consolidación de creencias** [Hano1] se aplica cuando una KB es inconsistente. La operación, denotada por $KB!$, se encarga de la remoción (*i. e.*, contracción) de suficientes sentencias de KB hasta llevarla a un estado de consistencia.

La operación de **combinación de creencias** se aplica en situaciones donde el conocimiento es provisto por diversas fuentes y es frecuente que aparezcan conflictos al unir la información. Se puede descomponer la combinación en dos suboperaciones: (a) reunir toda información provista en un único cuerpo, generalmente por unión de conjuntos; y (b) resolver los conflictos que aparezcan en el conocimiento unificado (*i. e.*, *consolidar* el mismo). Como trabajo destacado, Baral *et al.* [BKM91] ha establecido *operadores de combinación*

basados en la unión de todos los conjuntos de creencias, y en la selección de subconjuntos maximalmente consistentes mediante el uso de algún orden dado.

Finalmente, la operación de **integración de creencias** se propone unificar de manera consistente conocimiento provisto por diversas fuentes. Difiere del objetivo de la *combinación* en el sentido que, en los operadores de integración, siempre es posible identificar la procedencia de las distintas piezas de información con las que se cuenta; por lo tanto, en la resolución de conflictos es posible considerar no sólo las características intrínsecas a la información sino también otras informaciones que pueden ser de utilidad, como por ejemplo la cantidad de fuentes que soportan una pieza de información particular, o la credibilidad asociada a las mismas. El principal trabajo relacionado es el de Konieczny y Pino-Pérez [KP02], quienes sentaron las bases más importantes respecto de los operadores de integración. Además proponen una construcción posible para operadores de integración basados en el uso de pre-órdenes totales del conocimiento provisto por las diversas fuentes.

3. Líneas de Investigación y Desarrollo

Esta línea de I+D trata sobre la definición de procesos de integración de ontologías Datalog[±]. Para ello, se utilizan formalismos destinados a la resolución de incoherencias e inconsistencias provenientes de las áreas de Revisión de Creencias y Argumentación. Muchos de estos enfoques, basan su operación en la remoción de información como método de resolución de conflictos. Se propone una resolución alternativa de conflictos que surge de considerar varias ontologías en conjunto mediante la aplicación de *debilitamiento*. En lugar de remover completamente la información conflictiva, se modifica la misma de forma que los conflictos queden resueltos pero no haya que descartar (en su totalidad) la información.

Para esto, distintos ejes deben ser investigados, abarcando desde la definición de incoherencias e inconsistencia en el entorno de ontologías Datalog[±] hasta las posibles aplicaciones que podría tener un método automático de integración de estas ontologías.

3.1. Identificación de Incoherencias e Inconsistencias

A pesar de su creciente popularidad, no ha habido mucho estudio acerca de los aspectos de Representación de Conocimiento en Datalog[±]. Muy poco trabajo se ha hecho acerca de inconsistencias en ontologías Datalog[±]. Peor aún es la situación respecto del concepto de incoherencia.

Uno de los ejes de esta línea de I+D es la definición formal del concepto de incoherencia en Datalog[±], to-

mando como partida esfuerzos similares que han sido realizados para otros formalismos de representación de conocimiento, principalmente Description Logics. Adicionalmente, se procederá a identificar las propiedades que identifican a un conjunto de TGDs como incoherente, y las que hacen que una ontología Datalog[±] se vuelva inconsistente. De esta manera se podrán identificar tales casos, lo que será el primer paso para la posterior resolución de tales problemas.

3.2. Resolución por debilitamiento

Una vez definidos e identificados los conjuntos incoherentes de TGDs y aquellos conjuntos de átomos que provocan inconsistencias en la unión de varias ontologías Datalog[±], se debe proceder a la resolución de estos conflictos. En esta línea de I+D se hará mediante la aplicación de técnicas derivadas de Revisión de Creencias inspiradas en Kernel Contraction [Han94], donde el foco estará puesto en realizar debilitamiento, en lugar del enfoque tradicional que utiliza remoción de información.

En Kernel Contraction clásico, la resolución de incoherencia/inconsistencia se realiza tomando los conjuntos conflictivos mínimos y eligiendo de alguna forma que elemento remover de los mismos para solucionar el problema. En el caso de integración de ontologías Datalog[±] esto es la remoción de ciertos átomos y ciertas TGDs de la unión de todas las ontologías que se está integrando. Aquí se propone tomar ventaja de este enfoque modificando para que se utilicen métodos de debilitamiento de reglas.

Evidentemente esto no resulta trivial, ya que dado un conjunto conflictivo minimal no es claro cual de sus elementos deberíamos modificar. A su vez, incluso sabiendo cual modificar realizar tal modificación es una tarea compleja, donde surgen preguntas tales como qué modificación realizar. Por ejemplo, es necesario identificar si es mejor realizar debilitamiento mediante el agregado de átomos en el cuerpo de una TGD, o si es preferible modificar los cuantificadores que haya en la misma; o si hay características de las ontologías que hacen que una opción es mejor que otra.

3.3. Integración de ontologías Datalog[±]

En base a los métodos descritos de resolución de incoherencias e inconsistencias se procederá a definir operadores de integración de ontologías Datalog[±] basados en debilitamiento. Además, se darán las propiedades esperadas de tales operadores y se definirán métodos de obtención de operadores de tales características mediante Teoremas de Representación.

3.4. Integración de otras fuentes

Finalmente, se analizará el framework de integración de ontologías Datalog[±] como medio de integración de



otras fuentes de datos. Principalmente nos enfocaremos en la integración automática de Bases de Datos Relacionales.

Para esto, primeramente definiremos métodos para expresar bases de datos relacionales a través de ontologías Datalog[±], tanto los datos en sí como aspectos relacionados al esquema de las mismas, *e. g.*, las dependencias funcionales. Una vez logrado esto se podría utilizar los métodos de integración de ontologías Datalog[±] para obtener una federación de las bases de datos expresadas, ya que la ontología final resultante de la integración brindaría un esquema integrador de las mismas así como los datos que serían parte de la federación, manteniendo a su vez la coherencia de las restricciones de integridad respecto del esquema unificado y la consistencia de los datos almacenados respecto de las dependencias funcionales.

4. Resultados y Objetivos

El objetivo final de esta línea de investigación es el diseño y construcción de la infraestructura necesaria para la realización de procesos de argumentación en una escala masiva valiéndose de datos almacenados en Bases de Datos heterogéneas integradas en Federaciones.

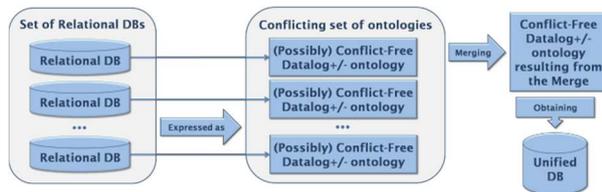
En particular, en el presente trabajo se propone la representación e integración de tales bases de datos a través de la utilización de ontologías Datalog[±], las cuales poseen una expresividad y tratabilidad adecuadas para tal fin. De esta forma, todos los desarrollos que se logren en Datalog[±] serían transferibles a bases de datos (y otras tecnologías estrictamente menos expresivas que Datalog[±] como ciertas Description Logics, *e. g.*, DL-Lite o *EL*).

Un aspecto sumamente importante para la consecución del objetivo final es la definición de métodos automáticos que permitan la creación de vistas unificadas de varias bases de datos como forma de construir las federaciones de bases de datos, de forma que los procesos de argumentación accedan a las mismas para dar soporte a los argumentos. En este trabajo proponemos la consecución de este mediante el uso de integración de ontologías a través de técnicas de debilitamiento.

- Definición de métodos de consolidación de ontologías Datalog[±] que permitan el manejo de inconsistencia e incoherencia. En particular, se propone utilizar un esquema de debilitamiento de reglas y átomos mediante la modificación de los mismos. Para esto, se aprovechará tanto enfoques clásicos

de Revisión de Creencias como Kernel Contraction [Han94, Han01] como así también refinamientos desarrollados en la presente línea de investigación (Cluster Contraction [Dea16, DMFS16]).

- En base a los mecanismos desarrollados procederemos a definir métodos que permitan la integración de ontologías Datalog[±], mediante la consolidación de la unión de ontologías Datalog[±]. Una vez logrado esto, se podría utilizar tales métodos para obtener una federación de las bases de datos expresadas, ya que la ontología final resultante de la integración brindaría un esquema integrador de las mismas, así como los datos que serían parte de la federación, manteniendo a su vez la coherencia de las restricciones de integridad respecto del esquema unificado y la consistencia de los datos almacenados respecto de las dependencias funcionales, como se muestra a continuación.



5. Formación de RR.HH.

En la presente línea de I+D se enmarca el desarrollo de una tesis doctoral para el Doctorado en Ciencias de la Computación de la Universidad Nacional del Sur.

Referencias

- [AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.
- [AK05] Leila Amgoud and Souhila Kaci. An argumentation framework for merging conflicting knowledge bases: The prioritized case. In *ECSQARU*, pages 527–538, 2005.
- [AM85] Carlos Alchourrón and David Makinson. On the Logic of Theory Change: Safe Contraction. *Studia Logica*, 44:405–422, 1985.
- [BKM91] Chitta Baral, Sarit Kraus, and Jack Minker. Combining multiple knowledge bases. *IEEE Trans. Knowl. Data Eng.*, 3(2):208–220, 1991.
- [BQL07] David A. Bell, Guilin Qi, and Weiru Liu. Approaches to inconsistency handling in description-logic based ontologies. In *OTM Workshops (2)*, pages 1303–1311, 2007.
- [CGL12] Andrea Calì, Georg Gottlob, and Thomas Lukasiewicz. A general datalog-based framework for tractable query answering over ontologies. *Web Semantics: Science, Services and Agents on the World Wide Web*, 14:57–83, 2012.
- [Dea16] Cristhian Ariel David Deagustini. *Consolidación de ontologías Datalog+*. PhD thesis, Departamento de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur., 2016.
- [DMFS16] Cristhian Ariel David Deagustini, Maria Vanina Martinez, Marcelo A. Falappa, and Guillermo Ricardo Simari. Datalog+-Ontology Consolidation. *J. Artif. Intell. Res. (JAIR)*, 56:613–656, 2016.
- [FHP⁺06] Giorgos Flouris, Zhisheng Huang, Jeff Z. Pan, Dimitris Plexousakis, and Holger Wache. Inconsistencies, negations and changes in ontologies. In *AAAI*, pages 1295–1300. AAAI Press, 2006.
- [Gär88] Peter Gärdenfors. *Knowledge in Flux: Modeling the dynamics of epistemic states*. MIT Press, 1988.
- [Gär03] Peter Gärdenfors. *Belief revision*, volume 29. Cambridge University Press, 2003.
- [GM88] Peter Gärdenfors and David Makinson. Revisions of knowledge systems using epistemic entrenchment. In *TARK*, pages 83–95, 1988.
- [Gro88] Adam Grove. Two modellings for theory change. *Journal of philosophical logic*, 17(2):157–170, 1988.
- [Han94] Sven Ove Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59(3):845–859, 1994.
- [Han01] Sven Ove Hansson. *A Textbook of Belief Dynamics: Solutions to Exercises*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.
- [KP02] Sébastien Konieczny and Ramón Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.
- [QH07] Guilin Qi and Anthony Hunter. Measuring incoherence in description logic-based ontologies. In *ISWC/ASWC*, pages 381–394, 2007.

Inteligencia Computacional Aplicada a la Optimización Multiobjetivo de Problemas de Scheduling con Restricciones

Pandolfi D., Villagra A., Valdez J.C., Molina D.
 Laboratorio de Tecnologías Emergentes (LabTEm)
 Instituto de Tecnología Aplicada - Unidad Académica Caleta Olivia
 Universidad Nacional de la Patagonia Austral
 {dpandolfi, avillagra, jvaldez, dmolina}@uaco.unpa.edu.ar

Resumen

Los problemas de Scheduling consisten en la asignación de tareas a recursos limitados donde ciertos objetivos deben optimizarse y varias restricciones deben cumplirse. La mayoría de los problemas del mundo real tienen varios objetivos que tratamos de optimizar al mismo tiempo. Estos se denominan problemas multiobjetivo y sus solución da lugar a un conjunto de soluciones que representan la posible negociación entre todos los objetivos. La negociación del conjunto óptimo de soluciones conforma el llamado frente de Pareto del problema de optimización multiobjetivo que se está resolviendo. Una aplicación muy interesante es la planificación del desarrollo de un yacimiento que requiere de un proceso altamente complejo e implica un número considerable de actividades. Estas actividades tienen que estar coordinadas, cumplir diversos tipos de restricciones y al mismo tiempo, optimizar criterios. El cronograma de las actividades en los pozos WAS (en inglés, Well Activity Scheduling) se ocupa de la coordinación para formar así cronogramas. Este proyecto propone desarrollar algoritmos metaheurísticos, que incorporen heurísticas y reglas de des-

pacho que sean competitivas con los algoritmos del estado del arte. Los enfoques propuestos tomarán como base algoritmos tales como cGA (celular Genetic Algorithm) y enfoques basados en algoritmos evolutivos multiobjetivo MOEA (Multi-Objective Evolutionary Algorithm).

Palabras clave: Metaheurísticas, Scheduling, Optimización Multiobjetivo, Manejo de Restricciones.

Contexto

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEm), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Caleta Olivia Universidad Nacional de la Patagonia Austral, en el marco del proyecto “Inteligencia Computacional aplicada a la optimización multiobjetivo de problemas de scheduling con restricciones”.

Introducción

Los problemas de planificación (Planing Problems) son una sub-clase de problemas de optimización combinatoria en campos tales operaciones de producción y despacho en la industria manufacturera y extrac-

tiva. Particularmente, los problemas de secuenciamiento de tareas conocidos como Scheduling Problems tienen su aplicación en la industria y por lo tanto un fuerte impacto económico y social, donde el objetivo principal de los investigadores es la reducción de los costos de producción en la industria [10]. Los problemas de Scheduling incluyen la combinación de recursos, tareas, objetivos y restricciones, donde un incremento en el tamaño del problema produce un incremento exponencial del espacio de soluciones. Muchos de los problemas de Scheduling son computacionalmente complejos y el tiempo requerido para calcular una solución óptima se incrementa con el tamaño del problema. Además, se ha demostrado, que muchos problemas de Scheduling pertenecen a la clase de NP-Hard [1].

El problema de secuenciamiento de tareas de Flow Shop (FSSP, en inglés Flow Shop Scheduling Problem) tiene como objetivo encontrar la secuencia óptima de n puestos de trabajo que pueden ser procesados en m máquinas. En [8] demostraron que este problema es NP-completo para instancias de más de dos máquinas, y por lo tanto esto muestra la dificultad de lograr soluciones óptimas según crece el tamaño del problema. Hejazi y Saghafian [9] realizaron un estudio sobre los trabajos publicados para resolver el FSSP y distintas variantes del mismo durante el período 1954-2004 donde se incluyen métodos exactos, como así también enfoques metaheurísticos. Framinan et al. [7] proporcionan una clasificación de las heurísticas más usadas para el FSSP.

En el problema de programación de tareas de Job Shop (JSSP, en inglés Job Shop

Scheduling Problems) existen m máquinas y n tareas. Cada tarea tiene una ruta predefinida que incluye a todas o a ciertas máquinas. Entre los JSSP la diferencia esencial radica en permitir o no la recirculación de un job en determinadas máquinas (es decir un job puede visitar a una máquina dada más de una vez). El JSSP no sólo es NP-duro, sino es una de las peores miembros de esta clase, Muth y Thompson [11].

La complejidad de muchos problemas de optimización discreta del mundo real está asociada con grandes espacios de búsqueda, demandas de rendimiento de tiempo real y ambientes dinámicos que no pueden ser resueltos por métodos exactos en tiempo razonable. Esto ha promovido en la comunidad científica el desarrollo de nuevos enfoques conocidos como metaheurísticas [2].

Los cGAs son una subclase de los GAs con una población estructurada espacialmente, es decir, los individuos de la población pueden aparearse solo con sus vecinos. Los cGAs, se diseñaron inicialmente para trabajar en máquinas paralelas, formadas por muchos procesadores que ejecutaban simultáneamente las mismas instrucciones sobre diferentes datos. El primer modelo de cGA fue propuesto por Robertson en 1987 y fueron inicialmente desarrollados para trabajar en máquinas paralelas, y más tarde se fueron adaptando para funcionar también en máquinas de procesamiento secuencial.

La mayoría de los problemas de optimización incluyen restricciones de cierta clase, que constituyen grandes desafíos a la resolución de problemas de optimización. Las

restricciones son límites impuestos a las variables de decisión y en general las restricciones son una parte integral de la formulación de cualquier problema [5]. El manejo de restricciones en algoritmos evolutivos puede abordarse de distintas técnicas y un amplio estudio del estado del arte es propuesto por Coello Coello [3]. La forma más común de incorporar restricciones en un algoritmo evolutivo ha sido a través de funciones de penalidad [12]. La idea de este método es transformar un problema de optimización con restricciones en un uno sin restricciones agregando (o sustrayendo) un cierto valor a la función objetivo basado en la cantidad de violaciones a las restricciones presentadas en una solución. Las técnicas basadas en algoritmos de reparación mapean una solución desde el espacio no factible al espacio factible. Los procesos de reparación que modifican genéticamente la solución no factible son conocidos como Lamarkianos y transforman una solución no factible en factible la que es evolucionada por el algoritmo. Un enfoque menos destructivo de la solución no factible, permite combinar el aprendizaje y la evolución, estrategia conocida como Baldwiniana. En este último enfoque las soluciones son reparadas sólo para su evaluación. Estudios analíticos y empíricos indican que esta técnica reduce la velocidad de convergencia del algoritmo evolutivo y permite converger a óptimos globales [13].

Los problemas con múltiples objetivos surgen de manera natural en la mayoría de las disciplinas y su solución ha sido un reto para los investigadores durante mucho tiempo. A pesar de la considerable variedad de técnicas desarrolladas en la investigación de operaciones para abordar estos problemas, la complejidad de su solución

exige enfoques alternativos. La formulación de un problema con un único objetivo es extendida a problemas multiobjetivo, donde no hay una función objetivo a optimizar, sino varias. Por lo tanto, aquí no hay una solución única, sino un conjunto de soluciones que se encuentra a través del uso de la teoría de Optimalidad de Pareto [6]. El uso de algoritmos evolutivos (AEs) para resolver problemas de esta naturaleza ha sido motivado principalmente debido a la naturaleza basada la población población de soluciones que permite la generación de varios elementos del óptimo de Pareto. Finalmente, Coello Coello [4], ofrece una descripción de trabajos de investigación más importante que se ha llevado a cabo en relación con la solución de problemas computacionalmente costosos de optimización multiobjetivo. Los enfoques que se discuten brevemente incluyen aproximación de funciones y aproximación evolutiva. Además, el uso de enfoques alternativos, tales como algoritmos culturales, pequeños tamaños de población y algoritmos híbridos que utilizan algunas soluciones para reconstruir el frente de Pareto con potentes motores de búsqueda locales.

Líneas de investigación y desarrollo

Las actividades petroleras de producción, explotación y transporte del petróleo son muy importantes para el desarrollo económico de la sociedad industrial moderna. Sin embargo, estas actividades son generadoras de riesgos que se traducen en contaminaciones accidentales que afectan directamente al ecosistema. El impacto ambiental ocasionado por la industria petrolera comprende los efectos de todas y cada una de las fases involucradas en las etapas tanto de exploración, explotación, transformación, distribución y comercialización. En muchas de las regiones donde se ha explotado el petróleo, se han gene-

rado contingencias de diferentes magnitudes como consecuencia de las operaciones desarrolladas en: tanques de almacenamiento, oleoductos, equipos de perforación, buques petroleros, etc. Además, de los impactos generados en el medio como consecuencia de la exploración y explotación del petróleo, existen también daños ocasionados a las personas que lo emplean o se encuentran en contacto con él dentro de su ámbito laboral. En la última década, las Empresas dedicadas a la explotación, producción y transporte del recurso han visto la necesidad de implementar medidas de prevención a fin de evitar y / o minimizar los daños ocasionados al medio ambiente, personas y bienes materiales.

Las empresas petroleras realizan visitas de mantenimiento preventivo a cada una de sus locaciones petroleras (pozos productores, inyectores, baterías y colectores). Un yacimiento está formado por bloques y a su vez éste por baterías. Cada batería está formada por pozos de producción que son en promedio entre 15 y 20. Cada pozo tiene diferente nivel de producción que es conocido a priori y varía en el tiempo. La producción del pozo define la categoría y la cantidad de veces que debe visitarse al mes. Los pozos no pueden ser visitados más de una vez al día y dependiendo del tipo de pozo existen ciertas tareas que se deben realizar. Cada tarea tiene asignado un determinado equipamiento necesario, una frecuencia de realización y un tiempo aproximado de su duración.

La planificación del desarrollo de un yacimiento de petróleo requiere de un proceso altamente complejo e implica un número considerable de actividades en un conjunto de pozos. Estas actividades tienen que estar coordinadas, cumplir diversos tipos de restricciones y al mismo tiempo, optimizar criterios. Tales criterios suelen ser objetivos como: el nivel de producción, tiempo de completación de las tareas, etc. El cronograma de las actividades

en los pozos WAS se ocupa de la coordinación de estas actividades para formar así cronogramas de actividad, así como el mantenimiento dinámico de estos planes en el tiempo.

El plan de perforación en un yacimiento es una tarea desafiante considerada altamente dinámica y así como con previsiones económicas y restricciones. Las actividades de explotación están relacionadas con el desarrollo de nuevos pozos de petróleo (well perforation), el mantenimiento correctivo de los pozos existentes (well workover) y el mantenimiento preventivo (well preventive maintenance). Estas actividades deben ser realizadas con recursos siempre limitados que operan sobre un área geográfica amplia y están sujetos a los tiempos de parada previstos relacionados con su mantenimiento de rutina.

Cada pozo inicia la producción de petróleo tan pronto como hayan concluido sus actividades de desarrollo. El plan de desarrollo de un yacimiento está destinado a maximizar la ganancia a corto y largo plazo que resulta de la producción de petróleo.

Este proyecto propone desarrollar un conjunto metaheurísticas basadas en inteligencia computacional que proporcionen información detallada para decisiones a corto y largo plazo, mono-criterio y multi-criterio. Esta herramienta permitirá determinar el orden óptimo de perforación, hora y lugar, tipos de pozos y por lo tanto el número de nuevos pozos inyectores y productores. Además, proporcionará planes de mantenimiento correctivo y preventivo. Todos los planes deben cumplir las restricciones de actividades y recursos dispuestos para los criterios de optimización establecidos.

Resultados obtenidos/esperados

En esta línea de trabajo se pretende:

- a) Estudiar estado del arte para algoritmos metaheurísticos, heurísticas y reglas de despacho para distintos problemas de scheduling (máquina única, flow shop, job shop), optimización multiojetivo y mecanismos para el tratamiento de restricciones.
- b) Implementar motores de planificación basados en metaheurísticas mono y multi-objetivos con restricciones.
- c) Estudiar los modelos de procesos para la planificación de actividades del desarrollo de reservorios de gas y petróleo (actividades, restricciones, y optimización multi-objetivo).

Formación de recursos humanos

Un integrante de este proyecto de investigación está desarrollando su Tesis de Maestría en temáticas afines y dos integrantes del proyecto comenzarán sus tesis doctorales en temáticas relacionadas.

Se cuenta con tres becarios alumnos de las carreras de Sistemas.

Referencias

- [1] Brucker. P. (2004) Scheduling Algorithms. Springer-Verlag 3rd ed.
- [2] Chiong, R., & Dhakal, S. (Eds.). (2009). Natural intelligence for scheduling, planning and packing problems (Vol. 250). Springer.
- [3] Coello, C. A. C. (2002). Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. *Computer methods in applied mechanics and engineering*, 191(11), 1245-1287.
- [4] Coello Coello C., (2015)., Multi-objective evolutionary algorithms in real-world applications: Some recent results and current challenges. In *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences* (pp. 3-18). Springer International Publishing.
- [5] Dhar, V., and Ranganathan, N. (1990). Integer programming vs. expert systems: an experimental comparison. *Communications of the ACM*, 33(3), 323-336.
- [6] Ehrgott M. (2005), *Multicriteria Optimization*. Springer, Berlin, second edition, 2005. ISBN 3-540-21398-8.
- [7] Framinan J. , Gupta J., and Leisten R.. (2004). A review and classification of heuristics for permutation flow-shop scheduling with makespan objective. *Operational Research Society*, 55(12):1243–55.
- [8] E. Garey and D. Johnson. (1979) *Computers and Intractability: A Guide to the Theory of NP Completeness*. Freeman and Co.
- [9] S. Hejazi and S. Saghafian. (2005). Flowshop scheduling problems with makespan criterion: a review. *Production Research*, 43(14):2895–2929.
- [10] Leung J. (2004) *Handbook of Scheduling: Algorithms, Models and Performance Analysis*. Computer and Information Sciences Series. Chapman and Hall.
- [11] Muth J.F. and Thompson.G.L., (1963). *Industrial Scheduling*. Prentice-Hall, Englewood Cliffs, N.J.,
- [12] Schwefel P. (1995), *Evolution and Optimum Seeking*. John Wiley & Sons, New York.
- [13] Whitley, D., Gordon, V. S., & Mathias, K. (1994, October). Lamarckian evolution, the Baldwin effect and function optimization. In *International Conference on Parallel Problem Solving from Nature* (pp. 5-15). Springer Berlin Heidelberg.

Librería de Metaheurísticas para Problemas de Optimización Continuos y Discretos

M.Bilbao(1,2), F.Sanchez,D.Ormachea,L.Sloboda (1)
D.Pandolfi, A.Villagra, Marta Lasso, Daniel Molina (2)
Guillermo Leguizamón (3)

(1) Grupo de Investigación en Inteligencia Artificial – Departamento de Informática Facultad de Ingeniería Universidad Nacional de la Patagonia San Juan Bosco

Ruta 1 Km 4 Comodoro Rivadavia
{martinbilbao,fsanchez,lsloboda,dormachea}@ing.unp.edu.ar

(2) Instituto de Tecnología Aplicada – Departamento de Ciencias Exactas Universidad Nacional de la Patagonia Austral

Ruta 3 Acceso Norte - Caleta Olivia - Santa Cruz
{dpandolfi,avillagra,mlasso,dmolina}@uaco.unpa.edu.ar

(3)LIDIC – Facultad de Ciencias Físico, Matemática y Naturales Universidad Nacional de San Luis

Ejercito de los Andes 950 - San Luis
{legui}@unsl.edu.ar

Resumen

Las Metaheurísticas [1] son técnicas estocásticas que permiten resolver problemas de optimización continuos y discretos en general. Cuando se aborda un problema muy complejo con muchas variantes, las técnicas matemáticas no alcanzan para obtener soluciones fiables ya que el tiempo computacional aumenta de manera exponencial con el tamaño del problema (Problemas NP)[2]. Las Metaheurísticas actúan a favor de este problema ya que explora mejor el espacio de búsqueda eligiendo lugares prometedores y descartando el espacio donde las soluciones no aportan al resultado final. Debido a que existen muchas técnicas Metaheurísticas y cada una puede ser mejor adaptada en problemas específicos debido a la representación de soluciones, se plantea el diseño y construcción de un framework

de Metaheurísticas genérico para trabajar con problemas de optimización continua y discreta. Dicho framework contará con las técnicas Metaheurísticas más utilizadas en la literatura e incorporará benchmarks de pruebas con problemas estándares de resolución NP. Dichos problemas serán obtenidos de las variantes CEC 2005 y CEC 2009 donde plantean diferentes formulas matemáticas de optimización de varias variables para poder ejecutar y probar los algoritmos diseñados de manera eficaz.

Palabras claves: Optimización, Metaheurísticas, sistemas bioinspirados

Contexto

El proyecto de investigación se desarrolla bajo convenio de actividad

interinstitucional del programa de mejora de carreras informáticas PROMINF entre la UNPA-UACO y la UNPSJB y las actividades se desarrollan en el Instituto de Tecnología Aplicada de la UNPA- UACO y el Grupo de Investigación en Inteligencia Artificial de la UNPSJB.

Introducción

Este proyecto se basa en la implementación de una biblioteca general de algoritmos evolutivos paralelos para la resolución de problemas continuos y discretos. Surge la necesidad de dicha implementación las escasas herramientas de optimización para éstas técnicas y la dificultad encontradas en cada una de ellas. Existen muchas bibliotecas de algoritmos evolutivos para resolver problemas de optimización continua y discreta. Éstas bibliotecas fueron diseñadas por universidades en marco de proyectos de investigación conjunta y son utilizadas y referenciadas por diferentes grupos de investigación en el mundo. Las bibliotecas más utilizadas y reconocidas son MALLBA [3], ParadisEO [4], JMetal[5] y JCell[6] que fueron diseñadas por universidades españolas y francesas respectivamente.

ParadisEO es un framework de algoritmos metaheurísticos paralelos diseñado por el INRIA [7] y el laboratorio informático de la Universidad de Lille (LIFL). Soporta algoritmos evolutivos poblacionales y de trayectoria, algoritmos metaheurísticos distribuidos, paralelos e híbridos y optimización multiobjetivo.

Está desarrollado en C++, es portable para las arquitecturas Windows, Linux y Mac y soporta computación distribuida y paralela con MPI y computación GRID mediante Globus / Condor. Actualmente está en su versión 2.0 y se distribuye bajo la licencia CeCill del INRIA. El sitio oficial es <http://paradiseo.gforge.inria.fr/>

Mallba es un framework de algoritmos metaheurísticos paralelos diseñado por tres universidades españolas (Universidad de Málaga, Universidad de la Laguna, y Universidad de Barcelona) en marco de un proyecto español coordinado. Es una biblioteca utilizada para resolver problemas de optimización combinatorial y contiene métodos exactos, heurísticos e híbridos, poblaciones y de trayectoria. Está implementado en C++, es portable y soporta computación distribuida y computación GRID. Actualmente está en su versión 2.0, es software gratuito y su sitio oficial se encuentra disponible en [4]

Nuestra Propuesta es diseñar una librería que incorpore algoritmos metaheurísticos paralelos poblacionales y de trayectoria para optimización continua y discreta, como así también soporte hibridización de las técnicas y optimización multiobjetivo. Dicha biblioteca debe estar adaptada para la portabilidad y permitir computación distribuida y GRID.

Líneas de Investigación, Desarrollo e Innovación

Los Objetivos Generales del proyecto son implementar una biblioteca paralela de algoritmos evolutivos para la resolución de problemas continuos y discretos con la participación de dos universidades que trabajan en la línea de Metaheurísticas y paralelismo y fortalecer las relaciones entre ambas universidades, capacitando recursos humanos (docentes y alumnos) favoreciendo la mejora en la enseñanza académica.

Objetivos Específicos:

- Diseñar un framework de técnicas Metaheurísticas poblacionales y de trayectoria para ser utilizada en los proyectos de investigación relacionados con las temáticas correspondientes.
- Interactuar con otros grupos de investigación en el desarrollo conjunto de la herramienta.
- Proveer una herramienta única para optimización continua, discreta y multiobjetivo.
- Fortalecer el intercambio de experiencias y avances realizados entre las instituciones participantes.
- Incrementar las actividades de investigación futura entre las universidades involucradas.

Resultados y Objetivos

Los objetivos cumplidos son los siguientes:

Estudio del estado de arte de las librerías de Metaheurísticas existentes

Se investigaron diferentes librerías existentes en la literatura (MALLBA, Paradiseo,

JCELL, JMETAL) y se adaptaron los modelos a las necesidades del proyecto.

Definir el modelo de diseño de la estructura de la biblioteca

Se definieron bajo estándares de ingeniería de software alternativas de modelos eficientes, extensibles y mantenibles para asegurar la calidad del framework a desarrollar

Diseñar algoritmos metaheurísticos de trayectoria y poblacionales

Se implementaron nuevos algoritmos basándonos en el modelo de diseño del punto anterior, tanto los algoritmos de trayectoria como los poblacionales deben ser genéricos y adaptables a problemas de optimización continuos y discretos. En nuestro caso se diseñaron doce técnicas Metaheurísticas para la biblioteca, las cuales ya están en proceso de evaluación con los benchmarks de prueba CEC 2005 y CEC 2009 [8]

Difundir los resultados en congresos y revistas nacionales e internacionales para compartir las experiencias y aceptar sugerencias o críticas a nuestros trabajos.

Se ha publicado un artículo referente a la temática en el WICC 15 [9]

Diseñar problemas genéricos para probar las técnicas implementadas y permitir al usuario analizar su algoritmo con problemas de resolución óptima o conocida

Se han utilizado benchmarks de prueba CEC 2005 y CEC 2009 [8] para ampliar los recursos de la librería y poder probar los algoritmos de manera rápida con conclusiones certeras.

Los objetivos en curso son los siguientes:

Análisis cualitativo y cuantitativo de los resultados obtenidos en nuestro estudio

Diseñar técnicas de análisis estadístico de resultados de manera automática en función a los resultados y obtener las gráficas de boxplot, análisis de evolución de fitness y tablas dinámicas para su posterior uso en el documento científico.

[6] <https://jcell.gforge.uni.lu>

[7] <http://www.inria.r>

[8] Benchmark Generation for CEC 2009 Competition on Dynamic Optimization, CEC 2009

[9] Frameworks de Metaheurísticas para Problemas de Optimización Complejos WICC 2015, Universidad Nacional de Salta, Argentina.

Formación de Recursos Humanos

Dos integrantes de esta línea de investigación están finalizando sus estudios doctorales en temáticas afines.

Un integrante ha finalizado sus estudios de Maestría y ha comenzado su formación doctoral.

Un integrante de esta línea ha finalizado sus estudios doctorales.

Cinco integrantes están formalizando sus anteproyectos de tesis de maestría

Se dirigen varias tesis de grado de la carrera Licenciatura en Informática (UNPSJB) y de la Ingeniería en Sistemas (UNPA) en temáticas afines.

Referencias

[1] Metaheuristics: Theory and application, Osman Ibrahim and Kelly James, 1996

[2] Stephen Cook, The complexity of Theorem-Provin Procedures, Third Annual ACM Theory of Computing, 1971

[3] <http://www.lsi.upc.edu/~mallba/>

[4] <http://paradiseo.gforge.inria.fr>

[5] <http://jmetal.sourceforge.net>

Modelo para Analizar Mensajes y Detectar Actitudes Peligrosas a través de Análisis de Sentimientos con Algoritmos de Aprendizajes

Juan Calloni, Sergio Paez, Javier Saldarini, Juan Cuevas,
Micaela Mulassano, Andrés Bianciotti, Eduardo Scarello,
Leandro Banchio, Federico Degiovanni, Lucia Scharff

Grupo GARLAN / Secretaría de Ciencia y Tecnología
Facultad Regional San Francisco / Universidad Tecnológica Nacional
Av. de la Universidad 501, 03564-421147
facultad@sanfrancisco.utn.edu.ar

Resumen

Es de amplio conocimiento la utilización masiva de las distintas redes sociales. Éstas han cambiado los hábitos y características de la comunicación, tal como la facilidad de intercambio de información, la existencia de receptores globales y la accesibilidad de éstas tecnologías a todos los sectores sociales.

Las actitudes violentas y peligrosas en redes sociales constituyen un campo de estudio objeto de varias disciplinas. Detectar este tipo de actitudes de manera temprana, colaboran a la prevención de los efectos que éstas podrían causar.

Uno de los problemas que trata de resolver este proyecto, es la seguridad, en específico, las actitudes sospechosas o violentas, como violencia de género o bullying.

En definitiva, se trata de brindar un modelo que permita realizar análisis, extrayendo información de redes sociales de acceso público, para demostrar, con métodos de aprendizaje, tanto supervisados como no supervisados y con métodos de análisis de sentimientos y minería de opiniones, qué mensaje o grupo de mensajes se aproximan a ser peligrosos.

Palabras clave:

Modelo, Machine learning, Análisis de sentimientos, NoSQL, Minería de datos, Cloud computing.

Contexto

La línea de investigación presentada se encuentra en el marco de las áreas prioritarias para el desarrollo de las actividades de I+D que se formalizaron a través de la Resolución de Consejo Directivo N 353/2016 de la Universidad Tecnológica Nacional Facultad Regional San Francisco cuyas áreas prioritarias son:

- Gestión de procesos de negocios
- Ingeniería de software
- Gestión y tecnologías de las organizaciones
- Calidad de Software
- Seguridad de la información
- Bases de datos

El área principal es Ingeniería de Software relacionándose con el desarrollo en la Nube, métodos de aprendizajes, inteligencia artificial y bases de datos NoSQL. Éstas líneas se encuentran insertas en el Proyecto de Investigación y Desarrollo (PID) tutorado Nro EITUNCO0004317 “Modelo de un

Analizador de Mensajes en Redes sociales para la detección de usuarios con actitudes peligrosas o violentas a través de análisis de sentimientos con algoritmos de aprendizajes”, el mismo se encuentra en desarrollo desde Abril de 2016. Dicho proyecto es ejecutado por el grupo de I+D GARLAN de la UTN Facultad Regional San Francisco.

El proyecto se encuentra homologado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Tecnológica Nacional, a través de la Disposición N° 52/16 y se lo reconoce bajo el código TUN4317, el mismo está incluido en el Programa I&D + i de Electrónica, Informática y Comunicaciones de la Universidad Tecnológica Nacional.

Introducción

Este proyecto tiene como objetivos:

1. Determinar un modelo analizador de mensajes para la detección de actitudes peligrosas de usuarios a través de métodos de aprendizajes.
2. Desarrollar un aplicativo orientado a servicios (SaaS) accesible vía API RESTful que sirva para validar el modelo.
3. Implementar el prototipo con una muestra representativa de expresiones reales elegidas al azar para verificar el comportamiento de los algoritmos y parámetros del modelo.

En particular a través de estos objetivos las herramientas de inteligencia artificial como, Support Vector Machines (SVM)[1], que pertenece a un conjunto de algoritmos de aprendizaje supervisado que están propiamente relacionados con problemas de clasificación y regresión a partir de un conjunto de ejemplos de entrenamiento (de muestras), es posible etiquetar las clases y entrenar el sistema para construir un modelo que prediga la clase de una nueva muestra.

Por otro lado, K-means[2], es un método de agrupamiento y pertenece a un conjunto de algoritmos de aprendizaje no supervisado que tiene como objetivo la partición de un conjunto de 'n' observaciones en 'k' grupos en el que cada observación pertenece al grupo más cercano a la media. En el aprendizaje de máquina y la ciencia cognitiva, las redes neuronales artificiales (RNA) son una familia de modelos de aprendizaje estadísticos inspirados en las redes neuronales biológicas (los sistemas nervioso central de los animales, en particular, el cerebro) y pueden ser utilizadas como algoritmos de aprendizaje supervisados y no supervisados[3]. Con la ayuda de un experto en el dominio, se obtiene un conjunto de datos, con el cual entrenar los algoritmos de aprendizaje. Entonces, con una herramienta de análisis de minería de datos, como Weka, analizar qué métodos brindan mayor certeza. Una vez realizado, es posible sacar conclusiones, sobre qué algoritmo de aprendizaje utilizar; supervisado, no supervisado o varios de ellos. Él o los métodos seleccionados son los candidatos a implementar para realizar el modelo.

Los siguientes pasos describen el proceso a realizar para la extracción de la información que será usada para el análisis posterior.

Son utilizadas las API's de las redes sociales para obtener información, y poder realizar estudios con datos reales. En específico, en este trabajo se obtendrán mensajes de la red social Twitter, y dejando a futuros proyectos la implementación con otras API's de redes sociales y sistemas de participación. La información obtenida es ingresada a una base de datos NoSQL (orientada a documentos, para almacenar metadatos) y así analizarlo de forma local. Es usado este tipo de base de datos ya que el prototipo deberá funcionar en la nube y una de sus características es la alta escalabilidad[4]. Al final, para validar el modelo, el proyecto planea construir un aplicativo orientado a servicio (SaaS) como prototipo

utilizando un método heurístico probarlo y validar el modelo propuesto. Se trabaja con algoritmos de aprendizaje no supervisado de clustering para agrupar mensajes con características necesarias del dominio de aplicación, y a partir de éste crear un modelo basado en técnicas de aprendizaje supervisado, lo cual requiere una tarea de etiquetado y validación por un experto en el dominio. En el caso de los modelos no supervisados, estamos trabajando con el algoritmo K-means. Los resultados positivos del primer agrupamiento son clasificados manualmente.

Luego, con métodos de regresión logística y RNA es posible determinar a qué conjunto pertenece un nuevo elemento de prueba. Durante las primeras etapas del proyecto fue contemplada y planificada la posibilidad de analizar documentos públicos provenientes de las filtraciones de cables que publicó el portal Wikileaks, sin embargo; los análisis posteriores determinaron que la profundidad semántica de la temática expuesta escapaba a los alcances este proyecto. El Idioma elegido es el castellano utilizando algoritmos de aprendizaje enfocados para determinar el modelo que mejor resuelve la clasificación. Resumiendo, los pasos de la investigación son:

Obtención del conjunto de datos y agrupamiento de los mensajes

Generación del dataset supervisado

Implementación del algoritmo que clasificará nuevos elementos de prueba. La cantidad de iteraciones depende del número de algoritmos a investigar hasta encontrar el que mejor se adapte

Análisis de los resultados

Despliegue y evaluación del prototipo SaaS en un entorno de cloud computing.

El proyecto busca corroborar nuestra hipótesis: Es posible clasificar mensajes de forma automatizada de acuerdo a diversos intereses sociales, centrándonos en actitudes violentas y ciberbullying.

Líneas de Investigación, Desarrollo e Innovación

El desarrollo e investigación se basa en encontrar técnicas de aprendizaje de máquina que puedan aportar resultados predictivos en la materia de seguridad social a través de aplicaciones que un usuario pueda utilizar. Con esto, nos referimos a implementar algoritmos de aprendizaje que detecten mensajes en español y tengan consecuencias negativas en la vida social, creando herramientas útiles que sean implementadas utilizando tecnologías de última generación.

Resultados Obtenidos/Esperados

Dentro de los principales avances que hemos obtenido hasta el momento se encuentra el desarrollo de modelos para realizar un SaaS que permita detectar mensajes de texto. En un principio vinculados a actitudes sospechosas de usuarios y actualmente enfocándonos en mensajes que puedan relacionarse con el ciberbullying[5, 6, 7, 8,9].

Para complementar, fue creada una interfaz de comunicación con las redes sociales para obtener el conjunto de datos a utilizar en el desarrollo del sistema. En la experimentación realizamos las pruebas con la API Search de Twitter.

Para poder almacenar y consultar el conjunto de datos obtenidos ha sido creada una interfaz para guardar los mensajes a estudiar en una base de datos y ejecutar el proceso offline. Esta interfaz nos permite experimentar con distintas estrategias de almacenamiento reutilizando la mayor cantidad de código.

Para no tener que programar todos los algoritmos de aprendizaje de máquina y optimizar los tiempos de análisis y clasificación de los datos se creó una interfaz para conectarse con una API que provea los algoritmos de aprendizaje necesarios en la investigación. Actualmente se centró específicamente en la arquitectura del sistema y se está

trabajando en el análisis y clasificación del conjunto de datos.

La arquitectura del sistema se creó utilizando la tecnologías de contenedores Docker[10] para montar la aplicación. La cual está formada básicamente por los siguientes contenedores:

- Contenedor web que aloja la lógica del dominio desarrollada con el lenguaje de programación Java.
- Contenedor de base de datos para alojar el conjunto de datos extraído de las redes sociales, utilizando MongoDB como sistema de almacenamiento.
- Un sistema de contenedores que implementa sharding para almacenar el conjunto de datos 1-Gram de Google[11], utilizando MongoDB como sistema de almacenamiento.

Nuestros objetivos a corto plazo involucran:

- Implementar un modelo similar al planteado en el paper TweetNorm[12] para realizar un preprocesamiento y normalización de los mensajes en el conjunto de datos.
- Terminar de definir el conjunto de datos que utilizaremos como entrenamiento para los algoritmos de aprendizaje de máquina.
- Analizar distintos tipos de algoritmos de aprendizaje, K-means, SVM, RNA, para evaluar el conjunto de datos y obtener el algoritmo que mejor se adapte a nuestro dominio

El objetivo final es evaluar cómo funcionan los distintos módulos del prototipo en un ambiente de cloud computing.

Formación de Recursos Humanos

La formación de recursos humanos en el área de AI (Artificial Intelligence) y

Machine Learning ha ganado campo en los últimos años en la ciencias de computación, además de que año a año los algoritmos y modelos se vuelven más complejos, por lo que tener recursos humanos enfocados en estas disciplinas es de vital importancia.

La investigación en la disciplina de NLP (Natural Language Processing), detección de imágenes, analizadores de video, etc., también ha tenido un crecimiento importante con las nuevas tecnologías de procesamiento de datos, por lo que conceptos de escalabilidad y desempeño también deben ser tenidos en cuenta a la hora de investigar, diseñar e implementar.

El proyecto plantea dar conocimientos sobre AI, Machine Learning, NLP, bases de datos NoSQL, así como también aplicar conceptos de arquitecturas en la nube para una escalabilidad horizontal y un buen desempeño.

Dentro del proyecto se contempla la formación de alumnos de las cátedras Inteligencia Artificial, Gestión de Datos, Diseño de Sistemas y Redes de Información en los aspectos referentes a métodos de aprendizaje supervisados o no supervisados, en bases de datos NoSQL, aspectos de cómo aplicar conceptos de arquitecturas en la nube y el desarrollo de aplicaciones SaaS.

Se pretende capacitar en estas tecnologías a docentes y alumnos avanzados de la carrera Ingeniería en Sistemas de Información mediante talleres que se dictarán a medida que avance el proyecto. Esto permitirá contar con personal de apoyo, captar nuevas ideas y opiniones para mejorar el proyecto.

Se llevará a cabo una serie de conferencias informativas en la Universidad Tecnológica Nacional Facultad Regional San Francisco para exponer el proyecto conceptualizado con las principales ventajas ofrecidas desde un marco teórico (previo al inicio, presentación oficial para la puesta en marcha y transición del proyecto). ■

Referencias

[1] Cortes, Corinna; Vapnik, Vladimir. Support-vector networks. *Machine learning*, 1995, vol. 20, no 3, p. 273-297.

[2] Kanungo, Tapas, et al. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence*, 2002, vol. 24, no 7, p.881-892.

[3] Russell, Stuart J and Norvig, Peter, 2004, *Artificial intelligence*. Englewood Cliffs, N.J. : Prentice Hall.

[4] Sadalage, Pramod J.; Fowler, Martin. *NoSQL distilled: a brief guide to the emerging world of polyglot persistence*. Pearson Education, 2012.

[5] Morales, J. and Arias Orduña, A. (2007). *Psicología social*. Madrid: McGraw-Hill Interamericana de España. p.457, p.459, p.489.

[6] Anderson, Janna; RAINIE, Lee. *Millennials will benefit and suffer due to their hyperconnected lives*. Washington DC, Pew Research Center, 2012.

[7] Hernández Prados, María Ángeles and Solano Fernández, Isabel María, 2007, *Ciberbullying, un problema de acoso escolar*. *Revista Iberoamericana de Educación a Distancia*. 2007.

[8] Cervantes Benavides, L. (2013). *Una propuesta para identificar, clasificar y tipificar el Bullying (Acoso Escolar)*. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*.

[9] Smith, P. (1999). *The nature of school bullying*. London: Routledge.

[10] "Docker". Docker. N.p., 2016. Web. 13 Julio 2016.

[11] Google Ngram Viewer, 2016. <http://storage.googleapis.com/books/ngrams/books/datasetsv2.html> [online]

[12] Tweet Norm. http://www.lrec-conf.org/proceedings/lrec2014/pdf/442_Paper.pdf Febrero de 2016.

Planificación Estocástica Basada en Markov para la Composición y Recomendación de Servicios Web

Cintia Ferrarini¹, Raymundo Forradellas²

¹Instituto de Informática / Facultad de Ciencias Exactas, Físicas y Naturales / Universidad Nacional de San Juan.

²Departamento de Informática/Facultad de Ciencias Exactas, Físicas y Naturales / Universidad Nacional de San Juan.

¹ ferrarini@iinfo.unsj.edu.ar, ² kike@usaps.net

RESUMEN

La planificación en la Composición de Servicios Web, posibilita la creación de nuevos servicios que brindan una solución aceptable cuando un servicio web simple no puede satisfacer la funcionalidad demandada por un usuario o sistema. Por lo tanto, implica una semántica particular en la definición, selección y combinación de estos. Los recientes progresos en este ámbito han posibilitado la localización e invocación de aplicaciones a través de la Web.

La capacidad para la selección e integración eficiente de servicios entre organizaciones heterogéneas en tiempo de ejecución se convierte en un requisito importante para la prestación de los mismos.

Este trabajo presenta el tema de investigación que está desarrollando la Mgter. Ferrarini para obtener el grado de Doctor en Ingeniería de la Universidad Nacional de Cuyo, Mendoza. En el que propone un método de planificación basado en el modelo de Procesos de Decisión de Markov Parcialmente Oculto (POMDP) para la Composición de Servicios Web, con el objetivo de entregarle alternativas de paquetes de

servicios web a un usuario teniendo en cuenta sus preferencias personales y atributos contextuales. El método propuesto se encuentra en proceso de validación y forma parte de la segunda capa de una arquitectura de software de tres capas.

Palabras clave: procesos estocásticos, composición de servicios web, planificación bajo incertidumbre.

CONTEXTO

El trabajo de tesis se desarrolla en el ámbito del Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan, para obtener el grado de Doctor en Ingeniería de la Universidad Nacional de Cuyo, bajo la dirección del Dr. Raymundo Forradellas. La doctoranda financia sus estudios con una beca cofinanciada entre CONICET y la Universidad Nacional de San Juan. Resol. N° O191

INTRODUCCIÓN

Las Tecnologías de la Información y la Comunicación (TIC) han transformado todos los ámbitos de la sociedad mediante la automatización de procesos y tareas; han favorecido el acceso a la información, a los servicios, a nuevas formas de comunicarse, de aprender y de trabajar. No obstante, este acceso ilimitado a diferentes recursos disponibles en la web, han vuelto complejo el proceso de análisis, selección y toma de decisiones en estos entornos dinámicos.

Es en este ámbito, donde la Inteligencia Artificial (IA), juega un rol fundamental. La IA intenta emular el razonamiento humano a través de automatizar tareas intelectuales con el propósito de brindar soluciones a problemas del mundo real [1]. Abarca una diversidad de áreas, entre las que se encuentran la Planificación Automática (PA). Esta última es de interés para la comunidad científica, ya que permite establecer computacionalmente rumbos de acción en los casos en que un ser humano no los pueda abordar adecuadamente por su complejidad. La Planificación Automática [2],[3] estudia computacionalmente el proceso de deliberación en la actuación e intenta construir herramientas de procesamiento de información que ofrezcan acceso a recursos de planificación eficientes para resolver problemas de abordaje complejos para un ser humano, permitiéndole establecer rumbos de acción posibles. Particularmente, en aquellos problemas de naturaleza dinámica, estocástica y en donde la gestión del tiempo juega un rol importante [4]. La automatización de planes contribuye a los procesos de toma de decisiones y a la realización de tareas complejas en una

gran variedad de dominios como: la industria, las telecomunicaciones, la logística de transporte, las fuerzas armadas, los sistemas de información, los servicios web [3][5], la robótica[6], los juegos, la gestión de la educación y la salud, entre otros.

Un problema de planificación en IA [3] consiste en encontrar una secuencia de acciones para transformar el estado actual en uno de los estados meta, contando con la descripción del estado actual del sistema, el conjunto de acciones que pueden realizarse sobre él y una descripción de un conjunto de estados meta a alcanzar.

La planificación en la Composición de Servicios Web, es un mecanismo de creación de nuevos servicios Web y por tanto, requiere de una semántica particular en su definición, selección y combinación [3][5][7]. Los recientes progresos en el ámbito de la Composición de Servicios Web han posibilitado la localización e invocación de aplicaciones a través de la Web. Contribuyendo a aquellos casos en que un servicio web simple no puede satisfacer la funcionalidad requerida por un usuario o sistema y por lo tanto debe ser capaz de coordinar y funcionar con otros servicios disponibles a fin de dar respuesta a la solicitud.

Este trabajo de tesis intenta hacer su aporte a la planificación automática, específicamente en la planificación de paquetes/ composición de servicios web, en tiempo real, en un entorno de recomendación. Para tal fin, se propone un método de planificación para la toma de decisiones basado en un algoritmo de iteración-valor de los Procesos de Decisión de Markov Parcialmente Observables (POMDP) para la

composición de servicios web en un Entorno de Recomendación, que contemple las preferencias y atributos contextuales de sus usuarios. El ámbito de aplicación se caracteriza como un dominio de espacio discreto, con el objetivo de entregarles alternativas de paquetes de servicios que respondan a las necesidades a sus usuarios y en el menor tiempo posible. Los POMDPs han sido ampliamente utilizados en otros ambientes de decisión, por proveer un buen modelo matemático para entornos con incertidumbre [8],[9],[10] pero aun no han sido incorporados en este tipo de contextos. Una de las razones es su alta complejidad computacional, lo que implica un esfuerzo en mejorar su capacidad de procesamiento reduciendo el tiempo de resolución, entre otras. En el marco de esta investigación se optó por un diseño metodológico que combina instancias exploratorio-descriptivas con instancias experimentales. Teniendo en cuenta las siguientes fases: Fase N°1 de Investigación y Estudio del estado del arte; Fase N°2 de Problematización y Formulación del problema. Identificación y Definición de variables. Formulación de Hipótesis; Fase N°3 de Construcción y Formulación de la Propuesta; Fase N° 4 de Validación – Experimentación; y Fase N°5 de Comunicación y Divulgación que implica la escritura de la tesis doctoral y la publicación en eventos y revistas científicas nacionales e internacionales.

AVANCES

Se estudio, por una parte, la problemática de la composición de servicios, donde se identificaron las características que conforman un paquete de servicios y los tipos de servicios

que lo componen. Y por otra, los diferentes mecanismos de asignaciones de recursos y planificación en Inteligencia Artificial, Investigación Operativa (IO) y Sistemas Multiagentes (SMA) existentes en la literatura. Esto permitió considerar la incorporación del modelo de POMDP para la toma de decisiones en la construcción de paquetes de servicios en tiempo real en un entorno de recomendación como un mecanismo para mejorar su tiempo de respuesta. Este entorno contará con una arquitectura de tres capas. En la primera, se contará con la especificación el modelo de usuario que incorpora sus preferencias mediante la valoración de atributos. En la segunda, se define el modelo de decisiones del que forma parte el método de planificación basado en POMDP propuesto y; la tercera capa, estará conformada por un modelo de extracción de datos basado en minería de datos preexistente.

En este momento se están realizando pruebas experimentales (con datos aleatorios) del método de planificación propuesto teniendo en cuenta su tiempo de respuesta, a fin de refinarlo. Una vez finalizado ese proceso se incorporará a un Sistema de Recomendación destinado al Turismo.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Este proyecto se encuadra en el Área Inteligencia Artificial de las Ciencias de la Computación. Las líneas de investigación que se abordan son la Planificación Automática en Sistemas de Tiempo Real y la Planificación Estocástica.

RESULTADOS Y OBJETIVOS

El trabajo de tesis tiene como objetivo general obtener un método para la generación paquetes de servicios adaptados a los requerimientos de los usuarios. Para ello se analizaron los métodos aplicados a sistemas con incertidumbre y de tiempo real. Se han comparado los métodos existentes para la creación de paquetes/composición de servicios web. Se está trabajando en un planificador estocástico que mejore el tiempo de respuesta en la construcción de paquetes/ composición de servicios web, teniendo en cuenta las preferencias de sus usuarios.

Los resultados alcanzados en el marco de esta investigación son la generación de conocimiento sobre el diseño y desarrollo de un planificador para la composición de servicios web. Mediante el desarrollo de un prototipo para la realización de pruebas experimentales con algunos usuarios para poder mejorar los prototipos.

Se han presentado hasta el momento los siguientes trabajos:

“Generación de Planes para la creación de Paquetes de Servicios”. Encuentro de Investigadores Jóvenes de San Juan y 2º Jornada de Becarios de Investigación de la UNSJ. Argentina, San Juan; Año: 2012. Autora: Cintia Ferrarini Oliver.

“Planificación con POMDP para la creación de Paquetes de Servicios”. Libro. Artículo Breve. Conferencia. 2ND IJCAI SCHOOL on Artificial Intelligence – Doctoral Consortium 2014. AAAI (SADIO), 43 JAIIIO. Autores: Cintia Ferrarini Oliver, Silvana Aciar,

Raimundo Forradellas. Argentina. Buenos Aires. 2014.

“Stochastic planning method for building service packages”. Colombia, Popayan. 2016.

Autores: Cintia Ferrarini Oliver, Raimundo Forradellas y Fernando di Sciascio.

FORMACIÓN DE RECURSOS HUMANOS

La Mgter. Cintia Ferrarini Oliver, ha cumplimentado los cursos de formación requeridos por el Doctorado en Ingeniería de la Universidad Nacional de Cuyo. Actualmente, se encuentra en la etapa de desarrollo de la tesis requisito para acceder al título de Doctor en Ingeniería con orientación en sistemas. En 2010 obtuvo el título de Magíster en Informática otorgado por la Universidad Nacional de la Matanza. En 2015 obtuvo el título de Especialista en Educación Superior y TIC, otorgado por Ministerio de Educación de la Nación. En 2015 obtuvo el título de Licenciada en Ciencias de la Educación otorgado por la Universidad de Congreso. En 2016 obtuvo el título de Profesora de Ciencias de la Educación otorgado por la Universidad de Congreso.

Ha dirigido o dirige las siguientes Tesis de Grado:

- *“Plan Estratégico de Gobierno Electrónico para la Municipalidad de 25 de Mayo”*, perteneciente a Héctor Rogelio Aravena Rodríguez, para obtener el título de Licenciado en Sistemas de Información. Directoras: María Inés Lund – Cintia Ferrarini Oliver. Año 2010. Calificación obtenida 10 (diez).

- “Sistema Web Gamificado (SiWGa) para Promover Competencias de Resolución de Problemas en Matemática en Nivel Primario”, perteneciente a Florencia Belchior, para obtener el título de Licenciado en Sistemas de Información. Directoras: Mónica Gilda González de Doña – Cintia Ferrarini Oliver. Año 2017. Presentación de tesis ante tribunal, a la espera de designación de fecha de Exposición.
- “Gestión de Recursos para videojuegos en Entornos Distribuidos”. Tesis de Grado en desarrollo, perteneciente a Mauro Gabriel Avendaño para obtener el título de Licenciado en Sistemas de Información. Directores: Emilio Ormeño – Cintia Ferrarini Oliver. Año 2017.

REFERENCIAS

- [1] Weiss, G (1999). “*Multiagent Systems. A Modern Approach to Distributed Modern Approach to Artificial Intelligence*”. MIT Press, Cambridge, Massachusetts Institute of Technology, London, England.
- [2] Durfee, E.H.(2008). “*Planning for Coordination and Coordination for Planning*,” Web Intelligence and Intelligent Agent Technology, WI-IAT '08. IEEE/WIC/ACM International Conference on, vol.1, no., pp.1-3, 9-12 Dec.
- [3] Guzmán Luna, J. (2009). “*Modelo de Planificación y Ejecución Concurrente para la Composición de Servicios Web Semánticos en Entornos Parcialmente Observables*”. Tesis Doctoral. Universidad Nacional de Colombia.
- [4] Bargiacchi E., Verschoor C. R., Li G. & Roijers, D. M. (2013). “*Decentralized Solutions and Tactics for RTS*”.
- [5] Ferrarini Oliver, C., Aciar, S., Forradella, R. (2014) “*Planificación con POMDP para la creación de Paquetes de Servicios*”. Jornadas Argentinas de Informática (JAIIO)
- [6] Pajarinen J. & Kyrki V. (2015) “*Robotic manipulation of multiple objects as a POMDP*”. Artificial Intelligence.
- [7] Braziunas D. (2003). *POMDP solution methods*. Department of Computer Science University of Toronto 2003. Disponible en: https://www.techfak.uni-bielefeld.de/~skopp/Lehre/STdKI_SS10/POMDP_solution.pdf
- [7] Uc-Cetina, V.; Moo-Mena, F. & Hernandez-Ucan, R. (2015). Composition of Web Services Using Markov Decision Processes and Dynamic Programming. Hindawi Publishing Corporation e Scientific World Journal Volume 2015, Article ID 545308, 9 pages. <http://dx.doi.org/10.1155/2015/545308>.
- [8] Cassandra, A. R. (1998). *Exact and approximate algorithms for partially observable Markov decision processes*. Ph. D. thesis. Brown University. Providence, Rhode Island (1998). The POMDP Page, <http://www.pomdp.org/>
- [9] Puterman M. L. (2005). “*Markov Decision Processes: Discrete Stochastic Dynamic Programming*”. Wiley Series in Probability and Statistics. Wiley John + Sons, Softcover. ISBN: 9780471727828 / 0471727822.
- [10] Shani, G., Pineau, J. & Kaplow, R.. “*A survey of point-based POMDP solvers*”. Journal Autonomous Agents and Multi-Agent Systems. Volume 27, Issue 1, pp 1-51. Cover Date 2013-07. DOI 10.1007/s10458-012-9200-2. Print ISSN 1387-2532. Online ISSN 1573-7454. Publisher Springer US.

Problemas Complejos Resueltos con Metaheurísticas

German Dupuy, Eber Bezzone, Juan Barbero, Natalia Stark, Fernando Sanz Troiani,
Hugo Alfonso, Carlos Bermudez, Carolina Salto¹, Gabriela Minetti¹
Laboratorio de Investigación en Sistemas Inteligentes (LISI)
Facultad de Ingeniería - Universidad Nacional de La Pampa

Calle 110 Esq. 9 (6360) General Pico - La Pampa - Rep. Argentina

Te. / Fax: (02302) 422780/422372, Int. 6302

e-mail: ¹{saltoc,minettig@ing.unlpam.edu.ar}

Resumen Este proyecto de investigación se enfoca en el estudio de técnicas metaheurísticas para resolver distintos problemas de optimización. Una de las líneas de investigación está abocada al estudio de nuevas técnicas metaheurísticas y su adaptación para resolver problemas complejos de planificación. En particular se pone especial énfasis en metaheurísticas que simulan comportamientos sociales de distintas especies, tales como *Cuckoo Search*, *Bee Colony Algorithm*, *Migration Bird Optimization*, entre otros. Otra de las líneas de investigación se enfoca en el desarrollo de estrategias adaptativas para modificar la probabilidad de mutación sin control externo en algoritmos genéticos. De esta manera, se reduce considerablemente el tiempo dedicado a la configuración paramétrica. Una tercera línea de investigación apunta a examinar si la metaheurística *Problem Aware Local Search* (PALS), un método eficiente inicialmente propuesto para el problema de ensamblado de fragmentos de ADN, puede ser usado en otros dominios de aplicación y con otros problemas de optimización. También se analizan alternativas de diseño de los principales componentes para construir una versión de PALS eficiente y exacta y así resolver los problemas del nuevo dominio de aplicación de una manera competitiva. Por último, una línea de investigación se orienta a la propuesta de una nueva metodología, denominada HAPA, para tratar la configuración y evaluación del desempeño de los algoritmos genéticos distribuidos ejecutados sobre plataformas heterogéneas, con el objetivo de obtener una implementación eficiente y eficaz de este tipo de algoritmos.

Palabras claves: Metaheurísticas, algoritmos genéticos distribuidos, paralelismo, mutación, estrategias adaptativas, problemas complejos

Contexto

Estas líneas de investigación se desarrollan en el marco de dos proyectos de investigación, llevados a cabo en el Laboratorio de Investigación de Sistemas Inteligentes (LISI) de la Facultad de Ingeniería de la Universidad Nacional de La Pampa. Uno de ellos acreditado por la Agencia Nacional de Promoción Científica y Tecnológica (PICTO 0278-2011) dirigido por la Dra. Salto, y el otro acreditado por dicha facultad y dirigido por la Dra. Minetti. Cabe destacar que desde hace varios años, los integrantes de estos proyectos mantienen una importante vinculación con investigadores de la Universidad Nacional de San Luis (Argentina) y de la Universidad de Málaga (España), con quienes se realizan publicaciones conjuntas.

Introducción

Un gran número de problemas de optimización de la vida real en ciencia, ingeniería, economía y negocios son complejos y difíciles de resolver. Ellos no pueden solucionarse de manera exacta en tiempo razonable. Usar algoritmos aproximados, tales como las metaheurísticas, es la principal alternativa para resolver esta clase de problemas.

El surgimiento en los últimos años de nuevas técnicas metaheurísticas para resolución de problemas de optimización generan un ámbito de análisis de experimentación interesante. La característica que comparten estas metaheurísticas es que están basadas en el comportamiento social de distintas especies, como la formación del vuelo en V de las

aves en su fase migratoria, el parasitismo obligado de algunas especies de cucos poniendo sus huevos en los nidos de otras aves (de otras especies), entre otras ideas. La mayoría de estas metaheurísticas tienen una formulación para problemas de optimización continuo. Por lo tanto, su aplicación en la resolución de problemas combinatoriales, como lo son los problemas de planificación de la producción o de enrutamiento, plantea un desafío interesante en la comunidad científica.

Por otra parte, las metaheurísticas cuando se les enfrenta a problemas de optimización complejos y/o alta dimensionalidad pueden presentar una degradación en su efectividad para encontrar el óptimo y, en muchos casos, reducir su desempeño. Esta degradación motiva la necesidad de desarrollar estrategias más eficientes para explorar las áreas más prometedoras del espacio de soluciones. Las alternativas pueden variar entre incluir adaptaciones en los procesos de búsqueda, incorporar búsquedas locales específicas al problema en cuestión, o establecer paralelizaciones de las tareas desarrolladas.

Determinar qué metaheurística es la más adecuada para la resolución de un problema, cuál es la configuración apropiada de la misma o sobre qué plataforma se ejecuta son interrogantes que intentan responder las diferentes líneas de investigación abordadas en nuestro proyecto.

Desarrollo

En las siguientes secciones se describen cada una de las líneas de investigación desarrolladas por este grupo de trabajo.

Nuevas técnicas metaheurísticas

En esta línea de investigación se trabaja con nuevas técnicas metaheurísticas surgidas en los últimos años para resolver problemas de relevancia en la industria. En particular, se ha elegido el problema de planificación de *job shop* flexible (FJSSP), el cual es una extensión del problema de planificación *job shop* clásico y considerado un problema NP-duro [1]. En este problema una misma operación se puede procesar en más de una máquina y de esta manera hay varias máquinas alternativas disponibles para procesar un trabajo. Por lo tanto, el FJSSP consiste de dos subproblemas: (i) asignar

cada operación a una máquina del conjunto de máquinas posibles y (ii) secuenciar las operaciones asignadas en las máquinas.

En particular se ha avanzado en la puesta a punto de un algoritmo de búsqueda cuco (CS) [2], [3] para resolver este problema. CS está basado en el comportamiento de parasitismo obligado de algunas especies de cucos en combinación con el comportamiento de los vuelos de Lévy de algunas aves y de las moscas de la fruta. Se ha trabajado especialmente en definir la generación de nuevas soluciones candidatas, las cuales se obtienen al perturbar una solución de la población siguiendo una distribución Lévy. Un vuelo Lévy se puede describir como un camino aleatorio en el cual el largo de paso (distancia entre dos soluciones) se decide por una función de distribución de probabilidades. Se usan tres movimientos u operadores diferentes los cuales son controlados por el valor generado por la distribución Lévy: intercambio, inserción e inversión. Estos operadores son tradicionales para el problema en cuestión y generan diferentes pasos en el movimiento. El movimiento de intercambio se usa para introducir pequeñas perturbaciones mientras que el movimiento de inserción realiza grandes saltos. El movimiento de inserción introduce perturbaciones intermedias.

Otro problema de relevancia en la industria es el problema de enrutamiento vehicular con capacidad (*Capacited Vehicle Routing Problem, CVRP*). Este problema se relaciona con la determinación de rutas óptimas (costos totales mínimos) usadas por una flota de vehículos para realizar entregas a los clientes desde un mismo depósito, considerando que todos los vehículos tienen una capacidad limitada e idéntica.

Para resolver CVRP se está estudiando la adaptación de la metaheurística conocida como optimización basada en la migración de las aves (MBO) al mismo [4]. MBO toma como referencia el comportamiento de las aves durante sus largos viajes migratorios. Cada solución desarrollada en la ejecución del algoritmo representa un ave en la formación de vuelo en V y se ve beneficiada por la solución que tiene enfrente. Por tal motivo al iniciar la ejecución del algoritmo el primer paso es establecer la organización de los integrantes de la bandada. De esta manera contamos con

una solución líder y con dos listas enlazadas que representarán a las ramas izquierda y derecha de la formación. Luego, el líder será el encargado de encabezar la búsqueda y de facilitar el trabajo a sus seguidores inmediatos, quienes harán lo mismo con los suyos, y así siguiendo hasta llegar al final de la formación.

Metaheurísticas con estrategias adaptativas

La segunda línea de investigación está relacionada con la propuesta de estrategias para adaptar la probabilidad de mutación (p_m) durante la búsqueda genética [5]. El objetivo de la estrategia propuesta, denominada APmGA, es incrementar la p_m si se pierde gradualmente la diversidad genética, con el fin de mantener una población distribuida en el espacio de búsqueda. Por el contrario, el valor p_m se reduce cuando se observa un incremento en la diversidad de la población. Por lo tanto, estos cambios en el valor de p_m son una fuente adicional para equilibrar exploración y explotación. Para lograr esto, la estrategia monitorea la diversidad genotípica presente en la población al usar la medida entrópica de Shannon [6]. A partir de esta propuesta [7] surge la necesidad de comparar diferentes métodos de adaptación de p_m durante la ejecución de un algoritmo genético (GA), y de esta manera determinar cuáles de ellos permite que el GA brinde resultados de alta calidad con un esfuerzo numérico reducido. El método propuesto es comparado con el presentado por Riff et al. [8], el cual se basa en las ideas de aprendizaje reforzado. Esto es, un operador recibe un premio en su probabilidad de uso cuando su aplicación genera un individuo mejor que sus padres. Análogamente, este recibe una penalidad cuando el hijo tiene un valor de fitness peor que los de sus progenitores. Tanto los premios como las penalidades dependen fuertemente de la mejora/degradación del resultado de la función de evaluación. El último método a comparar está basado en la idea de Srinivas et al. [9], donde las probabilidades de mutación son variadas dependiendo de los valores de fitness de la solución. Es decir que, las soluciones con alto valor de fitness son protegidas, mientras que las soluciones con fitness por debajo de la media sufren una variación genética importante. Estas tres estrategias adaptativas son consideradas debido a que difieren en la información de la búsqueda

para actualizar el valor de probabilidad y a su rendimiento.

Metaheurísticas diseñadas específicamente para resolver un determinado problema

Problem Aware Local Search es una metaheurística basada en trayectoria, propuesta por Alba y Luque [10], que fue inicialmente desarrollada para resolver el problema de ensamblado de fragmentos de ADN, un problema de optimización combinatoria NP-duro. PALS realiza una explotación exhaustiva y sistemática del vecindario de la solución tentativa a considerarse. Una característica clave de este algoritmo es que las soluciones vecinas son explotadas eficientemente por medio de un cálculo de fitness muy rápido basado en los cambios realizados en la solución, en lugar de la evaluación completa del nuevo valor de fitness. Esto permite que PALS examine un gran número de soluciones tentativas muy rápidamente, analizando más exhaustivamente el espacio de búsqueda y así obtener resultados de gran calidad.

Dado que el poder de esta metaheurística, aún, no ha sido completamente explotado, en esta línea de investigación se estudia si PALS puede ser eficientemente empleado para resolver otros problemas de optimización combinatoria [11]. Los problemas elegidos son importantes tanto en la ciencia como en el ámbito empresarial, a saber: el problema de asignación cuadrática, el problema de planificación de flujo y el problema de mochilas múltiples.

Por otra parte PALS ofrece un marco general de búsqueda el cual puede instanciarse de acuerdo a algún diseño alternativo, pero, actualmente, la literatura no muestra un análisis de los puntos de diseño clave de esta metaheurística. En consecuencia en esta línea de investigación también se proponen y analizan diferentes puntos de diseño para PALS. Los cuales incluyen diferentes estrategias para generar la solución inicial, distintos criterios para agregar y seleccionar movimientos al y del conjunto de soluciones candidatas, operadores de variación apropiados para generar movimientos en las soluciones candidatas y la definición de funciones de evaluación parcial del fitness según el problema. Esto da origen a un total de 81 configuraciones distintas de PALS.

Metaheurísticas distribuidas

Por último, una línea de investigación se orienta

a la propuesta de una nueva metodología, denominada *Hardware Aware Parallel Algorithms* (HAPA) [12], [13], para tratar la configuración y evaluación del desempeño de los algoritmos genéticos distribuidos ejecutados sobre plataformas de cómputo heterogéneas, con el objetivo de obtener una implementación eficiente y eficaz de este tipo de algoritmos.

Las investigaciones más recientes de algoritmos genéticos distribuidos sobre plataformas heterogéneas [14], [15], [16], [17], [18] se enfocan en la resolución de un determinado problema, y no en la metodología que podría aplicarse cuando es necesario usar hardware heterogéneo. En este sentido, nuestra propuesta consta de un nuevo procedimiento metodológico, y el consecuente diseño algorítmico. El objetivo de HAPA es facilitar una implementación eficiente y numéricamente precisa de algoritmos genéticos distribuidos (dGA) sobre un conjunto de máquinas con diferentes procesadores, tamaños de memoria principal, y sistemas operativos. HAPA consta de tres fases: (i) el cálculo de un ranking de los procesadores con el fin de conocer la plataforma (*offline*), (ii) el diseño de algoritmos derivados de la fase anterior, y (iii) la validación de un dGA (*online*). En particular se ha aplicado la metodología HAPA para regular las condiciones de parada del dGA y la frecuencia de la migración de dicho algoritmo.

Resultados obtenidos

A continuación se detallan los resultados obtenidos por cada una de las líneas de investigación abordadas.

Nuevas técnicas metaheurísticas

CS resultó ser una técnica fácil de trabajar para la resolución del FJSSP, ya que necesita de muy pocos parámetros para su funcionamiento. Los resultados obtenidos han sido prometedores para las instancias del problema analizadas, resultando un algoritmo competitivo en cuanto a cantidad de soluciones óptimas reportadas como en el esfuerzo computacional incurrido.

El análisis de los resultados obtenidos, al resolver CVRP usando MBO, establece que esta metaheurística es robusta a la hora de considerar

instancias de prueba de diferente tamaño. Además brinda soluciones de alta calidad, dado que los errores con respecto al óptimo son menores al 6 %. Por último, también se concluye que es una herramienta eficiente ya que encuentra las mejores soluciones en menos de 170 iteraciones.

Metaheurísticas con estrategias adaptativas

Las estrategias de control consideradas para actualizar la probabilidad de mutación reducen el tiempo previo de puesta a punto para determinar el valor de probabilidad que más favorezca al desempeño del GA. En tanto que, los resultados obtenidos por nuestra propuesta APmGA comparados con los de la literatura, AcGA y AGA, son muy alentadores, dado que permiten obtener soluciones de mayor calidad que los otros con un esfuerzo numérico razonable. Esto se debe a que el tiempo computacional dedicado al cálculo de la probabilidad no es significativo.

Metaheurísticas diseñadas específicamente para resolver un determinado problema

Los resultados obtenidos por PALS, bajo cada una de las 81 configuraciones, en la resolución de los diferentes problemas de optimización combinatoria prueban que tiene la habilidad para resolverlos eficientemente. Por lo tanto, las fronteras de PALS han sido exitosamente expandidas a otros dominios. Además, esto se logra bajo configuraciones similares de los puntos de diseño algorítmicos. Esto ayuda a reducir el tiempo requerido para determinar la mejor configuración para PALS.

A partir de este estudio, también, puede concluirse que PALS permite que el diseñador logre fácilmente un correcto equilibrio entre la intensificación y la diversificación. La primera se logra por medio del elitismo usado a la hora de generar la solución inicial y de agregar movimientos al conjunto de soluciones candidatas. En tanto que la diversificación se alcanza a través de la aplicación del operador de variación.

Metaheurísticas distribuidas

De las pruebas realizadas para evaluar el desempeño de HAPA respecto de un dGA ejecutado en un entorno homogéneo, se concluye que: la calidad de la solución es comparable con la obtenida usando una plataforma homogénea, pero lo más importante es que se logra una disminución del esfuerzo numérico y, consecuentemente, del tiempo de ejecución para hallar dichas

soluciones cuando HAPA se aplica en ambientes heterogéneos. En resumen, este trabajo aporta una forma de evitar el diseño *ad-hoc* de algoritmos para plataformas heterogéneas, también se abre una nueva línea de investigación sobre la forma de inyectar conocimientos de hardware en los parámetros del software de los algoritmos. Además de ser una propuesta innovadora, se ha demostrado que es numéricamente competitiva y que los tiempos de ejecución son reducidos.

Formación de recursos humanos

Cada año se incorporan al proyecto alumnos avanzados en la carrera Ingeniería en Sistemas, quienes trabajan en temas relacionados a la resolución de problemas de optimización usando técnicas inteligentes, con el objeto de guiarlos en el desarrollo de sus tesis de grado y, también, de formar futuros investigadores científicos. Actualmente, tres becarios alumnos desarrollan sus tareas dentro del marco del proyecto (1 beca CIN y las restantes de Universidad). Durante el transcurso del año 2016 se dirigió una tesis de grado (defendida en diciembre de 2016) y otros dos trabajos los cuales se encuentran en estado avanzado de desarrollo. Por otra parte, los docentes-investigadores que integran el proyecto realizaron diversos cursos de posgrado relacionados con la temática del proyecto, con el objetivo de sumar los créditos necesarios para cursar carreras de posgrado.

REFERENCIAS

- [1] E. Garey, D.S. Johnson, and R. Sethi, "The complexity of flow-shop and job-shop scheduling," *Mathematics of Operations Research*, vol. 1, pp. 117–129, 1976.
- [2] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *Proceedings of the World Congress on Nature & Biologically Inspired Computing (NABIC '09)*, 2009, pp. 210–214.
- [3] D. S. Yang X-S, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4, pp. 330–343, 2010.
- [4] E. Duman, M. Uysal, and A. F. Alkaya, "Migrating birds optimization: A new metaheuristic approach and its performance on quadratic assignment problem," *Inf. Sci.*, vol. 217, pp. 65–77, Dec. 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.ins.2012.06.032>
- [5] N. Stark, G. Minetti, and C. Salto, "Comparison of different approaches for adapting mutation probabilities in genetic algorithms," in *XXII Congreso Argentino de Ciencias de la Computación*, 2016, pp. 75–84.
- [6] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
- [7] E. Alba and B. Dorronsoro, "The exploration/exploitation tradeoff in dynamic cellular genetic algorithms," *Proceedings of IEEE Transactions on Evolutionary Computation*, vol. 9, no. 2, pp. 126–142, 2005.
- [8] E. Montero and M.-C. Riff, "On-the-fly calibrating strategies for evolutionary algorithms," *Information Sciences*, vol. 181, no. 3, pp. 552–566, 2011.
- [9] M. Srinivas and L. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 4, pp. 656–667, 1994.
- [10] G. Luque and E. Alba, "Metaheuristics for the dna fragment assembly problem," *International Journal of Computational Intelligence Research*, vol. 1, no. 2, pp. 98–108, 2005.
- [11] G. Minetti, Gabriela F. and Luque and E. Alba, "The problem aware local search algorithm: an efficient technique for permutation-based problems," *Soft Computing*, pp. 1–14, 2017. [Online]. Available: <http://dx.doi.org/10.1007/s00500-017-2515-9>
- [12] C. Salto and E. Alba, "Adapting distributed evolutionary algorithms to heterogeneous hardware," *Trans. Computational Collective Intelligence*, vol. 19, pp. 103–125, 2015.
- [13] —, "Propuesta metodológica para el diseño de metaheurísticas paralelas en ambientes de cómputo heterogéneo," in *Actas de la XVII Conferencia de la Asociación Española para la Inteligencia Artificial*, 2016, pp. 301–302.
- [14] J. W. B. Jr. and S. V. Kumar, "Asynchronous genetic algorithms for heterogeneous networks using coarse-grained dataflow," in *GECCO*, ser. Lecture Notes in Computer Science, E. Cantˆo-Paz, J. A. Foster, K. Deb, L. Davis, R. Roy, U.-M. O'Reilly, H.-G. Beyer, R. K. Standish, G. Kendall, S. W. Wilson, M. Harman, J. Wegener, D. Dasgupta, M. A. Potter, A. C. Schultz, K. A. Dowland, N. Jonoska, and J. F. Miller, Eds., vol. 2723. Springer, 2003, pp. 730–741.
- [15] M. García-Arenas, J. J. Merelo Guervós, P. Castillo, J. L. J. Laredo, G. Romero, and A. M. Mora, "Using free cloud storage services for distributed evolutionary algorithms," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '11. New York, NY, USA: ACM, 2011, pp. 1603–1610. [Online]. Available: <http://doi.acm.org/10.1145/2001576.2001792>
- [16] Y. Gong, M. Nakamura, and S. Tamaki, "Parallel genetic algorithms on line topology of heterogeneous computing resources," in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '05. New York, NY, USA: ACM, 2005, pp. 1447–1454. [Online]. Available: <http://doi.acm.org/10.1145/1068009.1068239>
- [17] K. Meri, M. G. Arenas, A. M. Mora, J. J. Merelo, P. A. Castillo, P. García-Sánchez, and J. L. J. Laredo, "Cloud-based evolutionary algorithms: An algorithmic study," *Natural Computing*, vol. 12, no. 2, pp. 135–147, 2012. [Online]. Available: <http://dx.doi.org/10.1007/s11047-012-9358-1>
- [18] S. Mostaghim, J. Branke, A. Lewis, and H. Schmeck, "Parallel multi-objective optimization using master-slave model on heterogeneous resources," in *IEEE Congress on Evolutionary Computation*. IEEE, 2008, pp. 1981–1987.

Proceso de Elicitación de Requerimientos para la Construcción de Modelos Predictivos Basados en Sistemas Inteligentes

Cinthia Vegega, Pablo Pytel, Hugo Ramón, Luciano Straccia, Anahi Bazet, Ariel Deroche, Sebastián Plawner & María Florencia Pollo-Cattaneo
 Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS)
 Programa de Maestría en Ingeniería en Sistemas de Información.
 Facultad Regional Buenos Aires. Universidad Tecnológica Nacional.
 cinthiavg@yahoo.com.ar; {flo.pollo, ppytel}@gmail.com

Resumen

¿Cuál opción es la más favorable? ¿Qué nos deparará a futuro? Estas son preguntas que nos hacemos cada vez que tenemos que tomar alguna decisión. Por esta razón, contar con mecanismos que permitan realizar predicciones acertadas es algo que la Humanidad siempre ha querido. Dicha necesidad no es propia únicamente de los individuos y también impacta en las organizaciones. Teniendo en cuenta estas cuestiones, el presente proyecto propone el desarrollo de un proceso que a partir de los deseos y necesidades de una organización, asista a la elicitación de los requerimientos para implementar exitosamente un Modelo Predictivo basado en Sistemas Inteligentes.

Palabras clave: Ingeniería de Software. Elicitación de Requerimientos. Modelos Predictivos. Sistemas Inteligentes. Aprendizaje Automático.

Contexto

En el marco de las actividades del Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS) se comienza una nueva línea de trabajo en el campo de la automatización de procesos orientados a la Toma de Decisiones en organizaciones, mediante la aplicación de

Modelos Predictivos, basados en Sistemas Inteligentes. De esta manera, se articula dentro de los objetivos de GEMIS en el campo de la Informática, la generación de nuevos conocimientos en el área de la Ingeniería de Software con la aplicación de tecnologías no convencionales provenientes del Aprendizaje Automático. Además, se desarrolla dentro del ámbito del PID con incentivos UTN ‘UTI3799TC’, y una Tesis de Maestría.

Introducción

La Toma de Decisiones es un proceso esencial en la vida de las organizaciones [1]. Como se indica en [2], las organizaciones son sistemas no lineales altamente sensibles a las innovaciones, a los eventos o al azar propio del ambiente empresarial. Esto lleva a que diariamente deben decidir cómo asignar sus valiosos recursos basados en predicciones [3] sobre el tiempo, el esfuerzo y/o los riesgos que implica llevar a cabo sus actividades. Esta situación se encuentra acentuada debido al entorno altamente complejo y de difícil predicción del siglo XXI [4]. Además genera que las organizaciones deban actuar más allá de los sistemas tradicionales de gestión e incorporar nuevos mecanismos para la “*creación y potenciación del conocimiento organizativo*” tales como

los provistos por la Inteligencia Artificial [5; 6]. Un caso de estos nuevos mecanismos se puede encontrar en el ámbito de los Modelos Predictivos.

A pesar de que tradicionalmente se han aplicado Técnicas Estadísticas y Modelos Paramétricos para generar predicciones [7], en las dos últimas décadas se han incorporado diversos métodos asociados al Aprendizaje Automático [3; 8]. Estos métodos son considerados normalmente como ‘cajas negras’ o modelos basados en datos [9] y permiten construir modelos que permiten encontrar la relación entre situaciones pasadas y futuras, utilizando los datos históricos disponibles.

No obstante, desafortunadamente los Modelos Predictivos suelen ser imprecisos o, en algunos casos, se equivocan, generando a menudo respuestas que son incomprensibles [3]. En este sentido, la calidad de la información requerida es sumamente importante para poder tomar decisiones bajo certeza [10]. Al reducir la ignorancia sobre el problema y su contexto es posible generar mejores predicciones. Como indica el físico Richard Feynman, *"lo que siempre deberíamos tener en cuenta para evitar grandes fracasos, catástrofes y pánicos no es lo que sabemos, sino lo que no sabemos que no sabemos"* [11]. Por consiguiente, además de recolectar los datos históricos que se aplicarán para construir el Modelo Predictivo, también es imprescindible identificar las características generales del dominio en donde se lleva a cabo la predicción, y así poder detectar situaciones o eventos de los que no se poseen datos, pero que el modelo deberá considerar.

En este contexto, la presente línea de trabajo propone un nuevo proceso ingenieril que asista al Ingeniero de Software en la difícil tarea de recolectar, entender, identificar y documentar la

información necesaria para implementar un Modelo Predictivo basado en Sistemas Inteligentes (también denominado como Modelo Predictivo Inteligente o MPI). Para ello, se consideran las particularidades del problema predictivo y de las arquitecturas más utilizadas para resolverlo, así como, las necesidades de las organizaciones argentinas.

Líneas de Investigación, Desarrollo e Innovación

De acuerdo a la meta establecida en la sección anterior, a continuación se exponen los motivos que justifican el proceso propuesto:

o **Modelos Predictivos Inteligentes**

Se considera al Aprendizaje Automático (o Machine Learning) como la disciplina que estudia y modela los procesos de aprendizaje con sus múltiples manifestaciones para poder ser trasladados a las computadoras [9]. De esta manera, se busca reproducir aspectos del comportamiento de los seres inteligentes para implementar sistemas software en ámbitos complejos [12]. Los sistemas software así generados se denominan Sistemas Inteligentes [13] y pueden ser aplicados en diferentes dominios para resolver diferentes tipos de problemas. Uno de estos problemas tiene que ver con la construcción de modelos capaces de inferir a partir de datos históricos las dependencias entre valores pasados y valores futuros a corto plazo [8], es decir, los Modelos Predictivos.

Debido al incremento de la cantidad de datos disponibles en los repositorios de las organizaciones, y mejoras de los rendimientos de procesamiento del hardware, es posible aplicar métodos computacionales intensivos para construir este tipo de modelo [14]. En este sentido se destacan las Redes Neuronales

Artificiales, o RNA, [15; 16] y las Redes Bayesianas, o RB, [17; 18] como las principales arquitecturas a ser utilizadas para este tipo de problema [14; 19; 20].

○ **Dificultades de la construcción de Modelos Predictivos Inteligentes**

Según indica [8], predecir el futuro es una de las tareas más importantes y difíciles de las ciencias aplicadas. En este sentido, resulta obvio que contar con datos representativos del problema le permitirá al modelo generar mejores predicciones. Sin embargo, obtener los datos no es una tarea trivial. Por ejemplo, tómesese la cuestión de determinar la cantidad de información histórica que se necesita para producir los mejores resultados [21]. Según [22], la respuesta habitual a la pregunta “¿cuántos datos se necesitan?” es “lo más posible”. Cuantos más datos se tenga, mejor se podrá identificar la estructura del modelo y los patrones que se utilizan para la predicción pero, en la práctica, es imprescindible ponerle algún límite. Entonces, para definir la cantidad de datos a utilizar es necesario primero identificar las fuentes de datos disponibles y comprender sus características. Lo mismo sucede al momento de determinar los atributos apropiados que deberán ser seleccionados como variables de entrada y salida del modelo [21]. Además de analizar las fuentes de información histórica, también se deberán utilizar otras fuentes, como son las opiniones de expertos [22].

Estas necesidades no escapan a los Sistemas Inteligentes. Como dice [12], “*el Aprendizaje Automático no es magia, no se puede conseguir algo de la nada*”. Una de las cuestiones principales de la implementación de estos sistemas tiene que ver con los pasos asociados a recolectar los datos, integrarlos, limpiarlos y pre-procesarlos. Para lo cual, es imprescindible identificarlos y

conocerlos primero [23]. Otro tipo de dificultad asociada a la construcción de los MPI tiene que ver con determinar el tipo de tecnología a aplicar y las parametrizaciones a ser consideradas (como por ejemplo, su topología). De acuerdo al Teorema de ‘No Free Lunch’ [24], no existe ningún algoritmo de Aprendizaje Automático que pueda ser aplicado para cualquier problema.

Del análisis anterior se desprende que, las principales dificultades asociadas a la construcción de un MPI dependen de los aspectos relativos al problema que se quiere resolver. En tal sentido, es imprescindible conocer las expectativas que el cliente tiene sobre el modelo que se traducirán en los objetivos que deberá cumplir. Dado que los mismos normalmente estarán muy relacionados con las metas estratégicas y tácticas de la organización, también se deben entender características propias de la organización, su contexto y los procesos de negocio realizados. Una vez que estos elementos son identificados, es necesario realizar un reconocimiento inicial de las fuentes de información disponibles en la organización identificando cuáles fuentes se encuentran informatizadas y cuáles no. En el caso de los repositorios, es muy importante también, determinar la estructura, naturaleza y calidad de los datos disponibles. Por consiguiente, se considera indispensable elicitación todos estos requerimientos en forma completa y detallada para poder construir así el modelo que satisfaga las necesidades predictivas del cliente.

○ **Elicitación de Requerimientos para Modelos Predictivos Inteligentes**

Dentro de la Ingeniería de Software existe la Ingeniería de Requisitos que es la encargada de recopilar, analizar y validar los requerimientos de los distintos interesados (o stakeholders) sobre el

sistema software a implementar [25]. Entre sus actividades se aplican diferentes técnicas de elicitación, cuyo principal objetivo es entender los deseos y las necesidades de dichos interesados [26].

Sin embargo, a diferencia de los proyectos de desarrollo de software tradicional, la problemática abordada en los proyectos de construcción de Sistemas Inteligentes es diferente. Por lo tanto, existe otra disciplina denominada Ingeniería del Conocimiento o INCO [27] la cual se encuentra orientada a la construcción de Sistemas Basados en Conocimientos en general y los Sistemas Expertos en particular. Estos últimos son Sistemas Inteligentes que tienen la capacidad de emular parte del comportamiento de un experto humano para realizar una tarea en un dominio muy complejo y poco estructurado. Aunque existen metodologías que guían este proceso de Adquisición de Conocimientos [28-30], éstas se concentran en definir los hechos y reglas que serán almacenados en una Base de Conocimiento.

Esto significa que las metodologías de la INCO no son aplicables para construir otros Sistemas Inteligentes que no tengan Base de Conocimiento, como es el caso de las RNA y las RB. En función de lo que se ha identificado anteriormente, los requerimientos de dichas arquitecturas se encuentran asociados a otros elementos del problema y su dominio. A partir de un relevamiento efectuado en el campo metodológico, se han encontrado diversas publicaciones vinculadas a la definición de estos elementos tales como [23; 31- 37]. No obstante, la mayoría de estas publicaciones se encuentran más orientadas a la identificación de aspectos relativos a la construcción de RB, dejando de lado las particularidades de las RNA. Además, dentro de las propuestas orientadas a las RB, cada una se

concentra en algún aspecto dejando de lado otros por lo que no existe una guía integral para el desarrollo de esta arquitectura. Por consiguiente, se considera de interés unificar las propuestas anteriores en un único proceso que asista a las primeras etapas de la construcción de una RNA y/o una RB como MPI. Para ello, además se reutilizará la experiencia existente en el uso de las técnicas de elicitación de requerimientos de la Ingeniería de Software y la INCO, lo cual dotará al proceso de desarrollo con las características del abordaje ingenieril contribuyendo así al avance del conocimiento científico mediante el uso de técnicas consistentes.

Resultados y Objetivos

La presente línea de trabajo propone un modelo de proceso definido y estandarizado para la elicitación de los requisitos necesarios para la implementación de Modelos Predictivos mediante RNA o RB. De esta manera, se procura aportar a los Ingenieros involucrados en la implementación de estos tipos de Sistemas Inteligentes, una guía que les brinde soporte durante las etapas iniciales del proyecto teniendo presente las particularidades de los mismos.

Dentro de los objetivos específicos considerados se encuentran:

- La identificación de los conceptos necesarios a ser elicitados para la construcción de MPI.
- La especificación de los procesos necesarios para la elicitación de los conceptos identificados.
- La selección de técnicas de extracción y educación de conocimientos a ser aplicadas en dichos procesos.
- La definición del tipo de arquitectura y topología a ser aplicado para la

construcción del MPI, teniendo en cuenta los requerimientos elicitados.

Formación de Recursos Humanos

El equipo de trabajo se encuentra conformado por tres investigadores formados, un tesista de maestría, dos graduados y un alumno en la carrera de grado. Asimismo, se prevé incorporar alumnos avanzados en la carrera de Ingeniería en Sistemas de Información con posibilidades de articular sus Trabajos Finales de Carrera de Grado. De esta manera se espera generar un verdadero espacio integrado de investigación en carreras de grado y posgrado.

Referencias

- [1]. Robbins, S. & Coulter, M. (2010). *Administración*. Décima Edición. Prentice Hall.
- [2]. Arévalo, L. E. B. (2013). *La organización empresarial como sistema adaptativo complejo*. Estudios Gerenciales, 29(127), 258-265.
- [3]. Mair, C., Kadoda, G., Lefley, M., Phalp, K., Schofield, C., Shepperd, M. & Webster, S. (2000). *An investigation of machine learning based prediction systems*. Journal of Systems and Software, 53(1), 23-29.
- [4]. García, F. J. M., Martínez, M. A. P. & García, J. S. (2003). *Gestión Estratégica del Conocimiento*. Asociación Universitaria Iberoamericana de Postgrado.
- [5]. Nilsson, N. J. (2014). *Principles of Artificial Intelligence*. Morgan Kaufmann.
- [6]. Russell, S. J., Norvig, P., Davis, E., Russell, S. J. & Russell, S. J. (2010). *Artificial Intelligence: a modern approach* (Vol. 2). Englewood Cliffs: Prentice Hall.
- [7]. Shepperd, M. & Kadoda, G. (2001). *Comparing software prediction techniques using simulation*. IEEE Transactions on Software Engineering, 27(11), 1014-1022.
- [8]. Bontempi, G., Taieb, S. B. & Le Borgne, Y. A. (2013). *Machine learning strategies for time series forecasting*. In Business Intelligence (pp. 62-77). Springer Berlin Heidelberg.
- [9]. Alpaydin, E. (2014) *Introduction to machine learning*. MIT press.
- [10]. Rodríguez, M. & Márquez Alegría, M. (2015). *Manejo de problemas y toma de decisiones* (Vol. 8). Editorial El Manual Moderno.
- [11]. Haldane, A. G. & Madouros, V. (2012). *The dog and the frisbee*. Revista de Economía Institucional, 14(27), 13-56.
- [12]. Domingos, P. (2012). *A few useful things to know about machine learning*. Communications of the ACM, 55(10), 78-87.
- [13]. Cohen, P. R. & Feigenbaum, E. A. (2014). *The handbook of Artificial Intelligence*. Vol. 3. Butterworth-Heinemann.
- [14]. De Gooijer, J. G. & Hyndman, R. J. (2006). *25 years of time series forecasting*. International journal of forecasting, 22(3), 443-473.
- [15]. Wang, S. C. (2003). *Artificial Neural Network*. In Interdisciplinary Computing in Java Programming (pp. 81-100). Springer US.
- [16]. Wang, L. & Fu, K. (2009). *Artificial Neural Networks*. Wiley Encyclopedia of Computer Science and Engineering. 181-188.
- [17]. Barber, D. (2012). *Bayesian Reasoning and Machine Learning*. The MIT Press.
- [18]. Premchaiswadi, W. (2012) *Bayesian Networks*. Ed. In-Tech.
- [19]. Zhang, G., Patuwo, B. E. & Hu, M. Y. (1998). *Forecasting with artificial neural networks: The state of the art*. International journal of forecasting, 14(1), 35-62.
- [20]. Chatfield, C. (2016). *The analysis of time series: an introduction*. CRC press.
- [21]. Walczak, S. (2001). *An empirical analysis of data requirements for financial forecasting with neural networks*. Journal of management information systems, 17(4), 203-222.
- [22]. Hyndman, R. J. & Kostenko, A. V. (2007). *Minimum sample size requirements for seasonal forecasting models*. Foresight, 6(Spring), 12-15.
- [23]. Trujillano, J., March, J. & Sorribas, A. (2004). *Aproximación metodológica al uso de redes neuronales artificiales para la predicción de resultados en medicina*. Med Clin (Barc), 122(s1).
- [24]. Wolpert, D. H. & Macready, W. G. (1997). *No free lunch theorems for optimization*. IEEE transactions on evolutionary computation, 1(1), 67-82.
- [25]. Nuseibeh, B. & Easterbrook, S. (2000). *Requirements engineering: a roadmap*. In Proceedings of the Conference on the Future of Software Engineering (pp. 35-46). ACM.
- [26]. Robertson, S. & Robertson, J. (2012). *Mastering the requirements process: Getting requirements right*. Addison-wesley.
- [27]. García Martínez, R. & Britos, P. (2004). *Ingeniería de Sistemas Expertos*. Editorial Nueva Librería.
- [28]. Buchanan, B. G., Barstow, D., Bechtal, R., Bennett, J., Clancey, W., Kulikowski, C., Mitchell, T. & Waterman, D. A. (1983). *Constructing an expert system*. Building expert systems, 50, 127-167. Addison Wesley.
- [29]. Gómez, A., Juristo, N., Montes, C. y Pazos, J. (1997). *Ingeniería del Conocimiento*. Editorial R. Areces. Madrid.
- [30]. Schreiber, G., Akkermans, H., Anjewierden, A., Hoog, R. D., Shadbolt, N. R. & Wielinga, B. (2000). *Knowledge engineering and management: the CommonKADS methodology*. MIT press.
- [31]. Chen, S. H. & Pollino, C. A. (2012). *Good practice in Bayesian network modelling*. Environmental Modelling & Software, 37, 134-145.
- [32]. Constantinou, A. C., Fenton, N., Marsh, W. & Radlinski, L. (2016). *From complex questionnaire and interviewing data to intelligent Bayesian network models for medical decision support*. Artificial intelligence in medicine, 67, 75-93.
- [33]. Henrion, M. (2013). *Practical issues in constructing a Bayes' belief network*. arXiv preprint arXiv:1304.2725.
- [34]. Masegosa, A. R. & Moral, S. (2013). *An interactive approach for Bayesian network learning using domain/expert knowledge*. International Journal of Approximate Reasoning, 54(8), 1168-1181.
- [35]. Mendes, E. (2011). *Knowledge representation using Bayesian networks - A case study in Web effort estimation*. In Information and Communication Technologies (WICT), 2011 World Congress on (pp. 612-617). IEEE.
- [36]. Pitchforth, J. & Mengersen, K. (2013). *A proposed validation framework for expert elicited Bayesian Networks*. Expert Systems with Applications, 40(1), 162-167.
- [37]. Zhou, Y., Fenton, N. & Neil, M. (2014). *Bayesian network approach to multinomial parameter learning using data and expert judgments*. International Journal of Approximate Reasoning, 55(5), 1252-1268.

Revisión de Creencias en Bases de Conocimiento Horn y su Complejidad Computacional

Néstor Jorge Valdez⁰

Marcelo A. Falappa⁰

⁰ Departamento de Ciencias de la Computación, Fac. de Ciencias Exactas y Naturales
Universidad Nacional de Catamarca (UNCa)

Av. Belgrano 300 - San Fernando del Valle de Catamarca
Tel.: (03834)420900 - e-mail: njvaldez@exactas.unca.edu.ar

⁰ Laboratorio de Investigación y Desarrollo en Inteligencia Artificial
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Av. Alem 1253, (B8000CPB) Bahía Blanca, Argentina
Tel: (0291)4595135 / Fax: (0291)4595136
e-mail: mfalappa@cs.uns.edu.ar

Resumen

Esta línea de investigación tiene como objetivo, dentro del marco de la teoría de cambio de creencias, la profundización del estudio de los operadores de cambio múltiple donde la lógica subyacente está conformada por cláusulas Horn. Se analizan los resultados alcanzados de estos operadores a los que hemos denominado *revisión múltiple priorizada Horn: Partial Meet y Kernel* en bases de creencias. Además, se estudian otros resultados presentados sobre operadores de *revisión múltiple no-priorizada*, de manera tal de posibilitar la extensión de estos operadores también bajo fragmento Horn. Así también, nos enfocamos en los aspectos computacionales de estas construcciones, donde estudiamos el problema de la complejidad computacional que presentan estos esquemas de revisión de creencias y quedificultan lograr su tratabilidad.

Palabras Claves Revisión de Creencias Múltiple Horn, Bases de Creencias, Complejidad Computacional.

1. Contexto

Esta línea de investigación se realizará dentro del ámbito del Laboratorio de Investigación del Departamento de Ciencias de la Computación de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de Catamarca. Esta investigación forma parte de las contribuciones de la tesis para la obtención del título de Doctor en Ciencias de la Computación de la Universidad Nacional del Sur por parte de uno de los autores del presente trabajo: Néstor Jorge Valdez. Así también, esta temática de estudio está asociado con el proyecto de investigación Bianual: “*Cambio de Creencias Múltiples en Sistemas Argumentativos Aplicado para Programación Lógica Rebatible*”, financiado por el programa de desarrollo científico y tecnológico de la Secretaría de Ciencia y Tecnología: Consejo de Investigación, de la Universidad Nacional de Catamarca. Periodo: 01/01/2016 al 31/12/2017.

2. Introducción

Un sistema de revisión de creencias constituye un marco lógico para modelar la dinámica de conocimiento, esto es, cómo modificamos nuestro estado epistémico cuando recibimos información nueva. Generalmente, la nueva información es asumida mediante una sentencia simple, como en la teoría AGM clásica [1], pero la entrada epistémica también podría presentarse a través de un conjunto de sentencias en lo que se conoce como cambio múltiple. En este contexto, existen diversos tipos de operadores de cambio, como los operadores de *cambio priorizada*, *no priorizadas*, *merging*, de *cambios selectivos*, y los operadores de *semi-revisión*, entre otros [4].

Debido a que las operaciones de cambio se llevan a cabo sobre el estado epistémico del agente, existen diferentes posibles construcciones para los distintos operadores de cambio. Dos de los modelos más comúnmente usados son sobre *conjuntos de creencias* y sobre *bases de creencias*. En los últimos años se han presentado diversos estudios que muestran un esfuerzo importante en la definición de las operaciones de contracción y revisión del estilo AGM que operan bajo lógica Horn [2, 12, 13]. El cambio de creencias bajo lógica Horn resulta importante por varias razones. En primer lugar, muchos sistemas de inteligencia artificial están expresados en lenguaje de cláusulas Horn y porque, además, han encontrado un amplio uso en bases de datos, programación lógica, sistemas de mantenimiento de la verdad, y bases de datos deductivas.

En nuestra investigación, consideramos un tipo de cambio múltiple bajo cláusulas Horn. Nos enfocamos en revisión de creencias en donde la nueva información es incorporada consistentemente en la *base de creencias*. Tratamos los operadores de cambio múltiple priorizadas, en donde se supone que las nuevas creencias deben ser plenamente

aceptadas. Otro objetivo de nuestro trabajo es abordar la complejidad computacional que presentan estos operadores de revisión de creencias con cláusulas Horn. Esto nos permitirá determinar el costo desde un punto de vista computacional de la generación de una base revisada bajo nuestros operadores con lógica Horn y demostrar que nuestros operadores son *representacionalmente factible* al manifestar un comportamiento polinomialmente aceptable.

Como principal contribución de esta investigación, caracterizamos a los operadores de revisión Horn priorizadas mediante conjuntos de postulados [11]. En cada caso, presentamos diferentes construcciones mediante técnicas de cambio parcial meet y cambio kernel. Además, tratamos el problema computacional de revisión de creencias analizando la complejidad computacional en el razonamiento de los operadores de cambio.

3. Línea de Investigación y Desarrollo

Esta línea de investigación considera como fuentes de información, nuestras publicaciones y aportes de investigadores realizados en la temática de operadores de revisión múltiple de la teoría de cambio de creencias bajo cláusulas Horn [9, 10, 11]. Principalmente, nuestro interés es estudiar las operaciones de cambio, en particular aquellos cambios de creencias que son inducidos a través de un conjunto (múltiple) de sentencias. Para ello, consideramos principalmente dos tipos diferentes de cambio múltiple. El primero, operadores de cambio múltiple priorizada en donde todas las nuevas creencias se supone que deben ser aceptadas, y el segundo, a operadores de mezcla (*merging*) que permite que creencias antiguas y nuevas jueguen roles simétricos dentro de un proceso de cambio. Para cada tipo de cambio, analizamos dos construcciones conocidas: una basada en kernels y otra basada en conjuntos de restos.

El término 'revisión múltiple' se utiliza para referirse a las operaciones de revisión que permite la revisión simultánea por más de una sentencia. Consideramos dos tipos de operadores directamente sin definirlos a partir de sendas operaciones de contracción. Definimos dos tipos de revisiones priorizadas, adaptadas de [4]:

- Revisión Kernel Múltiple.
- Revisión Partial Meet Múltiple.

La primera construcción de revisión múltiple por un conjunto de sentencias está basado en el concepto de un *A-inconsistent-kernels*. Aquí, se define una función de incisión que realiza un corte en cada *inconsistent-kernel*. Más precisamente, dicha función es una *función de incisión consolidada*, en el sentido de que todas las sentencias de A están protegidas y no pueden ser removidas por esta función. Esto es, *una función de incisión consolidada selecciona entre las sentencias de $K \setminus A$ que hace $K \cup A$ inconsistente'* [8].

La segunda construcción de revisión múltiple por un conjunto de sentencias está basado en el concepto de un *A-consistent-remainders*. En este caso, se define una función de selección que selecciona los 'mejores' *consistent-remainders*. La función es una *función de selección consolidada*, en el sentido de que todas las sentencias de $K \cap A$ están protegidas, y ellas están incluidas en la intersección de algún conjunto de restos. Luego, *'una función de selección consolidada selecciona un subconjunto del conjunto de $K \cup A$ cuyos elementos (todos) contienen el conjunto $K \cap A$ '* [8].

A diferencia de la preferencia absoluta para con la nueva información implícita en el modelo AGM, podemos cambiar nuestras creencias de una manera *no priorizada*. En los operadores más conocidos sobre cambio no priorizada en conjuntos de creencias están: *credibility limited revision*, *selective revision* y *screened revision*, cada uno con sus respectivas variantes. En lo que respecta a bases de creencias los operadores sobre cambio no priorizada podemos citar los ope-

radores de cambio basados en *explicaciones* [3] y en un tipo especial de cambio no priorizada en bases de creencias llamadas *merging* [4].

Por último, estudiamos la complejidad computacional de nuestros operadores de cambio múltiple de creencias Horn. Asumimos que el lector está familiarizado con las clases de complejidad y sus notaciones. Partimos de las siguientes presuposiciones: $P \subseteq NP$ y $P \subseteq coNP$. Un problema que está en P se dice que es *tratable*, mientras que un problema que está en NP , $NP-hard$ o $coNP-hard$ es *intratable* (en el sentido que cualquier algoritmo de resolución podría requerir una cantidad superpolinomial de tiempo en el peor caso). Hacemos uso de las clases de complejidad superior usando oráculos.

En particular $P^A(NP^A)$ correspondientes a las clases de problemas de decisión que son solucionados en tiempo polinomial por máquinas de Turing determinística (no determinística) usando un oráculo para A en tiempo polinomial [5]. Todos los problemas que analizamos reside en la *jerarquía polinomial*, introducida por Stockmeyer [6]. Las clases Σ_k^p , Π_k^p , y Δ_k^p de la jerarquía polinomial son definida por

$$\Sigma_k^p = \Pi_k^p = \Delta_k^p = P$$

y para $k \geq 0$,

$$\Sigma_{k+1}^p = NP^{\Sigma_k^p}, \Pi_{k+1}^p = co \Sigma_{k+1}^p, \Delta_{k+1}^p = P_{\Sigma_k^p}.$$

Se observa que:

$$\Delta_1^p = P, \Sigma_1^p = NP \text{ y } \Pi_1^p = coNP$$

Además, $\Sigma^p = NP^{NP}$, es la clase de problemas solucionables en un tiempo polinomial no determinística en una máquina de Turing que usa de forma libre un oráculo para NP . La complejidad de decisión $T * P \models Q$ (donde $*$ es un tipo de operador de revisión, T , P y Q son las entradas) fue estudiado por Eiter and Gottlob [7].

De allí se demuestra que, por ejemplo, en el enfoque de Dalal, el problema es de la clase $P^{NP} [O(\log n)]$. Esto significa que podemos decidir $P > Q$ con $O(\log n)$ a un NP oráculo, donde n representa el tamaño de T más P más Q . En los enfoques *basados en sintáxis* sus operadores permanecen en Π_2^P -hard (Π_2^P -complete para muchos de ellos) aún si el tamaño de P está limitado por una constante k . La complejidad de los enfoques *basados en modelos* decrece cuando el tamaño de P está limitado por una constante. La razón intuitiva para esto es que un cambio menor de P puede afectar solamente a pocos literales, ya que en este caso existen pocas alternativas para cambiar un modelo dado, desaparece la segunda fuente de complejidad. En el caso Horn, la complejidad de todos los enfoques (excepto el enfoque de Forbus) está en el primer nivel de la jerarquía polinomial, más precisamente, $co - NP - equivalente$. Intuitivamente, esta disminución de la complejidad se debe al hecho de que con cláusulas Horn el problema de inferencia resulta ser polinomial. Los enfoques basados en cardinalidad de Dalal y Forbus son los únicos cuya complejidad no disminuye bajo la restricción Horn. Esto se debe a que su medida basada en la cardinalidad de la proximidad entre modelos tiene un efecto perjudicial sobre las propiedades benignas de las cláusulas Horn [7].

4. Resultados y Objetivos Esperados

A través de esta línea de investigación se propone estudiar y definir nuevas construcciones de operadores de revisión

múltiples priorizadas y no priorizadas. Para cada nueva construcción se pretende establecer un conjunto de postulados que caractericen axiomáticamente a la operación de cambio y determinar relaciones entre las construcciones y los postulados. Permitir una eficiente implementación de estos nuevos algoritmos. Analizar a estos operadores de cambios múltiples bajo una lógica más restringida que la lógica proposicional clásica como es la *lógica Horn*. Estudiar la complejidad computacional que presentan estos operadores y determinar el costo desde un punto de vista computacional de la generación de una base revisada con estos operadores con cláusulas Horn y poder demostrar que los mismos son *representacionalmente factible*.

Por último, estos objetivos nos permitirán desarrollar nuevos procedimientos algorítmicos de cambios múltiples buscando no sólo alcanzar caracterizaciones axiomáticas sino también implementaciones computacionalmente tratables.

5. Formación de Recursos Humanos

El Docente-Investigador Néstor Jorge Valdez es dependiente del Programa de Desarrollo Científico y Tecnológico (adherido al Programa Nacional de Incentivos a los Docentes-Investigadores) de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca, es Magister en Ciencias de la Computación egresado de la Universidad Nacional del Sur, y aspira a alcanzar el título de Doctor en Ciencias de la Computación en esta última casa de estudios.

Referencias

- [1] Alchourrón, Gärdenfors, and Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50:510–530, 1985.
- [2] Richard Booth, Thomas Meyer, and Ivan Varzinczak. Next steps in propositional horn contraction. In Boutilier, C. (Ed.), *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 702–707, 2009.
- [3] Falappa, Kern-Isberner, and Simari. Belief revision, explanations and defeasible reasoning. *Artificial Intelligence Journal*, 141:1–28, 2002.
- [4] Falappa, Kern-Isberner, and Simari. Prioritized and non-prioritized multiple change on belief bases. *Journal of Philosophical Logic*, ISSN 0022-3611, New York: Springer, 41:77–113, 2012.
- [5] D. S. Johnson. A catalog of complexity classes, in *Handbook of Theoretical Computer Science*. J. van Leeuwen, Ed. A-Chap 2 Elsevier Science Amsterdam 1990.
- [6] L. J. Stockmeyer. The polynomial-time hierarchy. *Theoret. Comput. Sci.* 3, pages 1–22, 1976.
- [7] Eiter T. and Gottlob G. On the complexity of propositional knowledge base revision, updates and counterfactuals. *Artificial Intelligence* 57 (1992), pages 227–270, 1992.
- [8] Valdez and Falappa. Multiple revision on horn belief bases. *XXII Congreso Argentino de Ciencias de la Computación, XVII Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2016*.
- [9] Valdez and Falappa. Dinámica de conocimiento: Contracción múltiple en lenguajes horn. *XIX Congreso Argentino de Ciencias de la Computación, XIV Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2013*, pages–, 2013.
- [10] Valdez and Falappa. Dinámica de conocimiento: Contracciones horn a partir de ordenamientos epistémicos. *42 JAIIO Jornadas Argentinas de Informáticas, ASAI 2013*, 42 JAIIO'2013, pages 206–209, 2013.
- [11] Valdez and Falappa. Implementación para bases de creencias horn de operadores de contracción múltiple. *XX Congreso Argentino de Ciencias de la Computación, XV Workshop Agentes y Sistemas Inteligentes (WASI), CACiC'2014*, 2014.
- [12] Zhiqiang Zhuang. Belief change under the horn fragment of propositional logic. *PhD thesis, School of Computer Science and Engineering, Faculty of Engineering, University of New South Wales*, 2013.
- [13] Zhiqiang Zhuang, Maurice Pagnucco, and Yan Zhang. Definability of horn revision from horn contraction. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, Beijing, China*, 2013.

Ser Humano Virtual

Postay Juan Domingo, González Daniel

Unidad del Programa de Transferencia de Tecnologías Informáticas (PROTTI)
 Departamento de Sistemas - Área Tecnología Robótica
 Universidad John F. Kennedy, Argentina
 Bartolomé Mitre 1441, C1037ABA, CABA, Argentina.
 jpostay@kennedy.edu.ar - dgonzalez11@kennedy.edu.ar

RESUMEN

El trabajo Ser Humano Virtual (SHV) pretende mostrar que, mediante la conjunción adecuada de numerosos avances tecnológicos (Barrientos A, 1997, Bermejo S., 2003, Mataric M. J., 2007, Turkle, S., 2002) y el diseño del software de una serie de interfaces apropiadas para el ensamble de los mismos, es posible la construcción de un sistema robótico (Sabbatini, 1995, Sabbatini, 1997), provisto de inteligencia artificial (González, 2016), que pueda ser operado por cualquier individuo con capacidades diferentes en orden de permitirle virtualizar su humanidad en otros ámbitos reales distantes (VMFE, 2016), realizar tareas y actividades de socialización.

Palabras clave: Virtualización, Sistemas Tecnológicos, Inteligencia Artificial, Robótica, Discapacidades.

CONTEXTO

Ser Humano Virtual se inserta dentro del Programa de Transferencia de Tecnologías Informáticas (PROTTI), que es un programa promovido y subvencionado por la Universidad Kennedy para la transferencia de saberes emergentes en el área de la Informática y de Sistemas. A través de él se ponen a disposición de toda la comunidad los desarrollos realizados por alumnos y docentes de la Carrera de Sistemas de la Universidad. El objetivo del PROTTI es participar activamente en el desarrollo social y económico, brindando los conocimientos específicos generados en las investigaciones científicas del ámbito universitario, tanto a

particulares, empresas, como a distintos sectores del ámbito gubernamental.

1. INTRODUCCIÓN

Si bien al presente son muchos los adelantos tecnológicos y científicos que le permiten a un discapacitado tener una mejor calidad de vida, podríamos arriesgar a decir que no hay ninguna investigación anterior ni actual que marque un precedente en el tipo de hipótesis planteada, esto es, la implementación de un sistema que interrelacione integralmente los distintos avances tecnológicos existentes con el fin de que una persona con capacidades reducidas pueda sortear, de maneras eficientes, sus limitaciones en orden de socializarse como lo haría una persona normal.

Actualmente hay sistemas tecnológicos de virtualización que se emplean en distintos tipos de entornos. Por ejemplo, destacamos como novedoso los sistemas tecnológicos de virtualización a larga distancia en escenarios de conferencias comunicacionales, a través de los cuales un individuo puede ver, escuchar y hablar, mientras lo ven, lo escuchan y le hablan sus interlocutores. También podemos subrayar el uso de los sistemas tecnológicos mediante los cuales una persona opera acciones sobre una máquina, como ocurre en los laboratorios químicos en los que un individuo extiende su humanidad sobre los brazos de un robot para manipular elementos altamente radiactivos dentro de un recinto cerrado, o como se hace en algunas expediciones espaciales, al operar los mecanismos de movilidad de un robot para llevar a cabo determinadas tareas en el espacio exterior. Y en otro orden podemos

marcar, aún con más relevancia, que existen programas y dispositivos tecnológicos que le permiten a un discapacitado extremo introducirse como un avatar (representación gráfica, normalmente humana) en un metaverso (universo virtual) y socializarse, estableciendo vínculos con otros avatares que comparten su misma irrealidad.

En este contexto, y en base al conocimiento de todos esos desarrollos de las ciencias, es que nos propusimos desarrollar un sistema que permita a una persona con capacidades reducidas manejar un robot provisto de un agente inteligente, a través del cual logre, por ejemplo, realizar acciones e interactuar con otras personas en escenarios lejanos.

2. LINEAS DE INVESTIGACION Y DESARROLLO

A continuación, se muestra el esquema de construcción, o arquitectura, del sistema SHV, sus partes y las interfaces que se han construido para su ensamble.

En la Figura 1 encontramos, en el extremo izquierdo, al usuario humano con capacidades diferentes, el cual porta un dispositivo Emotiv Cerebro-Computador (C-C) que mediante la interfaz (1), representada en la gráfica por una flecha de ida y otra de vuelta (Emotiv, 2016), se comunica con una Pc. Dicha interfaz bidireccional está básicamente conformada por flujos de comandos de control e información entre la Pc y el dispositivo Emotiv que lleva puesto el usuario en su cabeza.

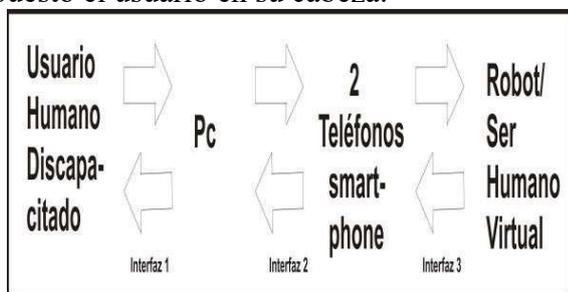


Figura 1. Esquema de interconexión del sistema Ser Humano Virtual

En el centro de la Figura 1 se aprecia otro juego de flechas que representa la interfaz (2) que vincula la Pc con dos teléfonos

Smartphone. Para el caso, esta interfaz (2) bidireccional lleva y trae comandos, además de señales de audio y video, entre la Pc y los teléfonos.

Finalmente, en el extremo derecho de la Figura 1 se visualiza la Interfaz (3), representada también por dos flechas, que permite el flujo bidireccional de comandos de control e información entre uno de los teléfonos Smartphone y la placa Arduino Uno (Arduino, 2016), controladora de un robot provisto de un Agente Inteligente, que bajo el nombre AASI (Aprendizaje Autónomo de Sistemas Inteligentes) fuera previamente desarrollado como parte de los proyectos PROTTI de nuestra misma Universidad (Gonzales, 2016). Detalles de la construcción del sistema:

- Interfaz (1).

La comunicación física entre el dispositivo C-C Emotiv y la Pc la realizamos vía Bluetooth- Ver Figura 2-



Figura 2. Interfaz Cerebro-Computador.

Mientras que el software utilizado para la comunicación lógica fue del propio fabricante de Emotiv. Mediante esta interfaz el dispositivo C-C recibe información de la Pc y envía instrucciones a la misma.

- Interfaz (2).

Para el diseño de la interfaz 2 lógica se analizaron distinto tipo de aplicaciones, algunas gratuitas y otras propietarias (trials). Muchas se instalaron, y se probaron sus performances en relación al uso destino.

En un comienzo se pensó que mediante un solo Smartphone del lado del robot AASI se podría llegar a manejar, desde la Pc, tanto la comunicación bidireccional de video y voz, como los comandos de control del robot (Romo, 2016). Luego de una selección previa,

que sirvió para descartar algunas aplicaciones que por especificaciones técnicas no contaban con todas las funcionalidades requeridas, se instalaron varias aplicaciones y, mediante el uso de una metodología de testeo (Gutiérrez J., 2005), se probaron con la idea de decidir el software a emplear en esta interfaz.

Mientras se trabajaba en esta tarea, como problemática se encontró que ninguna de ellas permitía dividir la pantalla del Smartphone en dos ventanas simultáneas; esto es, una para visualizar la comunicación de video y voz, y otra en la que se pudiera ver los comandos de la aplicación UKSHV que permite el control del robot AASI. Más tarde se concluyó que ésta era una limitación de las versiones actuales de Android, y que al presente sólo había un desarrollo de Samsung, el cual hacía algo parecido a lo que buscábamos, aunque la división de la pantalla era para aplicaciones propietarias del propio Samsung.

Este traspie nos llevó a la idea de instalar dos Smartphone, en vez de uno, en el robot.

A partir de esta decisión, casi todas las aplicaciones testeadas pasaban los casos de pruebas diseñados para operar las prestaciones de los Smartphone; entonces, por una cuestión de simplicidad en su instalación y manejo, se decidió que el Mobizen (Mobizen, 2016) era el más apropiado. En razón de lo expuesto, la interfaz (2) se dividió en dos partes, ambas implementadas mediante el doble uso de Mobizen -Ver Figura 3-.



Figura 3. Interfaz Pc-2 Smartphone.

El primer uso de Mobizen fue para establecer el enlace entre la Pc y el teléfono Smartphone 1 -ubicado en el robot-, dedicado a la comunicación bidireccional de video y voz. Dicha comunicación, en sí misma, se materializa a través de la aplicación Skype (Skype, 2016), previamente instalada en la Pc y el Smartphone 1; la que permite al usuario

escuchar en los parlantes y ver en la pantalla de su Pc aquello que acontece en la ubicación remota del Smartphone 1 -el audio es capturado mediante el micrófono y la imagen mediante la cámara frontal, del Smartphone 1-; además de permitirle ser visto en la pantalla y escuchado a través de los parlantes del Smartphone 1 -en los parlantes y la propia pantalla del Smartphone 1 se escucha y ve aquello captado por el micrófono y la cámara web ubicados en la Pc que opera el usuario-. Como se desprende de lo anterior, el teléfono Smartphone 1 mencionado lo utilizamos para que el usuario con capacidades diferentes sea visualizado en el escenario donde el robot se encuentra.

A través del mismo el usuario puede ver y escuchar, a la vez de ser visto y escuchado (Borenstein G., 2012, Igoe T., 2011, Somolinos Sanchez J. A., 2002). A posterior, pensamos cambiar la cámara integrada de este teléfono Smartphone del prototipo, ya que la misma presenta como limitación ser en dos dimensiones (2D).

- Interfaz (3).

Por un lado, la interfaz (3) física del enlace entre el teléfono Smartphone 2 y la placa Arduino Uno, que hace de cerebro operacional de la unidad robótica, se realizó mediante Bluetooth.

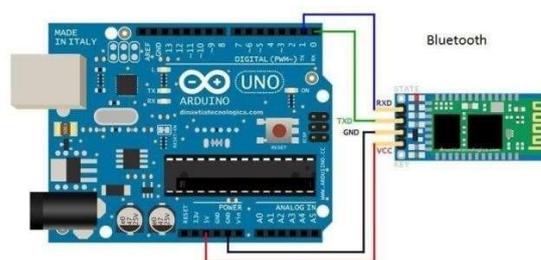


Figura 4. Plano de conexión de la placa Bluetooth HC-05 en la placa Arduino Uno.

Para establecerla, del lado del robot se instaló y configuró en su placa Arduino Uno una placa Bluetooth HC-05, cuyo plano de conexión se aprecia en la Figura 4.

En otro detalle, la instalación y puesta en funcionamiento de la placa HC-05 se puede apreciar en la Figura 5.

A su vez, para cerrar esa línea de comunicación física, del lado del Smartphone 2 se habilitó la conexión Bluetooth del mismo, y una vez detectada la señal emitida por el dispositivo HC-05, se configuró en conexión para poder transmitir bidireccionalmente con ella.

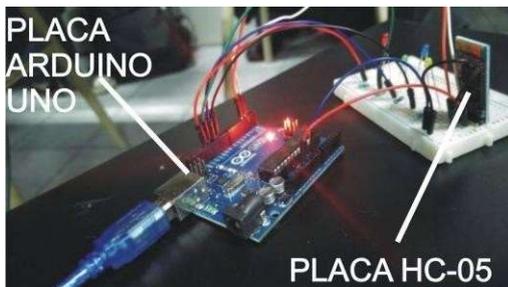


Figura 5. Placa Bluetooth HC-05 instalada en la placa Arduino Uno del robot AASI.

Por otro lado, para la conexión lógica, entre el teléfono Smartphone 2 y la placa Arduino Uno del robot, se desarrollaron dos aplicaciones. La primera, del lado de la placa Arduino Uno, una aplicación que llamamos SHV-UKino, y la segunda es la ya mencionada UKSHV, una app instalada del lado del Smartphone 2.

Para el desarrollo y programación de ambas aplicaciones se utilizó la metodología de prototipos (Fernández C., 2001), ya que esta permitía, a partir de una versión preliminar, ir avanzando con la incorporación de nuevas funcionalidades, a medida que las necesidades iban surgiendo.

La aplicación SHV-UKino se concretó mediante el software de desarrollo de Arduino, Por otro lado, para realizar el diseño de la aplicación UKSHV para el Smartphone 2, se utilizó el lenguaje de programación IDE (en español: Entorno de Desarrollo Interactivo) oficial de Androide, llamado Android Studio (Android, 2016); un lenguaje que permite el desarrollo de apps para Smartphone que cuenten con plataforma Android como sistema operativo. El programa resultante permite accionar los motores de tracción del robot AASI, como ser, para girarlo pausadamente a izquierda o a derecha, o para hacerlo levemente retroceder o avanzar, también para rotarlo 90 y 180 grados a

izquierda o a derecha, o hacerlo avanzar o retroceder a máxima velocidad, etc.



Figura 6. App UKSHV instalada en el Smartphone 2.

En la Figura 6 se aprecia la aplicación UKSHV ya instalada en el Smartphone 2.

3. RESULTADOS OBTENIDOS/ESPERADOS

A lo largo de todo el proceso de investigación se ha conseguido avanzar hasta llegar a la puesta en marcha del prototipo.

	Prototipo Actual	Versión Final
Comunicación	Wireless	4g
sensores adicionales	no	si
Sistema de movilización	ruedas	oruga
Cámara Kinect	no	si
apto exteriores	no	si
protocolo	http	https
Alta Disponibilidad	no	si

Tabla 1. Prototipo vs. Versión Final

Este hecho es el que nos ha permitido corroborar la hipótesis de trabajo que se expresara al inicio del proyecto. En cuanto los logros muestran con claridad que, mediante la

combinación de la tecnología presente, los avances de las comunicaciones y la implementación de las interfaces de software adecuadas se puede generar un sistema tecnológico robótico para que sea operado por un individuo con capacidades diferentes, en orden que éste pueda realizar distinto tipo de tareas y sociabilizarse.

En el cuadro de la Tabla 1 se muestran las dos instancias de evolución, que lleva desde el presente a la versión final. En ella se aprecian cuáles son los elementos del prototipo actual, que se proyectan cambiar en la versión final. Estos progresos se constituyen en motivo de un próximo avance en nuestra investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

En otro orden, independientemente de los logros en la propia investigación, debemos resaltar que SHV se ha convertido en una fuente de estímulo e inspiración para los alumnos que cursan las distintas materias que componen nuestra Carrera de Sistemas; pues, preocupados por la problemática en la que trabajamos, espontáneamente se han sumado a querer realizar aportes en distintas instancias y órdenes del proyecto.

Formaron parte del equipo estable del trabajo de Investigación SHV, el Lic. Postay Juan, como Director, el Ing. Gonzales Daniel, como Co-Director, y cinco alumnos de la Carrera de Sistemas, de los cuales uno de ellos está próximo a defender su trabajo de tesis de grado basando el mismo en una mejora de este desarrollo.

5. BIBLIOGRAFIA

- Android. (2016, 30 de agosto). Anrdroid Sudio: <https://developer.android.com/studio/index.html?hl=es-419>
- Arduino. (2016, 30 de agosto). Arduino Genuino.: <https://arduino.cc>
- Barrientos A, y otros (1997); Fundamentos de Robótica. Madrid, España: McGraw-Hill.
- Bermejo S., (2003) Desarrollo de robots basados en el comportamiento. Barcelona, España: Universidad Politécnica de Cataluña.
- Borenstein G. (2012). Making Things See, Sebastopol: Maker Media Inc, [Versión electrónica]: <https://play.google.com/store/apps/details?id=com.teamviewer.quicksupport.market>
- Emotiv. (2016, 30 de agosto). Desarrollos y Modelos: <https://www.emotiv.com>
- Fernández, C. (2001). Modelo de Prototipo. Ecuador: Universidad Estatal de Milagro. [Versión electrónica]: <http://es.slideshare.net/yanezcabrera/modelo-de-prototipo>
- González D., Postay J. Aprendizaje Autónomo en Sistemas Inteligentes, Universidad Kennedy (2016, 30 de agosto) <https://kennedy.edu.ar/investigacion/#1444754218661-948dda25-259f>
- Gutiérrez J., Escalona M. J., Mejías M., et al. (2005). “Analysis of Proposals to Generate System Test Cases from System Requirements”. Portugal, Porto: CAiSE’05 Forum.
- Igoe T. (2011). Making Things Talk, Sebastopol, O’Reilly Media Inc, [Versión electrónica]: https://proalias.com/books/-Making_Things_Talk_Second_Edition.pdf
- Mataric M. J. (2007). The Robotics Primer. EEUU: Massachusetts Institute of Technology.
- Mobizen. (2016, 30 de agosto). Expande tu Smart Life: www.mobizen.com/?locale=es
- Romo. The Smartphone Robot. (2016, 30 de agosto): <https://kickstarter.com/projects/peterseid/romo-the-smartphone-robot>
- Sabbatini, R.M.E. (1995). Um robô no seu futuro? Lisboa, Portugal: Correio Popular, Caderno de Informática.
- Sabbatini, R.M.E. (1997). Robô não é gente. Lisboa, Portugal: Correio Popular, Caderno de Informática.
- Skype. (2016, 30 de agosto). Skype: <https://skype.com/es/>
- Somolinos Sanchez J. A. (2002). Avances en Robótica y Visión por Computador. Murcia, España: Colección Ciencia y Técnica N°38.
- Turtle, S. (2005). The Second Self: Computers and the Human Spirit. Cambridge, EEUU: The MIT Press.
- VMFE -Virtual Machine for Everyone-. (2016, 30 de agosto): <http://vmlite.com>

Técnicas de Inteligencia Artificial Aplicadas a Problemas Científico-Tecnológicos

Nilda M. Pérez Otero, Javier Izetta Riera, Abigail Verazay, Virginia Battezzati, Juan Salinas, Susana G. Pérez Ibarra, Alejandra Arias

GIDIA / Facultad de Ingeniería / Universidad Nacional de Jujuy

Ítalo Palanca 10, +54 (388) 4221587

{nilperez, javierizetta, abigailrn, virginiavir, juansalinas90, gperezunju, alejandraarias8099 }@gmail.com

RESUMEN

Cada vez es más intensivo el uso de técnicas de inteligencia artificial en la mayoría de las actividades humanas y en particular, en el área científico-tecnológico. Algunas de las áreas, que muestran tanto la diversidad de campos a los que pueden aplicarse como la utilidad que pueden alcanzar, son: Bioinformática, Robótica, Tecnología de los alimentos y Recuperación de información.

Una de las tareas más requeridas es la generación automática de modelos que resuman las características que debe cumplir un elemento para pertenecer a una determinada categoría, tarea que se conoce como clasificación automática. Los métodos computacionales desarrollados para tal fin forman parte de lo que se conoce como Aprendizaje Automatizado (*Machine Learning*). Otra tarea importante es la optimización de funciones, en las últimas décadas diferentes tipos de algoritmos de optimización han sido desarrollados para resolver una gran cantidad de problemas. El principal desafío radica en problemas que presentan una función objetivo altamente no lineal y no convexa, esto dificulta garantizar la localización del mínimo global. Por lo tanto, la necesidad de encontrar nuevas técnicas que proporcionen un mejor desempeño en este tipo de problemas sigue aún vigente y hacen de esto un campo excitante para trabajar.

Palabras clave: Metaheurísticas,

Aprendizaje Automatizado, Inteligencia Artificial.

CONTEXTO

La línea de investigación aquí presentada se encuentra inserta en el proyecto *Técnicas de Inteligencia Artificial aplicadas a problemas científico-tecnológicos*, ejecutado a partir de 2016 por el Grupo de Investigación y Desarrollo en Informática Aplicada (GIDIA) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy.

El proyecto se encuentra acreditado y financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy (Cód. D/0150) y se encuentra bajo el Programa de Incentivos.

1. INTRODUCCIÓN

La Inteligencia Artificial (AI) propone nuevas herramientas computacionales que brindan solución a muchos problemas complejos del mundo real. Estas herramientas se utilizan en una variedad de aplicaciones que van desde el modelado, la clasificación y el reconocimiento de patrones hasta el análisis de datos. Las técnicas de inteligencia artificial se vienen utilizando en la ciencia y la tecnología de los alimentos para la clasificación, el modelado y la optimización de procesos, el control de calidad de los alimentos, la

predicción de propiedades reológicas de masa, la clasificación de vinos en función del contenido de antocianinas, entre otros [1]. Asimismo en el área de la ingeniería industrial se comenzaron a incorporar herramientas de inteligencia artificial, a tal punto que, en el año 2015, se realizó la Conferencia Internacional de Inteligencia Artificial e Ingeniería Industrial (AEII2015) [2].

Particularmente, una rama de la inteligencia artificial que en los últimos años ha tenido gran crecimiento es la del Aprendizaje Automatizado (*Machine Learning*), la cual propone nuevas herramientas computacionales que sirven de apoyo para tareas de investigación en distintas áreas.

Aprendizaje Automatizado

En las últimas décadas el incremento de métodos automáticos de medición y almacenamiento de datos produjo una revolución en las ciencias, donde la cantidad de nuevas observaciones disponibles supera ampliamente la capacidad actual de modelado. Esta situación originó la necesidad de desarrollar programas de computación que, a partir de un conjunto de datos, generen un modelo que produzca salidas aproximadas a las del proceso que está implícito en los datos observados. El aprendizaje automatizado intenta comprender los mecanismos por los cuales se adquiere el conocimiento a partir de los datos/experiencia y con ello lograr automatizar la etapa de modelado [3]. El conjunto de datos de entrada de un programa de aprendizaje automatizado consiste en descripciones de objetos de un universo acompañado de un valor de salida asociado. Cada objeto es considerado un ejemplo del universo que se desea conceptualizar. Generalmente estos ejemplos son tuplas de la forma (atributos; salida). El proceso de aprendizaje consiste entonces en la búsqueda, dentro de un espacio H , de

todas las hipótesis, la hipótesis f que mejor aproxime la función original F que dio origen al conjunto de datos de entrada.

Clasificación

Los problemas de clasificación son una parte del aprendizaje automatizado [4] en los que el objetivo es asignar a los datos $D = \mathbf{x}_1; \dots; \mathbf{x}_n$ un número de valores discretos llamados clases o categorías. Cada objeto tiene asignada una clase o categoría l \mathbf{x}_i conocida. Para un problema con c clases, l \mathbf{x}_i puede tomar c valores discretos distintos. El objetivo entonces es encontrar una función clasificadora f tal que para cada objeto \mathbf{x}_i sea $f \mathbf{x}_i = l \mathbf{x}_i$, donde $f \mathbf{x}_i$ es la clase o categoría asignada por f a un objeto.

Se han introducido muchos métodos de clasificación en los últimos años, algunos de éstos son: Redes Neuronales Artificiales [5], Máquinas de Vectores Soporte [6] y Random Forest [7] que se aplicaron a una amplia variedad de problemas científicos tecnológicos [8, 9, 10, 11, 12] demostrando buenos resultados. Esto lleva a proponer la utilización de estos métodos para resolver diferentes problemas en áreas como Bioinformática, Robótica, Recuperación de información y Tecnología de los alimentos, entre otros.

Metaheurísticas

Las metaheurísticas, en su definición original, son métodos de resolución que orquestan una interacción entre los procesos de mejora local y estrategias de mayor nivel para crear un proceso capaz de escapar de óptimos locales realizando una búsqueda robusta en el espacio de soluciones. Con el tiempo, estos métodos han llegado a incluir cualquier procedimiento que emplee estrategias para superar la trampa de la optimalidad local en espacios de soluciones complejos, especialmente aquellos procedimientos que utilizan una o más

estructuras locales como un medio para definir movimientos admisibles para la transición de una solución a otra, o para construir o destruir soluciones en procesos constructivos y destructivos [13].

Varias herramientas y mecanismos surgidos a partir de la creación de los métodos metaheurísticos demostraron ser eficaces, tanto es así que las metaheurísticas se convirtieron en el centro de atención en los últimos años como el enfoque preferido para resolver problemas de optimización complejos, en particular problemas de naturaleza combinatoria. Si bien, las metaheurísticas no son capaces de garantizar la optimalidad de las soluciones que encuentran, los procedimientos exactos o métodos de convergencia local (que en teoría podrían brindar tal garantía si se les permite correr el tiempo suficiente) a menudo son incapaces de encontrar soluciones cuya calidad sea similar a la obtenida por las principales metaheurísticas, en particular, en la resolución de problemas del mundo real donde se evidencia su eficiencia y eficacia para resolver problemas grandes y complejos. La aplicación de metaheurísticas comprende un gran número de áreas y campos disciplinarios, siendo algunos de ellos:

- Diseño de ingeniería, optimización de topologías y la optimización estructural en electrónica y VLSI, aerodinámica, dinámica de fluidos, telecomunicaciones y robótica.
- *Machine learning* y minería de datos en bioinformática y biología computacional, y finanzas.
- Modelado de sistemas, simulación e identificación en química, física y biología; control, señal, y procesamiento de imágenes.
- Planificación de problemas de enrutamiento, problemas de planificación, programación y

producción de robots, logística y transporte, gestión de la cadena de suministro, y otros.

En el diseño de una metaheurística se deben tener en cuenta dos criterios contradictorios: la exploración del espacio de búsqueda (diversificación) y la explotación de las mejores soluciones encontradas (intensificación) [14]. Las regiones más promisorias están determinadas por las “buenas” soluciones encontradas. En la intensificación, estas regiones se exploran más a fondo con la esperanza de encontrar mejores soluciones. En la diversificación, las regiones no exploradas deben ser visitadas para asegurarse de que todas las regiones del espacio de búsqueda se exploran de manera uniforme y que la búsqueda no se limita a sólo un número reducido de regiones.

Las metaheurísticas pueden clasificarse desde varios puntos de vista, [14] presenta las siguientes alternativas de clasificación:

- Inspiradas o no en la naturaleza: varias metaheurísticas se inspiran en procesos naturales, por ejemplo de la biología o de la física.
- Con o sin uso de memoria: algunos algoritmos metaheurísticos no emplean elementos de memoria, es decir, no se utiliza información extraída de forma dinámica durante la búsqueda mientras que otros usan una memoria que contiene parte de la información extraída en línea durante la búsqueda.
- Deterministas o estocásticas: una metaheurística determinista resuelve un problema de optimización mediante la toma de decisiones deterministas. Las metaheurísticas estocásticas aplican algunas reglas aleatorias durante la búsqueda.
- Búsqueda poblacional o búsqueda de única solución: los algoritmos basados en una sola solución manipulan y transforman una única solución

durante la búsqueda, mientras que en los algoritmos poblacionales es toda una población de soluciones la que evoluciona.

- Iterativas o avaras: en los algoritmos iterativos, se parte de una solución completa (o población de soluciones) y se transforman en cada iteración utilizando algunos operadores de búsqueda. Los algoritmos avaros, parten de una solución vacía, y en cada paso se asigna una variable de decisión del problema hasta que se obtiene una solución completa.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta línea de investigación consiste en analizar técnicas metaheurísticas y de aprendizaje automatizado y aplicarlas al desarrollo de algoritmos robustos para problemas de optimización global y clasificación, involucrados en tareas científicas tecnológicas de distintas áreas.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto, que se desarrolla en el bienio 2016-2017 tiene como objetivo general desarrollar algoritmos que implementen técnicas de inteligencia artificial, una combinación de las ya existentes u otras técnicas para resolver problemas científico-tecnológicos de interés. Actualmente se lograron los siguientes resultados:

- Clasificación Automática de Textos Periodísticos: en los últimos años el periodismo regional, al igual que en todas partes del mundo, pasó de su formato clásico de publicación al electrónico. Así las webs de noticias regionales también se ven obligadas a evolucionar y mejorar sus prestaciones a través de una mejor organización y categorización previa de toda la información disponible para el lector.

En este proyecto se propuso abordar la clasificación automática de textos periodísticos digitales a través del Aprendizaje Automatizado. Se desarrollaron cuatro clasificadores automáticos de textos periodísticos extraídos de páginas webs de noticias del NOA dos de ellos basados en *Random Forest* y los dos restantes basados en *Support Vector Machine*. Además se propusieron dos técnicas nuevas para la reducción de dimensionalidad del espacio de características. Estos clasificadores fueron evaluados con distintas colecciones de noticias, pudiéndose comprobar la robustez de los mismos al mantener un buen desempeño.

- Cálculos de equilibrio de fase: estos cálculos juegan un rol crucial en la simulación, diseño y optimización de procesos de separación. La dificultad de éstos radica en que la forma de la función objetivo, altamente no lineal y no convexa, no permite garantizar la localización del mínimo global. AEvol es un algoritmo evolutivo, con un operador genético sencillo, para la optimización de funciones continuas presentado por nuestro grupo de investigación en el año 2012. Una vez validado con funciones *benchmarks*, este algoritmo se aplicó en la minimización de la función de Gibbs, función termodinámica relacionada al equilibrio de fases. En los años siguientes se desarrollaron dos nuevas versiones de AEvol que agregaron a este algoritmo operadores de cruce con el fin de mejorar su desempeño. Estas nuevas versiones fueron validadas sobre funciones *benchmarks* logrando mejores resultados que AEvol. Esta mejora está relacionada a la ampliación de la búsqueda del espacio de soluciones, mejorando la tasa de éxitos al disminuir la probabilidad de caer en mínimos locales. Actualmente se pretende realizar un estudio comparativo del desempeño de AEvol

y sus distintas versiones utilizando tres sistemas que de acuerdo a la literatura son los más adecuados para la evaluación de nuevos algoritmos de optimización, debido a que presentan mínimos locales en el espacio de soluciones, situación que los hace desafiantes para el cómputo del equilibrio de fases.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes-investigadores, egresados y alumnos de la Facultad de Ingeniería de la Universidad Nacional de Jujuy. El equipo actual cuenta con 6 ingenieros en informática y 1 alumno. En el marco de este proyecto se finalizó una tesina de grado y los integrantes del equipo dirigen actualmente 3 tesinas más. Además se prevé la finalización de una tesis de doctorado y realización de una tesis de maestría.

5. REFERENCIAS

- [1] Lee, C. J. (Editor). International Conference on Artificial Intelligence and Industrial Engineering 2015 (AIIE 2015). Proceedings of a meeting held 26-27 July 2015, Phuket, Thailand. ISBN 9781510806450. Atlantis Press.
- [2] Sofu, B.A., Demir, N., Ekinçi, F.Y., 2007. *Gıda Bilimi Ve Teknolojisi Alanında Yapay Zeka Uygulamaları (Applications of Artificial Intelligence in Food Science and Technology Area)*. GIDA (2007) 32 (2): 93-99.
- [3] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- [4] Mitchell, T. *Machine Learning*. McGraw-Hill series in computer science. McGraw-Hill, 1997.
- [5] Hassoun, M. H. *Fundamentals of artificial neural networks*. MIT press, 1995.
- [6] Vapnik, V. *Statistical learning theory*. 1998, 1998.
- [7] Breiman, L. Random forests. *Machine learning*, 45 (1), 5-32, 2001.
- [8] Pavlidis, P., Wapinski, I., & Noble, W. S. (2004). Support vector machine classification on the web. *Bioinformatics*, 20(4), 586-587.
- [9] Byvatov, E., Fechner, U., Sadowski, J., & Schneider, G. (2003). Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *Journal of Chemical Information and Computer Sciences*, 43(6), 1882-1889.
- [10] Gardner, G. G., Keating, D., Williamson, T. H., & Elliott, A. T. (1996). Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *British journal of Ophthalmology*, 80(11), 940-944.
- [11] Lempitsky, V., Verhoek, M., Noble, J. A., & Blake, A. (2009). Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography. In *Functional Imaging and Modeling of the Heart* (pp. 447-456). Springer Berlin Heidelberg.
- [12] Miller, K., Huettmann, F., Norcross, B., & Lorenz, M. (2014). Multivariate random forest models of estuarine-associated fish and invertebrate communities. *Mar Ecol Prog Ser*, 500, 159-174.
- [13] Glover, F. W. & Kochenberger, G. A., editores (2003). *Handbook of Metaheuristics*, volume 114 of International Series in Operations Research & Management Science. Springer.
- [14] Talbi, E.-G. (2009). *Metaheuristics: From Design to Implementation*. Wiley Publishing.

Toma de Decisiones Individuales y Colectivas para Sistemas Multi-agente en Entornos Distribuidos

Martín E. Buron Brarda¹

Luciano H. Tamargo²

Alejandro J. García²

¹Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

²Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET),

Departamento de Ciencias e Ingeniería de la Computación,

Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina.

Tel: (0291) 459-5135 / Fax: (0291) 459-5136

e-mail: {martin.buron, lt, ajg}@cs.uns.edu.ar

RESUMEN

Esta línea de investigación se enfoca en mejorar las capacidades para la toma de decisiones individuales y colectivas de agentes en sistemas multi-agente. Dentro de este enfoque, se planea estudiar y desarrollar como mejorar en los agentes los siguientes aspectos: la capacidad de representación de conocimiento individual y colectivo, la capacidad de realizar inferencias, la capacidad de interacción e intercambio de información, y la capacidad de integrar esos elementos para tomar decisiones tanto individuales como colectivas. El aporte de esta investigación está orientado al desarrollo de formalismos y mecanismos para la toma de decisiones, por parte de agentes inteligentes deliberativos, en el contexto de un sistema multi-agente.

Palabras clave: Toma de Decisiones, Sistemas Multi-Agente, Argumentación

CONTEXTO

Esta línea de investigación se realizará dentro del ámbito del Instituto de Ciencias e

Ingeniería de la Computación (ICIC) de doble dependencia entre CONICET y la Universidad Nacional del Sur, y dentro del Departamento de Ciencias e Ingeniería de la Computación de dicha universidad.

El primer autor posee una beca para realizar un doctorado en Ciencias de la Computación financiada por la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Además, esta investigación está dentro del marco de los proyectos de investigación: “Formalismos para el tratamiento de confianza y reputación en sistemas multi-agente” (PGI 24/ZN32), y “Argumentación y dinámica de creencias para mejorar las capacidades de razonamiento y representación de conocimiento en sistemas multi-agente” (PGI 24/N035), ambos financiados por la Universidad Nacional del Sur.

1. INTRODUCCIÓN

Los agentes inteligentes ofrecen beneficios como: modularidad, velocidad, confiabilidad, facilidad de mantenimiento, reusabilidad y portabilidad. Por este motivo, se han convertido en un tema de gran interés para las

Ciencias de la Computación. Por eso mismo, el desarrollo de sistemas multi-agente [ABM05, DRM05, W09, Tam13], también ha ganado posición en área. En este tipo de sistemas, varios agentes interactúan para conseguir algún objetivo o realizar alguna tarea en común. En dicho escenario cada agente tiene información incompleta y capacidades limitadas, el control del sistema es distribuido, los datos están descentralizados, y la computación es asincrónica. Debido a las características enunciadas antes, los sistemas multi-agente constituyen un área en continuo crecimiento para el desarrollo de aplicaciones comerciales e industriales de gran escala. Esto se debe principalmente a que proveen de manera más natural soluciones a problemas complejos.

Uno de los aspectos centrales en el desarrollo de agentes de software y agentes físicos, es la capacidad de estos agentes para la toma de decisiones [W09]. Este aspecto influye directamente en el comportamiento del agente y de los agentes con los cuales interactúa al resolver problemas en forma conjunta. El tema de toma de decisiones ha sido abordado desde diferentes puntos de vista pero sigue siendo motivo de estudio con aspectos aún no desarrollados.

Uno de los aportes de esta investigación es mejorar la capacidad de los agentes para tomar decisiones, y de esta manera, poder resolver problemas de manera conjunta. Nuestra propuesta considera un conjunto de agentes deliberativos autónomos, que intercambian información y pueden aprovechar las capacidades individuales de cada miembro de la comunidad. Al intentar resolver un problema de manera conjunta, cada uno de los agentes puede brindar su aporte individual de varias maneras: aportar información relevante que solamente este agente dispone, dar razones a

favor o en contra de una decisión, o también aportando un plan o modificando sus planes a fin de lograr alguna meta que colabore en la decisión que se quiere tomar.

Pensar en resolver el problema de toma de decisiones de manera grupal, simplemente uniendo en un mismo repositorio las bases de conocimiento, los planes, las metas y las preferencias de todos los agentes participantes, es impracticable. Existen varias razones que impiden realizarlo de esa forma: una de las ventajas del modelo de agentes es que los agentes son entidades autónomas con sus propias metas y deben poder mantener ciertos niveles de privacidad sobre su conocimiento, metas y planes. Además, los agentes que tienen grandes volúmenes de información deberían poder aportar solamente lo que se considere relevante a la decisión a tomar, y por otro lado, los agentes no solo pueden colaborar con información, sino también con planes o realizando acciones que permitan resolver conflictos ante la presencia de elementos contradictorios.

Hoy en día la argumentación constituye un área de estudio de especial interés en el ámbito de toma de decisiones y sistemas multi-agente (ver por ejemplo [FECS14, GCRS13, GDS09, RS09]), principalmente, porque permite razonar con información incompleta e incierta, y manejar inconsistencias en los sistemas basados en conocimiento. Este tipo de razonamiento es particularmente atractivo para toma de decisiones, y dentro de la Inteligencia Artificial existe particular interés en abordar este tipo de problemas [FECS14]. La argumentación ha evolucionado como un mecanismo atractivo para formalizar el razonamiento de sentido común [PV02, FGKS11, CGS16]. En la literatura se evidencia un gran desarrollo tanto de la formalización de diferentes marcos de argumentación abstracta

[ABM05, AK07, NBD08], como así también de sistemas de argumentación estructurados (o basados en reglas) [GS14, MP14, Toni14].

En esta investigación se buscará extender los métodos de razonamiento automático basados en argumentación, para que puedan ser aplicados para la toma de decisiones individuales y colectivas de un conjunto de agentes. Y además se buscará desarrollar extensiones de alguna implementación de sistemas de argumentación estructurada con el fin de disponer de herramientas de especificación e implementación para entornos distribuidos donde puedan coexistir múltiples agentes en paralelo y compartir parte de su conocimiento. La programación lógica rebatible (DeLP por sus siglas en inglés) es un ejemplo de sistema de argumentación estructurado en el cual se puede representar conocimiento en un programa lógicorebatible, en forma de hechos, reglas estrictas y reglas rebatibles [GS04, GDS09, GS14]. Los objetivos particulares de esta investigación se orientarán a estos temas.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación toma como punto de partida varias publicaciones vinculadas a la misma. Por ejemplo, en [W09] se establecen bases para el estudio de sistemas multi-agente, y se enfatiza la importancia de contar con mecanismos formales de comunicación entre agentes. Este concepto también es respaldado por [GGS09, DRM05] que también consideran que todo lenguaje de especificación o implementación de agentes debe considerar primitivas para la interacción.

En [GS04] se presenta el sistema de argumentación estructurada DeLP, y en

[GS14] se introducen los DeLP-Servers que permiten utilizar la argumentación en ambientes distribuidos lo cual es imprescindible en aplicaciones de SMA.

En esta investigación se aplicarán los resultados obtenidos al desarrollo de agentes (tanto de software como robots físicos) que se desenvuelven en ambientes distribuidos y dinámicos, y pueden ejecutarse en paralelo. En particular, esta investigación buscará mejorar la capacidad de los agentes para razonar cuando se pretende tomar decisiones al resolver problemas de manera conjunta. Además, se intentará mejorar la capacidad de administrar de la mejor manera posible la información que los agentes reciben de sus pares.

La importancia de esta investigación radica en que los sistemas multi-agentes se han convertido en una herramienta natural para aplicaciones distribuidas en entornos dinámicos, y tienen aplicación directa en áreas como robótica cognitiva, comercio electrónico, y asistentes para toma de decisiones. El desarrollo de nueva tecnología en agentes inteligentes y sistemas multi-agente permitirá, además, lograr nuevos avances en áreas fundacionales de las Ciencias de la Computación como sistemas operativos distribuidos, bases de datos distribuidas, y lenguajes de programación en paralelo.

3. RESULTADOS ESPERADOS

En esta línea de investigación se espera poder concretar los siguientes objetivos específicos:

- Estudiar, desarrollar y formalizar métodos de toma de decisiones individuales y grupales en ambientes de múltiples agentes, considerando tanto escenarios de agentes

colaborativos como de agentes competitivos.

- Estudiar, y extender los métodos de razonamiento automático basados en argumentación, para que puedan ser aplicados para la toma de decisiones individuales y colectivas de un conjunto de agentes.
- Desarrollar y formalizar métodos para integrar a los sistemas de argumentación con la posibilidad de tener en cuenta información recibida de múltiples agentes en entornos distribuidos y dinámicos.
- Desarrollar extensiones de alguna implementación para sistemas de argumentación estructurada (como DeLP [GS14], ASPIC [MP14], o ABA [Toni14]) con el fin de disponer de herramientas de representación de conocimiento y razonamiento para la especificación e implementación de agentes que puedan coexistir en paralelo y compartir parte de su conocimiento.

4. FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación forma parte de los trabajos que se realizarán para la concreción de un Doctorado en Ciencias de la Computación por parte del primero de los autores. Los demás autores son quienes dirigirán el proyecto ya que han trabajado en temas relacionados con esta línea de investigación como se puede ver en los artículos [TGKS08, GGS09, Tam13, GCRS13, GS14, FECS14, TGGS16].

5. BIBLIOGRAFIA

[AK07] L. Amgoud, S. Kaci, An argumentation framework for merging conflicting knowledge bases, *Int. J. Approx. Reasoning* 45(2): 321-340, 2007.

[ABM05] K. Atkinson, T. J. M. Bench-Capon, P. McBurney, Multi-agent argumentation for e-democracy., in: *Proceedings of the Third European Workshop on Multi-Agent Systems*, Brussels, Belgium, Koninklijke Vlaamse Academie, 2005, pp. 35-46.

[CGS16] A. Cohen, A. J. García, G. R. Simari. A Structured Argumentation System with Backing and Undercutting. *International Scientific Journal Engineering Applications of Artificial Intelligence* v. 49 pp. 149-166. Elsevier science BV. March (2016) DOI: <http://dx.doi.org/10.1016/j.engappai.2015.10.001>

[DRM05] Dastani, M.; van Riemsdijk, M. B.; and Meyer, J.-J. C. 2005. Programming multi-agent systems in 3apl. In *Multi-Agent Programming*. 39–67.

[FGKS11] M. Falappa, A. García, G. Kern-Isberner, G. Simari. On the evolving relation between Belief Revision and Argumentation. *The Knowledge Engineering Review Journal* v26:1, pp. 35-43, Cambridge University Press, 2011.

[FECS14] E. Ferretti, M. Errecalde, A. J. García, G. R. Simari. A Possibilistic Defeasible Logic Programming Approach to Argumentation-Based Decision Making. *Journal of Experimental & Theoretical Artificial Intelligence* Vol. 26, Issue 4 pp 519-550. Taylor & Francis. (2014) DOI: 10.1080/0952813X.2014.921733.

- [GS04] A. J. García, G. R. Simari, Defeasible logic programming: An argumentative approach, *Journal of Theory and Practice of Logic Programming*, 4 (1), 2004, pp. 95-138.
- [GDS09] A. J. García, J. Dix, G. R. Simari. Chapter 8: Argument-based Logic Programming. Book: *Argumentation in Artificial Intelligence* Springer Science+Business Media, (2009), DOI 10.1007/978-0-387-98197-0_8.
- [GCRS13] A. J. García, C.I. Chesñevar, N. D. Rotstein, G. R. Simari. Formalizing dialectical explanation support for argument-based reasoning in knowledge-based systems. *Expert Systems with Applications* v40 issue 8 pp. 3233–3247, ISSN 0957-4174. Elsevier. (2013) DOI:10.1016/j.eswa.2012.12.036.
- [GS14] A. J. García, G. R. Simari. Defeasible Logic Programming: DeLP-servers, Contextual Queries, and Explanations for Answers. *Argument & Computation* v5. Taylor & Francis. (2014) DOI:10.1080/19462166.2013.869767
- [GGS09] S. Gottifredi, A. J. García, G. R. Simari, *Argumentation Systems and Agent Programming Languages*, In *AAAI Fall Symposium: The Uses of Computational Argument*, Washington D.C., USA, 2009.
- [MP14] S. Modgil & H. Prakken. The ASPIC+ framework for structured argumentation: a tutorial. *Argument & Computation* v5. Taylor & Francis. (2014) DOI:10.1080/19462166.2013.869766
- [NBD08] F. S. Nawwab, T. J. M. Bench-Capon, P. E. Dunne, *A Methodology for Action-Selection using Value-Based Argumentation*, *COMMA* 2008: 264-275.
- [PV02] H. Prakken, G. Vreeswijk, *Logical Systems for Defeasible Argumentation*, in: D. Gabbay, F. Guenther (eds.), *Handbook of Philosophical Logic*, Kluwer Academic Publishers, 2002, pp. 219-318.
- [RS09] I. Rahwan, G. Simari (eds.), "Argumentation in Artificial Intelligence", Springer-Verlag, 2009, ISBN 978-0-387-98196-3.
- [Tam13] L. H. Tamargo, M. Thimm, P. Krümpelmann, A. J. García, M. A. Falappa, G. Kern-Isberner, G. R. Simari. Chapter 7: Credibility-based selective revision by deductive argumentation in multi-agent systems. pp: 155-182. Book: *Trends in Belief Revision and Argumentation Dynamics*. E. Ferme G. D. Gabbay and G. Simari (Eds) ISBN: 978-1-84890-065-3. College Publications. 2013.
- [TGGS16] L. H. Tamargo, S. Gottifredi, A. J. García, G. R. Simari. Sharing beliefs among agents with different degrees of credibility. *Knowledge and Information Systems*. Springer. 2016. DOI: 10.1007/s10115-016-0964-6
- [TGKS08] M. Thimm, A. J. Garcia, G. Kern-Isberner, G. R. Simari. Using Collaborations for Distributed Argumentation with Defeasible Logic Programming. *Proc. Of the Twelfth Int. Workshop on Non-Monotonic Reasoning (NMR'08)*, pp 179-188. 2008.
- [Toni14] F. Toni. A tutorial on assumption-based argumentation. *Argument & Computation* v5. Taylor & Francis. (2014) DOI:10.1080/19462166.2013.869878
- [W09] Michael Wooldridge. *An Introduction to Multi-Agent Systems - Second Edition*. 2009 John Wiley & Sons Ed.

Uso de Metaheurísticas para el Diseño de Coberturas con Radio Frecuencia Aplicadas a la Telesupervisión de Yacimientos Petroleros.

Molina D., Villagra A., Villagra S., Valdez J.C., Rasjido J., Mercado V.,
Pandolfi D.

Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada (Caleta Olivia)
Universidad Nacional de la Patagonia Austral
{dmolina, avillagra, svillagra, jvaldez, jrasjido, vmercado, dpandol-
fi}@uaco.unpa.edu.ar

Leguizamón G.

Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LI-
DIC)
Departamento de Informática - Universidad Nacional de San Luis
legui@unsl.edu.ar

Resumen

El petróleo es la base de energética más importante de nuestra sociedad. Su obtención, involucra una serie de etapas, donde la extracción y el transporte son de extrema criticidad por su elevado costo y los posibles daños medio ambientales. La supervisión y el control en estas dos etapas son fundamentales y la tecnología es la piedra angular para cumplir los objetivos de producción y medio ambiente. Los sistemas *SCADA* (*Supervisory Control and Data Acquisition*), se distribuyen en puntos estratégicos dentro de los yacimientos, para monitorear de forma automática las variables del proceso de extracción y transporte y detectar los posibles puntos de falla. La interacción entre el usuario y el sistema es mediante la interfaz HMI (*Human Machine Interface*) Para lograr la comunicación entre tanta cantidad de dispositivos (miles de po-

zos petroleros por yacimiento) se necesita una cobertura de comunicaciones muy eficiente y la capacidad de procesar datos con un tiempo de respuesta adecuado. El objetivo de esta línea de investigación es realizar el análisis, estudio e implementación de metaheurísticas, utilizando paralelismo y un modelo de propagación de radio frecuencia real, en yacimientos petroleros en la zona norte de la provincia de Santa Cruz, Argentina.

Palabras clave: servicios inalámbricos, red de radio, paralelismo, metaheurísticas.

Contexto

La línea de investigación presentada en este documento se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEm) en el marco del Programa de Investigación en Ciencia y Tecnología de

la Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia. En el ámbito de una convocatoria de proyectos de I&D UNPA, proyecto denominado: “Metaheurísticas avanzadas aplicadas al diseño eficiente de redes de radio frecuencia en comunicaciones inalámbricas en locaciones petroleras”.

Introducción

Un medio de comunicación inalámbrico es del tipo *cordless* (sin cable) con lo cual no presenta la limitación de distancias y espacio que posee el uso de un cable. Para grandes extensiones, como las que se presentan en los yacimientos petroleros, es el medio de comunicación por excelencia. Para que la comunicación entre dispositivos pueda ser posible, el emisor y el receptor generan una señal con un nivel adecuado para que puedan intercambiar sus mensajes correctamente. La señal al desplazarse por el espacio se va atenuando por los obstáculos propios del terreno y los fenómenos atmosféricos (atenuación, reflexión, refracción y difracción). Esto va generando pérdidas en el nivel de la señal emitido, provocando interrupciones en el enlace. Según [1] el posicionamiento de antenas puede ser descrito de manera informal como: dado un conjunto de sitios candidatos, con distintos tipos de antenas, un área geográfica discretizada e información relativa a la estimación de tráfico, se debe seleccionar un subconjunto de sitios candidatos y sus valores de configuración para cumplir con la estimación de tráfico y cobertura de señal.

En el problema de diseño de la red de radio surgen tres aspectos:

Topográfico: Referido a las cotas y los accidentes del terreno, las coordenadas, el azimut y la distancia entre las estaciones.

Radio eléctrico: Vinculado a los fenómenos de difracción, refracción, absorción, etc. que afectan a la onda electromagnética en el espacio

Alto desempeño computacional: Relacionado con la performance necesaria para tratar una enorme cantidad de datos en un tiempo aceptable.

Estos tres aspectos son determinantes para el éxito de los sistemas SCADA [2] aplicados al proceso de producción de petróleo.

Los yacimientos son enormes superficies de terreno, donde se encuentran diseminados miles de puntos que supervisar (aparatos extractores, válvulas, sensores de oleoductos, etc.). En consecuencia se necesita una solución general, que integre los tres aspectos del diseño de la red y que a su vez, brinde una aproximación de la situación con un determinado grado de veracidad sin importar el tipo de geografía. Para ello se necesitara contar con:

- **Una herramienta de optimización:** Debe ser independiente, adaptable, no guiada y que pueda ser evaluada cuantitativamente para validar los resultados obtenidos. Las metaheurísticas [3, 4], son métodos que integran procedimientos de mejora local y estrategias de alto nivel para realizar una búsqueda robusta en el espacio del problema. El problema de selección de sitios para diagramas de cobertura, es de tipo NP-duro por lo que es apto para ser tratado con metaheurísticas.
- **Un modelo de propagación:** Debe ser adaptable de manera sencilla a un entorno urbano o rural, independiente de la geografía y que permita sacar conclusiones válidas sobre la predicción de señal sobre la zona a cubrir [5].

- **Alto desempeño computacional:** Es fundamental que los resultados obtenidos estén en el tiempo adecuado para la toma de decisiones de diseño. El uso del paralelismo como método de alto desempeño computacional además de reducir el tiempo de cómputo, también va a producir una mejora en la calidad de las soluciones encontradas [6].

El modelo de diseño de redes de radio propuesto por [7] plantea una arquitectura de diseño en capas y fue pensado para redes de celulares [8]. En cada una de las etapas se pueden aplicar metaheurísticas [9, 10, 11] que permitan generar datos para la capa superior [12]. Para este caso en particular, se propone trabajar en la capa de diseño de la red de radio utilizando paralelismo. Se aplicará el modelo explicado en [5] para el cálculo de cobertura de la señal.

Líneas de Investigación, Desarrollo e Innovación

En esta sección se describe la línea de investigación y desarrollo que se encuentra activa.

En [13] se realiza una revisión general de diferentes metaheurísticas resolviendo el problema de RND (*Radio Network Design*). El estudio ofrece una base de referencia confiable sobre un amplio espectro de algoritmos y medidas precisas de comparación de la eficiencia, confiabilidad y rapidez de las diferentes técnicas aplicadas a la resolución del RND. Este estudio establece que el algoritmo CHC (*Crossover elitism population, Half uniform crossover combination, Cataclysm mutation*) es un algoritmo que obtuvo buenos resultados.

Se remarca el hecho que el algoritmo CHC es poco utilizado en esta problemática y los resultados de su aplicación han

demostrado ser de mejor calidad con respecto a otras metaheurísticas. Tomando esta premisa se avanzará sobre la investigación, desarrollo y aplicación del algoritmo de diferentes versiones de CHC, denominadas, QCHC-RE, QCHC-TE y QCHC-ILS. El primero respeta la forma tradicional del algoritmo con un reemplazo elitista (RE) en la nueva población. El segundo varía el método de selección poblacional mediante un Torneo Elitista (TE). Padres e hijos compiten en un torneo seleccionando al de mejor valor objetivo. Y la tercera versión implementa un método de sacudida mediante una Búsqueda Local Iterada (*ILS, Iterated Local Search*). Todas las versiones se diferencian de la versión tradicional por el mecanismo de selección poblacional [14], la forma de realizar un proceso denominado cataclismo para salir del estancamiento [15] y el uso de la diversidad genética como parámetros de convergencia.

Para representar el área geográfica que se desea tratar, se necesita un método para obtener un mapa digital con los valores más representativos del terreno y que sea la referencia para el modelo de propagación seleccionado. La cartografía generada por la misión topográfica SRTM (*Shuttle Radar Topography Mission*) es una de las más apropiadas. Los archivos se encuentran disponibles en internet y puede modelarse cualquier parte del mundo. Para generar predicciones de comunicación confiables se necesita un modelo que debidamente probado en su uso con un fundamento estadístico de sus predicciones. El modelo de Longley – Rice, o ITM (*Irregular Terrain Model*) fue desarrollado en los años 60 y brinda lo que se denomina atenuación de referencia. Este valor indica la pérdida en espacio libre asociada a la zona donde se lo aplica. El modelo tiene como datos de entrada los parámetros de sistema (frecuencia, distancia, la altura de las antenas

y su polarización, vertical / horizontal) parámetros del entorno (coeficiente de irregularidad del terreno, refractividad del terreno, clima, etc.) y parámetros estadísticos para establecer un nivel de confianza sobre los datos obtenidos [16]. Aunque el uso de metaheurísticas permite reducir la complejidad de la búsqueda de una solución, el tiempo del proceso puede seguir siendo muy elevado. La proliferación de plataformas paralelas permite realizar cálculos cada vez más eficientes. La implementación del paralelismo a las metaheurísticas [19,20] aplicadas a la resolución de este problema surge de forma natural.

Resultados y Objetivos

En [14,15] se han desarrollado experimentos de los algoritmos QCHC comparándolos con AGs (Algoritmos Genéticos) utilizando operadores de un punto, dos puntos y cruzamiento uniforme para resolver el problema de RND [17]. La función objetivo utilizada en estos trabajos relaciona la maximización de la cobertura alcanzada por un conjunto de sitios y la minimización de uso de recursos en una sola expresión. La variante aportada es el uso de una función objetivo basada en la proporcionalidad de la cobertura de la superficie. Además, esta función objetivo minimiza las interferencias por superposición de lóbulos de radiación y el uso de radio bases como recursos de elevado costo. El CHC es un AG no tradicional [18] que combina una estrategia de selección conservativa que siempre preserva los mejores individuos encontrados y para salir de un estancamiento u óptimo local, se produce un reinicio de la población mediante un cataclismo. El operador de recombinación produce descendientes que maximizan sus diferencias genéticas con respecto a sus padres. La reproducción sólo se lleva a cabo si la distancia de

Hamming (diferencia genética entre los padres) es mayor al umbral establecido. El operador de cruzamiento HUX (*half uniform crossover*) es usado para maximizar la distancia genética entre individuos. La nueva población se genera con una selección de los mejores individuos, cuando converge la población se produce un reinicio de la misma conservando los mejores individuos.

Se propone analizar, evaluar y contrastar los resultados obtenidos con el algoritmo QCHC-RE aplicados a un área geográfica perteneciente a un yacimiento petrolero. Los resultados obtenidos serán validados a través de estudios experimentales y analizados bajo la teoría estadística apropiada. A continuación se enumeran los resultados esperados del proyecto:

- Estudio comparativo de los algoritmos propuestos contra técnicas tradicionales sobre el problema seleccionado.
- Aplicación de los algoritmos propuestos al problema de RND en un yacimiento petrolero
- Análisis de los algoritmos propuestos con el objetivo de estudiar formalmente las complejidades computacionales involucradas en cada uno de ellos.
- Dejar precedentes sobre el uso de metaheurísticas para resolver problemas de coberturas en una zona geográfica de la República Argentina

Formación de Recursos Humanos

En cuanto a la formación de recursos humanos cabe mencionar que en el marco de las actividades realizadas un integrante del proyecto está orientando el desarrollo de su tesis de Doctorado a esta línea de investigación. En tanto que, en el LabTEm se trabaja con alumnos avan-

zados en la carrera Ingeniería en Sistemas, en temas relacionados a esta línea de investigación, con el objeto de guiarlos en el desarrollo de sus trabajos de fin de carrera y también, de formar futuros investigadores.

Referencias

- [1] W. Corne, M. Oates, G. Smith Telecommunications Optimization: Heuristic and Adaptive Techniques. John Wiley & Sons Ltd, 2000.
- [2] A.R. Penin. Sistemas SCADA. Marcombo, 2012.
- [3] E. Talbi, Metaheuristics From Design To Implementation, John Wiley & Sons, Inc., July 2009.
- [4] Blum, C., & Roli, A. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3), 268-308, 2003.
- [5] G.A. Hufford, A.G. Longley, W.A. Kissick, A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Model, Ntia Report 82-100, 1982
- [6] Alba, E., Luque, G., & Nesmachnow, S. Parallel metaheuristics: recent advances and new trends. *International Transactions in Operational Research*, 20(1), 1-48, 2013.
- [7] K. Tutschku, N. Gerlich, and P. Trángia An integrated Approach to Cellular Network Planning, Institute of Computer Science, University of Wurzburg, 1995.
- [8] Th. Fritsch, K. Tutschku, K. Leibnitz, Field Strength Prediction by Ray Tracing for Adaptive Base Station Positioning in Mobile Communication Networks, August 1995.
- [9] M. Vega-Rodríguez, J. Gómez-Pulido, E. Alba, D. Vega-Pérez, S. Priem-Mendes, G. Molina, Evaluation of Different Metaheuristics Solving the RND Problem, *EvoWorkshops 2007, LNCS 4448*, pp. 101-110, 2007.
- [10] P. Calegari, F. Guidec, P. Kuonen, and D. Wagner. Genetic Approach to Radio Network Optimizations for Mobile Systems. In *Proceedings 47th IEEE Conference on Vehicular Technology*, volume 2, pages 755-759, 1997.
- [11] H.R. Anderson and J.P. McGeehan. Optimizing Microcell Base Station Locations Using Simulated Annealing Techniques. In *Proceedings 44th IEEE Conference on Vehicular Technology*, pages 858-862, 1994.
- [12] N. Erradi, F. Alami, N. Aknin, A. El Moussaoui, Genetic algorithms to optimize base station sitting in WCDMA networks (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol. 4, No. 3, 2013.
- [13] S. Mendes, G. Molina, M. Vega-Rodríguez, J. Gómez-Pulido, Y. Sáez, G. Miranda, C. Segura, E. Alba, P. Isasi, C. León, and J. Sánchez-Pérez, Benchmarking a Wide Spectrum of Metaheuristic Techniques for the Radio Network Design Problem. *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, October 2009.

[14] D. Molina, D Pandolfi, A Villagra, G. Leguizamón, Diseño eficiente de redes de radio frecuencia con algoritmos CHC en comunicaciones inalámbricas, CoNa IISI 2014.

[15] D. Molina, D Pandolfi, A Villagra, G. Leguizamón, Applying CHC Algorithms on Radio Network Design for Wireless Communication CACIC 2014. ISBN 978-987-3806-05-6.

[16] T. Rappaport, Wireless communications principles and practice. 1ra ed. New Jersey: Prentice Hall, 1996.

[17] D. Molina, D. Pandolfi, A. Villagra, Aplicación y evaluación de diferentes algoritmos genéticos canónicos en el diseño eficiente de redes de radio frecuencia en comunicaciones inalámbricas. ICT-UNPA-77-2013 Resolución Nro. 1121/13-R-UNPA.

[18] L.J. Eshelman. The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination. In Foundations of Genetic Algorithms, pages 265{283. Morgan Kaufmann, 1991.

[19] E. Alba, editor. Parallel Metaheuristics: A New Class of Algorithms. Wiley, 2005.

[20] F. Luna, A. J. Nebro, and E. Alba. Parallel evolutionary multiobjective optimization. In N. Nedjah, E. Alba, and L. de Macedo, editors, Parallel Evolutionary Computations, volume 22 of Studies in Computational Intelligence, chapter 2, pages 33 – 56. Springer, 2006.

**Arquitectura, Redes y
Sistemas Operativos**

Adaptación de un Middleware de Internet de las Cosas para Gestionar el Proceso de Fermentado en la Industria de Manufactura del Té

Eduardo O. Sosa^{1,2,a}, Diego Alberto Godoy^{1,2,b}, Edgardo A. Belloni^{1,2,c}, Juan de Dios Benítez^{1,2,d}, Hernán Bareiro^{1,e}, Fabián Favret^{1,f}, Paola Quiñones^{1,g}

¹Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.) - Universidad Gastón Dachary (UGD)

²Secretaría de Investigación y Posgrado (SECIP). Facultad de Ciencias Exactas, Químicas y Naturales; Félix de Azara 1552 - N3300LQH - Posadas
Universidad Nacional de Misiones

^aes@fceqyn.unam.edu.ar, ^bdiegodoy@citic.ugd.edu.ar, ^cebelloni@ugd.edu.ar, ^djuan.benitez@citic.ugd.edu.ar, ^ehbareiro@citic.ugd.edu.ar, ^ffabianfavret@citic.ugd.edu.ar, ^gpquiñones@citic.ugd.edu.ar

Resumen

En este trabajo se presenta la utilización de middleware para internet de las cosas adaptado a un sistema de monitoreo, que transforma un sistema tradicional de medición de temperatura y humedad del ambiente, en una implementación basada en WSN capaz de calcular la humedad del sólido que se está analizando.

Este trabajo presenta una línea de investigación dentro del proyecto denominado “Diseño de arquitecturas de soporte a la Internet del futuro y Ambientes Inteligentes para su Aplicación en Ciudades Inteligentes – Fase II” de la Universidad Gastón Dachary.

Palabras clave: Té negro, Redes de Sensores Inalámbricos, IoT

Contexto

El trabajo se enmarca en el proyecto de investigación denominado “Diseño de Arquitecturas de Soporte a la Internet del

Futuro y Ambientes Inteligentes” y “Diseño de arquitecturas de soporte a la Internet del futuro y Ambientes Inteligentes para su Aplicación en Ciudades Inteligentes – Fase II”, acreditados en la Secretaría de Investigación y Desarrollo de la UGD por RR.19/A/12 y RR.18/A/14 respectivamente.

Se relaciona y articula directamente con diversos proyectos de investigación acreditados en la Secretaría de Investigación y Posgrado (SECIP) de la Facultad de Ciencias Exactas, Químicas y Naturales de la UNaM, enfocados en las temáticas “Internet del Futuro” e “Internet de las Cosas”, entre los que se incluyen: Proyecto 16Q457 “Hacia la Programación de Sensores Inalámbricos en la Forma Web 2.0”; Proyecto 16Q474 “Simulaciones de Sistemas Modernos de Comunicación”; y Proyecto 16Q519 “Ambientes Inteligentes. Una Mirada a Internet del Futuro”.

Referente a este proyecto existen 7 tesinas de grado en curso, 5 tesinas de grado finalizadas, un trabajo final de Especialización finalizado (UNLP), una tesis de Maestría en Ingeniería de

Software finalizado (UNLP) y una Tesis Doctoral en Tecnologías de la Información y las Comunicaciones finalizada (Universidad de Vigo, España).

1. Introducción

La presente investigación parte de analizar los factores y plantear soluciones científico-tecnológicas para las fábricas elaboradoras de "Té Negro" [1] de la provincia de Misiones. La intención es mejorar la calidad de nuestro té. Teniendo en cuenta que éste se considera de calidad media-baja se espera a partir del análisis, estudio, pruebas y la construcción de un prototipo tecnológico demostrar que con soluciones de hardware-software estos valores pueden optimizarse.

Este proyecto tuvo como desafío diseñar y construir una solución tecnológica, que permite conocer los valores de humedad de la masa [2] que se encuentra en la cinta de fermentado de manera instantánea y, en el mismo momento que el proceso está ocurriendo. Pudiendo prescindir de muestreos, análisis de laboratorios, especialistas, conocimientos empíricos hasta hoy aplicados por idóneos; como de esperas de 12 a 48 horas para conocer dicho valor. Utilizando dispositivos diseñados para medir temperatura y humedad relativa [3]. El prototipo es una herramienta que expone valores que están apoyados en varias ciencias como ser la química básica, la química de alimentos, la ingeniería en electrónica con sus componentes de hardware innovadores y la ingeniería en informática logrando integrar todos estos saberes en un solo componente que brinda datos e información para la toma de las decisiones. Dichas decisiones determinan la buena, media o baja calidad del té negro elaborado.

2. Línea de Investigación

Para esta línea se han planteado los siguientes objetivos:

Como objetivo general se propone desarrollar una solución tecnológica de monitoreo de temperatura y humedad para determinar la humedad del sólido en el proceso de fermentado de té negro utilizando WSN. Asimismo los objetivos específicos son los siguientes:

- Caracterizar el proceso de industrialización del té negro enfatizando en los factores del proceso de fermentado que mejoran la calidad del té negro.
- Conceptualizar sobre el estado del arte de las Redes de Sensores Inalámbricos analizando las formas de medir las variables humedad relativa y absoluta con dicha tecnología.
- Describir el sistema de medición de la temperatura y humedad actual de los secaderos, identificando los requerimientos para un sistema basado en redes de sensores inalámbricos.
- Desarrollar un prototipo en base a un middleware de internet de las cosas que permita mostrar el monitoreo de la temperatura y humedad del proceso de fermentado [1] [4] del té negro.
- Evaluar los resultados obtenidos del conjunto de pruebas.

Los detalles del middleware para IoT [5] utilizado se pueden ver en [6] siguiendo los estándares de la ITU [7]. El estilo de aplicación conseguida es como las estudiadas en [8]. Para esta implementación se han utilizado sensores de la marca ISense [8]. Cabe destacar en este artículo se presentan los avances en la adaptación del middleware para el escenario de un trabajo anterior de los autores [10].

3. Resultados

En días de verano en la provincia de Misiones en la capital del departamento de Leandro N. Alem tanto en la zafra del año 2013, como del

2014 se desplegó en una fábrica secadora de té el prototipo del sistema de monitoreo de temperatura y humedad basado en una WSN. Centrados en el proceso de fermentado dentro de la planta fabril se repetían tres líneas, que recibían el material que provenía del proceso previo, con una cinta transportadora cada una de las líneas. Las dimensiones de las cintas transportadoras eran de un metro de ancho por seis metros de largo, construidas en acero inoxidable las zonas que están en contacto con el colchón de té. Cuenta con 2 rodillos responsables de remover el material los cuales tomaban las hojas que se encontraban en la parte inferior del colchón y lo trasladaban a la parte superior, de esta forma las hojas que estaban en contacto con la cinta pasaban a la superficie y viceversa. Cabe resaltar que el colchón de té se encuentra con un espesor de 10 cm en cada una de las sucesivas visitas.

Lo que se pudo observar con una primera impresión es que la masa de té proveniente del proceso de enrutado con un color verde claro, el aspecto que relucía era a un material foliar crudo, mojado, uniforme, carente completamente de aroma. En cambio una vez recorrida toda la cinta transportadora cuando la atención se fijó en la entrada del proceso de secado nos encontramos con una masa té de un color cobrizo producto de la oxidación producida en esta etapa, con un aspecto seco sin jugos en las uniones y uniforme.

Se cruzaron de lado a lado de la cinta transportadora alambres uno en cada punto a analizar y sobre los mismos se ubicaron los sensores. Los sensores se encontraban a 1 cm de la masa de té siendo el módulo que posee el transductor de temperatura y humedad el más próximo a la masa, el siguiente era el módulo principal y el último módulo era el encargado de alimentar a todo el sistema, el mismo se conectaba a las baterías que estaban formadas por cuatro pilas recargables tipo "triple A".

Todo el proceso de elaboración desde que entra la hoja recién cosechada hasta que se clasifica o tipifica el producto final dura aproximadamente 12 horas.

Como se mencionó anteriormente, la etapa del proceso donde la investigación se centró es el fermentado, el cual posee una duración que puede variar desde los 40 minutos hasta las 2 horas inclusive. La frecuencia con que se evaluaron las variables temperatura y humedad ambiental variaron en un rango que va desde 1 segundo hasta los 12 minutos. Logrando así tener una cantidad de muestras significativas del estado de las variables durante este proceso. Para la primera prueba se utilizaron cinco nodos sensores en total. Un único mote conectado a un computador con la función de puerta de enlace, recibiendo los datos de los otros cuatro sensores que medían temperatura y humedad, de los cuales, los tres primeros se encontraban en el lecho. El cuarto o último nodo se lo expuso a la intemperie para tener una referencia de las condiciones fuera de la nave.



Figura 1. Cinta transportadora al final del proceso de enrutado y principio del proceso de fermentado. (Fotografía tomada en enero del 2015)

La ubicación de los sensores corresponde a las necesidades de control que se tienen en el proceso que se está evaluando. El primer punto de control fue ubicado a la salida del proceso de enrutado, es decir a la entrada del proceso de fermentado para conocer las condiciones iniciales. De esta manera se pudo evaluar y modificar los procesos previos al fermentado, como ser: la recepción y clasificación de la materia prima, el enrutado y el proceso de marchitado. Esto se puede ver en la Figura 1.

El segundo nodo se ubicó exactamente en la mitad del proceso siendo éste el punto más importante pues en esta instancia aún se pueden

modificar las variables de humedad y temperatura del lecho (activando o desactivando los aspersores de agua, aumentando o disminuyendo la velocidad de la cinta transportadora, aumentando o disminuyendo el espesor de colchón de hojas o encendiendo y apagando los ventiladores).

Y por último se instaló un mote al final del proceso de fermentado, que permite obtener la información necesaria para determinar la temperatura del horno en la próxima etapa denominada secado.

Además se adicionó en esta prueba un sensor que no realizaba mediciones de manera continua. Fue introducido en la mitad del lecho en contacto directo con la masa de té que está en el fermentado y acompañó el recorrido de la cinta para determinar así la temperatura en el sólido. Exactamente a cinco centímetros de la cinta transportadora con té por encima y por debajo del sensor; midió la temperatura de la masa de té ya que el equipo se encontraba en equilibrio con dicha masa, cabe aclarar que no se tuvieron en cuenta los valores de humedad que éste identificaba por hallarse herméticamente protegido.

El sensor móvil antes mencionado pasó por debajo de los tres sensores fijos. Cuando el sensor que está en movimiento alcanzaba a los nodos fijos la diferencia que se obtuvo en la medición fue de 2 grados centígrados aproximadamente.

A continuación se en la Figura 2 se presenta parte de la vista terminada del front-end del middleware adaptado con el fin de exponer los datos de temperatura, humedad ambiente y humedad del sólido que se encuentra en el lecho de fermentado. Dicha vista está destinada al usuario final que puede ser el capataz de la fábrica u operarios que manejan los actuadores del proceso de fermentado (ventiladores, cuan alto se encuentra el colchón de té en la cinta, el humidificador y la velocidad del recorrido de la cinta).

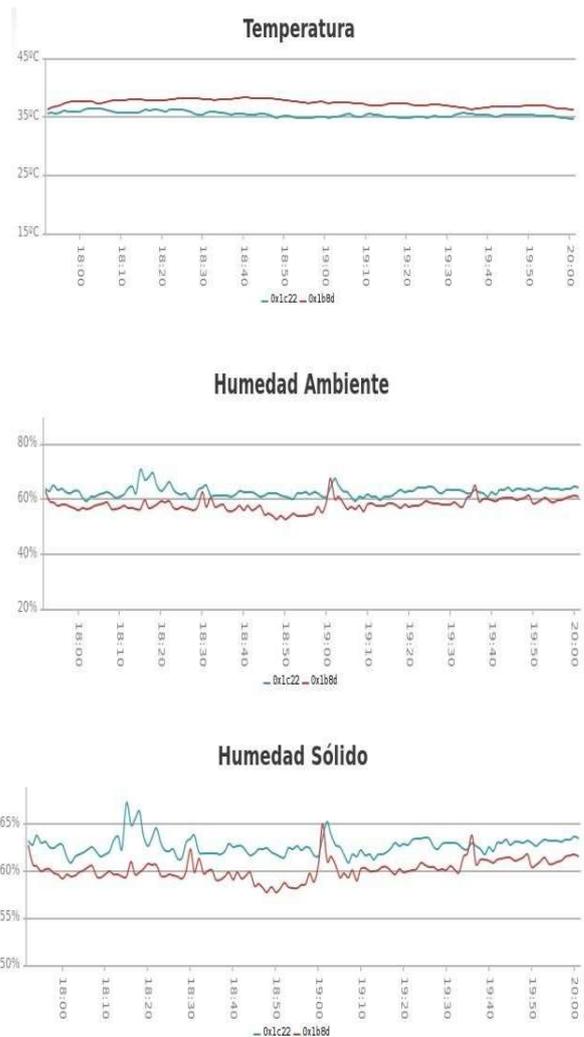


Figura 2. Vista de Usuario del Sistema de monitoreo de temperatura y humedad para el proceso de fermentado de té negro basado en una WSN. Elaboración propia.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por una Doctora en Ingeniería Telemática (España),

un Doctor en Ciencias Informáticas y Magister en Redes de Datos, un Doctor en Tecnologías de la Información y Comunicaciones, Magister y Especialista en Ingeniería de Software, un Maestrando de Ingeniería de la Web, dos Mastrandos de Redes de Datos, tres auxiliares de investigación graduado y ocho auxiliares de investigación en período de realización de trabajos de grado. El número de tesinas de grado en curso con proyecto aprobado es de cinco y el número de trabajos de especialidad finalizado es tres dentro de la línea de investigación. Los proyectos de grado se titulan “Diseño de un prototipo para monitoreo eficiente de iluminación basado en WSN utilizando HTML5”, “Contribución a la Gestión de Residuos Domiciliarios como una Aplicación en Ciudades Inteligentes” y “Análisis y comparación de modelos de propagación para optimizar la localización geográfica de Ganado”.

5. Bibliografía

- [1] Alimentos Argentinos,. [Online]. <http://www.alimentosargentinos.gov.ar>.
- [2] W BALTES, "Química de los alimentos," 2007.
- [3] A. Pezzutti and Crapiste. G.H., "Sorptional Equilibrium and Drying Characteristics of Garlic.," 1997.
- [4] FAO. (Organización de las Naciones Unidas para la Agricultura y la Alimentación). [Online]. www.fao.org
- [5] K Ashton, "'That 'Internet of Things' Thing'," 2009 (rev. 2011).
- [6] Diego Alberto Godoy, Eduardo Omar Sosa, and Rebeca Díaz Redondo, "Internet De Las Cosas: Middleware De Gestion De Datos De WSN," in *CONFERENCIA WWW/INTERNET 2015 e COMPUTAÇÃO APLICADA 2015 FLORIANÓPOLIS*, Florianópolis, 2015.
- [7] International Telecommunication Union ITU, *Service description and requirements for ubiquitous sensor network middleware.*: ITU-T Recommendation F.744, 2009.
- [8] Diego Alberto Godoy, Plataformas para la creación de mashups sensibles al contexto en entornos de inteligencia ambiental. Trabajo Final de Especialidad en Ingeniería de Software, 2013, Universidad Nacional de Plata.
- [9] Coalesense. (2015) iSense Wireless Sensor Network Software. [Online]. <http://bit.ly/1WMMiR0>
- [10] P. Quiñones, D. Godoy, and E. Sosa, "Redes Inalámbricas de Sensores: Una Experiencia En La Industria Del Té," in *42° Jornadas Argentina de Informática*, Córdoba, 2013.

Ampliando la Vida Útil de las WSN por Medio de los Protocolos de Ruteo, Modificación de AODV

Juan de Dios Benitez^{1,2}, Eduardo O. Sosa^{1,2}, Diego Alberto Godoy^{1,2}, Edgardo A. Belloni^{1,2}, Fabian Favret¹, Hernán Bareiro¹, Ricardo Urdinola¹, Micaela Olivera¹

¹Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.) - Universidad Gastón Dachary (UGD)

²Secretaría de Investigación y Posgrado (SECIP). Facultad de Ciencias Exactas, Químicas y Naturales - Universidad Nacional de Misiones

{juan.benitez, eduardo.sosa, diegodoy, ebelloni, fabianfavret, hbareiro, rurdinola, molivera}@citic.ugd.edu.ar,

Resumen

En este trabajo se presenta una modificación en el protocolo de ruteo multi salto reactivo AODV (Ad Hoc On-Demand Distance Vector) para WSN (Wireless Sensor Network) el cual permite por medio de la adaptación del cálculo de su métrica, aumentar la vida útil de una red al disminuir la probabilidad de aparición de segmentaciones. De esta manera, mediante la distribución de la carga energética de los nodos a lo largo de la red, se consigue una optimización en el consumo de las baterías de los nodos individuales, concluyendo con una vida útil de la red mejorada y aumentada. El presente estudio conforma una línea de investigación dentro del proyecto denominado “*Diseño de arquitecturas de soporte a la Internet del futuro y Ambientes Inteligentes para su Aplicación en Ciudades Inteligentes – Fase II*” de la Universidad Gastón Dachary.

Palabras clave: AODV, Protocolo Reactivo, Redes de Sensores Inalámbricos

Contexto

El trabajo se enmarca en el proyecto de investigación denominado “*Diseño de Arquitecturas de Soporte a la Internet del Futuro y Ambientes Inteligentes*” y “*Diseño de arquitecturas de soporte a la Internet del futuro y Ambientes Inteligentes para su Aplicación en*

Ciudades Inteligentes – Fase II”, acreditados en la Secretaría de Investigación y Desarrollo de la UGD por RR.19/A/12 y RR.18/A/14 respectivamente. Se relaciona y articula directamente con diversos proyectos de investigación acreditados en la Secretaría de Investigación y Posgrado (SECIP) de la Facultad de Ciencias Exactas, Químicas y Naturales de la UNaM, enfocados en las temáticas “*Internet del Futuro*” e “*Internet de las Cosas*”, entre los que se incluyen: Proyecto 16Q457 “*Hacia la Programación de Sensores Inalámbricos en la Forma Web 2.0*”; Proyecto 16Q474 “*Simulaciones de Sistemas Modernos de Comunicación*”; y Proyecto 16Q519 “*Ambientes Inteligentes. Una Mirada a Internet del Futuro*”. Referente a este proyecto existen 7 tesinas de grado en curso, 5 tesinas de grado finalizadas, un trabajo final de Especialización finalizado (UNLP), una tesis de Maestría en Ingeniería de Software finalizado (UNLP) y una Tesis Doctoral en Tecnologías de la Información y las Comunicaciones finalizada (Universidad de Vigo, España).

1. Introducción

El aumento de la capacidad de cómputo, la miniaturización de los componentes y el abaratamiento de los dispositivos, hacen posible la realización de la IoT (Internet of Things) [1] y de la Inteligencia Ambiental [2], en la que los

objetos cotidianos contarán con sensores conectados a Internet.

El desarrollo de estas tecnologías brinda oportunidades de aprovechamiento económico e industrial únicas, pero a la vez presentan un enorme desafío tecnológico. Como parte del desarrollo de las tecnologías de la IoT, las WSN [3] son un área de álgido estudio y desarrollo [4]. Hoy en día existen WSNs con miles de nodos, recolectando datos en infinidad de entornos, para todo tipo de fines y a un costo relativamente razonable.

2. Proyecto de Investigación

El objetivo general del proyecto marco es: "Diseñar arquitecturas de soporte a Internet del futuro y Ambientes Inteligentes para su aplicación a Ciudades Inteligentes". Como objetivos específicos se indican: A) Especificar el Estado del Arte de la Internet del de Futuro y los Ambientes y Ciudades Inteligentes. B) Determinar los protocolos de ruteo óptimos para implementaciones de WSN y ciudades inteligentes. C) Determinar el impacto en la vida útil de una red WSN a través de la utilización del protocolo AODV [5] estándar por medio de simulaciones. D) Analizar y diseñar una modificación a la métrica de dicho protocolo para que el mismo sea energéticamente eficiente aumentando la vida útil de una red. E) Simular escenarios de Ciudades inteligentes de redes de sensores inalámbricos. F) Determinar cuál es la métrica correcta que optimiza la vida útil de la red. G) Realizar las simulaciones de comparación entre los protocolos AODV estándar y AODV modificado, para analizar la mejora en la red.

Protocolo AODV

Una de las características más importantes que poseen los nodos en redes WSN es la habilidad de crear redes autónomas. Si se tiene en cuenta que usualmente son desplegados en grandes áreas de difícil acceso [6] [7], es fundamental contar con cierto grado de autonomía y auto

organización en los nodos a fin de facilitar el despliegue e implementación de la red WSN. Sin embargo, esto requiere que los nodos cumplan dos funciones muy diferentes: la recolección de información del medio para la cual fue distribuido y a su vez, el encaminamiento de los paquetes desde el origen hasta el destino seleccionado. [8] Generalmente, cuando se implementa una WSN, existe un nodo denominado Gateway o nodo Sink [9], el cual posee mayores recursos tanto energéticos como de hardware en comparación a los otros nodos. Este nodo Sink suele tener acceso a internet (si fuera necesario), paneles solares, etc.

Por otro lado, los demás nodos que conforman la red son realmente limitados en lo que se refiere a: potencia de cómputo, almacenamiento y, esencialmente, a la capacidad de energía de sus baterías internas.

Ésta última limitación es un factor crítico a considerar. El tiempo de vida de la red depende directamente de la energía residual con la que cuenta cada nodo, definida como la energía remanente que poseen las baterías a ser utilizada. A medida que los nodos pierden energía, la red puede llegar a sufrir fragmentaciones como se puede ver en la fig. 1, causando lo que se denominan islas aisladas, las cuales son inalcanzables por el resto de la red a raíz de la limitada potencia de transmisión que poseen los nodos. Esto implica que el deceso de un solo nodo puede ocasionar una degradación operacional significativa de la red [10].

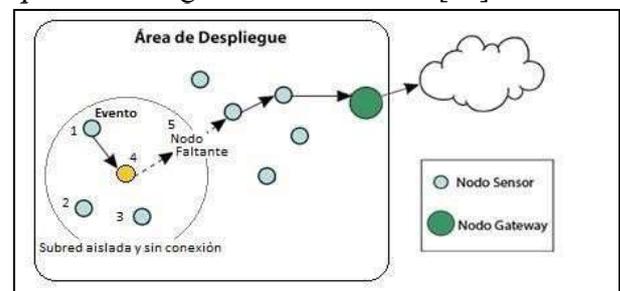


Figura 1: Fragmentación de la red

La complejidad de la red se incrementa cuando los paquetes se reenvían utilizando métodos

multi saltos, donde múltiples retransmisiones son necesarias para poder enviar un paquete desde el origen hasta el destino. Como consecuencia, el mayor consumo de energía se lleva a cabo en los momentos de transmisión y recepción de información por medio de los nodos. Esto implica que los protocolos y algoritmos de ruteo son claramente esenciales y decisivos en el tiempo de vida de una WSN. Sin embargo, uno de los algoritmos más utilizados por redes WSN como es el AODV – Un protocolo de Camino más corto - [11], sigue un enfoque tradicional con una búsqueda del camino con el menor costo a través del conteo de saltos, causando que los nodos participantes de la red concentren su tráfico a través de un grupo de nodos aumentando el consumo de energía en los mismos y de esta forma segmentando la red de manera más precipitada. El protocolo AODV [5] es un protocolo de ruteo reactivo, lo que implica que solo actúa cuando hay la necesidad de transmitir a un destino en particular a través de una ruta desconocida. El proceso se activa iniciando el descubrimiento del camino, lo que básicamente se lleva a cabo por medio de una inundación de la red a través de un paquete de RREQ (Route Request). Cuando el RREQ llega al nodo deseado, un paquete de RREP (Route Replay) es enviado al nodo origen de por medio de una transmisión en unicast como se puede ver en la fig. 2.

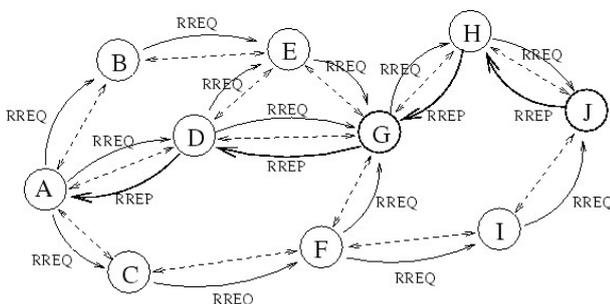


Fig. 2: Esquema de mensajes

El paquete RREQ, ver fig. 3, cuenta con un campo de 8 bits denominado “Hop Count” (conteo de saltos), el cual se incrementa en cada

retransmisión del paquete por todos los nodos participantes.

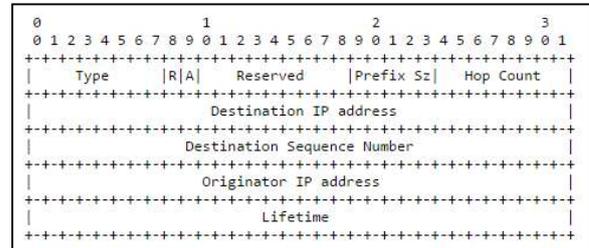


Fig. 3 Paquete RREQ

Cuando el paquete RREQ llega al destino, el campo de conteo de saltos tiene el valor necesario para transmitir el paquete entre el nodo origen y el destino. Esta métrica de saltos, representa la distancia física entre ambos nodos. Este costo es almacenado en el nodo, en su tabla de ruteo y determina el camino que se debe utilizar para alcanzar dicho destino

Modificación del Protocolo AODV

La modificación que se ha realizado en el algoritmo de cálculo de la métrica del protocolo AODV estándar se centra en mejorar la vida de la red. Durante el proceso de descubrimiento, los nodos incrementan el costo en el campo de hop count en el RREQ. Si en lugar de incrementar el valor en una unidad por cada salto, se reemplaza dicho valor por una métrica que tenga en cuenta la energía residual que tiene el nodo que retransmite el paquete -por medio de una aproximación matemática- el resultado sería un protocolo sensible a la distribución de la energía de los nodos en la red. Si se implementa una métrica basada en la siguiente función matemática por partes donde al llegar el nivel de batería al 25% la función asigna el valor máximo de métrica igual a 255, mientras que en la franja $25% < \text{Energía} < 100\%$ asignará valores acordes a la siguiente función cuadrática:

$$M = \begin{cases} 255 & \text{if } E \leq 0.25 \\ \dots & \text{if } 0.25 < E < 1.0 \end{cases} \quad (1)$$

La función propuesta para determinar este costo es:

$$f(x_n) \begin{cases} 255 & 0 \leq x \leq 0.25 \\ -\frac{3200}{9}x^2 + \frac{940}{9}x + \frac{2260}{9} & 0.25 < x \leq 1 \end{cases} \quad (2)$$

Donde “x” es la energía residual relativa en el nodo.

En la fig. 4 se puede apreciar de manera gráfica la ecuación 2 ejemplificando como irá variando la métrica a medida del decrecimiento de la energía relativa que cuenta el nodo.

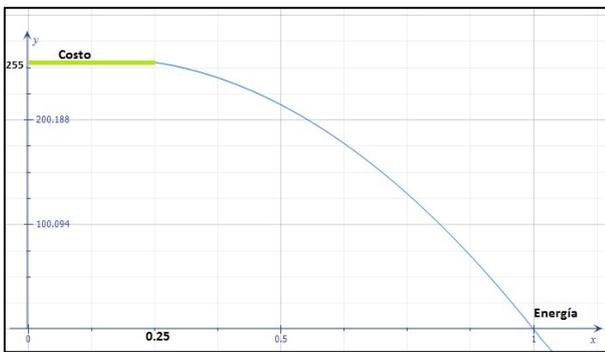


Fig. 4 Función de la Métrica

Proceso de Simulación

El entorno usado para realizar los análisis y simulaciones se compone de un PC de uso general con un sistema operativo Linux distribución Ubuntu [12], en el cual se corre el simulador Omnet++ [13], un simulador de eventos discretos orientado a objetos [14] distribuido bajo licencia publica académica. Una de las características más importantes del simulador, proviene de la capacidad del uso adicional de frameworks, lo que permite simplificar la implementación de los módulos necesarios.

El equipo de desarrollo de Omnet++, mantiene el framework llamado INET [13]. Sin

embargo, para este trabajo se ha utilizado una extensión denominada INETMANET [15], la cual provee la implementación de MANETS,

incluyendo el protocolo AODV, desarrollada por la Universidad Upsala [16]. Esta Extensión

ha sido usada y testada por la comunidad Omnet++ por alrededor de 10 años, por lo que se considera una extensión estable.

El proceso de simulación consiste en dos series

de simulaciones. En la primera serie se utiliza el protocolo AODV estándar, mientras que la

segunda serie se realiza bajo las mismas características que la anterior, pero utilizando el protocolo AODV modificado. Cada serie cuenta con 50 corridas del simulador configuradas con diferentes semillas. Para poder mantener la coherencia entre las series de simulaciones, se configura al simulador para mantener el mismo juego de semillas entre las dos series, permitiendo comparar ambos protocolos.

Cada corrida del simulador consiste en la generación aleatoria de un escenario de 1Km² el cual cuenta con 25 nodos sensores distribuidos con una probabilidad uniforme. La densidad de nodos resultante permite una alta probabilidad del uso de ruteo multi salto, para realizar el encaminamiento de los paquetes desde un nodo origen hasta el nodo gateway.

Como se desea que la simulación represente un escenario lo más real posible, se han configurado los nodos de manera que tengan ciclos de transmisión de datos aleatorios en un intervalo entre 50 y 120 segundos.

Todos los nodos están equipados con un modelo de consumo de batería lineal, implementado en el framework INETMANET, y el consumo de energía es el único factor determinante en la muerte de los nodos. Los nodos se encuentran equipados con baterías de 3800mAh con un voltaje de 12V.

En la tabla 1, se puede observar el consumo de energía original configurado en la simulación.

Tabla 1 Consumo de Energía.

<i>Node's State</i>	<i>Value</i>
Idle	0.5mA
Receiving	1mA
Sleeping	1mA
Sleeping	0.001mA

Para poder reducir los tiempos de simulación, y debido a las limitaciones de hardware para las simulaciones se ha decidido incrementar cada uno de los consumos en diez mil veces

Como el interés de este trabajo es la medición del tiempo de vida de la red, se ha decidido que cada simulación se corte al momento que el primer nodo se queda sin batería y muere. En ese momento se realiza la medición de la energía residual de la red para comparar con el protocolo original.

3. Resultados

Usando los resultados de las simulaciones, se han comparado los parámetros de tiempo de vida y distribución de energía en la red utilizando ambos protocolos, AODV estándar y AODV modificado.

Como se puede apreciar en la fig. 5, el protocolo AODV modificado tiene un tiempo de vida de la red mayor que el AODV estándar.

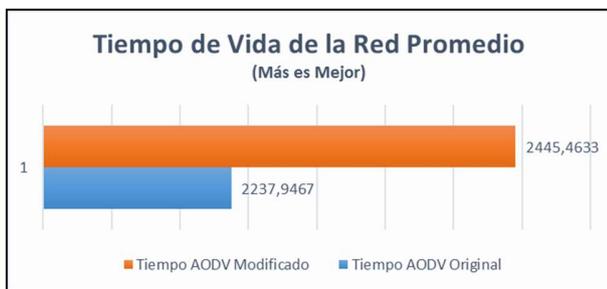


Fig. 5 Comparación de tiempo de vida en segundos

El resultado muestra una mejora del 9,2726% en la utilización del AODV modificado contra el AODV estándar.

Con la utilización del protocolo AODV modificado, se ha buscado obtener una mejor distribución del consumo de energía en la red. El resultado esperado se muestra a través del uso del desvío estándar de la energía residual promedio de las 50 corridas de simulación, representada en la fig. 6, donde la misma fue utilizada en la representación de una distribución normal fig. 7



Fig. 6 Comparación de Desvío Estándar promedio.

Efectivamente al realizar el cambio de protocolo se logró una reducción del 38,72% en el desvío estándar del promedio de la energía residual del protocolo modificado sobre el protocolo estándar. Esto significa que los nodos distribuyeron su energía de una manera más homogénea en la red.

En la práctica esto es equivalente a decir que los paquetes utilizaron diferentes rutas, evitando que se concentre el tráfico sobre una misma ruta lo que sería afirmar que se estaría concentrando sobre un mismo grupo de nodos agotando de manera más rápida su energía.

Éste resultado se puede observar en el estudio de la distribución normal, donde el protocolo AODV con la métrica modificada es más estrecho

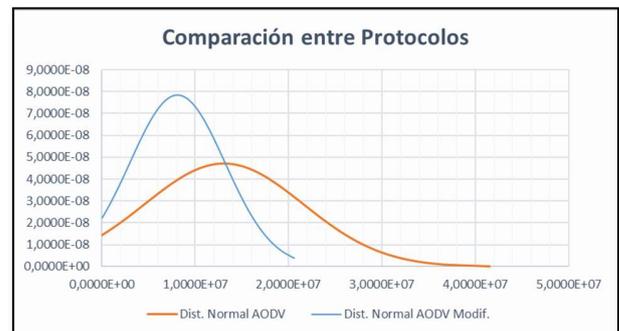


Fig. 7 Comparación de Distribución Normal.

De otra forma, se puede observar que el protocolo modificado, tiene una curva más hacia la izquierda que el AODV estándar, con lo que se puede deducir que la curva del protocolo modificado posee una energía residual media menor que el protocolo estándar.

Este es el efecto esperado y deseado en la comparación de ambos protocolos.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por una Doctora en Ingeniería Telemática (España), un Doctor en Ciencias Informáticas, Magister en Redes de Datos, un Doctor en Tecnologías de la Información y Comunicaciones, Magister y Especialista en Ingeniería de Software, un Maestrando de Ingeniería de la Web, dos Maestrandos de Redes de Datos, tres auxiliares de investigación graduado, y ocho auxiliares de investigación en período de realización de trabajos de grado. El número de tesis de grado en curso con proyecto aprobado es de cinco y el número de trabajos de especialidad finalizado es uno dentro de la línea de investigación. También se ha finalizado una Tesis de Doctorado. Los proyectos de grado se titulan “Diseño de un prototipo para monitoreo eficiente de iluminación basado en WSN utilizando HTML5”, “Contribución a la Gestión de Residuos Domiciliarios como una Aplicación en Ciudades Inteligentes” y “Análisis y comparación de modelos de propagación para optimizar la localización geográfica de Ganado”.

5. Bibliografía

- [1] K Ashton, "'That 'Internet of Things' Thing'," 2009 (rev. 2011).
- [2] Ahola J., Ambient Intelligence, 2001.
- [3] W. Dargie and C. Poellabauer, "*Fundamentals of Wireless Sensor Networks - Theory and Practice*." Reino Unido: Wiley, West Sussex, 2010.
- [4] Eduardo O. Sosa, Contribuciones al establecimiento de una red global de Sensores Inalámbricos. Tesis Doctoral, Junio 17, 2011.
- [5] C. Perkins, Elizabeth M. Belding-Royer, S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing," 2003.
- [6] Ruben Hidalgo, José Ignacio Moreno, "Routing Design in Wireless Sensor Networks and a Solution for Healthcare Environments," *IEEE LATIN AMERICA TRANSACTIONS*, vol. 9, no. 3, 2011.
- [7] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, "Deploying a wireless sensor network on an active volcano," 2006.
- [8] Rafael Marin Perez, "Algoritmos Fiables y Eficientes basados en Enrutamiento Geográfico para Redes Realistas de Sensores Inalámbricos," Murcia, 2012.
- [9] Fatma Bouabdallah, Nizar Bouabdallah, Raouf Boutaba, "On Balancing Energy Consumption in Wireless Sensor Networks," Rennes, France; Waterloo, Canada, 2008.
- [10] S. Ganesh, R. Amutha, "Efficient and Secure Routing Protocol for Wireless Sensor Networks," Chennai, Tamil Nadu, India, 2013.
- [11] Huarui Wu, Chunjiang Zhao, Li Zhu, "Study on an Energy-aware Routing Algorithm for Agriculture WSN," *TELKOMNIKA*, vol. 11, no. 7, p. 3576~3584, 2013.
- [12] Canonical Ltd. Ubuntu. [Online]. <http://www.ubuntu.com/download/desktop>
- [13] OMNeT++ developers. [Online]. <https://inet.omnetpp.org/>
- [14] OpenSim Ltd. OMNeT++. [Online]. <https://omnetpp.org/>
- [15] Alfonso Ariza Quintana. GitHub. [Online]. <https://github.com/aarizaq/inetmanet-2.0>
- [16] Uppsala University. [Online]. <http://www.uu.se/>

Análisis Comparativo de Prestaciones de Tráfico de Video Multicast en Redes IPv6

Higinio Facchini, Santiago Pérez, Fabian Hidalgo, Adrian Cárdenas
 CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería)
 Facultad Regional Mendoza, Universidad Tecnológica Nacional
 Rodriguez 273, Mendoza, Argentina - 0261-5244576
 (higiniofac,santiagocp)@frm.utn.edu.ar, (fabianhdlg,adriancard.mza)@gmail.com

RESUMEN

Actualmente existe un continuo y creciente número de usuarios de tráfico multimedia en la red. Este hecho involucra que se ha incrementado exponencialmente los requisitos de ancho de banda. Los consumidores son cada vez más exigentes en cuanto a la calidad y rendimiento de los productos basados en el tráfico de vídeo, y, por lo tanto, hay un fuerte incentivo para la mejora continua en las tecnologías multimedia. Comprender la estructura de los datagramas de vídeo y los requisitos que imponen a la red, ayudarán a los administradores de red a mejorar el tráfico en la misma.

Siguiendo una línea de investigación de tráfico multicast en el protocolo IPv4 anterior, se pretende continuar con la misma sobre el protocolo avanzado IPv6, comprender las diferencias de estructura de los datagramas de vídeo y los requisitos que imponen a la red, y obtener conclusiones sobre las mejores opciones de tráfico de video multicast sobre IPv6 en redes de laboratorio, midiendo variables de tráfico en la red como velocidad, cantidad de paquetes, retardos, jitter, etc. que ayudarán a mejorar el tráfico en la misma.

Palabras clave: multicast, codecs, tráfico de video, IPv6

CONTEXTO

La línea de investigación está inserta en tres proyectos de análisis de tráfico multimedia, en

el ámbito del Centro UTN CeReCoN (Centro de Investigación y Desarrollo en Computación y Neuroingeniería), del Departamento Ingeniería en Electrónica, de la Facultad Regional Mendoza, de la Universidad Tecnológica Nacional.

Los mismos se dividen en análisis de tráfico de video en redes cableadas e inalámbricas reales bajo las modalidades multicast y unicast, y en análisis similares con simulación de tráfico de video en redes Wi-Fi bajo distintas normas 802.11 con el análisis correspondiente de QoS (calidad de servicio).

1. INTRODUCCIÓN

De acuerdo con un estudio realizado por la multinacional tecnológica Cisco (Visual Networking Index), se puede observar la proyección del tráfico en la red Internet desde el año 2015 al 2020, de acuerdo a:

- a. El tráfico global crecerá de 72.5 exabytes mensuales en 2015 a 194.4 exabytes en 2020. Esto significa que en 2020 el tráfico será de 511 Terabytes/s (el equivalente a que 142 millones de personas realicen streaming de video HD en forma simultánea) (figura 1)
- b. El tráfico de video crecerá del 69% al 81%, o desde 50 exabytes mensuales a 160 exabytes (figura 1).
- c. La cantidad de dispositivos que soporten y realicen tráfico en IPv6, crecerá de 4 mil millones en 2015 a 13 mil millones en 2020, calculando que para ese año el 34% del tráfico global sea en IPv6 (o sea unos 65 exabytes mensuales) (figura 2).

Estas cifras nos brindan la información suficiente para tener en cuenta que el estudio del tráfico de video y de IPv6 son tendencias actuales y futuras para su estudio y análisis.

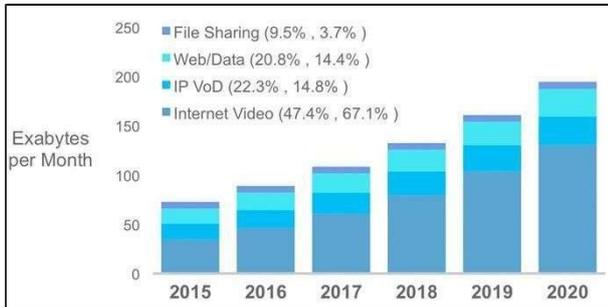


Figura 1: Tráfico IP global por tipo tráfico



Figura 2: Tráfico IPv6 global

Un componente importante de la creación de redes multimedia es la compresión de datos (codificación de la fuente) de las fuentes de datos multimedia (voz, audio, imagen y vídeo).

La compresión de vídeo o codificación de vídeo es el proceso de reducir la cantidad de datos requerida para representar una señal de vídeo digital, antes de la transmisión o almacenamiento. Una vez que los datos están comprimidos, el flujo de bits se paquetiza y se envía a la red. En la operación complementaria, la descompresión o decodificación, se recupera una señal de vídeo digital de la representación comprimida antes de la visualización.

El control de velocidad es una parte esencial de la mayoría de los codificadores de vídeo. Se determina el número de bits o el nivel de calidad de la trama codificada. Hay dos tipos de control de la frecuencia: tasa de bits constante (CBR) y velocidad de bits variable (VBR). Para

CBR la codificación de vídeo, los diseñadores de controles de tasa se centran en la mejora del juego de precisión entre la velocidad de bits de destino y la real velocidad de bits y el cumplimiento de las restricciones de baja latencia y de buffering. Como resultado vemos fluctuaciones en la calidad del video debido a cambios de escena y otros contenidos de vídeo. En los casos en que la restricción de la tasa no es tan estricta como en video en tiempo real, VBR se puede utilizar para mantener la calidad constante.

El salto del video SD (definición estándar) a HD (alta definición), la explosión de dispositivos y de herramientas originadoras y productoras de contenido, junto con la posibilidad de integrar las diferentes tecnologías y el acceso universal a estas, son los motores de la adopción y el acelerado crecimiento del tráfico de video sobre las redes. Además se considera que la adopción del video en los teléfonos móviles será la mayor responsable del crecimiento del tráfico en las redes.

Las aplicaciones como Videoconferencias, streaming en vivo, juegos on-line grupales, etc requieren codificación en tiempo real; mientras que aplicaciones como video on-demand requieren una precodificación del video. Además dependiendo del tipo de video, varían los requisitos mínimos de ancho de banda, espacio de almacenamiento, transmisión uni o bidireccional, etc.; y, por lo tanto, podemos conocer las variables que deben ser cuantificadas para cualquier implementación de vídeo y/o multimedia: direccionalidad, rendimiento, latencia y tolerancia de jitter, así como el número de canales y usuarios. Otra métrica clave es la tolerancia de error; generalmente se requiere una baja tasa de error de paquete, pero puede ser un problema para el usuario del Protocolo de Datagramas de Multidifusión (UDP).

Por otro lado, el espacio de direccionamiento IPv4 está técnicamente agotado; por lo que hay un esfuerzo importante en ir migrando hacia el protocolo IPv6 que tiene un espacio de

direccionamiento mayor, y además, se han adicionado mejoras en las características intrínsecas al protocolo, como en los temas de seguridad, protocolos de ruteo unicast y multicast, seguridad, calidad de servicio, etc.

En este trabajo se desea obtener conclusiones sobre las mejores opciones de tráfico de video en redes IPv6 en laboratorio reales.

Desarrollo y Trabajos de Experimentación

Los trabajos experimentales se realizarán sobre una red pilotode Laboratorio. La topología tendrá una composición mixta de redes cableadas e inalámbricas, routers, switches, access points, y equipos intermedios para dar soporte a los distintos tráficos. Además, equipos finales tanto cableados como inalámbricos que permitan simular una red empresarial de cierta envergadura.

El núcleo de la red tendrá como funcionalidad principal el ruteo de todos los tráficos, sobre el cual se realizarán las configuraciones de los distintos caminos del tráfico multicast, mientras que la conectividad primaria IP se realizará sobre cualquier protocolo de ruteo unicast para IPv6. En los bordes de esta red estarán los posibles usuarios que se podrán unir a los grupos multicast generados.

Se consideran distintas opciones de configuración, teniendo en cuenta la topología básica de la figura 3. Para el caso de clientes inalámbricos, los mismos se reemplazarán por los cableados, conectados a un Access Point correspondiente.

La generación del tráfico de video multicast se realizará de maneras diferentes:

- Con un generador sintético como es el software IPTraffic, que permite generar distintas sesiones de multicast, sobre archivos de videos capturados.
- Con un emisor real, como una cámara de video IP que realice Streaming multicast y/o servidores de video multicast

Una vez realizadas las configuraciones básicas se generará tráfico de video multicast con distintos codecs y de distintos tipos de video; y

se realizarán mediciones de Throughput, cantidad de paquetes, jitter, errores, consumo de ancho de banda, retardo, etc.; para obtener los datos generales y sacar las conclusiones de performance buscadas.

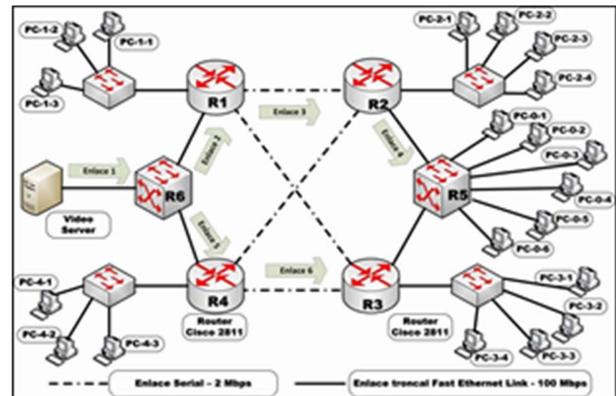


Figura 3: Topología de red

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los temas y líneas de investigación, que se tratarán durante el desarrollo del proyecto son:

- Tráfico multicast
- Direccionamiento multicast en IPv6
- Protocolos de ruteo multicast en IPv6
- Generación de streaming de video
- Códecs de video, especialmente en los actuales, como VP9, H265, etc.
- Generación y análisis de métricas de resultados directas e indirectas
- Análisis estadístico del tráfico de video

3. RESULTADOS OBTENIDOS/ESPERADOS

En proyectos anteriores de análisis de distintos tipos de tráficos, y específicamente de tráfico de video, se obtuvieron datos importantes en cuanto al funcionamiento y rendimiento de multicast frente a unicast en diferentes escenarios, contemplando redes cableadas e inalámbricas bajo el protocolo IPv4.

Para el tráfico de video se utilizaron archivos de video bajo distintos codecs, como MPEG4/2, H.264/3, obteniendo como datos primarios los siguientes:

- Cantidad de bytes y paquetes por códec,
- Tasa de bits,
- Tamaño de paquetes promedio,
- Espacio intertrama,
- Distribución estadística de paquetes por orden de llegada y de espacio intertrama, y
- Comportamiento frente a requisitos mínimos de QoS.

Siguiendo la línea se tiene en cuenta para el presente proyecto los siguientes objetivos:

- Analizar y comparar los distintos tipos de codecs de video actuales, como H265, VP9, etc
- Analizar y comparar las diferentes formas de generar tráfico de video multicast en IPv6
- Comparar el rendimiento del tráfico de video multicast frente a unicast en la medida que la cantidad de miembros multicast aumenta, variando las condiciones de la red

Obtenidos los datos primarios mencionados anteriormente, se buscarán los siguientes objetivos finales:

- Obtener conclusiones sobre la conveniencia de tráfico multicast para el transporte de video a múltiples usuarios dependiendo de tipos de tráfico de video, y
- Obtener conclusiones sobre el rendimiento de los distintos codecs de video de acuerdo a las distintas características de la red
- Obtener conclusiones sobre el rendimiento de los distintos tipos de tráfico de video según los codecs utilizados.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está integrado por docentes investigadores, y becarios graduados y alumnos del Centro UTN CeReCoN (Centro de

Investigación y Desarrollo en Computación y Neuroingeniería) de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.

En el año anterior se presentó y aprobó una tesis de Maestría. Además está la relación de la materia Proyecto Final de la carrera de Ingeniería en Electrónica, en la cual se incentiva que los proyectos finales de los alumnos estén enmarcados dentro de los proyectos de investigación y desarrollo del CeReCoN. Las actividades se llevan a cabo en el ámbito de las instalaciones del Centro, que cuenta con sus propias áreas de trabajo, 1 oficina técnico- administrativa, 2 Laboratorios con 11 computadoras cada uno, con material y con el siguiente equipamiento:

- 4 Routers CISCO 2811,
- 6 Routers CISCO 1721,
- 3 Switchs CISCO 2950,
- 2 Switchs CISCO 2960,
- 2 Switchs CISCO 3560,
- 1 ASA CISCO 5505,
- 2 routers Mikrotik,
- 4 Access Point Cisco y 2 Mikrotik,
- Placas inalámbricas de red,
- 2 cámaras de video IP con soporte de streaming multicast IPv4/IPv6,
- 1 Servidor de streaming de video
- 22 Computadoras con Sistemas Operativos Linux y Windows 7.
- Software IP Traffic de ZTI – Generador de tráfico IPv4/IPv6 unicast/multicast/broadcast y Medidor de performance (throughput, cantidad de paquetes, jitter, número de errores, tanto enviados como recibidos)
- Software Analizador de tráfico Wireshark
- Hardware Air Pcap para captura de tráfico wireless
- Conexión a Internet por IPv4 e IPv6
- Servidor HP Proliant con Linux base y Máquinas Virtuales

5. BIBLIOGRAFÍA

- Hatem BETTAHAR, “Tutorial on Multicast Video Streaming Techniques”, SETIT 2005, <https://pdfs.semanticscholar.org/0336/8dcef3301b9e7f1b5d6aa4bf2aece8cd83ce.pdf>
- IPv6 Multicast Technology, White Paper Hangzhou H3C Technologies Co., Ltd.
- S. Kasera, J. Kuri, “Reliable Multicast in Multi-access Wireless LANS”, Proceedings of INFOCOM99, March, 1999.
- S. Choi and K. Choi, “Reliable multicast for wireless LAN,in Resource, Mobility, and Security Management in Wireless Networks and Mobile Communications, Y. Zhang, H. Hu, andM. Fujise, Eds., CRC Press, Boca Raton, Fla, USA, 2006.
- Estado del arte de IPTV y consideraciones técnicas para su migración a IPv6 en Colombia – Torres, Ramirez, Lopez – <http://ingenieria1.udistrital.edu.co/digital/index.php/redesdeingenieria/article/view/48>
- Estudio de IPTV para la Universidad de Don Bosco – Escobar, Treminio, Eliseo http://rd.udb.edu.sv:8080/jspui/bitstream/123456789/205/1/38873_tesis.pdf
- Asignación equitativa de capacidad para tráfico unicast-multicast en redes IPTV con almacenamiento distribuido de contenidos – Valencia Berrio – Revista en Telecomunicaciones e Informática –V1 – N1 – Junio 2011 – Colombia
- A Resilient Multicast Protocol for Digital TV Over 802.11 Wireless Networks – R.Akester - Department of Computer Science - University College London
- <http://www.docstoc.com/docs/150678614/Empirical-Analysis-of-Video-Multicast-over-WiFi---MWNL> - Yeonchul Shiny, Munhwan Choi, Jonghoe Kooy, Young-Doo Kimz, Jong-Tae Ihmz, and Sunghyun Choi - Department of EECS and INMC, Seoul National University, Seoul, Korea
- Performance Analysis of Multicast Video Streaming in IEEE802.11 b/g/n Testbed Environment – Aleksander Kostuch, Krzysztof Gierlowski, and Jozef Wozniak – 2009 – Gdansk, Poland - http://link.springer.com/chapter/10.1007/978-3-642-03841-9_9#page-2
- Análisis y modelado de multicast interdominio para el soporte de servicios de video – J.Cachinero Pozuelo – 2009 – Universidad Politécnica de Madrid
- Measurement, Modelling and Simulation of Videoconference Traffic from VBR Video Encoders - S. Domoxoudis, S. Kouremenos, V. Loumos and A. Drigas
- S. Winkler, P. Mohandas, “The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics”, IEEE Transactions on Broadcasting, vol 54, no. 3, September 2008

Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos

Jorge R. Osio^{1,2}, Juan. E Salvatore¹, Eduardo Kunysz¹, Diego Montezanti^{1,4}, Daniel Alonso¹, Valentin Guarepi¹, Daniel Martin Morales^{1,3}

¹Instituto de Ingeniería y Agronomía - UNAJ

²UIDET CeTAD –Fac. de Ingeniería - UNLP

³Laboratorio LINES - Universidad Tecnológica Nacional FRLP

⁴Instituto de Investigación en Informática LIDI - Fac. de Informática – UNLP

{josio, jsalvatore, ekunysz, dmontezanti, dalonso, vguarepi, martin.morales}@unaj.edu.ar

Resumen

Las principales líneas de investigación de este proyecto están orientadas a la utilización de arquitecturas multicore para cómputo paralelo en aplicaciones de Telemedicina, Procesamiento de Imágenes Médicas, robótica y de sistemas ópticos basados en MIOC, (Multifunctional Integrated Optical Chip), que requieren procesamiento y transmisión de datos de forma remota.

Paralelamente a estos lineamientos, dentro del marco del proyecto se exploran nuevas tendencias en cómputo paralelo, siendo las tecnologías de computadoras de alta performance reconfigurable (HPCR) una alternativa interesante para optimizar la relación costo-prestaciones.

En relación a las comunicaciones, se investiga principalmente la utilización del protocolo UDP para la transmisión de imágenes y datos de audio y video, (donde es fundamental el envío de grandes cantidades de datos en tiempo real, pero no es crítica la pérdida de algunos datos), en combinación con TCP para la transmisión de pequeños paquetes de datos críticos.

Palabras clave: *arquitecturas paralelas, aplicaciones de telemedicina, sistemas embebidos, multicores, computadoras de alto desempeño reconfigurables.*

Contexto

Las líneas de Investigación descriptas en este trabajo forman parte del Proyecto de Investigación Científico-Tecnológico “Factibilidad y Eficiencia de Aplicaciones de Telemedicina sobre Sistemas Embebidos. Aplicaciones en Arquitecturas Multiprocesador”, que se desarrolla en la Universidad Nacional Arturo Jauretche (UNAJ).

El proyecto cuenta además con financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de los proyectos “kit de desarrollo para SW embebido” y “Sistema de Optimización en el consumo energético”.

Parte de las líneas de investigación desarrolladas se encuentran enmarcadas en los convenios de colaboración en Actividades de Investigación firmados por la UNAJ con la UIDET-CeTAD.

1. Introducción

En los últimos años, se ha buscado expandir el concepto del procesamiento paralelo con computadoras basadas en multicores hacia la utilización de plataformas de procesamiento más específicas. Para obtener mayor eficiencia, los fabricantes de computadoras de altas prestaciones, han introducido unidades FPGA(arreglo de compuertas programables en campo) en su diseño como soporte para el cómputo ([1-3]). Las mejoras con dichos dispositivos se pueden explorar en los siguientes parámetros:

Comunicaciones: dado que los dispositivos modernos permiten interfaces de alta velocidad como Gigabit Ethernet, o PCI Express, se ensayan diversas topologías conocidas y se exploran nuevas alternativas para optimizar el rendimiento general. Para ello, puede utilizarse la capacidad de reconfiguración parcial de los dispositivos de lógica programable [4].

Memoria: la flexibilidad permite que para cada problema específico se pueda utilizar una configuración de memoria determinada. Se cuenta con una cantidad limitada de memoria interna que en general puede utilizarse para procesos locales, y se puede combinar con distintas tecnologías de memoria existentes. Los dispositivos más modernos tienen incorporados módulos de control para memorias del tipo DDR3[5].

Software: es el aspecto más difícil de estandarizar en este tipo de plataformas, debido a que está constituido por drivers que manejan el hardware combinado con lenguajes de alto nivel. El desafío es identificarlos fragmentos de alto nivel que generan mayor *overhead* sobre el procesamiento, para luego implementarlos mediante compuertas lógicas [5].

Si bien el estudio de sistemas paralelos con múltiples procesadores, es un campo bien desarrollado, la utilización de múltiples dispositivos reconfigurables para HPC es un terreno que tiene múltiples posibilidades de exploración [6]. En el presente proyecto se exploran mejoras en comunicaciones y en SW (lenguaje de descripción de hardware VHDL).

Plataformas FPGAs para procesamiento paralelo

La implementación de paralelismo en plataformas FPGAs consiste en el uso de procesadores embebidos para ejecutar aplicaciones y en la utilización de las características que provee la lógica programable para manejar las porciones de código que se ejecutan concurrentemente [6].

La facilidad de implementar procesadores embebidos en forma rápida [5], junto con la posibilidad de proveer concurrencia mediante la programación en HW permiten combinar las FPGAs con el paralelismo obtenido mediante sistemas multicore para alcanzar la máxima eficiencia, que es un gran desafío en la búsqueda de optimizar el procesamiento de imágenes médicas y las señales MIOC, entre otras aplicaciones. Esto se debe a que cada algoritmo es un caso de estudio, en el que se debe decidir qué parte se implementa de manera concurrente y cuál de forma secuencial dentro el sistema multicore. Con las mejoras constantes que aporta la evolución de la tecnología sobre las FPGAs pueden lograrse diseños de gran magnitud, a tal punto que la tendencia actual es implementar microprocesadores de propósito general, conjuntamente con todo el hardware de propósito específico que requiere la aplicación, dentro de una FPGA.

Aplicaciones de transmisión y procesamiento de datos

La transmisión de grandes cantidades de paquetes de datos se realiza mediante el protocolo UDP [7]. Los datos se empaquetan y se envían a destino desde el sistema embebido basado en el LPC1769 y la EDU-CIAA (LPC4337), siendo recibidos en un servidor. Se utiliza un analizador de protocolos para determinar los tiempos de transmisión y estimar la frecuencia de envío mínima para las aplicaciones de interés [8].

La eficiencia del método de transmisión es fundamental para el envío masivo de datos, que principalmente consisten en imágenes médicas de alta calidad, hacia un servidor de almacenamiento exclusivo para telemedicina [9].

Previo a su envío, dichas imágenes requieren ser modificadas mediante la aplicación de diferentes algoritmos de procesamiento de imágenes que tienen características paralelizables, y que son los que posibilitan evaluar la eficiencia y el rendimiento de las arquitecturas paralelas estudiadas.

Adicionalmente, se utilizará la arquitectura paralela para el procesamiento de las señales obtenidas mediante un sistema MIOC, cuya salida (tensión eléctrica) es proporcional a la fase de la señal luminosa que lo atraviesa. Este dispositivo permite medir ángulos de inclinación [10].

Antecedentes del Grupo de Trabajo

Los antecedentes del grupo de trabajo parten de la dirección de tesis de grado sobre la implementación de un web server embebido orientado a telemedicina [11] y de un sistema embebido de transmisión de voz sobre IP [12]. Posteriormente se ha desarrollado investigación acerca de procesamiento de imágenes sobre diferentes arquitecturas paralelas [13].

Las líneas de investigación se enmarcan fundamentalmente en la búsqueda de técnicas innovadoras de procesamiento paralelo, sobre diferentes arquitecturas, que optimicen las prestaciones de sistemas de telemedicina. En particular, resulta de interés su incorporación y aplicación en el Hospital “El Cruce” de Florencio Varela. Adicionalmente, se pretenden encausar temas de estudios de posgrado actuales y futuros de docentes y profesionales surgidos de la UNAJ dentro de estas líneas de investigación. El director del proyecto participa en proyectos de investigación desde el año 2005 en la Facultad de Ingeniería de la UNLP como colaborador y a partir de 2010 como investigador. Actualmente, el grupo de trabajo investiga en los temas: Procesamiento Digital de Imágenes Médicas sobre plataformas FPGA, Procesamiento de Imágenes en arquitecturas FPGA multiprocesador, Envío de datos Ethernet para telemedicina, procesamiento de señales MIOC y Robot Humanoide multicore; pertenecientes a estudios de Especializaciones, Maestrías y Doctorados que se realizan en el marco de acuerdos de colaboración entre el Laboratorio CeTAD, el Lidi y la UNAJ a través del proyecto de investigación.

Como resultado del trabajo realizado recientemente en la UNAJ se han publicado artículos en diferentes congresos nacionales de Ingeniería Informática y Ciencias de la Computación [13 - 15].

2. Líneas de Investigación y Desarrollo

El grupo de investigación que se ha constituido recientemente en la UNAJ es multidisciplinario, y sus miembros

cuentan con experiencia en sistemas multiprocesador, procesamiento de imágenes, obtención y procesamiento de señales provenientes de MIOC, protocolos de comunicación TCP y UDP y robótica.

Temas de Estudio e Investigación

- Implementación de un sistema multiprocesador en Dispositivos Lógicos Programables (FPGAs).
- Análisis y determinación de la eficiencia lograda en el procesamiento de imágenes mediante la combinación de cómputo paralelo y concurrencia.
- Estudio de la frecuencia de refresco de datos para mostrar el estado actual de las variables sobre sistemas basados en microprocesadores ARM Cortex.
- Investigación relacionada con los protocolos TCP y UDP para el envío de paquetes, donde, en condiciones de red conocidas, se deben determinar las tasas de transmisión mínimas para asegurar el funcionamiento óptimo de un sistema de telemedicina.
- Estudio de las características paralelizables de las señales obtenidas de dispositivos MIOC y su procesamiento.
- Determinación de la eficiencia en la ejecución de algoritmos de control de un robot humanoide sobre un sistema multicore.

3. Resultados Obtenidos/Esperados

Investigación experimental

Hasta el momento se han obtenido resultados satisfactorios en relación a la primera parte de los objetivos principales:

- Se obtuvieron resultados en la implementación de un sistema que

permite la transmisión de datos y voz mediante internet utilizando el protocolo TCP, logrando fluidez en la emisión de voz en tiempo real.

- Se implementó un servidor web embebido en un sistema de procesamiento basado en un procesador Cortex M3, el cual permite mostrar variables vitales en tiempo real, como frecuencia cardíaca y temperatura.
- Se incorporó la plataforma CIAA (Computadora Industrial Abierta Argentina) para procesamiento paralelo, la cual contiene dos procesadores, un Cortex M4 y un Cortex M0 y cuenta con un sistema operativo de tiempo real (OSEK). Sobre esta plataforma pudieron probarse las primeras aplicaciones de cómputo paralelo, obteniendo resultados satisfactorios respecto a la ejecución del mismo algoritmo en un sistema monoprocesador.
- Por último, se encuentra en proceso de desarrollo un sistema que permite optimizar el consumo energético en instituciones educativas mediante la ejecución de algoritmos que toman en cuenta diferentes factores. Este proyecto se deriva de los objetivos principales y pretende minimizar el consumo en sistemas multiprocesador.

Para el año en curso, se esperan alcanzar importantes resultados en el área de cómputo paralelo y concurrencia posibilitados por los sistemas basados en FPGAs. Con esta arquitectura, constituida por varios procesadores implementados en una misma FPGA se espera lograr mayor eficiencia que en el sistema basado en la CIAA.

Se pretende medir la eficiencia de ejecución lograda con el sistema

multiprocesador mediante las siguientes aplicaciones:

- Ejecución de algoritmos de procesamiento de imágenes médicas.
- Ejecución de algoritmos de procesamiento de señales provenientes de sistemas ópticos basados en MIOC.
- Ejecución de algoritmos orientados a eficiencia energética.
- Ejecución de algoritmos utilizados en el control de un robot humanoide.

Para lograr una implementación eficiente de este sistema complejo se pretende que el procesamiento paralelo y la transmisión UDP se produzcan conjuntamente en tiempo real.

4. Formación de Recursos Humanos

Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

En este proyecto existe cooperación a nivel nacional. Hay dos investigadores realizando Doctorados y cuatro realizando Maestrías en temas relacionados con simulación de sistemas multiprocesador, sistemas embebidos y software embebido, sistemas multicore en HPC, y transmisión de datos biomédicos mediante UDP.

Adicionalmente, se cuenta con la colaboración de estudiantes avanzados.

5. Bibliografía

- [1] O. Mencer, K. Tsoi, S. Cramer, T. Todman, W. Luk, Ming Wong and P. Leong, "CUBE: a 512-FPGA Cluster", Dept. of Computing, Imperial College London, Dept. of Computer Science and Engineering The Chinese University of Hong Kong. (2009)
- [2] Keith Underwood, "FPGAs vs. CPUs: Trends in Peak Floating-Point Performance", Sandia National Laboratories. (2011)
- [3] D. Encinas, E. Kunysz, Szymanowski, Morales M., "Performance de arquitecturas multiprocesador: técnicas de simulación y plataformas reconfigurables", Instituto de Ingeniería, UNAJ, (2014)
- [4] Dong Yin, Deepak Unnikrishnan, Yong Liao, Lixin Gao and Russell Tessier, "Customizing Virtual Networks with Partial FPGA Reconfiguration", Dept. of Electrical and Computer Engineering University of Massachusetts (2010)
- [5] E. Kunysz, J. Rapallini, J. Osio, "Sistema de cómputo reconfigurable de alta performance (Proyecto HPRC)", 3ras Jornadas ITE - 2015 -Facultad de Ingeniería - UNLP
- [6] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Pearson Addison Wesley, 2003.
- [7] Richard Stevens, "The Protocols TCP/IP Illustrated, Volume 1", 1993
- [8] Jan Axelson, "Embedded Ethernet a internet complete", LLC, 2003
- [9] IEEE, A Unified Multimedia Database System to Support Telemedicine. Transactions of Information Technology in Biomedicine, Vol. 2, No. 3, 1998.
- [10] Rabus, "Integrated Ring Resonators", Springer, Berlín, 2007.
- [11] L. Iogna Pratt, "Implementación de un web server embebido", Fac. Ing., UNLP, 2013.
- [12] J. Salvatore, "Desarrollo de un Sistema de voz sobre IP", Fac. Ing., UNLP, 2012.
- [13] J. Osio, D. Montezanti, M. Morales, "Análisis de Eficiencia en Sistemas Paralelos", Ushuahia, Tierra del Fuego, WICC 2014
- [14] J. Osio, J. Salvatore, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador en aplicaciones de Telemedicina", UNS, Ciudad de Salta, WICC 2015.
- [15] J. Osio, J. Salvatore, E. Kunysz, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", UNER, Ciudad de Concordia, WICC 2016.

Automatización de Sistemas de Cultivos Hidropónicos

Marcelo Marinelli, Graciela Lombardo, Carlos Kornuta, Guillermo Wurn, Lisandro Solonezen.
Cichanowski, Miguel Alejandro

Departamento de Informática, Facultad de Ciencias Exactas Químicas y Naturales. Universidad
Nacional de Misiones
Felix de Azara 1552, Posadas, Misiones
Te: 376-4422186

marcelomarlinelli@fceqyn.unam.edu.ar, gracielalombardo@gmail.com, ckornuta@gmail.com
guillermow77@gmail.com, lsolonezen@gmail.com, miguel.cicha@gmail.com

RESUMEN

El objetivo de esta línea de investigación es desarrollar, implementar y evaluar sistemas de control de distintos dispositivos de cultivos hidropónicos, utilizando sistemas embebidos y minicomputadores de placas de hardware libre, de bajo costo. El software del sistema utiliza técnicas de lógica difusa para evaluar las variables del proceso y generar señales de control. Se evaluaron distintos tipos de sensores de humedad, temperatura, conductividad eléctrica y pH. Se seleccionaron distintas metodologías para el diseño de los controladores difusos. Se desarrollaron sistemas de adquisición de imágenes para verificar en forma remota la evolución del crecimiento de los cultivos.

CONTEXTO

Este proyecto se enmarca en el “Programa de Investigación en Computación” del Centro de Investigación y Desarrollo Tecnológico de la Facultad de Ciencias Exactas Químicas y Naturales (F.C.E.Q. y N.) de la Universidad Nacional de Misiones (U.Na.M.); también vinculado con el Doctorado en Ciencias Aplicadas de la U.Na.M.

Dentro del proyecto se desempeñan docentes, tesistas y becarios de las carreras de Analista en Sistemas de Computación, Licenciatura en sistemas de Información,

Profesorado en Física, Profesorado en Matemática, Doctorandos de la Facultad de Ciencias Exactas, de la Universidad Nacional del Centro de la Provincia de Buenos Aires (U.N.C.P.B.A.), y Doctorado en Ciencias Aplicadas de la F.C.E.Q. de la U.Na.M.

1. INTRODUCCIÓN

La Teoría de los conjuntos borrosos iniciada por Zadeh (1) a diferencia de la lógica matemática permite utilizar términos lingüísticos de fácil interpretación. Una rama de la lógica difusa se ocupa del estudio de sistemas de control (2). Los controladores difusos tienen la ventaja de poseer mayor interpretabilidad para poder ser diseñados en función de un experto en una temática determinada. A tal efecto, se pueden utilizar etiquetas lingüísticas para las variables intervinientes como: “muy bajo”, “bajo”, “alto”, “muy alto”. Por otra parte, los conjuntos difusos que definen las variables, tienen valores de pertenencia flexibles. Estas características hacen que los sistemas de controles difusos tengan una mejor respuesta que los sistemas reactivos convencionales. En las cámaras de germinación hidropónicas una variable importante, en la germinación de las semillas y el desarrollo de los cultivos, es la ventilación, no solo por el aporte de CO₂ sino también por el control de temperatura (3) y (4). Por estas razones se propone el desarrollo de un

sistema de control que tome muestras de temperatura y humedad dentro de la misma, y por medio de inferencias difusas controle el sistema de ventilación forzada y el aporte de temperatura provisto por una lámpara incandescente.

SISTEMA DE ADQUISICIÓN DE DATOS

Las variables de humedad y temperatura se adquieren con sensores DHT22 que permiten tomar muestras de estas variables por una conexión I2C.

Para el procedimiento de adquisición de los datos, software de controlador difuso tipo Sugeno embebido y control de las señales de salida, se utilizó una plataforma de hardware de código abierto *Raspberry PI* (Fig. 1).

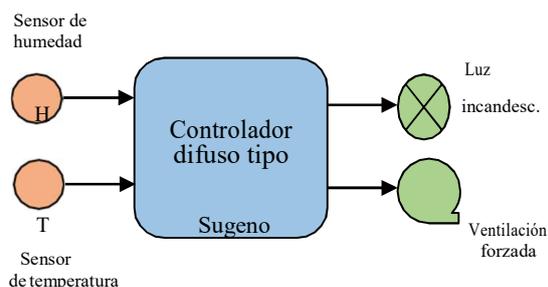


Fig.1: Diagrama del sistema de control

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se está trabajando en distintas líneas de investigación y desarrollo relacionadas con la automatización, utilizando sistemas embebidos y técnicas de I.A., a continuación, se detallan las más importantes:

- Controladores difusos aplicados al proceso de la elaboración de yerba mate.
- Sistemas de control de cultivos hidropónicos.

- Aplicación de controladores difusos desarrollados con la herramienta FIS de Matlab.
- Telemetría con tecnología *bluetooth* para el control de procesos y variables ambientales.
- Desarrollo de aplicaciones para dispositivos móviles y tecnología *bluetooth*, con el objeto de usar los dispositivos móviles para el control de procesos.

3. RESULTADOS OBTENIDOS Y ESPERADOS

En el marco de este proyecto se concluyó con la tesis doctoral “**Control Automatizado de cultivos Hidropónicos mediante Lógica Difusa**” (Marinelli 2015), del Doctorado en Ciencias Aplicadas de la Universidad Nacional de Misiones. También se concluyeron dos tesis de grado de licenciatura en Sistemas de Información: **Desarrollo de un Sistema de Navegación y Telemetría en Tiempo Real para Un Robot Móvil desde un Smart Phone Via Wifi (IEEE 802.11)**, Autor: Cichanowski, Miguel Alejandro Y **Desarrollo e implementación de un controlador difuso para la navegación de robots móviles utilizando el algoritmo de Wang & Mendel**, Autor: Lisandro Solonezen.

Se realizaron las siguientes publicaciones con referato:

- Marinelli, M., Acosta, N., Toloza, J. M., & Kornuta, C. (2016). Control difuso de una cámara de germinación. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- Marcelo Marinelli, Guillermo Wurm. (2016). Sistema de adquisición en tiempo real de conductividad eléctrica y pH en nutrientes hidropónicos. IV Seminario Argentina-Brasil de Tecnologías de la Información y la

Comunicación (SABTIC 2016). ISBN 978-987-3619-15-1.

Presentaciones a congresos:

- XXII CACIC2016 (Congreso Argentino de Ciencias de La Computación) 3 al 7 de octubre de 2016 U.N.S.L, San Luis.
- IV Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación (SABTIC 2016). 4 y 5 de noviembre de 2016, UNNE, Corrientes.

Como resultados esperados se desarrollará un sistema de control de variables basado en *Internet of things*, abreviado (IoT) (5), para el control de las variables ambientales y actuadores en el proceso de control de cultivos hidropónicos. Este sistema permite integrar la representación de los datos, los sistemas de telemetría y comunicación de alarmas en forma integrada con los servicios de Internet.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto se forman dos doctores, un Doctorando de CONICET y se dirigen dos tesis de grado de licenciatura en Sistemas de Información de la F.C.E.Q. y de la Universidad Nacional de Misiones. F.C.E.Q. y N. de la U,Na,M.

Se prevén incorporar tesis de la Maestría en Tecnología de Información y del Doctorado en Ciencias Aplicadas de la F.C,E,Q, y N. de la U.Na.M.

5. BIBLIOGRAFIA

1. Zadeh L.: The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 1(8), 199–249 (1975).
2. Nguyen H., Prasad N.: *Fuzzy Modeling and Control: Selected Works of Sugeno*. CRC Press, Boca Raton, Estados Unidos. (1999).

3. Domingues D., Hideaki W. Takahashi B., Camara C. & Nixdorf. S.: Automated system developed to control pH and concentration of nutrient solution evaluated in hydroponic lettuce production. *Computers and Electronics in Agriculture*. 84, 53–61. (2012).

4. Mohamed M. & Alsadon A. (2010). Influence of ventilation and sucrose on growth and leaf anatomy of micropropagated potato plantlets. *Scientia Horticulturae*, 123, 295–300.

5. Campoverde, A. M., Hernández, D. L., & Mazón, B. E. (2015). Cloud computing con herramientas open-source para Internet de las cosas. *Maskana*, 6(Supl.), 173-182.

Diagnóstico y Metodología para la Implementación de "Internet of Things" en el Planeamiento y Desarrollo de Ciudades Inteligentes

Gustavo Mercado¹, Luis Álvarez², Luis Bocaccini³, Marcelo Ledda¹, Javier Membrives³, Maximiliano Muros³,
Lucas López³, Damián Juárez³, Laura Guillén³,
Florencia Gómez³, Gonzalo Favaro³, Gonzalo Ciperiani³, Sofía Dumé³, German Berra³.

¹GridTICs – Grupo en Tecnologías de la Información y las Comunicaciones
Departamento de Electrónica / Facultad Regional Mendoza / UTN
gustavo.mercado@gridtics.frm.utn.edu.ar

²IRESE – Instituto Regional de Estudios de Energía
Departamento de Electromecánica/Facultad Regional Mendoza/UTN
alvarez.irese@frm.utn.edu.ar

³CIDER: Centro de Investigación y Desarrollo Regional
Facultad Regional San Rafael / UTN
libocaccini@frsr.utn.edu.ar

Resumen

El paradigma de “Ciudades Inteligentes”, como aquellas que adoptan soluciones novedosas para el mejoramiento de la vida de los ciudadanos, tienen cada vez más relevancia y protagonismo. Desde el punto de vista “tecnológico”, normalmente se emplea el concepto de “Internet de las Cosas” (IoT Internet of Things) para el desarrollo de las Ciudades Inteligentes. Este proyecto pretende entender y analizar las distintas alternativas de IoT en ciudades de la provincia de Mendoza y determinar, de acuerdo a las características y necesidades propias de cada ciudad, un modelo para el desarrollo de la Ciudad Inteligente. Se debe abordar el problema analizando las necesidades de las ciudades, con un enfoque a la sustentabilidad, habitabilidad y desarrollo urbano. Se llevarán a cabo acciones para determinar las necesidades y proponer las soluciones que utilicen tecnologías de IoT y de Smart Cities. Se cuenta con la colaboración, mediante convenios, de las autoridades y del personal técnico de los municipios intervinientes. (Mendoza y San Rafael), También se llevan a cabo encuestas ciudadanas con el propósito de verificar el nivel de apropiación de las tecnologías de IoT, por parte de los ciudadanos. Al final se genera y se propone una solución tecnológica de IoT y se determina un orden de prioridades y costos. Las ciudades intervinientes, podrán utilizar estos conocimientos para avanzar en el desarrollo propio de “Ciudad Inteligente”.

Palabras Clave: Smart Cities, Smart City,

Internet of Things, Internet de las Cosas, Ciudades Sustentables, Desarrollo Urbano

Contexto

El presente trabajo está inserto en el proyecto de investigación acreditado por la Universidad Tecnológica Nacional código ENTUNME0004275 denominado “Mendoza(s) Smart Cities”: *Diagnóstico y metodología para la implementación de "Internet of Things" en el planeamiento y desarrollo de ciudades inteligentes: Caso de estudio ciudad de Mendoza y ciudad de San Rafael*. El proyecto es del tipo “Tutorado” y llevado adelante por investigadores y becarios de la Facultad Regional Mendoza y Facultad Regional San Rafael.

Introducción

Una de las principales particularidades que llevo al hombre a ser la especie dominante en el planeta tierra fue su capacidad de sociabilizar. Esta capacidad lo llevo a vivir en sociedad y agruparse desde principios de la especie, en lo que hoy denominamos ciudades. Incluso, es tan común la vida en las diferentes ciudades del mundo, que plantearse el hecho de una vida en completo aislamiento de nuestros pares, es cada vez más complicado.

Sin embargo, no fue hasta los tiempos de la evolución industrial, cuando esas ciudades se convirtieron en los motores del crecimiento económico. Esta revolución aportó prosperidad a muchos países, pero su constante desarrollo no siempre fue de la manera “más inteligente” [1]. En

muchos casos, la superposición de un gran número

de individuos en una zona en particular llevó a tener problemáticas complejas, como por ejemplo la expansión directa de enfermedades, generando plagas que, en varias oportunidades, disminuyeron notablemente el número de habitantes.

La vida moderna, ha llevado al aglutinamiento de muchos individuos en ciudades, generando esto mismo, diferentes desventajas, como lo son: mayor consumo de energía, emisión de mayor cantidad de residuos, congestiones en las llamadas zonas céntricas, etc. Para solucionar este tipo de inconvenientes se han creado planes de desarrollo para el crecimiento sustentable de las ciudades. El objetivo es tratar de generar y desarrollar elementos tecnológicos, en innovación permanente, que no sólo mejoren la calidad de vida de los ciudadanos, sino que también mejoren los estándares de eficiencia y sustentabilidad de recursos para mejorar la productividad y reducir emisiones de desechos contaminantes.

Es así como surge el concepto de Smart city, o ciudad inteligente, la cual se puede describir como [2] aquella ciudad que aplica las tecnologías Internet de las Cosas (IoT – Internet of Things) con el objetivo de proveerla de una infraestructura que garantice:

- Un desarrollo sostenible.
- Un incremento de la calidad de vida de los ciudadanos.
- Una mayor eficacia de los recursos disponibles.
- Una participación ciudadana activa.

Por lo tanto, son ciudades que son sostenibles económica, social y medioambientalmente. La Smart City nace de la necesidad de mantener una armonía entre estos aspectos.

En Argentina el crecimiento de Ciudades Inteligentes es dispar, pero se denota una consistente tendencia hacia la modernización y sustentabilidad en determinadas ciudades [3].

El objetivo particular de la implementación de IoT es la resolución de problemáticas modernas las cuales mencionaremos brevemente a continuación:

Estacionamiento inteligente, monitoreo de aparcamientos disponibilidad en la ciudad.

El smartparking o estacionamiento inteligente [4], permite la ayuda a conductores para reducir el tiempo de búsqueda de lugares de estacionamiento en las zonas más congestionadas de capitales y ciudades más importantes del mundo. Esto contribuye a evitar la contaminación del aire, ya que en muchos casos el tiempo de búsqueda es

excesivo. Zonas importantes, como la de Westminster en Londres, son pioneras en la aplicación de estas tecnologías. No sólo les ha permitido reducir la cantidad de congestionamientos en zonas urbanas, sino que su aplicación viene acompañada de una reducción de emisiones contaminantes. Los más famosos ejemplos están enumerados en [5], entre los que se destacan Autostadt CarTowers - Wolfsburg, Alemania, el Smart Tower Europa o el estacionamiento flotante Umi-Hotaru - Tokio, Japón.

Bicycling, gestión de transporte servicios ciclisticos.

El BID (Banco Interamericano de Desarrollo), en el ciclo de inclusión, en las ciudades de América Latina y el Caribe (ALC) [6], deja constancia que uso de la bicicleta como medio de transporte urbano está creciendo significativamente y está ayudando a mejorar la movilidad, la equidad y el acceso a oportunidades socioeconómicas. El ciclismo urbano es una opción de movilidad con alto potencial para reducir algunos de los problemas de nuestras ciudades: la congestión de tráfico, la mala calidad del aire y las emisiones de gases que contribuyen al cambio climático. La bicicleta como opción de transporte contribuye al desarrollo y a la competitividad de las ciudades.

Estos beneficios sólo pueden materializarse si se construye o adapta infraestructura ciclo-incluyente, si los ciudadanos se involucran en los procesos de planeación de movilidad, si se adoptan políticas que regulen el uso y si las secretarías de movilidad cuentan con la información necesaria para monitorear el desempeño del uso de la bicicleta.

Mapas de ruido urbano: monitoreo de sonido en zonas de bar y zonas céntricas en tiempo real. La contaminación sonora no es un problema menor. Nuestros oídos son muy sensibles al ruido, y cualquier ruido superior a 85 decibelios (DB) se considera perjudicial para el oído humano. Según el Instituto Nacional de la Sordera y Otros Trastornos de la Comunicación de los Estados Unidos (National Institute on Deafness and Other Communication Disorders) [7], nos exponemos a por lo menos 85 decibelios de ruido cuando oímos los sonidos del tráfico de la ciudad. Y pérdida de la audición no es la única consecuencia potencialmente peligrosa de entornos ruidosos. La Organización Mundial de la Salud de las Naciones Unidas informa que la hipertensión arterial, aumento del nivel de estrés y ataques de ansiedad más frecuentes son parte de una larga lista de posibles efectos. En respuesta a todo esto surgen los mapas de ruido urbano, los cuales monitorean

en tiempo real los niveles de ruido de zonas conflictivas de distintas ciudades mediante el empleo de teléfonos inteligentes [8].

Congestionamientos de tráfico: control de vehículos y peatones niveles para optimizar las rutas de conducción y caminar.

Se espera que número de vehículos en las carreteras del mundo se duplique a alrededor de 2,5 millones en 2050. Sin embargo, en lugar de aumentar la congestión, la contaminación, las colisiones y la angustia de cercanías como era de esperar, todo lo contrario puede llegar a pasar [9].

¿Cómo puede haber más vehículos, pero menos tráfico en el futuro? Ciertamente, los gobiernos no están poniendo dinero - y concreta - en la construcción de más carreteras. Hoy en día, la tecnología está revolucionando casi todo lo imaginable sobre cómo llegar del punto A al punto B.

Armado con datos de los sensores en las carreteras y los vehículos, las ciudades ya están utilizando el análisis predictivos para anticipar y reducir la congestión del tráfico. Los conductores utilizan las redes sociales para detectar y evitar el estancamiento. Los urbanistas analizan datos para establecer claramente que las nuevas rutas de autobús más se necesitan. Incluso la forma en que parque está siendo reinventado.

Ciudades de todo el mundo se enfrentan a un rápido crecimiento y las crecientes dificultades de transporte. Pero ese crecimiento también proporciona oportunidades para construir sistemas de transporte inteligentes que mejoren fundamentalmente cómo las ciudades y los ciudadanos gestionar sus rutas de transporte.

Iluminación inteligente: adaptación en luces de la calle y edificios.

En las ciudades y edificios inteligentes [10], los niveles de iluminación pueden variar en función de variación de la luz natural, evitando así el derroche de energía, reduciendo los costos, y aumentando el confort. Atenuaciones de la iluminación en función de nuestra actividad: luces tenues de orientación en zonas de paso o con actividades como ver una película, niveles y configuraciones relajantes para una iluminación ambiental de descanso o tertulia, zonas de luz intensa para la lectura o una iluminación más viva cuando necesitamos que el ambiente nos transmita energía. Esto hace de su vivienda un lugar de pleno confort en cada ocasión.

Gestión de residuos: detección de los niveles de basura en contenedores para optimizar las

rutas de recolección de basura.

Desde la década de los ochenta, la importancia de la política de la Unión Europea sobre protección del medio ambiente y los recursos naturales ha ido aumentando sin cesar. La razón se debe a que las amenazas de daños y deterioro que pesan sobre el medio ambiente distan mucho de estar controladas. Por fortuna, cada vez son más las personas que, conscientes de los peligros latentes, exigen medidas de protección más decididas a nivel nacional y mundial.

En el siglo XX, cuando la población mundial creció y se tornó más urbana y próspera, la producción de residuos aumentó en diez veces. Al día de hoy, una persona media en Estados Unidos desecha el equivalente a su peso corporal en la basura cada mes. Por ello, la gestión de los residuos sólidos es uno de los mayores gastos en los presupuestos municipales.

Según la Agencia Europea de Medio Ambiente (AEMA) [11], Austria, Alemania y Bélgica son los países que reciclan mayor proporción de residuos urbanos. En total, en Europa se reciclaron el 35% de los residuos urbanos en 2010, una mejora significativa respecto al 23% registrado en 2001. Sin embargo, el estudio indica que a muchos países les resultará extremadamente complicado alcanzar los objetivos impuestos por la UE de reciclar el 50% de los residuos domésticos y similares para el año 2020.

De acuerdo con el informe, el Reino Unido aumentó el porcentaje de reciclado de residuos urbanos de un 12% a un 39% entre 2001 y 2010, mientras que Irlanda incrementó las tasas de reciclado de un 11% a un 36% durante el mismo período. Eslovenia, Polonia y Hungría también han mejorado notablemente las tasas de reciclado desde su adhesión a la UE. Las tasas de reciclado más elevadas se registran en Austria, con una tasa del 63%, seguida de Alemania (62%), Bélgica (58%) y los Países Bajos (51%), que son los únicos miembros que ya superan el objetivo de reciclaje de residuos en un porcentaje del 50%. En cuanto a Suiza, aunque no es miembro de la Unión Europea, está considerado como un país ejemplar en el trato de residuos, con un porcentaje del 51% de basuras urbanas tratadas y una tasa de reciclaje supera el 95% en materiales como el vidrio. España, por su parte, se sitúa próxima al 35%.

Red Inteligente de energía de seguimiento y gestión del consumo “Smart Grids”

Desde su descubrimiento, la electricidad fue un movilizador de avances tecnológicos en todas las ramas de la actividad humana. La llegada de la red

eléctrica fue por sí un símbolo de progreso y mejora en la calidad de vida de las personas. Esa filosofía de progreso y de mejora continua se encuentra vigente, y cada vez con mayor fuerza en la actualidad. Sin embargo, estructuralmente no ha sufrido una transformación sensible que la adapte a las nuevas necesidades del mercado eléctrico.

Hasta poco tiempo atrás, el paradigma dominante consistió en la universalización del servicio. La excepción fue, por su criticidad, la red eléctrica de alta tensión. La red eléctrica del futuro requiere un salto cualitativo, no cuantitativo. Debido a la necesidad de administrar mejor los recursos energéticos, favorecer la protección del medioambiente y responder a los requerimientos cada vez más exigentes de calidad de servicio y producto, surge el concepto denominado Redes Eléctricas requerimientos cada vez más exigentes de calidad de servicio y producto, surge el concepto denominado Redes Eléctricas Inteligentes [12] (del inglés, Smart Grids). Este concepto se basa en incorporar a la red eléctrica tradicional dispositivos electrónicos tales como medidores, sensores o mandos; vinculados a través de distintas tecnologías de comunicación, logrando la centralización y uso de la información para provecho de todos los actores involucrados; optimizando el funcionamiento del sistema eléctrico. De esta forma es posible que las empresas de servicios puedan administrar eficientemente sus activos y que el usuario final gestione su consumo en forma racional.

Objetivos y Resultados

Objetivo Principal

Diagnosticar, evaluar y ponderar soluciones de sustentabilidad urbana, utilizando tecnologías inteligentes (IoT), aplicadas en casos de estudio locales.

Objetivos secundarios

- Diagnosticar y determinar los problemas de sustentabilidad urbana que puedan ser solucionados con tecnologías inteligentes.
- Evaluar los problemas y soluciones y estimar y valorar las tecnologías inteligentes disponibles.
- Ponderar y determinar prioridades de necesidades/soluciones.
- Diseñar e Implementar prototipo de tecnología inteligente como modelo de concepto y prueba tecnológica
- Documentar y difundir los resultados.

Avances y resultados preliminares

Al momento se está llevando adelante la primera tarea de la metodología del proyecto:

DIAGNOSTICAR:

Relevamiento de datos existentes: se relevarán con los actores intervinientes los datos correspondientes a actividades y trabajos relacionados con la sustentabilidad urbana y suburbana. A tales efectos se realizarán entrevistas con los encargados de planificación urbana y territorial de los municipios de interés, con funcionarios responsables de las áreas, y personas consideradas clave o expertas a los efectos de cubrir la mayor amplitud posible de datos existentes y de planificaciones acordadas previamente.

Se relevarán como ejes temáticos los siguientes: conectividad, movilidad, calidad ambiental, energía, habitabilidad y conectividad.

Las tareas que se realizarán son:

- Identificación de actores clave en los cuales relevar
- Formulación de convenios con instituciones - Caso de estudio Ciudad de Mendoza y Ciudad de San Rafael
- Relevamiento de antecedentes documentales y personales
- Estudio, en detalle, de tecnologías aplicables en smart cities, incluyendo disponibilidad y costos

Formación de Recursos Humanos

Uno de los principales objetivos del proyecto es la capacitación de los recursos humanos.

La meta como investigadores es fortalecer la capacidad para realizar investigación científica, generar conocimientos y facilitar la transferencia de tecnología que permita el desarrollo humano.

Este proyecto de investigación posibilita la colaboración inter-institucional y la ejecución de proyectos conjunto entre grupos I+D de diferentes disciplinas.

Para lograr estos objetivos se dispondrá del siguiente personal:

- 4 Investigadores formados
- 1 Investigador de apoyo
- 1 Becario doctoral (beca UTN)
- 2 Becarios graduados (Beca BINID UTN)
- 4 Becarios alumnos (Beca alumno UTN) 1
- Tesista de posgrado de especialización 2
- Tesistas de carrera de grado

Adicionalmente se realizarán:

- Dictado de Cursos, Seminarios y Conferencia para público especializado de la región.

- Promoción, coordinación y asistencia técnica de tesis de grado para alumnos de Ingeniería de Sistemas de Información e Ingeniería Electrónica de la FRMza
- Promoción, coordinación y asistencia técnica a pasantes alumnos, cursantes de carreras de grado y de pre-grado en el ámbito de la UTN FRMza
- Promoción, coordinación, dirección y asistencia técnica a Tesis doctorales, postgrado y/o maestría.
- Presentación de Trabajos en Congresos y Reuniones Técnicas/Científicas.
- Publicación de Trabajos en revistas con/sin referato.

Referencias

- [1] Carol L. Stimmel, "Building Smart Cities: Analytics, ICT, and Design Thinking", Auerbach Publications.
- [2] Peter Waher, "Learning Internet of Things", 1st Edition, O'Reilly Ed
- [3] Capellán, Norberto, "Ciudades Inteligentes EL APORTE DE LAS TIC A LA COMUNIDAD" 1a ed, 2016, ED Autores de Argentina
- [4] Stephen Goldsmith, Susan Crawford "The Responsive City: Engaging Communities Through Data-Smart Governance", Jossey-Bass
- [5] Aparcamientos más extraños del mundo, <http://listas.20minutos.es/lista/aparcamientos-mas-extranos-del-mundo-365003>, visitado 23/2/2017
- [6] BID, "Ciclo de inclusión, en las ciudades de América Latina y el Caribe - Guía para impulsar el uso de la bicicleta", Banco Interamericano de Desarrollo, 2015.
- [7] National Institute on Deafness and Other Communication Disorders (NIDCD) <https://www.nih.gov/about-nih/what-we-do/nih-almanac/national-institute-deafness-other-communication-disorders-nidcd>, visitado 5/3/2017
- [8] Mike Barlow. "Smart Cities, Smarter Citizens", O'Reilly Ed.
- [9] GSMA Connected Living Programme "Intelligent Transportation Systems Report for Mobile", 2015 GSM Association
- [10] Universidad Externado de Colombia, "Guía de territorios y ciudades inteligentes" Segunda edición 2013
- [11] Agencia Europa de Medio Ambiente. <http://www.eea.europa.eu/es>, visitado 23/2/2017.
- [12] Gustavo Mercado et.al, "SG-SM - Smart Grid San Martin Red de Distribución y Generación de Energía Inteligente en Ciudad Gral San Martin – Mendoza" Anales del WORKSHOP DE INVESTIGADORES EN CIENCIAS DE LA COMPUTACION WICC 2015 16 y 17 de abril de 2015 - Salta - ARGENTINA

Estudio de Herramientas de Simulación en Redes Definidas por Software

D. Bolatti, R. Calcagno, C. Cuevas, S. Gramajo, R. Scappini,
J. Torre, M. Acuña, M. Marturet

Grupo Ingeniería en Sistemas de Información / Universidad
Tecnológica Nacional / Facultad Regional Resistencia
French 414, (3500) Chaco, Argentina. Tel. 362-4432683
{dbolatti, rcalcagno, cac, sergio, rscappini}@frre.utn.edu.ar,
{marilau.ml, juliana.itorre, matiasmarturet}@gmail.com

Resumen

A medida que aumentan los servicios y el software, las exigencias de usuarios para contar con disponibilidad, garantía, calidad y seguridad de la información son mayores. Las arquitecturas de redes existentes no fueron diseñadas para satisfacer dichas necesidades, provocando que las organizaciones de hoy en día, se vean limitadas en posibilidades. Por este motivo, se impulsa a reexaminar las arquitecturas actuales, buscando responder con nuevas técnicas y tecnologías.

La llegada de la virtualización ha flexibilizado la definición y el uso de sus recursos, permitiendo definir y modificar en tiempo real, a nivel software, una infraestructura completa basada en perfiles de aplicaciones, en necesidades de rendimiento, etc. El análisis de infraestructura de redes y DataCenters orientados a servicios se vuelve un elemento crítico en las empresas de telecomunicaciones.

En este contexto, se desarrolla un proyecto de investigación que estudia nuevas tendencias de redes de información como las Redes Definidas por Software y propone un modelo basado en esta tecnología, a fin de cubrir nuevas demandas de usuarios y organizaciones.

Palabras clave: SDN, Data Center Virtual, Controlador, GNS3, Mininet, OpenDayLight.

Contexto

Este proyecto está inserto en una línea de I/D presentada en la Universidad Tecnológica Nacional con código: UTN-2422. Título: “*Modelo para la evaluación de performance mediante identificación de tráfico y atributos críticos en Redes Definidas por Software*”. Dicho proyecto se lleva a cabo en el ámbito del Dpto. de Ingeniería en Sistemas de Información perteneciente a la Facultad Regional Resistencia de la Universidad Tecnológica Nacional.

Introducción

Un datacenter es un centro de procesamiento de datos, una instalación empleada para albergar un sistema de información de componentes asociados, como telecomunicaciones y los sistemas de almacenamientos. Generalmente incluyen diferentes dispositivos de seguridad para permitir que los equipos tengan el mejor nivel de rendimiento con la máxima disponibilidad del sistema.

Un datacenter ofrece varios niveles de resguardo, en forma de fuentes de energía de backup y conexiones adicionales de comunicación, que puede no ser utilizada hasta que pase algún problema en el sistema primario. El principal objetivo de estas instalaciones es ejecutar las aplicaciones centrales del negocio y almacenar datos operativos. En él residen los sistemas de softwares corporativos más tradicionales, como ser Enterprise

Resource Planning (ERP) y Customer Relationship Management (CRM), entre otros.

En el diseño e implementación de un DataCenter, es de particular interés la utilización racional y eficiente de los recursos disponibles, y si esto se puede hacer en forma dinámica y adaptable a la demanda, resulta de una importancia central. La tecnología emergente denominada **SDN (Redes Definidas por Software, -Software Defined Networking)** aplica de manera ajustada este concepto separando el plano de control del plano de reenvío, donde el plano de control controla varios dispositivos. Esta separación permite al control de la red volverse directamente programable y abstrae la infraestructura a las aplicaciones y servicios de red [1].

En este sentido se comporta como una arquitectura dinámica, manejable, económica y adaptable; ideal para las aplicaciones de naturaleza dinámica de hoy en día que usan gran ancho de banda con características destacables como:

- Directamente programable: el control de la red es directamente programable porque se desacopla de las funciones de reenvío.
- Ágil: la abstracción del control del reenvío permite a los administradores ajustar dinámicamente el flujo de la red ante las necesidades cambiantes.
- Configuración programable: SDN permite a los administradores de red configurar, administrar y optimizar los recursos de red rápidamente de manera dinámica con programas SDN automatizados.
- Estándar abierto y neutral: SDN es un estándar abierto que simplifica el diseño y las operaciones de red, donde las instrucciones son proveídas por el controlador SDN en lugar de múltiples dispositivos y protocolos de vendedores específicos.

La arquitectura SDN, especificada por

la ONF (Open Networking Foundations) que se ilustra en la Figura 1, muestra a alto nivel, los puntos de referencia e interfaces al controlador [1-2].

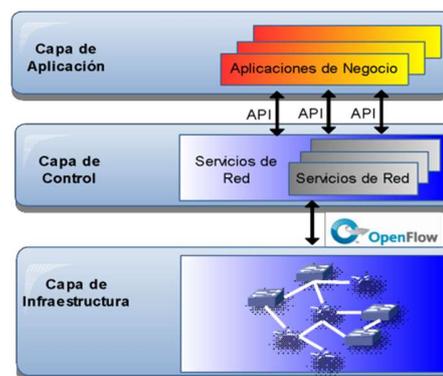


Figura 1. Arquitectura de Software-Defined Network

Para poder concretar esta nueva arquitectura fue necesario crear y estandarizar una interfaz de comunicaciones entre el control y el reenvío de datos. Para ello se creó el protocolo OpenFlow que permite el acceso directo a la gestión de datos de reenvío en dispositivos de red como switches y routers, tanto físicos como virtuales y de un modo abierto.

Ningún otro protocolo estándar tiene la funcionalidad y finalidad de OpenFlow que transfiere el control de los dispositivos de red a la lógica del software de control

OpenFlow utiliza el concepto de flujos para identificar el tráfico de red basado en reglas estáticas o dinámicas programadas por el software de control de SDN. También permite definir el modo en que el tráfico debe rutearse a través de los dispositivos. Debido a que OpenFlow permite ser programado sobre una base de flujo, una arquitectura basada SDN-OpenFlow proporciona un control muy granular de la infraestructura de red, lo que permite responder a los cambios en tiempo real en la aplicación, el usuario e inclusive en los niveles de sesión.

Una característica particular de SDN

basada en OpenFlow es que se puede implementar en las redes existentes, tanto físicas como virtuales, lo que hace simple la introducción progresiva de esta tecnología, incluso en los entornos de red de múltiples proveedores.

Líneas de investigación y desarrollo

En el proyecto marco “*Modelo para la evaluación de performance mediante identificación de tráfico y atributos críticos en Redes Definidas por Software*”, en el que se encuentra incluido el presente trabajo, se propone (i) el análisis de la arquitectura y estándar de SDN, (ii) el desarrollo de la hipótesis de trabajo: “implementación de SDN en un Data Center Virtual”, (iii) la propuesta de un escenario de red para implementar un modelo de estudio para una red overlay SDN en un Datacenter, y finalmente (iv) la transferencia de conocimientos a organizaciones de la región a fin de incursionarlas en las nuevas técnicas y tecnologías.

Con el objeto de llevar a un entorno de laboratorio la arquitectura mencionada se utilizaron un conjunto de productos base, listados a continuación:

- Simulador de Red Gráfico GNS3

Es un simulador de red gráfico que permite virtualizar, planificar, probar y solucionar problemas de una red entera a través de la plataforma escalable de cualquier vendor, sin la necesidad de interactuar directamente con el hardware de red. Lo que diferencia a GNS3 de los demás simuladores, es la habilidad para emular ruteo y conmutación como así también incorporar verdaderas máquinas virtuales y conectarlos entre sí a través de un sistema de túneles lógicos (Overlay Network)[3].

- Mininet

Es un emulador de red que crea una red de hosts, switches, controladores y

enlaces virtuales. Los hosts corren un software de red Linux estandarizado, los switches soportan OpenFlow con gran flexibilidad para la configuración de ruteo y SDN. Mininet permite investigar, desarrollar, aprender, armar prototipos, testing, debugging, y otras tareas con el beneficio de tener una red experimental funcionando dentro de cualquier PC.[4]

- OpenDayLight (ODL)

Es una plataforma de código abierto para redes programables definidos por software, ideal para obtener una amplia gama de opciones de configuración, y hay varias nuevas aplicaciones construidas sobre ella que hacen la transición a SDN aún más fácil.

En particular para este trabajo se utilizó la distribución Helium, desarrollada con una serie de complementos cuyo objetivo es permitir la interacción con el controlador de forma más transparente. De esta forma se trabajó con Karaf; una plataforma genérica que proporciona funciones y servicios de alto nivel diseñados específicamente para la creación de servidores basados en OSGi. Esta distribución viene con una serie de proyectos incluidos que son fáciles de instalar gracias a Karaf.[5]

- Wireshark

Es uno de los más famosos analizadores de protocolos de red en el mundo (estándar de facto). Permite ver lo que está sucediendo en la red con un gran nivel de detalle. Como características principales podemos destacar que es mantenido bajo la licencia GPL, es muy robusto, puede capturar datos de la red o leer datos almacenados en un archivo, basado en la librería pcap, con una interfaz muy flexible, con gran capacidad de filtrado, admite entre otros el formato estándar de archivos tcpdump, permite la reconstrucción de sesiones TCP, se ejecuta en diferentes plataformas, es un opulento analizador VoIP, y los datos

pueden ser leídos en tiempo real, dependiendo de la plataforma, desde Ethernet, IEEE 802.11 u otros.

Resultados y objetivos

La topología elegida para la propuesta y estudio se definió en base a la infraestructura de RadioBash.com, una empresa dedicada al servicio de streaming, que requiere grandes flujos de datos. [6]

La topología para el escenario de simulación (Figura 2) está implementada en GNS3 para realizar pruebas de conectividad. Incluye 3 routers, siendo R1 y R2 los routers de núcleo vinculados a R3, permitiendo redundancia y balanceo de carga, y habilitando así la posibilidad de simulaciones más complejas en el futuro.

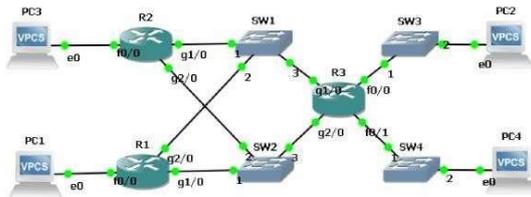


Figura 2. Topología implementada en GNS3

El acceso está representado por dos redes externas indicadas aquí por f0/0 del R1 200.45.10.0 y 200.45.11.0 por f0/0 del R2. Las PC 1 y 3 representan las nubes de entrada y en esta figura están a los efectos de prueba de conectividad; las PC 2 y 4 representan los servidores.

Las pruebas iniciales se efectuaron implementando protocolo de ruteo RIP2[7] a los efectos de comprobar el funcionamiento básico. Posteriormente se implementó el protocolo de ruteo OSPF[8] que ofrece la suficiente flexibilidad de configurar el escenario para simulaciones de mayor complejidad.

Para la implementación de SDN en el escenario, se incluyen dos máquinas virtuales:

1. Controlador SDN OpenDayLight, consistente en una MV corriendo Ubuntu

Server 64 bits con la distribución Heliium.

2. Mininet, para la creación de la red virtual con soporte OpenFlow, consistente en una imagen importada en Virtual Box de MiniNet / Ubuntu VM[9]. Esta máquina virtual incluye en sí MiniNet, todos los binarios OpenFlow, herramientas pre-instaladas, y ajustes para la configuración del kernel para soportar las redes Mininet. La topología de prueba puede verse en la Figura 3.

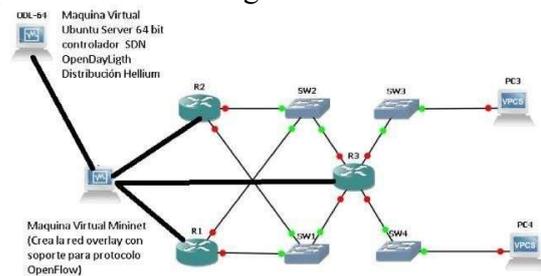


Figura 3. Escenario de prueba para SDN

En trazo grueso se indica la red virtual con soporte para el protocolo OpenFlow que posibilita la implementación de SDN en la topología de estudio.

El Controlador ODL puede ser iniciado desde una terminal conectada a la MV ODL-64, ubicado en el directorio de la distribución Karaf ingresando el comando `./bin/karaf`.

Luego para arrancar el proceso de la red overlay, desde una terminal conectada a la MV Mininet:

```
~$sudo mn -controller=remote, ip=(ip correspondiente a la MV ODL) -topo=linear,3 -mac -
```

```
~$switch=ovsk,protocols=OpenFlow13
```

Este comando crea la red SDN con tres switch y tres host que se conectarán a los tres routers de la topología de estudio. Además desde esta terminal se comprueba la conectividad de la red, y se asocian las interfaces de cada router con las interfaces de los host virtuales de mininet.

Para visualizar la topología y los

respectivos nodos, se puede ingresar desde el navegador a la dirección *http://(ip correspondiente a la MV ODL):8181/index.html*, donde como usuario administrador se pueden observar los nodos (Node Connectors) y estadísticas de cada conexión.

El presente trabajo describe solamente la elección de las herramientas y la propuesta del modelo para el estudio de implementación de SDN sobre una red IP convencional. Esta instancia de pruebas iniciales no incluye la simulación con tráfico sino pruebas de conectividad que nos permitirán ajustar las configuraciones del controlador SDN en conjunto con la MV Mininet para la inclusión de la red overlay con soporte OpenFlow. En esta etapa del trabajo se pretende alcanzar con el modelo la suficiente funcionalidad que permita; en una etapa posterior, hacer pruebas de tráfico específicas para diferentes perfiles del mismo.

Los trabajos futuros de simulación forman parte de otra etapa dentro del proyecto que busca la identificación de atributos que puedan utilizarse en modelos de decisión. Con el conocimiento adquirido de la arquitectura y estándar de red SDN, y de las herramientas disponibles a fines de simularla, se espera poder instruir a organizaciones de la región en esta nueva tendencia en infraestructura, a fin de que busquen suplir con nuevas técnicas y tecnologías las demandas crecientes de los usuarios.

Formación de Recursos Humanos

El equipo de trabajo está integrado por un director, Dr. Ing. Sergio Gramajo, dos investigadores realizando su tesis de maestría con temáticas afines a la del proyecto, Ing. Carlos Cuevas e Ing. Diego Bolatti, dos investigadores, Ing. Ricardo Calcagno e Ing. Reinaldo José Ramón Scappini, y tres becarios.

Dos becarios son alumnos avanzados de la carrera de ingeniería en sistemas de información y un becario graduado de iniciación a la investigación. Esto hará posible fomentar la actividad de investigación en alumnos que están próximos a recibirse y graduados jóvenes estimulando la actividad de investigación.

Por otro lado el equipo cuenta con la ayuda de colaboradores externos, el Dr. Luis Martínez López (Catedrático Universidad de Jaén, España), y la Dra. Macarena Espinilla Estévez (Investigadora Universidad de Jaén).

Referencias

- [1] O. N. F. (ONF). SDN Architecture.
- [2] O. N. F. ONF, "Software-Defined Networking: The New Norm for Networks," ed, 2012.
- [3] GNS3. *Frequently Asked Questions*. Available: <https://www.gns3.com/software/faq>
- [4] M. Park. *ON.LAB Mininet*. Available: <https://www.sdxcentral.com/projects/on-lab-mininet/>
- [5] R. Hill. (2016, *Karaf: Step by Step Guide*. Available: https://wiki.opendaylight.org/view/Karaf:Step_by_Step_Guide
- [6] RadioBash.com. *About RADIOBASH.COM*. Available: <http://radiobash.com/about-us/>
- [7] T. I. E. T. F. (IETF®), "RIP Version 2," ed.
- [8] T. I. E. T. F. (IETF®), "OSPF Version 2," ed.
- [9] *Mininet VM Images*. Available: <https://github.com/mininet/mininet/wiki/Mininet-VM-Images>

Implementación de Middleware Publicador/Subscriber para Aplicaciones Web de Monitoreo

Defossé N., López R. A., Marcelo E. Gómez, Konstantinoff P., Wahler S., Castro L., Harris G.

Departamento de Informática Trelew, Facultad de Ingeniería, Universidad Nacional de la Patagonia

Mitre 655 – CP (9100) Trelew Tel.: (+54280) 4428402

{nahuel.defosse, lopez.ricardo, gomezmarcelo, pkonstantinoff, sebastian.wahler} @gmail.com

RESUMEN

En la actualidad, millones de clientes se conectan a la nube utilizando el protocolo HTTP (Protocolo de Transferencia de Hipertexto). Históricamente, el estudio de los sistemas distribuidos ha propuesto diversas técnicas de optimización de acceso; como ser Sun RPC, CORBA, SOAP. Sin embargo, en el ámbito de las aplicaciones web una tendencia muy generalizada es la abstención a este tipo de *middlewares*, en favor de principios arquitectónicos propuestos como REST (*Representational State Transfer*) [1].

Uno de los problemas que presenta esta mecánica de comunicación, es la imposibilidad que un cliente reciba actualizaciones de un recurso remoto sin iniciar un requerimiento. Una técnica actual para mitigar el problema consiste en la utilización de WebSockets.

Paralelamente, las bases de datos relacionales han ganado capacidades de notificación a través de canales asincrónicos. Estas pueden ser aprovechadas para recuperar cambios en los datos de tiempo real.

En este trabajo se presenta el desarrollo de una aplicación web con requerimientos de actualizaciones de estado en tiempo real. Inicialmente se plantea su diseño exponiendo recursos bajo REST, para luego abordar un enfoque Publicador Subscriber sobre los mismos recursos utilizando *middleware* basado en WebSockets.

Palabras Clave: Websocket, REST, MQTT, Real Time Web

CONTEXTO

El presente trabajo se encuentra dentro de las líneas de investigación del proyecto “Computadoras Industriales con control WEB”, del Departamento de Informática de la Facultad de Ingeniería de la Universidad Nacional de la Patagonia San Juan Bosco.

1. INTRODUCCIÓN

Una API REST se conforma de un conjunto de URIs (Uniform Resource Identifier), las cuales dependiendo del método HTTP que el cliente utilice tienen una semántica diferente. Un ejemplo de definición de métodos HTTP ante URIs puede apreciarse en Tabla 1.

Por ejemplo, si el sistema en el servidor se tratase de un RDBMS (*Relational DataBase Management System*), una petición GET podría asociarse a una instrucción SQL SELECT, con POST a un INSERT, con PUT a un UPDATE y DELETE a su homónimo SQL.

Ciertos encabezados HTTP cobran relevancia como el caso de *Content-Type*, definiendo la codificación o representación del estado del recurso a ser transferido, siendo popular la utilización de Javascript Object Notation o JSON por su versatilidad de operación del lado del cliente.

URI	GET	PUT	POST	DELETE
Colección, como <code>/resources</code>	Listar las URIs de colección.	Reemplazar toda colección.	Crear un nuevo elemento. La nueva URI se asigna automáticamente.	Borrar la colección entera
Elemento, como <code>/resources/item</code>	Recuperar representación del miembro de la colección, expresado en el tipo de medio apropiado (Content Type)	Reemplazar el elemento actualizar. Crearlo si existe.	No utilizado generalmente por ya incluir identificación. Se utiliza para la creación de la identificación de colección.	Borrar el elemento referenciado en la colección.

Tabla 1. Descripción de métodos HTTP bajo REST.

En nuestro estudio se desarrolló una aplicación web de monitoreo de estaciones transformadoras eléctricas. La cual está conformada por una base de datos, un módulo de recolección de estados y eventos y un módulo de visualización.

El módulo de visualización está compuesto por gráficos vectoriales redimensionables (SVG) encargados de mostrar el estado del campo en tiempo real al operador. Estos gráficos son almacenados en una tabla de la base de datos llamada PSVG (por Pantalla SVG).

Cada uno de los archivos SVG está constituido por varios elementos o nodos -que representa algún estado de la estación-. Estos elementos poseen un identificador (TAG) que los relaciona con una fila de una tabla llamada EG (Elemento Gráfico). En cada columna de esta tabla se almacenan el estado de los atributos relleno (*fill*), contorno (*stroke*), texto (*text*) y un *timestamp* que indica la marca temporal de la última actualización.

La tabla EG es actualizada periódicamente por un proceso de recolección de estados y eventos.

La aplicación web descarga un documento HTML con su correspondiente código JavaScript. Este accede a las tablas previamente mencionadas como recursos REST. El diagrama de secuencias de esta operación puede observarse en la Figura 1. En función de la pantalla (o diagrama SVG) seleccionado, se realiza una petición asincrónica GET a `/api/v1/eg/?`

`psvg_id=<identificador_pantalla>`, que recupera el estado inicial de la estación y lo aplica a cada nodo del SVG haciendo *queries* al DOM, utilizando el TAG y modificando atributos (*fill*, *stroke* y *text*). Durante este proceso guarda el *timestamp* mayor. Luego cada 5 segundos se repite la petición, pero agregando como parámetro a la misma, el último *timestamp* recuperado, para filtrar solo las novedades y aplicando los cambios que sean recuperados a los nodos SVG referenciados.

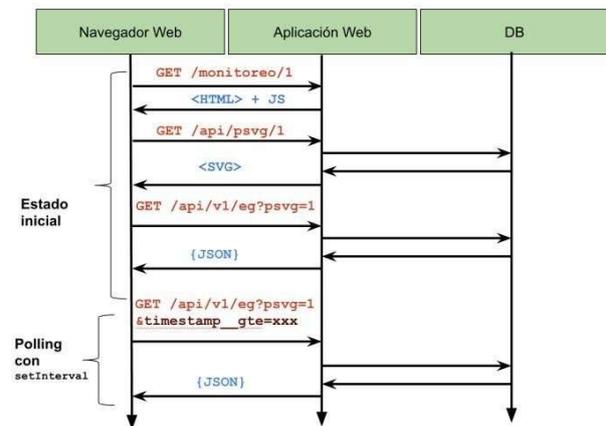


Figura 1: Secuencia de recuperación de estado.

Dado que la carga de 160 elementos iniciales rondaba los 240Kb y su actualización era de 16Kb por petición, detectamos un *overhead* innecesario en el módulo de visualización y consecuentemente en la base de datos. Sumado a esto, el uso de peticiones cronometradas genera desfases y retrasos del estado real.

Para solucionar estos problemas abordamos la detección de cambios dentro del dominio de la base de datos. Para esto, PostgreSQL provee las instrucciones `NOTIFY` y `LISTEN` [6], estas permiten publicar y recibir mensajes a través de un canal. Basándonos en estas funcionalidades, se desarrolló una utilidad de línea de comandos llamada `pg_notifications`, que agrega un *stored procedure* (SP) y una serie de *triggers* ante los eventos `INSERT` o `UPDATE` de las tablas que se indiquen como argumento. Para la publicación de datos, el RDBMS provee la información acerca de cuál tabla y fila activaron el *trigger* que invoca el SP encargado de convertir la fila afectada a formato JSON y publicar con `NOTIFY` al canal.

Para la recepción de datos del canal se desarrollaron una serie de *callbacks*, entre los cuales se encuentran el *logging* por salida estándar o invocación de un comando.

Para llevar estas actualizaciones al módulo de visualización se requiere una vía hacia el navegador web donde el cliente no deba realizar peticiones cada cierto tiempo. Para esto se investigó el protocolo WebSocket, el cual consiste una comunicación bidireccional entre un servidor web y un cliente. A diferencia de HTTP, donde se transporta texto, WebSocket está orientado a tramas. El tipo de tramas a transmitir se negocia al establecimiento de la conexión. El formato de estas tramas se conoce como sub-protocolos de WebSocket y se encuentran establecidos por la Autoridad de Nombres de Internet (o IANA por sus siglas en inglés).

Message Queuing Telemetry Protocol MQTT[3] fue el sub-protocolo seleccionado para el envío de notificaciones de nuestra aplicación web.

MQTT es un *middleware* orientado a mensajes (MoM)[2] basado en el paradigma Publicador Subscriptor [3]. Se compone de un broker central al cual se conectan los cliente.

Al conectarse un cliente, si éste desea recibir publicaciones bajo un tópico, lo hace saber al broker mediante un mensaje *Subscribe*. Una vez que el *broker* acepta la suscripción, las publicaciones que se correspondan con el tópico serán enviadas al cliente suscrito.

Finalmente, se vincularon los mensajes recuperados por la utilidad `pg_notifications` y se los envió al broker como publicaciones.

Del lado del código HTML, se procedió a adaptar el código Javascript -que utilizaba un *timer* para disparar la llamada AJAX-, sustituyéndolo por la utilización de una biblioteca para realizar comunicaciones MQTT. Una característica de los mensajes *Publish*, es que aceptan una bandera llamada *retain*, que le indica al *broker* que debe guardar el mensaje. Cualquier cliente que se suscriba posteriormente al tópico de un mensaje retenido, recibirá éste último. La utilización de este flag permitió completar en la suscripción, gran parte del estado inicial de la estación, pero se debió agregar además un proceso que publique el estado de los elementos que no hubieren cambiado.

Finalmente, las llamadas a `/api/v1/svgelement/` se reemplazaron por suscripciones al tópico `/api/v1/EETT/+`, donde EETT es la PSVG y el + representa el comodín mono-nivel.

El Topic Mapper se encarga de publicar las actualizaciones de `fill`, `stroke` y `text` de los EG con el tópico correspondiente.

En la figura 3 pueden observarse los componentes del sistema y su interacción. La línea punteada identifica el flujo de los eventos asincrónicos que son enviados al cliente, atravesando sus diferentes fases: detección en el trigger, publicación en el canal de la base de datos, recolección por el Topic Mapper, asignación de tópico y publicación al broker

MQTT y por último, recepción por parte de los clientes suscritos.

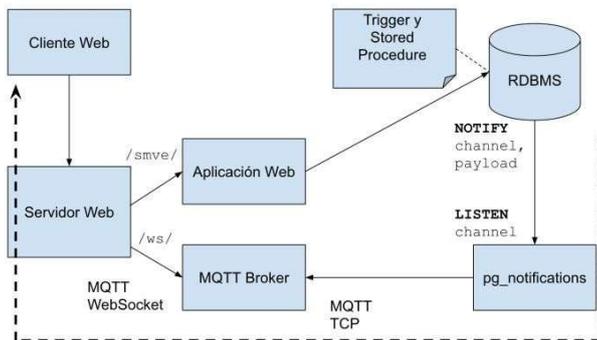


Figura 2: Aplicación adaptada a MQTT-WC.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El estudio de aplicaciones para monitoreo de centrales eléctricas involucra una serie de elementos, cada uno de ellos con diferentes desafíos específicos. Actualmente contamos con las siguientes líneas de investigación dentro del grupo.

- Sistemas SCADA [7].
- Protocolos de comunicación industrial.
- Actualizaciones asincrónicas con MQTT-WS.
- Microcontroladores DSP y aplicaciones WEB.
- Protocolos IoT.

3. RESULTADOS OBTENIDOS

En este trabajo se muestra una mejora en el desempeño y la experiencia de usuario de una

aplicación de monitoreo de centrales eléctricas, puntualmente un proceso de *polling* de 5 segundos pasó a requerir fracciones de segundo para observar cambios en un diagrama SVG. La eliminación del *polling* redujo la cantidad de recursos requeridos por la aplicación web en estado de reposo. Además hemos aprovechado características reactivas del RDBMS PostgreSQL, haciendo innecesarias las consultas basadas en ventana de *timestamp*.

Una captura de pantalla del sistema en funcionamiento puede observarse en la Figura 3.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por docentes y alumnos del grupo de investigación relacionado con el proyecto enunciado, dirigido por el Mg. Ing. Ricardo López y el Lic. Marcelo Gómez.

El Lic. Nahuel Defossé se encuentra desarrollando su trabajo de tesis de Maestría dirigido por el Dr. Fernando Tinetti y co-dirigido por el Mg. Ing. Ricardo López. Los alumnos Lucas Luis Castro y Gonzalo Harris realizaron su tesina de licenciatura implementando actualizaciones asincrónicas con MQTT-WS.

El alumno Pedro Konstantinoff se encuentra desarrollando su tesina de grado en relación a temáticas de IoT, reutilizando estos resultados.

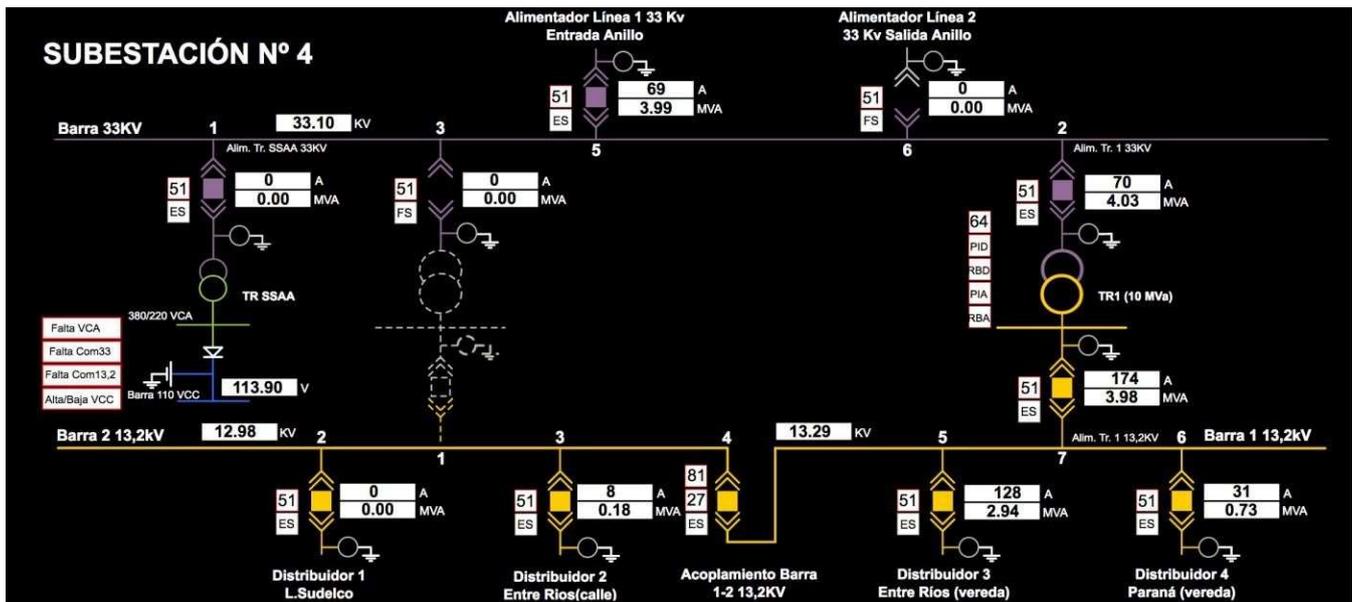


Figura 3: Pantalla SVG en funcionamiento

5. BIBLIOGRAFÍA

- [1] "Fielding Dissertation: CHAPTER 5: Representational State Transfer"
https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm. Fecha de acceso 25 feb.. 2017.
- [2] Coulouris, George F. Distributed Systems: Concepts and Design. Boston, MA: Addison-Wesley, 2011.
- [3] Especificación MQTT de OASIS. Sitio Web: <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html> (fecha de acceso)
- [4] Architectural impact of the SVG-based graphical components in web applications. Moreno, Olivera. 2008.
- [5] Real Time Water Supply System Hydraulic and Quality Modeling – A Case Study. Sadok , Gomes , Eisenhauer , Kelner.
- [6] Listen/Notify en PostgreSQL. Sitio Web: <https://www.postgresql.org/docs/9.1/static/sql-notify.html> (fecha de acceso)
- [7] Defossé, N. , van Haaster, D. M. , Pecile, L., Tinetti, F. G. , & Lopez, R. A. (2014). Implementación de Sistemas SCADA Utilizando Lenguajes de Alto Nivel. *WICC 2014 XVI Workshop de Investigadores en Ciencias de la Computación* (Vol. XVI, pp. 779–784). Ushuaia. [8]López, R. A., Pincioli, E., & Tinetti, F. G. (2014). Microcontroladores asociados a medición y comunicaciones en sistemas SCADA de energía. *Workshop de Investigadores En Ciencias de La Computación, XVI*, 767–773. [9] Gómez, M. E., Wahler, S.

P., Tinetti, F. G., & López, R. A. (2014). Implementación de mensajes rápidos y valores de muestreo IEC61850 sobre Ethernet con microcontroladores. *Workshop de Investigadores En Ciencias de La Computación*, 774–778.

Implementando IoT en la Universidad Nacional de Misiones

Milton E. Sosa, Eduardo O. Sosa, Diego A. Godoy

Secretaría de Investigación y Posgrado (SECIP). Facultad de Ciencias Exactas,
Químicas y Naturales; Félix de Azara 1552 - N3300LQH - Posadas

Universidad Nacional de Misiones (UNaM)

{mesosa; es; dgodoy}@fceqyn.unam.edu.ar

Resumen

La mayoría de los aparatos y de oficina en la Universidad de Misiones son controlados manualmente, ergo la iluminación, ven-tilación y acondicionamiento de aire de los ambientes de los escenarios elegidos también lo son. Estas acciones conducen al despilfarro de energía. Se ha comprobado en la UNaM que los dispositivos permanecen conectados inclusive en ausencia de alumnos/docentes, no existiendo a la fecha contralor y relación alguna de temperaturas ambiente y/o iluminación artificial/natural. El presente trabajo en el seno de la Facultad, pretende iniciar un sistema de control automatizado inteligente de equipos y aparatos, pretendiendo alcanzar un ahorro económico, al mismo tiempo que minimizar el efecto invernadero del consumo de energía. La solución planteada se lleva a cabo con componentes económicos utilizando la infraestructura física existente. El sistema consiste en un módulo de ahorro de energía por contralor de iluminación y de acondicionamiento de aire considerando la presencia de personas en los ambientes. La central permite el accionamiento/control de los módulos a distancia por medio a de una aplicación en la red desarrollada al efecto.

Palabras clave: Internet de las Cosas (IoT), Smart Buildings, Sistemas embebidos, Redes de Sensores

Contexto

El proyecto se relaciona con proyectos sobre aplicaciones de conceptos de Internet de las Cosas (IoT), Internet del Futuro y sistemas inteligentes. Es el producto de proyectos de la Secretaría de Investigación y Posgrado de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones Proyecto 16Q519 “Ambientes Inteligentes. Una Mirada a Internet del Futuro” y que deviniera en el proyecto 16Q581 “Eficiencia Energética y Ambientes Inteligentes. Investigación y Desarrollo Experimental en la UNaM.

La descentralización de la Facultad de Ciencias Exactas Químicas y Naturales, y conjuntamente con ello la necesidad de realizar un uso eficiente de la energía, ha llevado a las autoridades a plantear un trabajo acciones puntuales en la Facultad, para lo cual se ha desarrollado íntegramente un prototipo, pretendiendo la optimización del uso de la energía eléctrica, ya sea por acción u omisión de los docentes, administrativos y/o alumnos. Se han coordinado actividades y consultas al Centro de Investigación sobre Tecnologías de la Información y Comunicación (CITIC) de la Universidad Gastón Dachary (UGD) [1], de manera de avanzar en el tema realizando las consultas pertinentes a otros involucrados en la cuestión.

Introducción

La Internet de las Cosas (IoT, Internet of Things) representa un nuevo ambiente de cosas u objetos unívocamente identificables, los cuales pueden comunicarse entre ellos formando una red dinámica de alcance global [2]. La idea de IoT ha evolucionado y hoy en día las cosas no sólo se limitan a objetos identificables por RFID, sino que también involucran sensores, actuadores, o cualquier otro tipo de elemento conectados a una red física e identificable unívocamente

Los entornos inteligentes representan escenarios avanzados de aplicación de las TIC's, mediante las cuales interaccionan objetos de uso cotidiano en continua evolución como herramientas transparentes para usuarios no experimentados. Los entornos inteligentes han surgido rápidamente como un paradigma nuevo y emocionante que tiende a incluir diferentes campos de investigación de la computación ubicua o pervasiva y en red. Podemos definir a los ambientes inteligentes como *“aquellas tecnologías significativas que se mimetizan entretejiéndose en la trama de la vida cotidiana hasta que se tornan indistinguibles de la vida misma”* [3]. Ya en año 1999 el *Information Society Technologies Advisory Group* ha utilizado el término "Inteligencia Ambiental"[4] para describir una visión donde *“las personas estarán rodeadas de interfaces inteligentes e intuitivas embebidas en objetos cotidianos de nuestro alrededor y un entorno que reconoce y responde a la presencia de individuos de manera invisible”*. Un entorno inteligente es entonces la combinación de un entorno físico con una colección de sistemas embebidos dedicados a la recopilación de información de un escenario determinado en un contexto heterogéneo, existiendo una infraestructura capaz de almacenar, compartir, entender y manejar esta información. Al espacio integrador de los elementos mencionados se lo denomina entorno inteligente.

En los edificios inteligentes se utilizan microcontroladores y sensores para automatizar

diversos parámetros, entre ellos la iluminación y acondicionamiento de aire, constatando la presencia de individuos tendiendo al ahorro energético y una utilización más eficiente de la energía y al confort del individuo. Domótica es un término de la tecnología de la información y la comunicación que se utiliza en edificios donde diversos componentes se comunican a través de la red local. Esta tecnología se puede aplicar para controlar, prevenir y actuar de acuerdo con los criterios seleccionados. Esta tecnología es capaz de comunicarse automáticamente con otras “cosas” del entorno a través de Internet, teléfonos fijos o teléfonos móviles. *Un edificio inteligente es un tipo de edificio con tecnología instalada propia para hogares inteligentes* [5].

La integración de las redes de sensores inalámbricos con la red cableada tradicional plantea diversos retos técnicos [6], centrados fundamentalmente en el desarrollo de sensores y de la infraestructura de redes de sensores.

Una de las cuestiones más difíciles de considerar tiene que ver con las respuestas a las siguientes cuestiones: ¿Cómo evaluar, construir y mantener un servicio en un ambiente inteligente de manera sustentable? ¿Cómo puede la sociedad donde habitamos disfrutar de éste pretendido estándar de vida elevado sin afectar a la naturaleza y el ambiente? Para hallar una respuesta válida a las cuestiones precedentes, debemos hacer converger a los diferentes campos tecnológicos y sociales para que cooperen en el establecimiento del servicio pretendido. La evaluación de la sustentabilidad de las aplicaciones ha comenzado a ocupar un lugar preponderante en la concurrencia de mejorar la tecnología y reducir el consumo de recursos naturales existentes [7].

La década pasada surgió en Alemania un interesante concepto: las oficinas inteligentes. Gran parte de nuestra vida la pasamos en la oficina, por lo que la misma debería ser un sitio amigable para con el personal y el entorno. ¿Qué es una oficina inteligente? Se considera así a aquella que conoce su actividad, siendo capaz de mensurar consumo, gasto y desperdicio de los recursos que son indispensables para su

operación. Debemos considerar que los edificios, en forma global son responsables del consumo de 17% del agua dulce, 15% de productos maderables, 33% de emisiones de CO₂ y 40 % de energía y materiales [8].

Cook [9] define a los ambientes inteligentes como aquellas tecnologías que se presentan como sensibles, receptivas, adaptables, transparentes, omnipresentes e inteligentes. Todos estos ambientes deben permitir la posibilidad interacción entre los diferentes elementos que conforman los distintos entes. Como interoperabilidad [10], se conoce a la capacidad de las entidades participantes en una comunicación de: 1) compartir cierto tipo de información, 2) procesar esa información de acuerdo a una semántica operacional previamente acordada.

Los hogares y oficinas inteligentes unifican el hecho de utilizar sensores inteligentes y las tecnologías de la información de forma transparente para electrodomésticos y elementos de oficinas [11]; tratando de administrar el uso eficiente de la energía a través de redes inteligentes. Otro término en continua evolución es *Smart Grid*, siendo básicamente la construcción de redes de energía avanzadas fusionando las redes eléctricas existentes y los sistemas de comunicación para controlar los dispositivos de la red de alimentación. Los beneficios de la red inteligente no se limitan a la

red pública, sino también las instalaciones residenciales, industriales y comerciales.

Una arquitectura basada en multi-agente al que se aplican la inteligencia y la colaboración para actuar hacia objetivos específicos distribuidos (definidos por el usuario) puede resolver los procesos de decisión complejos, dinámicos, tal como los objetivos planteados en esta esta implementación para la gestión de energía en un ambiente inteligente.

En el dominio emergente de las redes de sensores existe una amplia variedad de dispositivos con diferentes capacidades, desde sistemas relativamente potentes con procesadores de clase PC y las interfaces inalámbricas de banda ancha (IEEE 802.11), a otros de mucho menor capacidad, siendo los nodos de baja potencia. Los consumidores en ese entorno precisarán indefectiblemente de herramientas de automatización avanzadas, las cuales deberán ser capaces de implementarse por medio de tecnologías avanzadas de sensores [12].

En el presente trabajo se pretende avanzar en el contralor y accionamiento remoto de diversos correctivos/preventivos ante situaciones que se puedan presentar en escenarios remotos de la FCQyN.

Como primer escenario se han seleccionado 2 aulas en el campus universitario ubicado en Miguel Lanús.

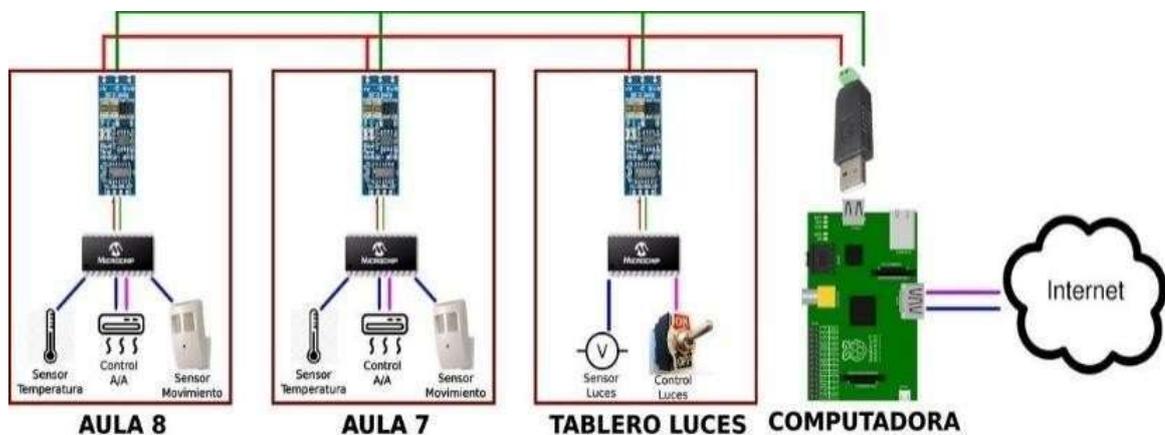


Ilustración 1. Esquema de instalación en escenarios elegido

Avanzar hacia la gestión inteligente de la energía requerirá cambios no sólo en la forma en que se suministra la energía a los domicilios y comercios, sino también en los mecanismos necesarios para control de dispositivos y para la gestión del fluido provisto.

Líneas de Investigación

En el seno de la Facultad de Ciencias Exactas, Químicas y Naturales se desarrollan actividades en proyectos de investigación que presentan un ordenamiento disciplinar más acotado a procesos constitutivos del campo de la Tecnologías de la Información y Comunicaciones, con una lógica de trabajo que ha nutrido de experiencias, prácticas y saberes centrales al efecto de dar cuenta de las mejores formas de avanzar sobre temas del dominio especificado. Por ello los estamentos involucrados proponen el desarrollo de líneas de investigación que, transversalmente, den cuenta de los objetivos perseguidos originando compilación de ideas proyecto, desarrollos, captura de información, conceptualización, análisis y producción de elementos utilizables en escenarios reales, con posibilidades ciertas de extrapolación trabajos del futuro.

Hemos establecido una sucesión continua y ordenadora de actividades de estudio, reflexiones sistemáticas y creativas, tormentas de ideas y discusiones alrededor del tema Internet de las Cosas, enlazando entre si uno o varios proyectos de trabajo para así desarrollar actividades académicas e intelectuales dinámicas, en medio de búsquedas, hipótesis, logros, metas y objetivos para producir, construir y aumentar conocimientos sobre el tema AMI. De esta manera actuamos colectivamente a situaciones que demandan procesos sistemáticos fundamentalmente interdisciplinarios.

Materiales, Métodos e Implementación

El hardware ideado, desarrollado e implementado forma parte de una red de

topología bus (maestro/esclavo), para lo cual el medio de comunicación es una adaptación híbrida, por lo que es apta para la implementación de comunicaciones por medio de cables ó en forma inalámbrica.

El nodo maestro, que coordina al resto de los nodos, está basado en una computadora de bajo costo Raspberry Pi [13] con un adaptador USB-EIA/485 [14] desde el cual inicia el bus de comunicaciones RS-485 [15], realizando múltiples saltos de manera que cada nodo esclavo forme parte de la red. Este nodo se encuentra solidariamente configurado para formar parte de la red IPv4 de módulo Campus de la Facultad de Ciencias Exactas, Químicas y Naturales.

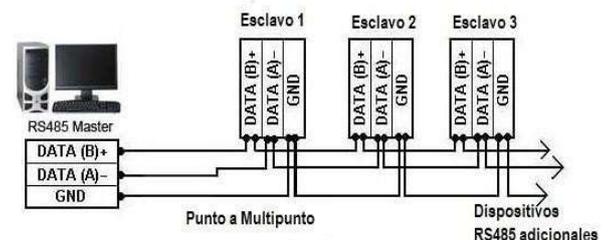


Ilustración 2. Bus de datos EIA RS-485

Los nodos esclavos desarrollados específicamente para esta aplicación y situados en las aulas, se basan en microcontroladores de bajo costo PIC 18F2550 [16]. Se ha elegido este modelo principalmente debido a que se dispone de todas las herramientas, tanto de compilación como de programación. Su arquitectura de 16 bits es suficiente para desarrollar las operaciones que requiere este proyecto y dispone de un port USB para la conexión con un ordenador. El micro ha sido programado para medir la temperatura ambiente del aula por medio de un sensor de bajo costo LM35 [17].

La condición de trabajo de los equipos de aire acondicionado se releva a través de un sensor de falta de fase ubicado en bornes de la alimentación del mismo. La alimentación de la energía eléctrica de los acondicionadores es controlada por medio de contactores del tipo normal cerrado. Todos los componentes son controlados y accionados por el microcontrolador; del cual deben recibir un

comando específico desde el nodo maestro. A fin de realizar control presencial de individuos en los ambientes controlados e intervenidos, el microcontrolador releva el estado de una señal proveniente de sensores de movimiento standard convencional, del tipo utilizado frecuentemente en sistemas de alarmas domiciliarias.

Físicamente el nodo se ubica en el tablero de luces, cumpliendo la función de monitorear el estado de las luminarias de las aulas e interrumpir la corriente eléctrica (al igual que en el caso de las aulas) en el único caso de que se envíe un comando específico desde el maestro.

Para lograr una alternativa de comunicación híbrida, se incluyó un nodo que adapta los niveles de señal del estándar EIA485 a TTL de 3.3V utilizados por los módulos de radio XBEE los cuales operan en la banda de 900 MHz. Estos nodos no participan de forma activa de la red generando información ni actuando sobre otros parámetros, únicamente cumplen la función de cambiar de medio de comunicación.

La configuración de estos nodos de RF es de tipo "transparente", es decir la red de sensores inalámbrica se comporta como si continuase siendo de tipo cableada. Esto permite generar saltos en forma inalámbrica y continuar con una red cableada en el otro extremo.

Formación de Recursos Humanos

Dps de los integrantes realizan su Maestría en el ámbito de las Tecnologías de la Información y Comunicación. En el año 2016 el director del proyecto a dirigido la tesis de Doctorado de uno de los integrantes obteniendo la calificación sobresaliente "Magna Cum Laude" en la Universidad de Vigo en España.

Referencias

- [1] Universidad Gastón Dachary, [En línea]. Available: www.ugd.edu.ar.
- [2] E. Borgia, «The Internet of Things vision: Key features, applications and open issues,» *Computer Communications*, vol. 54, pp. 1-31, 2014.
- [3] M. Weiser, «The computer for the 21st century,» vol. 265, pp. 66-75, Septiembre 1991.
- [4] Information Society Technologies Advisory Group, «Orientations for Workprogramme 2000 and beyond,» 1999. [En línea]. Available: <http://ow.ly/uRncG>.
- [5] H. Aspelund, T. Laberg y H. Thygesen, «Planning and management in municipal services, Norway: Directorate for Social and Health».
- [6] M. Gaynor, S. Moulton, M. Welsh, E. LaCombe, A. Rowan y J. Wynne, «Integrating wireless sensor networks with the grid,» *Internet Computing, IEEE*, vol. 8, nº 4, pp. 32-39, 2004.
- [7] J. Kjeldskov, M. B. Skov, J. Paay y R. Pathmanathan, «Using mobile phones to support sustainability: a field study of residential electricity consumption,» de *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012.
- [8] The World Green Building Council (WorldGBC), «Better places for People,» [En línea]. Available: <http://www.worldgbc.org/>.
- [9] D. Cook, J. Augusto y V. Jakkula, «Ambient intelligence: Technologies, applications, and opportunities,» *Pervasive and Mobile Computing*, vol. 5, nº 4, pp. 277-298, 2009.
- [10] L. Brownsword, D. Carney, D. Fisher, G. .. Lewis, E. Morris, P. Place, J. Smith, L. Wrage y B. Meyers, «Current Perspectives on Interoperability,» Pittsburgh, 2004.
- [11] U. Premarathne, I. Khalil y M. Atiqzaman, «Trust based reliable transmissions strategies for smart home energy consumption management in cognitive radio based smart grid, Ad Hoc Networks,» 2016.
- [12] R. J. ROBLES y T.-h. KIM, «Applications, Systems and Methods in Smart Home Technology: A Review,» *International Journal*

of Advanced Science & Technology, vol. 15,
2010.

- [13] raspberry Pi Foundation, [En línea]. Available: <https://www.raspberrypi.org/>.
- [14] Future Technology Devices International Limited, [En línea]. Available: <http://bit.ly/1TeEIC5>.
- [15] Maxim Integrated, «Guidelines for Proper Wiring of an RS-485 (TIA/EIA-485-A) Network,» [En línea]. Available: <http://bit.ly/1L09YoK>.
- [16] Microchip Technology Inc., «PIC18F2550 in Production,» [En línea]. Available: <http://bit.ly/1RyVeBe>.
- [17] National Semiconductor Corporation, «LM35 Precision Centigrade Temperature Sensors,» [En línea]. Available: <http://bit.ly/1VL3YaW>.

Las Comunicaciones en las Zonas Rurales: Posibles Soluciones para su Desarrollo

Antonio Castro Lechtaler^{1,2y3}; Alejandro Arroyo Arzubi¹; Fernanda Carmona³;
Antonio Foti⁴; Rubén Fusario¹; Alejandro Oliveros⁵; Jorge García Guibout⁶, Germán Kurt Grin².

¹Laboratorio de Redes (RedLab), Escuela Superior Técnica - Facultad de Ingeniería del Ejército. Universidad de la Defensa, CABA, C1426; ²CISTIC/FCE - Universidad de Buenos Aires, CABA, C1120; ³Universidad Nacional de Chilecito, Chilecito, Provincia de La Rioja, F5360; ⁴Universidad Nacional de 3 de Febrero, Sáenz Peña, Provincia de Buenos Aires, B1674; ⁵Universidad Nacional el Oeste, San Antonio de Padua, Provincia de Buenos Aires, B1718 e ⁶Instituto Tecnológico Universitario - Universidad Nacional de Cuyo, M5500 Mendoza, Provincia de Mendoza.

antonio.castrolechtaler@gmail.com; aarroyo_arzubi@hotmail.com; fcarmona69@gmail.com; fo-ti.antonio@gmail.com; rfusario@gmail.com; aoliveros@untref.edu.ar; german.grin@gmail.com; jgarcia@itu.uncu.edu.ar

RESUMEN

En la Argentina, como en otros países en vías de desarrollo, muchas regiones rurales con baja densidad poblacional carecen de servicios de comunicaciones por falta de interés comercial de las empresas públicas de telecomunicaciones en brindar estas prestaciones.

Esta situación dificulta a la población rural acceder a una adecuada educación, a la actividad económica, la atención de la salud en casos de urgencia, y provoca todo tipo de problemas a las comunidades que tienen estas carencias.

El presente *Grupo de Investigación organizado en Red de Universidades* tiene por objetivo buscar una solución a la problemática de la falta de conectividad en las zonas rurales con el objeto de que se puedan brindar servicios isócronos y de datos de banda ancha con acceso a la Red Internet.

La idea central de la investigación es buscar distintas alternativas, que seguramente diferirán de las utilizadas en los países centrales, pero que pueden constituir una solución a este problema. En el caso particular de Argentina las distancias son condicionantes por su importancia, y la densidad poblacional es sustancialmente muy inferior al que se puede encontrar en otros países, especialmente los desarrollados.

Es por ello que se orientó el estudio hacia las técnicas digitales inalámbricas *wireless technologies* -en especial aquellas de largo alcance tales como: microondas, 802.11, WiMax, CDMA450, 802.22 y otras similares- podrían dar solución al problema planteado.

Palabras Clave:

PLC, CSMA/CA, WLAN, 802.11, 802.22

CONTEXTO

Este Grupo de Investigación trabaja, los problemas de conectividad que se presentan en las comunidades rurales, con el apoyo de Entes Nacionales que están vinculados con esta problemática, en particular, en comunidades muy pequeñas que carecen de comunicaciones de banda ancha u otras de tamaño algo mayor que cuentan con facilidades, pero que hasta el momento son claramente insuficientes para el desarrollo de actividades rentables o para participar adecuadamente en Redes Sociales.

Estas áreas son claramente no rentables para las empresas proveedoras de servicios de telecomunicaciones y consecuentemente no son de interés comercial para esas empresas, las cuales no invierten en el desarrollo de la infraestructura adecuada.

Los trabajos que se realizan con el grupo de investigadores tienen como objetivo principal

hallar soluciones técnicas, a costos razonables, para los problemas de conectividad descritos.

El Grupo tiene su sede en el Laboratorio de Redes (**RedLab**) la Escuela Superior Técnica “Gral. Div. Manuel N. Savio” (EST) de la Universidad de la Defensa (UNDEF) y en las Universidades de Buenos Aires y las Nacionales de Chilecito, Tres de Febrero y del Oeste.

Los trabajos realizados hasta el presente han incluido pruebas de campo, en base a subsidios obtenidos por parte de la Agencia Nacional de Promoción Científica y Tecnológica [1] y fondos aportados por las universidades participantes.

Actualmente, luego de varias pruebas efectuadas, utilizando diversas tecnologías existentes en el mercado, se está trabajando sobre la base de analizar y probar las posibilidades que brindan los equipos que responden a la Recomendación 802.22 de la IEEE; y desde el punto de vista de su posible implementación en las instalaciones existentes que posee el Sistema Argentino de Televisión Digital de la Empresa ARSAT.

CITEDEF y el Consejo Profesional de Ingeniería de Telecomunicaciones, Electrónica y Computación (COPITEC) dieron su aval para este proyecto, dado su interés en ser aplicado en beneficio de los productores rurales, sus propios sistemas y redes.

1. INTRODUCCIÓN.

El problema de las comunicaciones rurales ha movilizado a distintos grupos de investigación y empresas de telecomunicaciones a buscar soluciones a este tipo de problemas, por cuanto estas zonas geográficas proporcionan cantidades significativas de productos alimenticios en sus diferentes etapas de fabricación y constituyen una trascendente fuente de productos básicos de exportación e ingresos de divisas. Para muchos países participan generando un porcentaje significativo del producto bruto interno de ellos.

La serie de Recomendaciones 802.XX incluye un conjunto de normas que regulan el funcionamiento de las comunicaciones inalámbricas.

Luego de evaluar con resultados poco satisfactorios los equipos que utilizaban la norma 802.11 fueron apareciendo distintas tecnologías que permitieron ampliar esta Serie merced al trabajo de distintos grupos de investigación [2], [3], [4], [5], [6].

Estos trabajos culminaron el 1 de julio de 2011 cuando finalizó el proceso de aprobación, del estándar “IEEE 802.22 - “IEEE 802.22: Cognitive Wireless Regional Area Network - Medium Access Control (MAC) and Physical Layer (PHY). Specifications: Policies and Procedures for Operation in the TV Band¹”.

La misma fue aprobada con el apoyo del Comité LAN/MAN² de la IEEE [7].

Este nuevo estándar proporciona una opción que permite establecer enlaces inalámbricos full dúplex a distancias de entre 30 a 70 km entre antenas, utilizando frecuencias no restringidas por las regulaciones gubernamentales.

La norma que pertenece a la serie 802.XX³ tiene por objeto establecer los criterios para el despliegue de múltiples productos interoperables de la misma, ofreciendo acceso a la banda ancha fija en diversas áreas geográficas, incluyendo especialmente los de baja densidad de población en las zonas rurales, y evitar la interferencia a los servicios que trabajan en la televisión de radiodifusión.

La misma es conocida actualmente como Red Inalámbrica de Área Regional y está pensada para operar principalmente como una forma de poder acceder a servicios de banda ancha a redes privadas de datos ubicadas en Zonas Rurales.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

¹ “IEEE 802.22 - Red de Área Regional Cognitiva Inalámbrica de Control de Acceso al Medio MAC) y la Capa Física (PHY). Especificaciones, Políticas y Procedimientos para la Operación en las Bandas de Televisión”.

² LAN: Local Area Network; MAN: Metropolitan Area Network.

³ Redes Inalámbricas.

Las líneas de Investigación y desarrollo son las siguientes:

21. Estudio de todas las Recomendación de la Serie 802.X para determinar la viabilidad de su uso en las comunicaciones rurales.

22. Estudio de casos concretos de uso actual de equipamientos que utilicen la Norma 802.22 para comunicaciones en distancias de hasta 100 km.

23. Ventajas y posibilidades del uso de los Espacios Blancos que existen en el ancho de banda que se utiliza para los Sistemas de Televisión Digital Terrestre.

24. Posibilidad de aprovechamiento de las instalaciones de las estaciones de Televisión Digital Terrestre instaladas a lo largo del país, para la instalación del equipamiento necesario para el funcionamiento de los equipos necesarios para brindar comunicaciones rurales.

25. Estudio práctico y teórico [8, 9], de las interferencias entre canales debido al uso intensivo de las comunicaciones inalámbricas para todo tipo de servicios de comunicaciones.

26. Determinación de la posible utilización de las frecuencias asignado para la transmisión del Sistema Nacional de Televisión Abierta para combinarlo con un sistema basado en estas normas, para la utilización en las comunicaciones rurales utilizando los espacios blancos.

3. RESULTADOS OBTENIDOS / ESPERADOS.

3.1. Se han obtenido los siguientes resultados:

3.1.1. Del estudio del estándar IEEE 802.22 y del equipamiento analizado se ha determinado que tanto por el alcance como por las prestaciones puede resultar una solución a las comunicaciones rurales, en las condiciones descritas.

3.1.2. El mismo está teóricamente dentro de las distancias requeridas para satisfacer los objetivos del proyecto.

3.1.3. Las frecuencias asignadas al espectro de la Televisión Digital Abierta en la modali-

dad de radio cognitiva [10, 11] son adecuadas a las características de este proyecto.

3.1.4. Se ha considerado que el Sistema de TDA puesto en marcha puede ser una oportunidad para que esta norma sea considerada en la reasignación del espectro, actualmente en estudio, por parte de la Comisión Nacional de Comunicaciones.

3.1.5. Al trabajar en las mismas frecuencias que otros servicios, pero protegiendo las transmisiones de los operadores principales habilitados en las mismas, el estándar posee un conjunto de capacidades que incluyen:

- Detección del espectro;
- Servicios de geolocalización;
- Acceso a base de datos con información sobre el estado del espectro;
- Registro y seguimiento de la gestión del conjunto de los canales que están operando en un determinado momento en una zona geográfica determinada [12].

Estas capacidades lo hacen también que pueda ser una solución al problema planteado.

3.2. En lo que respecta a los resultados esperados se estima ellos podría sintetizarse en los siguientes:

3.2.1. La 802.22 podría estar capacitada para:

- Explotar y detectar canales operativos que podrían producir interferencias tales como: transmisiones de televisión;
- la emisión de micrófonos inalámbricos;
- Las transmisiones de dispositivos de protección como podrían ser faros inalámbricos u otras transmisiones como por ejemplo la telemetría médica (que requiere ser protegida por la autoridad regulatoria local).

3.2.2. La capa de enlace toma elementos de la norma 802.3, de amplia difusión y probada eficiencia. Estas características deben ser verificadas mediante trabajos de campo.

3.2.3. Se estima que la existencia de un sistema de televisión por radiodifusión ya instalado en un gran porcentaje evitaría tener que usar una porción adicional del espectro de

frecuencias, cada vez más escaso y congestionado.

3.2.4. Se analizará y se buscarán resultados sobre la utilización de los espacios Blancos con el objeto de buscar reducir el uso del Espectro de Frecuencias.

Como resultado de estas actividades se presentó un trabajo [13] el XXII Congreso Argentino de Computación - CACIC 2016, el que **fue seleccionado para ser publicado en el libro de los mejores artículos** que se publica anualmente, en este caso del CACIC 2016 (31 trabajos), en base a los resultados realizados por los evaluadores sobre el total de artículos presentados.

Se estima que se deberá continuar con el estudio de esta recomendación en sus aspectos técnicos, para determinar fundamentalmente sus limitaciones, si ellas existieran, todo ello sin perjuicio de buscar otras opciones.

3.2.5. Se efectuará un relevamiento del equipamiento que el mercado está ofreciendo sobre esta norma y un análisis de las capacidades del mismo; buscando obtener una idea de las capacidades y costos de este tipo de equipamientos para cubrir distintas zonas del territorio nacional.

3.2.6. Se continuarán las actividades de campo para verificar el verdadero rendimiento del equipamiento y la dificultad que requerirá su despliegue, tal como el equipo de investigación efectuó sobre el terreno el Proyecto Corral de Lorca.

Las instalaciones de las estaciones base del sistema de televisión digital terrestre, instaladas sobre *shelters*, sin duda pueden ser útiles para el despliegue de parte de los equipos requeridos por la 802.22.

4. FORMACIÓN DE RECURSOS HUMANOS.

Desde el año 2015 en este grupo trabajan Investigadores en Formación y alumnos de las carreras de grado y posgrado vinculadas con los temas que hacen tanto a las comunicaciones, como a la seguridad de los sistemas que podrían ser utilizados.

Durante el año 2017 se han sumado al proyecto nuevos investigadores en formación y alumnos de las distintas universidades participantes en especial de las Carreras de Ingeniería en Informática y Electrónica.

Algunos de ellos han recibido las becas *Estímulo a las Vocaciones Científicas*, perteneciente al *Programa Estratégico de Investigación y Desarrollo, Plan de Fortalecimiento (Componente de Formación de Recursos Humanos) del Consejo Interuniversitario Nacional*.

Cabría la posibilidad incluso algunos de ellos, realicen su Trabajo Final de Carrera en algún tema de los que aborda la presente línea de investigación.

Los integrantes son docentes y alumnos de las asignaturas las siguientes asignaturas en las distintas Universidades participantes: Tecnología de las Comunicaciones; Sistemas de Comunicaciones I y II; Comunicaciones Inalámbricas; Redes de Computadoras, entre otras.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.-.

5. BIBLIOGRAFÍA.

[1] Proyecto FONCyT - ANPCyT. PICTO 11-18621. Redes Privadas Comunitarias. Proyecto finalizado y aprobado. Antonio Castro Lechtaler (Director).

[2] García Guibout, J., García Garino C., Castro Lechtaler, A. R., Fusario R., y Sevilla, G., (2007) Physical and Link Layer in Power Line Communications Technologies. *Proceedings of 13th of Argentine Congress on Computer Science*. ISBN 978 - 950 - 656 - 109 - 3. Pág. 56 a 67.

- [3] García Guibout, G., García Garino, C., Castro Lechtaler, A. R., Fusario R. y Sevilla, G. (2007) Power Line Communications in the Electric Network. *Proceedings of 13th of Argentine Congress on Computer Science*. ISBN 978 - 950 - 656 - 109 - 3. Pág. 68 a 79.
- [4] García Guibout, J., García Garino. C., Castro Lechtaler, A. y Fusario, R., (2008). Transmission voice over 802.11. *Proceedings of 14th of Argentine Congress on Computer Science*. ISBN 978 - 987 - 24611 - 0 - 2. Pág. 307 a 318.
- [5] Castro Lechtaler, A., Foti, A., Fusario, R., García Garino, C., y García Guibout, J., (2009) Communication Access to Small and Remote Communities: The Corral de Lorca Project. *Proceedings of 15th of Argentine Congress on Computer Science*. ISBN 978 - 897 - 24068 - 4 - 1. Pag. 1.117 a 1.126.
- [6] A. Castro Lechtaler, A. Foti, C. García Garino, J. García Guibout, R. Fusario and A. Arroyo Arzubi. Proyecto Corral de Lorca: Una solución de conectividad a grupos poblacionales pequeños, aislados y distantes de centros urbanos. *Proceedings de la Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática: CИСCI 2010 - Volume III - ISBN - 13: 978 - 1 - 934272 - 96 - 1*. pp. 121 a 127. Orlando, USA. June 2010.
- [7] IEEE 802.22 - Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) *Specifications Policies and Procedures for Operation in the TV Bands*.
- [8] Gómez, C., (2013). Spectrum Regulation and Policy Officer Radiocommunication ITU. Apia, Samoa. www.itu.int/ITU-D/asp/CMS/Events/.../ITU-APT-S3_Cristian_Gomez.pdf
- [9] CEPT Report 24. A preliminary assessment of the feasibility of fitting new/future applications/services into non-harmonized spectrum of the digital dividend (namely the so-called "white spaces" between allotments. Report C from CEPT to the European Commission in response to the Mandate on: Technical considerations regarding harmonization options for the Digital Dividend. 1 July 2008.
- [10] Mitola, J. and Maguire, G., (1999) Cognitive radio: making software radios more personal. *IEEE Personal Communications Magazine*, Volume 6 Issue 4. pp. 13 to 18.
- [11] Mitola, J. (2000) Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio. Dissertation submitted in partial fulfillment of the degree of Doctor of Technology. Royal Institute of Technology (KTH) - Sweden. Teleinformatics. ISSN 1403 - 5286.
- [12] Cordeiro, C., Kiran Challapali, and Dagnachew Birru, Sai Shankar (2006) N. IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive *Radios Journal of Communications*, Vol. 1, N° 1.
- [13] Castro Lechtaler, A.; Foti, A.; Arroyo Arzubi, A.; García Guibout, J.; Carmona, F.; Fusario, R. y Oliveros, A. Using White Spaces: A solution for frequency spectrum overloading. XI Workshop of Architecture, Networks and Operating Systems (WARSO). *Proceedings of the 22th Argentinean Congress on Computer Science*. ISBN 978-987-733-072-4. pp. 903 to 913. San Luis. October. 2016.

Sistema de Acceso y Autenticación en Redes Definidas por Software

Andres Peñasco^{1,3}, Miguel Méndez-Garabetti^{1,2}

¹Universidad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería
penasco.andres@gmail.com, miguel.mendez@um.edu.ar

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Universidad Nacional de Cuyo, Facultad de Ciencias Aplicadas a la Industria

RESUMEN

Actualmente vivimos en una sociedad digital, donde casi todo está conectado y accesible desde cualquier lugar a través de Internet. Sin embargo, a pesar de su amplia adopción, las redes IP tradicionales de gran envergadura son complejas y muy difíciles de gestionar. Configurar grandes redes según políticas predefinidas, o reconfigurarla para responder a fallas, cambios en la infraestructura y cargas de trabajo, se torna una tarea laboriosa y complicada. Las Redes Definidas por Software (Software-Defined Networking, SDN) son un paradigma emergente que promete mejorar las falencias de las redes convencionales, introduciendo la capacidad de programar la red. Para lograrlo, SDN separa el plano de control del plano de datos, promoviendo la centralización del control de la red y convirtiendo a los routers y switches subyacentes en dispositivos de reenvío de datos simples. Este trabajo de I+D consiste en realizar una revisión bibliográfica de investigaciones recientes, como así también el análisis de los diferentes tipos de soluciones disponibles, incluyendo controladores y tipos de switches, con el objetivo de desarrollar una aplicación que

permita administrar el ingreso y la asignación de privilegios de usuarios a una red IP, mediante el uso de flujos manejados por un controlador SDN. De esta forma, será posible demostrar que el desarrollo de aplicaciones que gestionen el plano de control, podrán ofrecer soluciones completas capaces de solucionar los problemas y/o debilidades de las redes tradicionales.

Palabras clave: SDN, openflow, control de acceso, seguridad, redes programables.

CONTEXTO

El presente trabajo de I+D se desarrolla como proyecto de tesis de posgrado de la Maestría en Teleinformática, Dirección de Posgrado, perteneciente a la Facultad de Ingeniería de la Universidad de Mendoza, (Ciudad, Mendoza). El presente proyecto fue aceptado como propuesta de tesis en diciembre de 2016 y aún se encuentra en fase de desarrollo.

1. INTRODUCCIÓN

Las redes de datos se han convertido en uno de los componentes esenciales de toda red

corporativa, siendo de suma importancia que éstas operen de forma eficiente [1]. Por ello, es necesario realizar una adecuada gestión de las redes actuales, considerando la evolución de los sistemas informáticos y las tecnologías emergentes, tales como servicios de computación en la nube [2], sistemas distribuidos [3], Big Data [4], entre otros. En este contexto, las redes de datos tradicionales se enfrentan a una gran cantidad de limitaciones de diseño, que impide la rápida adaptación de la red a cambios, como la reacción ante nuevas vulnerabilidades, implementación de nuevos servicios con requisitos especiales, ampliación de la infraestructura o incorporación de nuevos dispositivos.

Para mantener grandes redes en funcionamiento, es necesario configurar individualmente cada uno de los dispositivos de red, como routers, switches y puntos de acceso, utilizando comandos de bajo nivel, o específicos de cada proveedor. Lo que conlleva un alto grado de complejidad, generando grandes costos operacionales, debido a la especificidad de los conocimientos necesarios para llevar a cabo esta gestión.

Además de la complejidad de configuración mencionada anteriormente, los entornos de red tienen que soportar la dinámica de las fallas y adaptarse a los cambios de carga. En este sentido los mecanismos de reconfiguración y respuesta automática en las redes IP actuales son prácticamente inexistentes. El plano de control (que decide cómo manejar el tráfico de red) y el plano de datos (que reenvía el tráfico de acuerdo con las decisiones tomadas por el plano de control) se agrupan dentro de los dispositivos de red, reduciendo la

flexibilidad y obstaculizando la innovación y la evolución de la infraestructura de red. Por ejemplo, la transición de IPv4 a IPv6, iniciada hace más de una década y aún en gran parte incompleta [5], da testimonio de este desafío, mientras que de hecho IPv6 representaba simplemente una actualización de protocolo. Las redes tradicionales deterministas, que utilizamos hoy en día, en las que el comportamiento de los dispositivos depende de su configuración previa, necesitan evolucionar a una arquitectura de red dinámica, transformándose en entornos escalables, flexibles, fáciles de gestionar y securizar. En respuesta a estas necesidades, se considera que las Redes Definidas por Software (Software-Defined Networking, SDN) [6] y su evolución marcan el camino que ofrece soluciones óptimas a las debilidades planteadas.

Las SDN, son un concepto que viene evolucionando hace varios años y se encuentra en constante desarrollo y expansión. Este tipo de redes divide el plano de datos del plano de control, logrando de esta forma infraestructuras programables, automatizadas, adaptables a las necesidades y problemas futuros [7]. Al separar el plano de control es posible gestionar la red de forma centralizada, gracias a la incorporación de un controlador que mantiene una visión global de la red y del contenido de la misma, proporcionando la capacidad de insertar, modificar o eliminar flujos de datos según sea necesario. Permitiendo programar directamente sobre arquitectura SDN, utilizando módulos de software instalados en el controlador, agilizando los procesos de configuración.

Además, las arquitecturas SDN pueden ser implementadas bajo estándares abiertos, de modo que no dependen de dispositivos de fabricantes específicos o protocolos propietarios. La ONF [8], ha definido el primer estándar abierto, denominado OpenFlow [9], un protocolo que se encuentra en continuo desarrollo y permite manejar directamente el plano de reenvío de dispositivos de red como switches y routers, ya sean físicos o virtuales [7].

Debido a los inconvenientes mencionados de las redes tradicionales, y las ventajas que presenta la arquitectura SDN, es que este trabajo pretende realizar un estudio de la misma, con el propósito de desarrollar un prototipo de aplicación, capaz de gestionar el control de acceso de usuarios en una arquitectura SDN. Para ello se evaluarán los últimos avances registrados en la literatura, tal como los trabajos [10] y [11], donde los autores proponen una solución para el control de acceso a redes en una arquitectura SDN híbrida, utilizando switches virtuales y el protocolo de autenticación RADIUS [12]. El desarrollo aplicación propuesta en este trabajo permitirá administrar el ingreso y la asignación de privilegios de usuarios a la red, mediante el uso de flujos manejados por el controlador SDN. De esta forma, será posible demostrar que el desarrollo de aplicaciones que gestionen el plano de control, permiten ofrecer soluciones completas capaces de solucionar los problemas o debilidades de las redes tradicionales. Además, se realizará un análisis e investigación de diferentes tipos de soluciones disponibles, incluyendo controladores y tipos de switches, como así también la comunicación y configuración de

los mismos a través del estándar abierto OpenFlow en sus diferentes versiones.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Redes definidas por software: Se investigará esta tecnología emergente con el objetivo de aplicar las ventajas de la misma, en el desarrollo de la aplicación. Además se analizarán los controladores SDN disponibles para determinar cuál se adapta a las necesidades requeridas por el proyecto.

Protocolos de autenticación: Se realizará un estudio sobre los protocolos y estándares de autenticación existente, tratando de utilizarlos en la aplicación a desarrollar y de este modo brindarle seguridad a la misma.

Virtualización de redes: Se utilizarán herramientas de simulación y virtualización de dispositivos, para generar escenarios de pruebas y poder depurar la aplicación desarrollada antes de utilizarla en un escenario real.

Control de Acceso a redes: Se realizará un análisis de las aplicaciones de control de acceso a redes tradicionales, con el objetivo de comparar ventajas y desventajas con respecto a la aplicación propuesta.

3. RESULTADOS ESPERADOS

El presente proyecto, se centra en el desarrollo e investigación de un método de acceso y autenticación para redes definidas por software, capaz de resolver problemas complejos presentes en las redes actuales.

Particularmente se espera lograr:

1. Estudiar los diferentes tipos de controladores y dispositivos SDN existentes.
2. Evaluar últimos desarrollos e investigaciones realizadas para control de acceso sobre SDN.
3. Implementar una SDN.
4. Configurar y gestionar de forma centralizada los dispositivos de la SDN implementada.
5. Desarrollar un prototipo de aplicación para el control de acceso de usuarios a una arquitectura SDN y evaluar su funcionamiento en contraste con metodologías similares.
6. Documentar características, arquitectura, ventajas y desventajas del protocolo Openflow y las SDN.
7. Evaluar el alcance de la solución y su viabilidad.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de posgrado por parte del Lic. Andrés Peñasco, quien es estudiante de la Maestría en Teleinformática de la Universidad de Mendoza, la cual tiene como director al Mg. Ing. Miguel Méndez-Garabetti.

5. BIBLIOGRAFÍA

1. W. Stallings, Data and computer communications, 8° ed., Prentice Hall, 2006.
2. NIST, «National Institute of Standards and Technology,» 2011. [En línea]. Disponible: <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>.
3. G. Coulouris, J. Dollimore y T. Kindberg, «Distributed Systems: Concepts and Design,» 5th ed., Addison Wesley, 2011.
4. ISO, «Big Data: Preliminary Report 2014,» 2014. [En línea]. Disponible: http://www.iso.org/iso/big_data_report-jtc1.pdf. [Último acceso: septiembre 2016].
5. O. De León y LACNIC, «Análisis detallado de la información cuantitativa relevante relativa a la transición hacia una red IPv6,» 2014. [En línea]. Disponible: <http://portalipv6.lacnic.net/caf-lacnic/anexo-iii/>. [Último acceso: Marzo 2016].
6. D. Kreutz, F. Ramos, P. Verissimo, C. Rothenberg, S. Azodolmolky y S. Uhlig, «Software-Defined Networking: A Comprehensive Survey,» IEEE, 2015.
7. Open Networking Foundation, «Software-Defined Networking: The New Norm for Networks,» 2012. [En línea]. Disponible: <https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf>. [Último acceso: Agosto 2016].
8. Open Networking Foundation, «Software-Defined Networking (SDN) Definition,» 2016. [En línea]. Disponible: <https://www.opennetworking.org/sdn->

- resources/sdn-definition. [Último acceso: 18 Agosto 2016].
9. N. McKeown, T. Anderson, H. Balakrishnan, P. Guru y L. Peterson, «Openflow,» 2008. [En línea]. Disponible: <http://archive.openflow.org/documents/openflow-wp-latest.pdf>. [Último acceso: septiembre 2016].
 10. V. Dangovas y F. Kuliesius, «SDN-Driven Authentication and Access Control System,» 2014. [En línea]. Disponible: <http://sdiwc.net/digital-library/web-admin/upload-pdf/00001098.pdf>. [Último acceso: 2016].
 11. V. Dangovas y F. Kuliesius, «SDN Enhanced Campus Network Authentication and Access Control System,» 2016. [En línea]. Disponible: <http://ieeexplore.ieee.org/document/7536925/>. [Último acceso: 10 Septiembre 2016].

Un Prototipo de Plataforma de Ciudades Inteligentes

Aplicación para la protección ciudadana en el Departamento de Godoy Cruz

Lucas Iacono^{1,2,3,4}, Elina Pacini^{1,2}, Pablo D. Godoy^{1,2,3}, Carlos G. García Garino^{1,2,4}, David Monge^{1,3}, Carlos Catania^{1,2}

¹ITIC, Universidad Nacional de Cuyo, Mendoza, Argentina

²Facultad de Ingeniería, Universidad Nacional de Cuyo, Mendoza, Argentina

³Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Mendoza, Argentina

⁴Universidad de Mendoza, Mendoza, Argentina

RESUMEN

Como objetivo general se propone implementar una plataforma de Ciudad Inteligente en el departamento de Godoy Cruz, Mendoza, destinada a la seguridad ciudadana. Se propone recolectar datos a través de la interconexión de redes de sensores inalámbricos (WSN), que serán utilizados para integrar la plataforma de Ciudad Inteligente.

La integración de WSN y Cloud Computing da lugar a infraestructuras denominadas Cloud de Sensores, las cuales pueden aplicarse a ciudades inteligentes. En este proyecto se propone aplicar una plataforma de gestión de Cloud de Sensores a aplicaciones de monitoreo y seguridad en ciudades inteligentes. Para este propósito, se adaptará una plataforma de gestión de Cloud de Sensores denominada Sensor Cirrus, la cual está destinada actualmente a agricultura de precisión.

Esta plataforma ofrecerá nuevas soluciones a la comuna de Godoy Cruz, ya que brindará información para estudiar la problemática delictiva y mejorar la seguridad en la comuna. En general se espera contar con un modelo para la implementación de ciudades inteligentes apto para ser implementado en el departamento de Godoy Cruz o similares, que contribuya a solucionar problemas tales como la inseguridad.

Además, se prevee contar con un prototipo que permita realizar experimentos de performance, analizar mejoras y pruebas

piloto. Este prototipo tendrá la posibilidad de ser escalado a un sistema de mayor envergadura que pueda ser implementado en todo el departamento.

Palabras claves: Ciudades inteligentes, redes de sensores, Cloud de sensores.

CONTEXTO

a - Proyecto

El presente proyecto fue presentado en la convocatoria: Proyectos de Vinculación Tecnológica "Universidades Agregando Valor", financiados por la Secretaría de Políticas Universitarias (SPU), en octubre de 2016.

b - Instituciones

Las siguientes instituciones forman parte del proyecto:

- 1) ITIC - Instituto para las Tecnologías de la Información y las Comunicaciones, Universidad Nacional de Cuyo.
- 2) La Cooperativa, Empresa Eléctrica de Godoy Cruz. Edificación, Servicios Públicos y Consumo Ltda.
- 3) Municipalidad de Godoy Cruz.
- 4) SOFTNUVO SA.

c – Proyectos relacionados

El proyecto se encuentra relacionado con dos proyectos financiados por la Universidad Nacional de Cuyo, que se mencionan a continuación:

- 1) Título del proyecto: Desarrollo de Cloud de

sensores con nodos sensores móviles aplicados a predicción de heladas. Director: Lucas Iacono. Tipo de proyecto: Proyectos de Investigación SeCTyP Tipo 1. Institución que financia: Universidad Nacional de Cuyo.

2) Título del proyecto: Implementación de laboratorios remotos basados en cloud computing. Director: Pablo Daniel Godoy. Tipo de proyecto: Proyectos de Investigación SeCTyP Tipo 1. Institución que financia: Universidad Nacional de Cuyo.

1. INTRODUCCION

Una Ciudad Inteligente es aquella que incorpora TICs (Tecnologías de la Información y las Comunicaciones) en la gestión urbana y usa estos elementos como herramientas para estimular la formación de un gobierno eficiente que incluya procesos de planificación colaborativa y participación ciudadana [1]. Actualmente en el país, debido al crecimiento rápido y no planificado de las ciudades se han generado desafíos que no pueden ser resueltos de manera tradicional, por ejemplo el problema de inseguridad, en el cual tanto las estrategias policiales, como la justicia y sistemas penales ya no son suficientes para atacar la complejidad del tema, ni tiene el carácter preventivo e integral que el ciudadano demanda.

La incorporación de las TICs en la gestión urbana puede brindar las siguientes ventajas: a) facilitan la integración de procesos en la administración pública, y a su vez aportan información necesaria y transparente para una mejor toma de decisiones y gestión presupuestaria; b) generan procedimientos comunes que aumentan la eficiencia del gobierno; c) optimizan la asignación de recursos y ayudan a reducir gastos innecesarios; d) eleva el grado de satisfacción de los habitantes pues permiten prestar una mejor atención a los usuarios de servicios y mejorar la imagen de los órganos públicos; e) permiten una mayor participación de la sociedad civil organizada y de los ciudadanos en la administración por medio del uso de herramientas tecnológicas que ayudan a monitorear los servicios públicos, identificando problemas, informando e

interactuando con la administración municipal para resolverlos; f) producen indicadores de desempeño que son útiles para medir, comparar y mejorar las políticas públicas [1][2].

En este proyecto se propone aplicar una plataforma de gestión de Cloud de Sensores a aplicaciones de monitoreo y seguridad en ciudades inteligentes. Para este propósito, se adaptará una plataforma de gestión de Cloud de Sensores denominada Sensor Cirrus, la cual está destinada actualmente a agricultura de precisión. La plataforma Sensor Cirrus fue diseñada por uno de los integrantes del presente proyecto como parte de su tesis doctoral [3].

2. LINEAS DE INVESTIGACION Y DESARROLLO

2.1 Ciudades Inteligentes

Estas ciudades incorporan TICs para su gestión urbana, como se describió en la sección 1.

La arquitectura de una ciudad inteligente requiere la construcción de una infraestructura de redes de datos sólida, basada en estándares tecnológicos que le permita crecer de manera segura y que, al mismo tiempo, garantice que las inversiones se conserven a lo largo del tiempo. A esta red deben estar integrados los millones de sensores esparcidos por las ciudades para los más variados fines.

2.2 Redes de sensores inalámbricos

Las WSN son redes formadas por nodos, los cuales poseen sensores de distinto tipo (temperatura, infrarrojos, cámaras, etc.), un microprocesador, memoria, fuente de alimentación y un transmisor de señales de radiofrecuencia. Los nodos transmiten los datos a un dispositivo denominado estación base, el cual los almacena y exporta a Internet. La característica fundamental de las WSN es el bajo costo de los nodos, lo que permite construir WSN con un gran número de nodos, y por consiguiente, un gran número de puntos de medición [4].

Se han desarrollado un número muy grande de aplicaciones para las WSN en distintos campos, por ejemplo: monitoreo de variables

agrícolas, domótica o edificios inteligentes, cuidado de la salud, aplicaciones industriales y militares, etc.

2.3 Cloud Computing y Cloud de sensores

Cloud Computing ofrece al usuario recursos de procesamiento y almacenamiento en forma de servicios, accesibles desde cualquier navegador web en dispositivos con conexión a Internet [5].

Las WSN pueden generar grandes volúmenes **Things-IoT**:

En la actualidad, Internet comunica una diversidad tan amplia de dispositivos que surge el término Internet of Things (IoT) para referirse a esta red global de objetos interconectados. Las WSN conectadas a Internet que permiten el acceso a sus datos remotamente son un ejemplo de IoT [5][6].

Dentro de la literatura referente a IoT, las ciudades inteligentes se consideran como uno de los tres grandes campos en los que se divide IoT. Por este motivo, este proyecto requerirá incluir el estudio de bibliografía referente a IoT para la solución de problemas o propuestas de implementación.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo general de este proyecto es implementar un prototipo para la gestión de datos de ciudades inteligentes en el municipio de Godoy Cruz, orientado a problemas de seguridad ciudadana. Además, este prototipo tendrá la posibilidad de ser escalado a un sistema de mayor envergadura que pueda ser implementado en todo el departamento.

Además de un prototipo, se espera contar con un modelo para la implementación de ciudades inteligentes apto para ser implementado en el departamento de Godoy Cruz o similares, que contribuya a solucionar problemas tales como la inseguridad.

Para alcanzar este objetivo general, se han propuesto objetivos intermedios o etapas. Se espera que cada uno de estos objetivos intermedios produzca resultados útiles. Estos objetivos se enumeran a continuación:

3.1 Estudio bibliográfico

de datos, por lo que se requieren infraestructuras con recursos computacionales de altas prestaciones como Cloud Computing para almacenar y procesar dichos datos.

La integración de WSN y Cloud Computing da lugar a infraestructuras denominadas Cloud de Sensores, las cuales pueden aplicarse a ciudades inteligentes.

2.4 Internet de las Cosas (Internet of

El objetivo de este estudio bibliográfico es conocer diferentes alternativas de solución para la implementación de ciudades inteligentes. Se tendrán en cuenta propuestas implementadas en otras ciudades del mundo, y se estudiará la posibilidad de adecuar algunas de estas propuestas al municipio de Godoy Cruz.

Resultado esperado: artículo científico que sirva de base en el desarrollo del proyecto. El mismo será publicado en algún congreso relacionado con la temática.

3.2 Adecuar una plataforma de Cloud de sensores al ámbito urbano

Se cuenta con una plataforma de Cloud de sensores destinada a sensar variables en el sector agroindustrial, destinada al estudio de las heladas, implementada por uno de los integrantes del grupo de trabajo de este proyecto [3]. Se propone adaptar este prototipo para ser empleado en el ámbito urbano, donde las condiciones ambientales y los requisitos de la aplicación son diferentes. Resultado esperado: Contar con un prototipo parcial de Cloud de sensores apto para ser empleado en el ámbito urbano.

3.3 Analizar tecnologías para la capa de transporte

Los datos recolectados por el Cloud de sensores deben transportarse hacia los servidores de procesamiento y almacenamiento a través de la infraestructura disponible. Se deberá analizar la tecnología disponible y su utilización.

El grupo de trabajo ha realizado trabajo de investigación sobre la integración de redes de sensores inalámbricos a redes TCP/IP [7] y a

Cloud [8], con propósitos de monitoreo agrícola.

También ha realizado trabajos sobre la integración de redes de sensores inalámbricos a laboratorios remotos a través de Internet, y un prototipo que incluye dicha integración [8]. Resultado esperado: Reporte que permita tomar decisiones sobre la tecnología a emplear en la capa de transporte de datos.

En el presente proyecto, la Cooperativa Eléctrica de Godoy Cruz proporcionará en todo el departamento una red de fibra óptica para realizar el transporte de datos. Dicha fibra será utilizada en el proyecto para acceder a los datos proporcionados por los diferentes componentes que integren la plataforma de Ciudades Inteligentes, como por ejemplo, alarmas comunitarias, sensores de alarmas, etc.

3.4 Desarrollar un módulo de análisis de datos

Se estudiarán los requisitos del sistema en cuanto al análisis de datos, y se desarrollará un módulo que permita procesar dichos datos y proporcionar indicadores y estadísticas que resuman los aspectos principales relacionados con la seguridad ciudadana.

Resultado esperado: Contar con un módulo de análisis de datos funcional.

3.5 Desarrollar un módulo de adquisición de datos orientado a la seguridad ciudadana

Se estudiarán los requisitos que este módulo debe cumplir y se implementará este módulo. El mismo funcionará como parte del Cloud de sensores desarrollado.

Resultado esperado: Prototipo de un módulo de adquisición de datos para seguridad ciudadana.

3.6 Implementación de un prototipo del sistema completo

Este prototipo debe integrar todos los módulos del sistema, varios de los cuales pueden desarrollarse de manera independiente. Resultado esperado: Contar con un prototipo del sistema completo.

3.7 Realización de pruebas de campo

Realizar pruebas de campo del prototipo, en un ambiente controlado, con la finalidad de detectar y corregir posibles fallas y proponer mejoras.

Resultado esperado: Reporte de funcionamiento, de fallas detectadas y correcciones realizadas, y posibles mejoras a realizar en el futuro.

4. FORMACION DE RECURSOS HUMANOS

El grupo de trabajo está formado por 6 docentes investigadores, algunos de ellos poseen doctorado en ciencias de la computación, otros en ingeniería.

También se prevee la participación de alumnos de grado.

Una de estas tesis doctorales, relacionada con laboratorios remotos [8] y que fue defendida en el año 2016, emplea varias de las técnicas de acceso remoto que se usarán en la implementación de este proyecto. Otra tesis doctoral de reciente defensa, propone e implementa un Cloud de sensores destinado al monitoreo de variables agrícolas. Este Cloud de sensores será una parte central de este proyecto [3].

5. BIBLIOGRAFIA

- [1] Cocchia, A. (2014). Smart and digital city: A systematic literature review. In Smart city (pp. 13-43). Springer International Publishing.
- [2] K. Su, J. Li and H. Fu, "Smart city and the applications," 2011 International Conference on Electronics, Communications and Control (ICECC), Zhejiang, 2011, pp. 1028-1031. doi: 10.1109/ICECC.2011.6066743
- [3] Gestión Remota de Redes de Sensores Inalámbricas Mediante Tecnologías de Cloud Computing, Lucas Iacono, tesis doctoral, Universidad de Mendoza, 2015.
- [4] Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer networks*, 52(12), 2292-2330.
- [5] Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation computer systems*, 25(6), 599-616.

[6] Botta, Alessio, Walter de Donato, Valerio Persico, and Antonio Pescapé. 2016. Integration of cloud computing and internet of things: a survey. *Future Generation Computer Systems* 56: 684-700

[7] Estudio de la Integración entre WSN y redes TCP/IP, Lucas Iacono, Pablo Godoy, Osvaldo Marianetti, Carlos García Garino, Cristina Párraga, En: Memoria de Trabajos de

Difusión Científica y Técnica, vol. 10, 2012

[8] Plataforma de Desarrollo de Laboratorios Remotos de Redes de Sensores Inalámbricos basados en Cloud Computing, Pablo Godoy, tesis doctoral, Universidad de Mendoza, 2016.

Un Sistema de Virtualización Distribuida

Pablo Pessolani

Departamento de Ingeniería en Sistemas de
Información
Facultad Regional Santa Fe
Universidad Tecnológica
Nacional Santa Fe, Argentina
ppessolani@frsf.utn.edu.ar

Toni Cortes*

Barcelona Supercomputing Center y
Departamento de Arquitectura de Computadores
Universitat Politècnica de Catalunya
Barcelona - España
toni.cortes@bsc.es

Silvio Gonnet

INGAR
(CONICET, Universidad Tecnológica Nacional)
Santa Fe - Argentina
sgonnet@santafe-conicet.gov.ar

Fernando G. Tinetti

III-LIDI
Facultad de Informática
Universidad Nacional de La Plata
La Plata, Argentina
Comisión de Inv. Científicas, Prov. Bs. As.
fernando@info.unlp.edu.ar

RESUMEN

Este trabajo refiere a los avances, logros alcanzados y planes de investigación futuros sobre el proyecto del Sistema de Virtualización de Recursos Distribuidos presentado en WICC 2012 [1].

En las tecnologías de virtualización actuales, la potencia de cómputo y el uso de recursos de las máquinas virtuales se limitan a la máquina física donde se ejecutan. Para alcanzar otros niveles de rendimiento y escalabilidad, las aplicaciones Cloud suelen estar particionadas en varias VMs o Contenedores ubicados en diferentes nodos de un cluster de virtualización. Los desarrolladores a menudo usan ese modelo de procesamiento porque sus aplicaciones no pueden hacer uso de los servicios de la *misma instancia* del Sistema Operativo (OS) en todos los nodos donde se ejecutan sus componentes.

El sistema propuesto combina e integra tecnologías de Virtualización y Sistemas Distribuidos que puede proporcionar la *misma instancia* de un Sistema Operativo Virtual (VOS) en cada nodo del clúster. El resultado es un Sistema de Virtualización Distribuida (DVS) con los beneficios de ambos mundos, adecuado para ofrecer servicios Cloud de alto rendimiento y con posibilidades de ofrecer otras características que son requeridas por los proveedores de IaaS. Un DVS es capaz de ejecutar concurrentemente múltiples instancias de diferentes VOS, asignando un subconjunto de nodos para cada instancia, y a su vez compartiendo nodos entre ellas.

Palabras claves: Virtualización, Máquinas Virtuales, Sistemas Operativos Distribuidos.

* Este trabajo está siendo subvencionado parcialmente por parte del Ministerio de Ciencia y Tecnología de España por la ayuda TIN2015-65316-P y del Gobierno Catalán por la ayuda 2014-SGR-1051.

CONTEXTO

Este proyecto de I/D involucra a investigadores de varios laboratorios y centros de investigación (ver afiliaciones de los autores), en un área que por su amplitud requiere de múltiples enfoques. Más específicamente, se deriva de las tareas de investigación en el área de virtualización realizadas en el contexto de PIDs de la UTN en la Facultad Regional Santa Fe y de los programas de postgrado de la Facultad de Informática de la UNLP.

INTRODUCCION

La tecnología de virtualización desarrollada a fines de la década del '60 [2] ha resurgido a fines de los años '90. Esto se debe a que han cambiado los modos de procesar la información, a los avances en el hardware y a las atractivas características que esta tecnología ofrece tales como: la capacidad para consolidar múltiples servidores virtuales en una misma computadora física y la implementación de entornos aislados para la ejecución de aplicaciones con mayor seguridad y protección.

En los diferentes sistemas de virtualización actuales una Máquina Virtual (abreviado VM en inglés), o en su caso un Contenedor están contenidos en un computador, por lo que su poder de cómputo se encuentra limitado por éste.

Por otro lado, los Sistemas Operativos Distribuidos de Imagen Única (abreviado SSI-DOS en inglés) ejecutan procesos en forma distribuida virtualizando recursos abstractos de software. Estos recursos son idénticos o de similares características a los que brinda un OS tradicional (centralizado) tales como archivos, tuberías (pipes), sockets, colas de mensajes, memoria compartida, semáforos o mutexes. Los usuarios y aplicaciones tienen la visión de un único computador virtual conformado por los recursos computacionales, procesos, recursos abstractos y servicios que se encuentran distribuidos en los distintos nodos que componen el cluster. Podría interpretarse que un SSI-DOS

representa una forma de virtualización *reversa*. A diferencia de la virtualización tradicional que permite crear múltiples computadores virtuales en un *único* computador físico, los SSI-DOS integran en un único OS los recursos de múltiples computadores físicos.

El Sistema de Virtualización Distribuida propuesto combina e integra tecnologías de Virtualización y de SSI-DOS con los beneficios de ambos mundos. Por sus características se presenta como adecuado para brindar servicios Cloud de alto rendimiento. En este sentido, dispone de la posibilidad de ofrecer otras características atractivas para proveedores de servicios IaaS tales como alta disponibilidad, replicación, elasticidad, balanceo de carga y migración de procesos. Un DVS es capaz de ejecutar concurrentemente múltiples instancias de diferentes VOS, asignando un subconjunto de nodos para cada instancia, y a su vez compartiendo nodos entre ellas (Fig. 1).

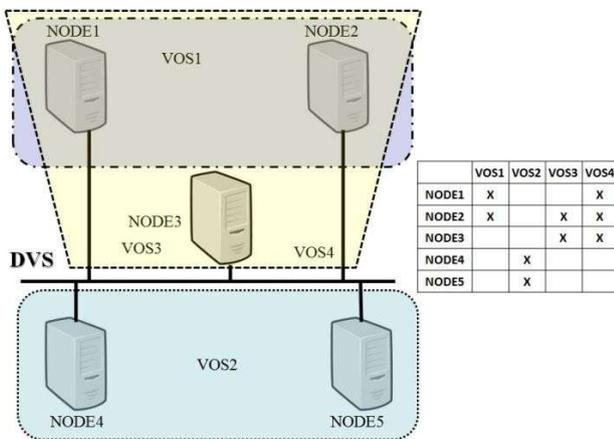


Figura 1. Sistema de Virtualización Distribuido (DVS)

Bases Tecnológicas del Proyecto

El proyecto se nutre de las tecnologías de sistemas operativos tales como: 1) Multiservidor, 2) SSI-DOS y 3) Unikernel. En tanto que las siguientes tecnologías de virtualización aportan conceptos y estrategias al proyecto: 1) Paravirtualización, 2) Virtualización *basada en Sistema Operativo*, 3) Virtualización *de Sistema Operativo*. En particular, desarrollaremos un resumen sólo de esta última por cuestiones de espacio y porque la misma no suele ser considerada en las taxonomías de virtualización.

Virtualización de Sistema Operativo

Un OS es la capa de software que se encuentra entre las aplicaciones y el hardware. Gestiona los recursos de hardware y proporciona servicios (llamadas al sistema) para los programas y aplicaciones.

En la tecnología de virtualización de Sistema Operativo, el OS-guest no opera directamente sobre el hardware, sino que utiliza los servicios ofrecidos por el OS-host. Como en la tecnología de paravirtualización, el OS-guest debe ser modificado, no para utilizar servicios de un hipervisor sino

los de un OS-host. User Mode Linux (UML) [3] y Minix Over Linux (MoL) [4] se basan en esta estrategia. UML permite que múltiples Linux-guest puedan ejecutarse como una aplicación en un sistema Linux-host. Como cada OS-guest es un proceso que se ejecuta en espacio de usuario, UML proporciona a los usuarios una forma de ejecutar múltiples instancias de Linux en un único computador sin requerir de privilegios de *root*.

De manera similar, MoL es un sistema de Virtualización de OS desarrollado en UTN-FRSF que permite ejecutar múltiples instancias de Minix [5] sobre un Linux-host en forma aislada y segura. En MoL se emulan las APIs, System Calls y servicios de Minix. Manteniendo la arquitectura de Minix, MoL se compone de un conjunto de servidores, de tareas y de un micro-kernel (virtual) que se ejecutan enteramente en modo usuario dentro de un Contenedor. Sus procesos no tienen acceso a ningún recurso del Linux-host ni de otros procesos MoL, a excepción de aquellos que le fueron específicamente asignados.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de I/D de este proyecto refiere al desarrollo de un modelo de Sistema de Virtualización Distribuida como tecnología para brindar servicios del tipo IaaS. Para ello se requiere el conocimiento y dominio de tecnologías de Virtualización, de Sistemas Operativos Distribuidos de Imagen Única y de Sistemas Operativos Multiservidor.

Durante el tiempo transcurrido desde el inicio del proyecto (año 2012) se implementó un prototipo que se ejecuta en un cluster de computadores x86 y Linux como OS-host. Se desarrollaron dos VOS sencillos para ser ejecutados como *guests* sobre el prototipo como prueba de concepto. Uno de ellos es un VOS multiservidor y el otro es un unikernel, ambos en condiciones de brindar servicios de Internet (servidor web).

Sistema de Virtualización Distribuida

En un SSI-DOS se implementan dentro del propio sistema las políticas y mecanismos de balanceo de carga, migración de procesos, sincronización, locking, elección de líder, consenso, exclusión mutua, censado de parámetros de rendimiento, replicación, monitoreo y administración de recursos. El resultado es que los módulos de software que lo componen se encuentran fuertemente acoplados entre sí y dentro del kernel del sistema. En el modelo de DVS propuesto se relaja el acoplamiento entre los diferentes módulos de software separando los servicios y funcionalidades en diferentes componentes de la arquitectura. Estos componentes son los siguientes (ver Fig. 2):

- *Guest Virtual Operating System*: Es un VOS multiservidor constituido por un conjunto de procesos que pueden estar distribuidos en los distintos nodos de un cluster físico. Un VOS es una emulación de un OS adaptado para trabajar utilizando los servicios e interfaces que le ofrece el DVS [6]. Un VOS presenta a las aplicaciones las APIs y llamadas al sistema de un OS original.

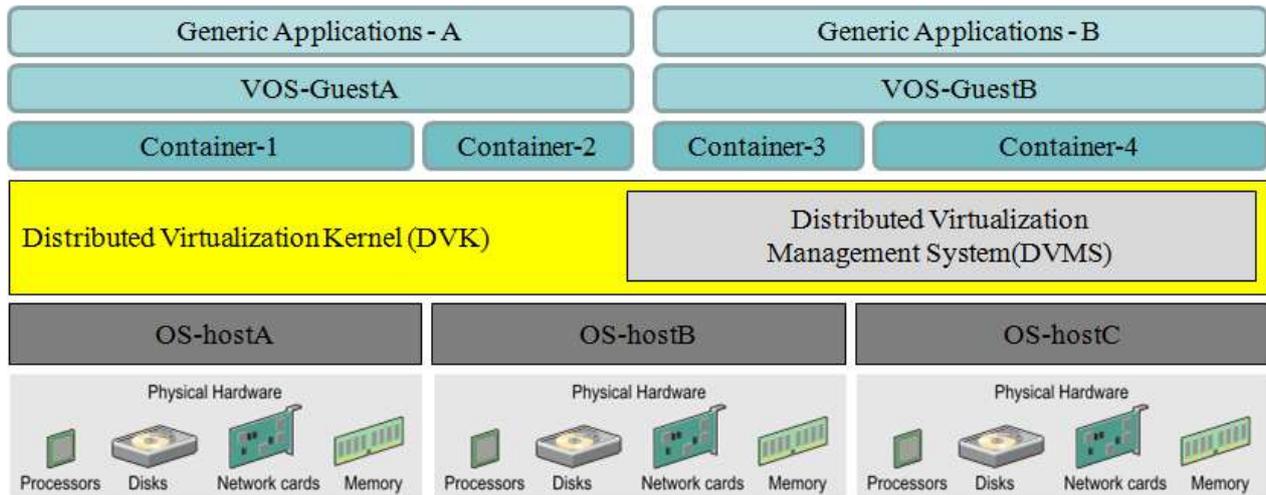


Figura 2. Arquitectura del DVS

- *Containers*: Es el entorno o subsistema en donde se ejecutan componentes de un VOS sobre un OS-hosts. Existe un Container por cada VOS en cada nodo del cluster.
- *Distributed Virtualization Kernel (DVK)*: Es el software que integra todos los recursos del cluster, gestiona y limita los recursos asignados a cada VOS. Es responsable de brindar a los VOS las interfaces para los protocolos de bajo nivel tales como IPC, comunicaciones de grupo, sincronización, locking, elección de líder, detección de fallos, consenso, exclusión mutua, censado de parámetros de rendimiento, mecanismo de migración de procesos, servicios de clave-valor y virtualización de dispositivos.
- *Distributed Virtualization Management System (DVMS)*: Es el software que le permite al supra-administrador (admitiendo que hay un administrador por cada VOS) gestionar los recursos del cluster, asignar recursos a los VOS, realizar diversos tipos de monitoreo del sistema y tomar acciones sobre el DVS.

La unidad de asignación de recursos a un VOS no es el nodo sino los recursos virtuales que brinda el OS-host en cada nodo. Esta granularidad en la asignación de recursos permite una mejor explotación de los mismos brindando mayor elasticidad y facilitando su gestión.

La contribución esperada de esta línea de investigación es un nuevo modelo de Sistema de Virtualización que permita el *particionado* de los recursos de un cluster entre diferentes VOS y la *agregación* de conjunto de sus recursos distribuidos en múltiples nodos en un único VOS.

Comunicación entre Procesos (IPC)

Los OS multiservidor representan una base conceptual muy importante en el proyecto y en ellos son fundamentales los mecanismos de comunicación entre procesos (IPC). Por esta razón, se dedicó una línea de investigación y desarrollo al tratamiento de este crítico componente. En principio, se desarrolló un mecanismo de IPC embebido en el kernel de

Linux que emulaba las primitivas de IPC de un OS-Minix al que se denominó M3-IPC. En su primera fase las comunicaciones se limitaban a un único host, aunque ya soportaba el aislamiento de comunicaciones entre procesos de diferentes VOS [7]. Esto significa que el mismo kernel de IPC permitía comunicar procesos de un mismo VOS, pero adicionalmente permitía ejecutar múltiples VOS en forma concurrente. Esta característica preserva el aislamiento requerido a cualquier sistema de virtualización.

El paso siguiente en el diseño y desarrollo fue llevar las comunicaciones a procesos en diferentes nodos en forma transparente para la aplicación. Para ello, se decidió realizar las comunicaciones inter-nodos mediante procesos *proxies* ejecutando en modo usuario. Si bien esto representa una reducción del rendimiento por el agregado de más cambios de contextos necesarios para las comunicaciones, tiene la ventaja de otorgar flexibilidad al momento de seleccionar el protocolo de comunicaciones entre nodos. Aún utilizando esta estrategia, el rendimiento de M3-IPC supera a RPC tanto en la transferencia de mensajes (11805 msg/s vs. 9900 msg/s) como en la de bloques de datos (58 Mbytes/s vs. 49 Mbytes/s) entre nodos conectados a un switch dedicado con interfaces de 1Gbps.

Durante el transcurso del proyecto se desarrollaron *proxies* utilizando protocolos TCP, UDP, TIPC, UDT y Raw Sockets demostrando la versatilidad que otorga el diseño. Por otro lado, éste también permitió incorporar de manera sencilla facilidades tales como la compresión de datos (LZ4), cifrado de datos (OpenSSL), o priorización de mensajes de acuerdo a criterios tales como el tipo de mensaje a transferir, el VOS origen/ destino o cualquier otro criterio, sin necesidad de modificar el kernel del sistema. La arquitectura permite utilizar *proxies* con protocolos diferentes para comunicar diferentes pares de nodos. Si bien son atractivas las ventajas que presentan los *proxies* en modo usuario, se dispone de *proxies* como módulos de kernel Linux utilizando protocolos TCP y TIPC (en desarrollo) para evaluar comparativamente el rendimiento.

Otra característica a destacar de M3-IPC es la transparencia en la localización, en la migración de procesos y en la replicación. Un proceso de un VOS no necesita conocer en que nodo se ejecuta el destinatario de su mensaje, solo necesita conocer su *endpoint*. De igual forma, si el supra-administrador o el DVMS deciden migrar un proceso de un nodo a otro, los otros procesos del VOS no necesitan realizar ninguna operación para adecuarse a ese cambio.

Para aplicaciones que requieren tolerancia a fallos se suelen utilizar esquemas de redundancia basados en *State Machine* [8] o en *Primary-Backup* [9]. M3-IPC soporta este último modelo asignándose a múltiples procesos de un mismo VOS localizados en diferentes nodos el mismo *endpoint*, aunque solo uno de ellos será el *endpoint* Primario o Activo. En caso de fallos, la aplicación podrá reasignar la función de Primario a uno de los procesos Backup. Todos estos cambios resultan transparentes a los otros procesos del VOS, particularmente a aquellos que utilizaban los servicios del servidor fallido.

Virtual Operating System

Con el desarrollo previo de MoL que utilizaba el IPC provisto por Linux, el paso siguiente fue migrar estos mecanismos a M3-IPC para conformar un VOS multiservidor. Se presentan a continuación algunos de los proyectos que resultaron de este trabajo.

Replicated Disk Server (RDISK)

Uno de los principales modelos de servicio de la computación en la nube es el de Infraestructura como Servicio (IaaS). En ella se ofrecen recursos computacionales y de almacenamiento a través de tecnologías de virtualización (ejemplo: SAN- Storage Area Network).

Si bien existen diversas tecnologías de virtualización de almacenamiento, el DVS ofrece transparencia en la ubicación, replicación y migración de procesos para mejorar la disponibilidad del servicio. Para ello se desarrolló RDISK [10], que es un servidor de almacenamiento que soporta dispositivos físicos, imágenes de discos en archivos o imágenes de discos en memoria o cualquier otro tipo de archivo que el OS-host pueda brindar, tal como un archivo remoto a través de NFS.

Para comunicarse con RDISK se requiere de un protocolo que debe utilizar M3-IPC para la transferencia de datos y mensajes. De todos modos, está planteado un proyecto para desarrollar un *wrapper* entre el protocolo NBD y M3-IPC para que su utilización pueda ser aprovechada por Linux sin necesidad de soportar M3-IPC.

RDISK puede utilizarse con o sin replicación. En éste último caso, para sacar provecho de las características de M3-IPC se debe utilizar el esquema *Primary-Backup*. Para la comunicación entre estos procesos replicados en diferentes nodos (los cuales tienen asignado el mismo *endpoint*) se utiliza un sistema de comunicaciones grupales denominado Spread Toolkit [11]. El proceso RDISK Primario realiza un multicast atómico de las actualizaciones que se realizan sobre el dispositivo a todas las réplicas de RDISK. La replicación de datos entre el Primario y las réplicas soporta la

compresión de datos (utilizando LZ4) y se planea incorporar el soporte de encriptación utilizando OpenSSL.

En caso de fallo del Primario, Spread notifica a los servidores Backup. Estos realizan una elección del nuevo Primario en base a la información entregada por Spread y asignan el *endpoint* Primario al servidor electo. Spread le permite a RDISK detectar fallos de procesos y de nodos, particiones/reconstitución de red y rearranque de nodos.

MoL Filesystem Server(MoL-FS)

Otros de los componentes del VOS multiservidor es el servidor de sistema de archivos MoL-FS [12] que, en principio es la migración del FS server de Minix. A este servidor que se tomó como base para el desarrollo se le incorporaron otras facilidades tales como la posibilidad de utilizar diferentes tipos de dispositivos donde gestionar el sistema de archivos. Los tipos de dispositivos previstos actualmente son: imagen de dispositivo en memoria RAM, imagen de dispositivo en archivo Linux (cualquier tipo de archivo, incluso accesible vía NFS), dispositivos RDISK, y está previsto desarrollar el soporte de dispositivos con protocolo NBD.

MoL-FS, utiliza M3-IPC para recibir peticiones por parte de las aplicaciones y para acceder a los servicios RDISK. A los fines de extender su uso a aplicaciones Linux estándares, se creó un módulo FUSE (Filesystem in USErspace).

VOS Multiserver

Como se mencionó anteriormente, se tomó a MoL como base para el desarrollo de un VOS multiservidor. La arquitectura de este VOS está compuesta de los siguientes procesos que tienen su equivalencia con Minix 3:

- *Tarea del Sistema (SYSTASK)*: Es responsable de realizar operaciones de bajo nivel tales como la creación/terminación de procesos, realizar la copia de datos entre los espacios de direcciones de diferentes procesos, gestionar privilegios, temporizadores y alarmas. Dispone de funcionalidades relacionadas a la ejecución en múltiples servidores tales como la migración de procesos, y la replicación.
- *Process Manager Server (PM)*: Es responsable de ofrecer a las aplicaciones de usuario las APIs y System Calls POSIX. Gestiona además la jerarquía de procesos y sus PIDs.
- *Reincarnation Server (RS)*: Es responsable de someter a la ejecución procesos de tipo Tareas o Servidores. Los mismos pueden ponerse a ejecución en cualquier nodo del cluster donde esté previsto que el VOS ejecute, no solo en el nodo donde ejecuta el RS.
- *File System Server (FS)*: Se utiliza MoL-FS.
- *Disk Task*: Se utiliza RDISK.
- *TTY Task*: Es responsable de gestionar la interacción entre el usuario y el VOS. Para ello utiliza dispositivos TTY virtuales (PTYs) que brinda el host Linux.
- *Ethernet Task (ETH)*: Es responsable de gestionar las interfaces virtuales de red del VOS en uno de los nodos. Para ello utiliza dispositivos virtuales (TAP) que brinda el hostLinux.

- *Internet Server (INET)*: Es el servidor donde se implementan los protocolos de red TCP,UDP,IP, etc.
 - *Information Server (IS)*: Es el servidor que permite obtener información de los componentes del sistema.
- Actualmente restan por terminar los trabajos de desarrollo y pruebas de TTY, ETH e INET.

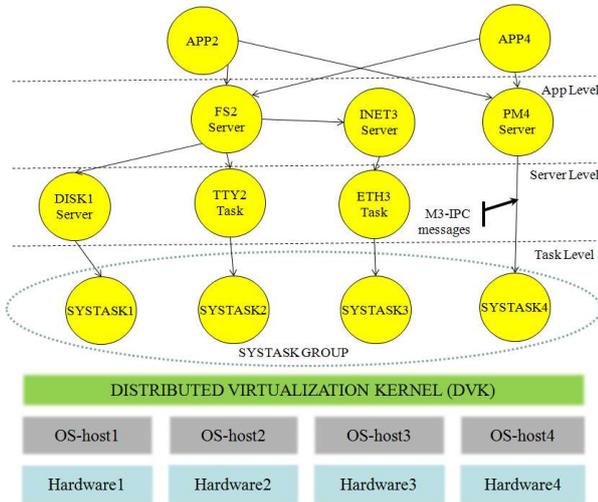


Figura 3. Arquitectura del VOS multiservidor ejecutando en DVS.

VOS Unikernel

Paralelamente al VOS multiservidor se desarrollaron tres variantes de VOS Unikernel [13] utilizando como base para el desarrollo a lwIP-tap [14]. En la Tabla I, se encuentran subrayados los componentes externos al Unikernel los que son accedidos utilizando M3-IPC.

TABLA I. VARIANTES DE VOS UNIKERNEL

Modelo	Web Server	TCP IP	Network Interface	File system	Storage
1	nweb	lwip	TAP	<u>MoL-FS</u>	Linux image file
2	nweb	lwip	TAP	FAT library	<u>RDISK</u>
3	nweb	lwip	<u>ETH</u>	FAT library	Linux image file

Queda demostrada la versatilidad que ofrece el modelo de arquitectura de virtualización propuesto para el desarrollo de diferentes VOS.

RESULTADOS Y OBJETIVOS

El objetivo primario del proyecto aquí presentado es desarrollar un modelo de DVS y acompañado de un prototipo básico como prueba de concepto. Se espera que con esta tecnología los VOS dispongan en forma transparente (para las aplicaciones y usuarios) del poder de cómputo y demás recursos de los nodos que los constituyen.

La tecnología DVS se presenta como apta para brindar servicios Cloud de tipo Infraestructura como Servicio (del inglés IaaS) porque admite balanceo de carga para mejorar el

rendimiento, facilita el mantenimiento del hardware sin interrupción del servicio y permite reducir el consumo energético del cluster desconectando nodos ociosos utilizando la migración de procesos individuales y de VOSs.

El DVS permite el particionado de los recursos de un cluster para conformar múltiples VOS-guests que podrán ser utilizados y administrados por diferentes comunidades de usuarios. Los nodos del cluster físico pueden ser asignados para uso exclusivo de un VOS en particular o compartir sus recursos entre múltiples VOS.

FORMACIÓN DE RECURSOS HUMANOS

El núcleo de la línea de I/D presentada se inició como una tesis de doctorado (aún no finalizada), pero dada la amplitud de la temática se desarrollaron varios trabajos en el ámbito de PIDs y de la cátedra de Virtualización y Sistemas Operativos Avanzados de la Facultad Regional Santa Fe de UTN. Se realizaron trabajos de fin de cátedra de varias promociones, tesis de grado y proyectos relacionados a cargo de otros docentes-investigadores.

Se prevén para el futuro un número importantes de proyectos, algunos ya mencionados en otras secciones, a ser desarrollados por integrantes del PID o de la cátedra de acuerdo a la complejidad que los mismos representen.

REFERENCIAS

- [1] Pablo Pessolani, Toni Cortes, Silvio Gonnet, Fernando G. Tinetti; "Sistema de Virtualización con Recursos Distribuidos". WICC 2012. Pág. 59-43. Argentina, 2012.
- [2] Galley S., "PDP-10 Virtual Machines". In Proc. ACM SIGARCH-SIGOPS Workshop on Virtual Computer Systems, Cambridge, MA, 1969.
- [3] Dike J., "A user-mode port of the Linux kernel", USENIX Association, Atlanta, Oct 10-14, 2000.
- [4] Pessolani Pablo; Jara Oscar, "Minix over Linux: A User-Space Multiserver Operating System", Computing System Engineering (SBESC), Florianópolis-Brazil, November 2011.
- [5] Tanenbaum A., Woodhull A., "Operating Systems Design and Implementation, Third Edition", Prentice-Hall, 2006.
- [6] D. Hall, D. Scherrer, J. Sventek, "A Virtual Operating System", Journal Communication of the ACM, 1980.
- [7] Pablo Pessolani, Toni Cortes, Silvio Gonnet, Fernando G. Tinetti. "Un mecanismo de IPC de microkernel embebido en el kernel de Linux". WICC 2013. Argentina, 2013.
- [8] Fred B. Schneider, "Implementing Fault-Tolerant Services Using the State Machine Approach: A Tutorial", ACM Computing Surveys, 1990.
- [9] Navin Budhiraja, Keith Marzullo, Fred B. Schneider, and Sam Toueg. "The primary-backup approach". In Distributed systems (2nd Ed.), Sape Mullender (Ed.). ACM Press/Addison-Wesley Publishing Co., NY, 199-216, 1993.
- [10] Mariela Alemandi, Oscar Jara, "Un driver de disco virtual tolerante a fallos", Venado Tuerto, JIT 2015.
- [11] The Spread Toolkit. <http://www.spread.org>.
- [12] Diego Padula; Mariela Alemandi; Pablo Pessolani; Silvio Gonnet; Toni Cortes; Fernando Tinetti, "A User-space Virtualization-aware Filesystem", Buenos Aires, CoNaISI 2015.
- [13] Anil Madhavapeddy and David J. Scott. "Unikernels: Rise of the Virtual Library Operating System.", Queue - Distributed Computing December 2013.
- [14] <https://github.com/takayuki/lwip-tap>

**Bases de Datos y
Minería de Datos**

Análisis de Información de Redes Sociales (Twitter)

Angélica Urrutia¹ y Carolina Nicolás²

¹Universidad Católica del Maule Talca Chile

²Universidad Santo Tomás Santiago Chile

RESUMEN

El vertiginoso desarrollo de la tecnología y con ello la globalización del conocimiento, ha generado un alto interés de las redes sociales en las organizaciones, y su presencia se ha multiplicado exponencialmente en los últimos años. Es por ello, que aquí se propone analizar los datos extraídos de redes sociales, específicamente Twitter, con la finalidad de obtener diferentes elementos que permitan la gestión y el análisis relacionado con las opiniones que los usuarios proporcionan sobre las distintas empresas que utilizan a Twitter como una herramienta social en sus páginas web. Esta información es de gran utilidad para la gestión de clientes y su preferencia a marcas u organizaciones.

CONTEXTO

Llevar un registro de las opiniones y comentarios que se están compartiendo en la red y así dar respuestas directas a dudas o críticas de los clientes, o bien analizar los datos, o promocionar productos, realizar estudios de mercado, entre otras opciones es fundamental para las organizaciones [1,5, 10].

Propuesta perfil del usuario.

El análisis de los datos extraídos de Twitter, de esta investigación, es focalizado en perfiles de organizaciones: bancarias, telefónicas, retail y supermercados, seleccionando las cuentas vigentes. El interés de la información extraída de cada página se centra en la *visualización del perfil del tipo de tuitero*, considerando las siguientes características:

- **Liderazgo**, entendiéndose con ello que se desea visualizar qué tweet es el que se está hablando con más recurrencia y a quién pertenece. Para ello es necesario contar con los mensajes emitidos y con información como **la cantidad de “retweet”** y **“Me gusta”** que poseen los tweets que hacen referencia a los diferentes perfiles.
- **Experiencia**, este enfoque busca visualizar que tan participativo o experto es el usuario que realiza el tweet, ya sea porque posee gran cantidad de mensajes emitidos, alta cantidad de personas a las que les interesa su perfil, por la cantidad de perfiles de interés o la cantidad de **“Me gusta”** que realiza a diferentes tweets. Para esto, la información que se requiere es la relacionada con el perfil del **usuario**, donde los datos relevantes para este análisis corresponden a la **fecha de creación de dicha cuenta, la cantidad de tweets, cantidad de followers, cantidad de following** y **cantidad de likes** que posee el usuario.
- **Origen**, que utiliza la información relacionada con la **localidad** del usuario, que es ingresada por él mismo al momento de crear su cuenta. Esto restringe a obtener un análisis efectivo, ya que la localidad se presenta en diversos formatos o incluso de manera incorrecta, así como también debido a que no es información obligatoria, nos encontramos con usuarios sin definir su localidad.

A modo de resumen, se presenta a continuación la Figura 1, que muestra un diagrama con los requerimientos mencionados y las características requeridas en el análisis del perfil del usuario de una red social, caso Twitter.

Para el análisis de los datos, se considera la recopilación de diferentes tweets de las páginas web de las empresas seleccionadas [2, 7], se extraen datos asociados a las características para cada perfil, y posterior implementación utilizando la herramienta, como por ejemplo Qlik Sense Desktop, con la finalidad de generar visualizaciones personalizadas e interactivas, dashboard, de datos que busquen responder las siguientes interrogantes de investigación:

- ¿Cómo el tuitero se comporta en las diferentes marcas?
- ¿Cómo afectan las opiniones en la red, al perfil de Twitter de la empresa?

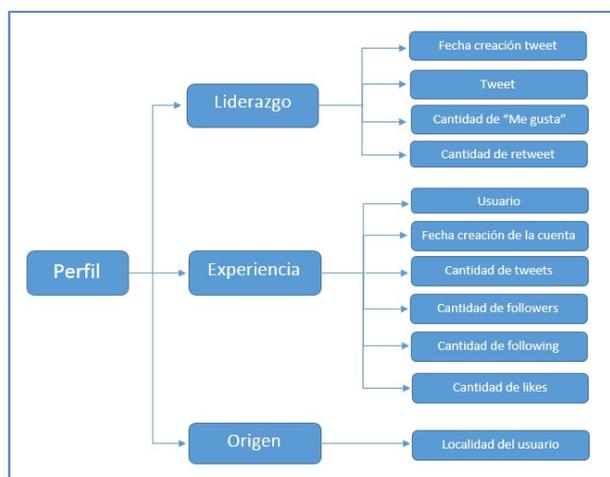


Figura 1: Características propuestas para perfil del usuario Twitter.

1. INTRODUCCIÓN

Para la implementación y obtención de análisis de resultados del perfil del usuario, es de vital importancia la existencia de una arquitectura de software definida adecuadamente, que permita extraer y analizar datos desde Twitter [3, 4, 9, 12].

Propuesta de arquitectura para analizar perfil twitter.

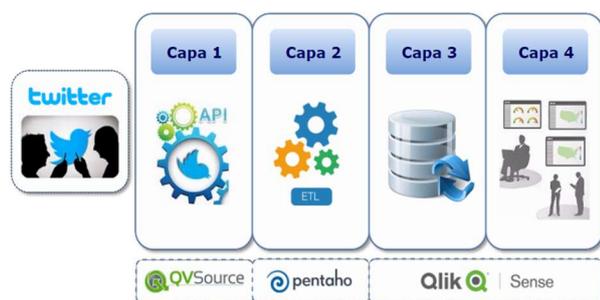


Figura 2: Propuesta de arquitectura de software para analizar perfil Twitter.

La Figura 2 muestra una propuesta de una arquitectura de software/herramientas compuestas por cuatro capas de procesos descritas a continuación, cada una de ellas requiere de herramientas específicas asociada

a diferentes tareas y procesos, como por ejemplo: QVSource, Pentaho y Qlik Sense Desktop.

Capa 1: Extracción de la fuente de datos. Es la primera capa la que realiza la extracción de datos desde Twitter haciendo referencia a los tres perfiles propuestos. Para el proceso de extracción se utiliza la herramienta QVSource, que permite conectar Qlik Sense Desktop a la gran cantidad de APIs sociales y empresariales que se encuentran disponibles en la web. En este caso, se utiliza QVSource como conector con la API Search de Twitter de forma interactiva, permitiendo así obtener los datos de todos los tweets que hacen referencia a cada usuario e información específica del usuario que realiza dichos tweets.

Capa 2: Procesos ETL. En esta segunda capa se utiliza la herramienta Spoon de Pentaho Data Integration, para llevar a cabo los procesos de transformación y carga de los datos. Los datos obtenidos, son transformados con el propósito que sean lo más claro y limpios posible, para luego generar un solo archivo con los datos de los doce perfiles seleccionados para su posterior análisis.

Capa 3: Selección base de datos. Esta capa adquiere importancia, ya que, en ella es donde se comienza a utilizar el software interactivo Qlik Sense Desktop, para el análisis del perfil de usuario Twitter. Como primera tarea, se carga el archivo generado en la Capa 2, luego seleccionar los atributos o datos de interés, para llevar a cabo el modelo de análisis de la información, ya propuesto en la Figura 1. De esta forma se obtiene la base de datos con los elementos de análisis que se requieren.

Capa 4: Visualización de los resultados. En esta última capa es donde se visualizan los resultados obtenidos desde las capas anteriores de la arquitectura propuesta. A partir del análisis de la base de datos generada y las diferentes opciones de visualización que entrega Qlik Sense Desktop, se construyen dashboard dinámicos y de fácil comprensión, con la finalidad de dar respuestas a los indicadores de gestión propuestos de manera

que sean útiles para la toma de decisiones con respecto a lo que opinan los tuiteros [6, 8, 11].

Producto de esta investigación.

Se tiene una arquitectura que permite realizar la implementación para el análisis de datos extraídos desde herramientas sociales (Twitter), relacionado con el valor de la experiencia de clientes, a través del modelo de “perfil de usuario”, para lo cual se utiliza como herramienta de visualización de datos, el programa QlikSense Desktop. Este tipo de solución presenta una gran ventaja que es su bajo costo y, además, posee la particularidad que puede ser utilizada, por cualquier persona que tenga la necesidad de resolver este modelo para el desarrollo y optimización de su organización, sin la necesidad de contar necesariamente con expertos en el área de Inteligencia de Negocios.

La solución propuesta, cumple con las etapas fundamentales de un modelo de Inteligencia de Negocios, puesto que se proporciona un modelo de análisis de información, diferentes visualizaciones, que permiten entregar respuestas a las inquietudes planteadas por las empresas para tomar las mejores decisiones basadas en información relevante.

Obtener información que proporcionan los usuarios a través de las redes sociales, constituye un elemento de alto valor y un punto fundamental para las organizaciones, ya que es a través de ellas donde las personas exponen reclamos, opiniones, experiencias o necesidades que son relevantes para ellos y que constituye una información valiosa para todo tipo de entidades que deseen permanecer conectadas con el mercado.

Para llevar a cabo este análisis de información, la red social tomada como ejemplo es Twitter, ya que, es una de las redes más utilizadas, con una gran cantidad de usuarios activos. Por esta razón, la elección de esta red para desarrollar el modelo de perfil de usuario se consideró como la más adecuada e ideal, principalmente por su característica particular de poseer mensajes de texto limitados a 140 caracteres, y gracias a que la información entregada por Twitter es de acceso público, al igual que las aplicaciones

que permiten conectarse a la API de esta red social para así obtener la información deseada. De esta manera, la posibilidad de acceder a datos generados por el mundo exterior a través de las redes sociales, está al alcance de cualquier organización, procurando así una gran y valiosa oportunidad para ellas de acercarse en forma adecuada y con información relevante a sus clientes, ya que esta posibilidad les ha permitido conocerlos con mayor profundidad.

Esta alternativa, es hoy día crucial en el mundo globalizado y con organizaciones y clientes cada vez más exigentes e informados. Por ello, las organizaciones se ven enfrentadas de manera permanente a desafíos, en que la toma de decisiones que sean correctas puede hacer la diferencia entre lograr el éxito o simplemente el fracaso. En este sentido, el acceso a información útil, relevante y confiable permite a la empresa tomar decisiones estratégicas y orientar sus operaciones de negocio enfocándose en la gestión comercial de sus clientes que impactará positivamente en la imagen corporativa de la empresa y con ello, ser competitiva y consolidar su posición en el mercado.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el transcurso de la investigación realizada se han identificado algunos temas interesantes que pueden surgir a partir de esta. A continuación se proponen algunos trabajos futuros que se pueden realizar.

Algoritmo para realizar análisis de sentimiento de los tweets: Esta investigación entrega la base para obtener diferentes análisis del valor de la experiencia de clientes de Twitter, el cual puede ser complementado relacionado otros factores asociados a los mensajes para así obtener estudios más específicos.

Gestión de la Información: Con los datos almacenados por cada tweet se puede focalizar preguntas de investigación más específicas, asociadas a la gestión de satisfacción de marca de cada cliente que

accede a la página web de las organizaciones y entrega su opinión en ella. Una tarea pendiente es ampliar las herramientas sociales como facebook entre otras.

3. RESULTADOS OBTENIDOS O ESPERADOS

El análisis de los datos y sus representaciones gráficas son relevantes en esta investigación, es así como la Figura 3 muestra el dashboard, que da respuesta a la experiencia y localidad indicados en la propuesta de perfil del usuario.



Figura 3: Dashboard en relación al enfoque de la experiencia.

El dashboard permite complementar información entre las visualizaciones, siendo un tema relevante la fecha de creación de la cuenta del usuario, de esta forma se analiza la cantidad de tweets, likes, followers y following, que el usuario posee durante el tiempo que lleva interactuando en Twitter. Por ejemplo, no es lo mismo un usuario con miles de seguidores que creó su cuenta hace 10 años atrás, que un usuario con la misma cantidad de seguidores incorporado a la red social hace un par de años, pues tiene mayor participación y popularidad este último que posee una gran cantidad de usuarios a quienes le interesa su perfil en un período más corto. Para un análisis más completo se utiliza la información que indica la cantidad de tweets que el usuario ha realizado, visualizando de esta manera los usuarios que tienen mayor participación en la red social en un corto plazo.

A pesar de que la mayoría de los usuarios no ingresa su localidad o lo hace de manera

incorrecta, se decide de igual manera, realizar un análisis donde se visualiza la cantidad de tweets realizados por la localidad del usuario, para tener una mirada de esta información y generar la posibilidad de analizar si será de utilidad según las necesidades. Para este dashboard, se ha utilizado un panel de filtrado, que permite visualizar la información, según el año en que el usuario creó su cuenta de Twitter, y también más específicamente, por día y mes de la creación de dicha cuenta. Finalmente, la información y análisis de datos tienen un fin de utilidades y formar de visualizar que apoyarán la gestión de las empresas en base al perfil de usuario propuesto en la Figura 1..

4. FORMACION DE RECURSO HUMANO

En esta investigación trabajan académicas de la Universidad Santo Tomás en Chile con el proyecto “Twitted, its effect on the experience value of the Brand”, y académicas de la Universidad Católica del Maule con el grupo de investigación en bases de datos TRICAHUE.

En Ambas universidades participan alumnos de pregrado y postgrado de las carreras de Ingeniería Comercial e Ingeniería Civil Informática, desarrollando tesis de título y grado de licenciado y magíster.

5. BIBLIOGRAFIA

- [1] GHIASSI, M.; SKINNER, J.; ZIMBRA, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with applications*, vol. 40, no 16, p. 6266-6282.
- [2] ESULI, F. SEBASTIANI (2006): Determining term subjectivity and term orientation for opinion mining. In Proceedings of EACL-06, 11th Conference of the European Chapter of the Association for Computational Linguistics, pp. 193-200.
- [3] JANSEN, B.J.; ZHANG, M.; SOBEL, K. (2009): Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, Vol. 60, No. 11, pp. 2169-2188.
- [4] KAZUSHI IKEDA, GEN HATTORI, CHIHIRO ONO, HIDEKI ASOH, TERUO HIGASHINO

- (2013). Twitter user profiling based on text and community mining for market Analysis Japon. Knowledge-Based Systems Volume 51, Pages 35-47.
- [5] LIN J., RYABOY D. (2012): Scaling big data mining infrastructure: the twitter experience. Newsletter ACM SIGKDD Explorations Newsletter Volume 14 Issue 2. pp. 6-19 New York USA.
- [6] LEWIS D.D. (1998). Naive (Bayes) at forty: The Independence assumption in information retrieval. In Proceedings of ECML-98, 10th European Conference on Machine Learning, Chemnitz, Germany, pp. 4-15.
- [7] MARTÍNEZ E., M. MARTÍN T., M. PEREA J. M., L. UREÑA A. (2011): Técnicas de clasificación de opiniones aplicadas a un corpus en español. Procesamiento del Lenguaje Natural, Revista nº 47 septiembre de 2011, pp. 163-170. Universidad de Jaén.
- [8] MOSTAFA, MOHAMED M. (2013). More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, vol. 40, no 10, p. 4241-4251.
- [9] NIGAMY K., LAERTY J., MCCALLUMZ A. (1999). Using Maximum Entropy for Text Classification. In IJCAI-99 Workshop on Machine Learning for Information Filtering.
- [10] STRATEBI (2013). Nuevas tendencias en Bussiness Intelligence del Big Data al Social Intelligence.
http://www.stratebi.es/todobi/May13/Nuevas_Tendencias_BI.pdf Madrid, España.
- [11] SAXENA, A., & GADHIYA, S. (2014). A Survey on frequent pattern mining methods-Apriori, Eclat, FP growth. In *International Journal of Engineering Development and Research* (Vol. 2, No. 1) IJEDR.
- [12] SEBASTIANI F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, 34-1, pp. 1-47.

Análisis y Elaboración de Datos para el Desarrollo de un Sistema de Indicadores de Ayuda Social

Daniel Xodo, Gustavo Illescas, Moisés Bueno, María Rosa Dos Reis

Instituto de Investigación en Tecnología Informática Avanzada, Informática de Gestión. Departamento de Computación y Sistemas, Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires.
Paraje Arroyo Seco, Tandil (7000), Argentina. +54 249 4385680
dxodo@exa.unicen.edu.ar, illescas@exa.unicen.edu.ar, bueno@econ.unicen.edu.ar, dosreis@econ.unicen.edu.ar

RESUMEN

El uso de indicadores como herramienta de control y gestión estratégica, operativa y financiera se ha generalizado a diversas áreas de acción humana. Las dificultades de su determinación son múltiples y diversas en función de las variables a vincular, valores, rangos, relación entre las mismas, relevancia de los indicadores, como así también las consecuencias de la modificación entre ellas. Las acciones de ayuda social y su administración son particularmente sensibles por su impacto en la comunidad, la variación de las condiciones de aporte y recepción de ayudas, la variación en los requerimientos, cambios en la composición poblacional, condiciones de salubridad, alternativas económicas de empleo, vivienda, morbilidad, mortalidad, preñez, natalidad y otras que afectan a la población asistida. El uso de métodos multicriterio (AHP/ANP y similares) puede mejorar significativamente la gestión, y los análisis de su aplicación en una organización funcionando (Proyecto Koinonía) son el objetivo central de este proyecto.

Palabras clave: Indicadores, Ayuda Social, Multicriterio, ANP, AHP

CONTEXTO

La propuesta emerge del proyecto de incentivos actual (03/C282, Análisis y Elaboración de Datos para el Desarrollo de un Sistema de Indicadores de Ayuda Social. UNCPBA (2017-2020), dentro del Instituto de Tecnología Informática Avanzada (INTIA) de la Facultad de Ciencias Exactas (EXA), Universidad Nacional del Centro de la provincia de Buenos Aires (UNCPBA).

Colaboran con el proyecto miembros de Universidades extranjeras, entre otros:

- Laboratorio en Ingeniería de Software de la Universidad Carlos III (UC3M), Madrid, España.
- Centro de Investigación en Matemáticas (CIMAT), A.C. Unidad Zacatecas, México.
- Universidad de Piura. Ingeniería industrial y de sistemas. Piura, Perú.
- Universidad Autónoma de Coahuila, México.

1. INTRODUCCIÓN

La gestión de asistencia a grupos sociales mediante la búsqueda de bienes y servicios y su asignación a partir de las variables definitorias de requerimientos, su ponderación cualitativa y cuantitativa y la asignación con equidad y oportunamente requiere una gran labor de seguimiento cuya realización y

utilización en la ejecución de las tareas propias del sistema es, permanentemente, post-facto.

El sistema de ayuda gestado por Koinonía ha realizado importantes avances en la estructuración sistémica de la ayuda social, con relevantes vinculaciones con otros sistemas y modelos de ayuda. No obstante, subsisten dificultades operativas y de cuantificación de recursos y necesidades que PROYECTO KOINONIA busca superar.

Al problema confluye la variación casi permanente de los recursos que el medio social aporta (variaciones en cantidad o características de los bienes disponibles a asignar) y también variaciones en las necesidades del grupo social debido a cambios en la composición poblacional, condiciones de salubridad, alternativas económicas de empleo, vivienda, morbilidad, mortalidad, preñez, natalidad y otras que afectan a la población asistida.

El control de las operaciones y la gestión del sistema puede ser establecido mediante un sistema de indicadores que, vinculando las variables representativas de la actividad y relaciones cuantitativas determine el nivel de funcionamiento de los distintos sub-sistemas y su contribución al funcionamiento del sistema general.

Así, por ejemplo, las relaciones cuantitativas y nominales del abastecimiento, y por ende la disponibilidad de bienes para asignar tendrá su propio sistema de indicadores operativos, de gestión y estratégicos.

Por otra parte, la población asistida poseerá un sistema de indicadores que refleje sus necesidades en base a las cuales, serán las asignaciones, cuya efectividad será reflejada como vínculo, mediante indicadores de causa e indicadores de efecto.

La determinación y seguimiento de estos sistemas de indicadores conlleva dificultades de diverso orden cuya resolución mediante algoritmos computacionales son el objetivo del proyecto que puede ser resumido:

1.- Análisis de demanda en el tipo de asistencia en función de los indicadores establecidos sobre la población asistida.

2.- Determinación de las fuentes prioritarias de origen de bienes acorde a la demanda establecida en el punto anterior. (Priorizar fuentes en función de la demanda, rapidez, facilidad de acceso)

3.- Evolución de la composición cuantitativa de la demanda de los bienes a asignar.

4.- Análisis de la significación de las variables poblacionales consideradas en el sistema y su evolución en la medida que estas afecten al sistema de ayuda.

En numerosos trabajos se ha encontrado evidencia que apoya la importancia del capital social en el bienestar de la sociedad (Narayan and Cassidy, 2001; Nakagawa and Shaw, 2004), incluso la forma en que el capital social es incrustado en las estructuras sociales puede contribuir al bien público (Woolcock and Narayan, 2000) y en el ámbito económico, ayudar a disminuir los costos de transacción (Fukuyama, 2001). Diversos elementos del capital social tales como la confianza, la participación comunitaria, el compromiso social, el voluntariado, parecieran influir en las actitudes y comportamientos de la sociedad.

Tsai y Ghoshal (1998), examinan las relaciones entre las dimensiones estructurales, relacionales y cognoscitivas del capital social, y entre éstas dimensiones y los patrones de intercambio de recursos e innovación del producto dentro de una compañía multinacional. En sus hallazgos más importantes encontraron que las interacciones sociales, una manifestación de la dimensión estructural, y la confianza, manifestación de la dimensión relacional, están significativamente vinculadas al intercambio de recursos. La forma de medir estas dimensiones es a través de indicadores que forman cada uno de los constructos (dimensiones) del capital social.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes del tema que se está investigando están relacionados a la aplicación de métodos de multicriterios y de la metodología Respondent Driven Sampling (RSD).

Los métodos de evaluación y decisión multicriterio comprenden la selección de entre un conjunto de alternativas factibles, la optimización con varias funciones objetivo simultáneas, un agente decisor y procedimientos de evaluación racionales y consistentes. Es decir, en un problema de decisión multicriterio se trata de identificar la mejor o las mejores soluciones considerando simultáneamente múltiples criterios en competencia. Durante la aplicación de algunos de estos métodos se identifican problemas en casos en los que se manejan numerosas alternativas y gran variedad de criterios. Lo anterior implica que el proceso de aplicación de los métodos en esos casos es complejo y su confección de forma manual se torna en ocasiones impracticable, debido a la gran cantidad de información que se maneja y la cantidad de cálculos matemáticos que se realizan, muchos de ellos complejos. La presente investigación tiene como objetivo implementar 2 métodos de toma de decisiones Proceso de Análisis Jerárquico (AHP) y Proceso Analítico en Red (ANP) en un sistema de soporte a la toma de decisiones multicriterio. (Sanchez Gonzales, 2015).

Por otra parte, el proyecto propone aplicar la metodología RDS que, consiste en un análisis teórico basado en cadenas de Markov el cual permite mostrar que este método reduce los sesgos generalmente asociados a las muestras por cadenas referenciales, además de producir estimadores asintóticamente insesgados.

Teniendo en cuenta que la muestra obtenida no es una muestra aleatoria, se hace necesario un método de estimación especial, específicamente diseñado para obtener estimadores asintóticamente insesgados de la población y que tenga en cuenta las características del proceso de muestreo utilizado.

De esta manera, en lugar de intentar estimar directamente a partir de la muestra los estimadores de la población como lo hacen los métodos de muestreo y estimación usuales, la técnica RDS utiliza la información disponible en la muestra acerca de la red

social, para obtener estimadores asintóticamente insesgados de las proporciones de la población en los distintos grupos en los que se quiere caracterizar. Estos estimadores se llaman estimadores de prevalencia de la población (Moreno, 2010).

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados alcanzados a la fecha han sido publicados en numerosos artículos entre los que se encuentran (últimos tres años):

- Dos Reis, María R.; Bueno, Moisés E.; Xodo, Daniel: “Conocimiento en acción: asignación de recursos a familias carentes mediante la aplicación de un algoritmo genético - proyecto Koinonia”. ISSN 1853-9777. Revista de la Escuela de Investigación Operativa. Año XXII N° 35. Editorial: EPIO – Págs. 107 a 127 - Mayo 2014.

- Xodo, D; Dos Reis, M; Bueno, M; Bengochea, J; Vidal Porcel, É; Suarez, D: “Asignación de Prioridades en la Ayuda Social Mediante ELECTRE”. VII Congreso de Ingeniería Industrial COINI 2014. UTN FRCH, Puerto Madryn, Octubre de 2014.

- Dos Reis, M.; Xodo, D.; Bueno, M.: “Transferencia de conocimientos a la sociedad: una experiencia de Ingeniería en proyectos de carácter social. Proyecto Koinonia”. VII Congreso de Ingeniería Industrial COINI 2014 UTN FRCH, Puerto Madryn, Octubre 2014.

- Illescas Gustavo, Sanchez-Segura María Isabel (UC3M, Madrid, España), Xodo Daniel. “Una aproximación a la aplicación de métodos matemáticos en el control de gestión por indicadores”. Revista EPIO 35. Mayo 2014.

- Illescas Gustavo, Sanchez-Segura María Isabel, Cansina Graciela Ana “Comprobación de métodos de pronóstico de indicadores dentro de la gestión del conocimiento organizacional” 3er Congreso Internacional de Mejora de Procesos Software (CIMPS 2014). Centro de Investigación en Matemáticas (CIMAT, Zacatecas A.C.). Octubre 2014, Zacatecas, México

- Trotti, M; Dos Reis, M.; Bueno, M; Xodo, D: “Algoritmo Genético para la Asignación Eficiente de Recursos en catástrofes Climáticas”. XVII WORKSHOP DE INVESTIGADORES EN CIENCIAS DE LA COMPUTACION. ISBN 978-950-34-1084-4. Salta, 2015

- Illescas, G; Bueno, M; Dos Reis, Rosa; Recofsky; Weimann, R. H.: “Herramienta de soporte para la toma de decisiones en organizaciones de asistencia social. Proyecto Koinonia”. Anales del XXVIII ENDIO – XXVI EPIO y VIII RED M M ISBN 978- 987-24267-6-7. B Blanca, 2015.

- Dos Reis, Rosa; Minvielle, M.; Garrido, J.; Xodo, D.; Bueno, M.: “Diseño de ontología para sistemas de gestión del conocimiento aplicado a una red de ONGs de ayuda social. Proyecto Koinonia”. Anales del XXVIII ENDIO - XXVI EPIO y VIII RED M M ISBN 978-987-24267-6-7. Bahía Blanca, Argentina. 20 al 22 de Mayo de 2015.

- Bueno, M.; Dos Reis, M R.; Xodo, D.: “Mejora en la calidad de la educación universitaria: sistemas inclusivos”. COINI 2015 UTN FRC, Córdoba, ISBN 978-987-1896-50-9.

- Dos Reis, María Rosa; Bueno, Moisés: “Conocimiento en Acción: Asociación Civil Proyecto Koinonia en la Universidad”. V Jornadas de Extensión del Mercosur. Eje: Educación, Comunicación y Cultura. Ponente en mesa: Educación y nuevas tecnologías. 19 y 20 de mayo de 2016. Tandil. UNCPBA.

- Dos Reis, María Rosa; Trotti, Marcos; Bueno, Moisés; Illescas, Gustavo: “Un enfoque Genético como Asistente para la toma de decisiones en situaciones de Emergencia – PROYECTO KOINONÍA”. XXIX ENDIO - XXVII EPIO. Libro Digital PDF ISBN 978-987-24267-7-4. Buenos Aires, junio de 2016.

4. FORMACIÓN DE RECURSOS HUMANOS

La estructura del equipo de trabajo se muestra en la siguiente tabla:

Apellido y nombre	Título	Cargo	Funciones
Xodo, Daniel	MBA.	Prof. UNCPBA	Director
Illescas, Gustavo	Dr.	Prof. UNCPBA	Co-director
Bueno Moisés	MBA.	Prof. UNCPBA	Integrante
Dos Reis Rosa	MBA.	Prof. UNCPBA	Integrante
Tripodi Gustavo	MBA.	Prof. UNCPBA	Integrante
Martinez, Mariano	Ing.	Aux. UNCPBA	Integrante
Montero, Norma	Ing.	Aux. UNCPBA	Integrante
Etchepare, Federico	Tesis en curso	Alumno de grado	Colaborador
Servat, Agustín	Tesis en curso	Alumno de grado	Colaborador
Figini, Iris	Tesis en curso	Alumno de grado	Colaborador
Weimann, Mariano	Tesis en curso	Alumno de grado	Colaborador
Fanjul, Luciano	Tesis de grado finalizada 2016	Alumno de grado	Colaborador
Corvi, Emiliano	Tesis de grado finalizada 2016	Alumno de grado	Colaborador

Contempla la participación docente de Facultades de Ciencias Exactas y Cs Económicas; y alumnos de ambas

Tesis de Doctorado en Administración. La Mag. Dos Reis, es doctorando del Programa Doctoral en Administración de la Facultad de Ciencias Económicas-UNCPBA.

Tesis de Doctorado en Matemática Computacional e Industrial. El Ing. Mariano Martinez, es doctorando en Matemática Computacional e Industrial, Fac. de Cs. Exactas, UNCPBA, Tesis en el tema “Modelos predictivos aplicados a indicadores”, Dir: Dr. G. Illescas.

Tesis de Doctorado en Administración: El Mg. Ing. Tripodi, integrante del proyecto, se encuentra en fase de redacción de la propuesta de tesis doctoral para el Programa Doctoral en Administración de la Facultad de Ciencias Económicas-UNCPBA.

Tesis de Maestría en Administración de Negocios: UTN-FRTL. Maestrando: Lic. Marcelo Matassa. Dir. Mag. Daniel Xodo.

Tema: “Caracterización del sector micro-empresario de Trenque Lauquen en términos de parámetros socioeconómicos”. En Elaboración Resolución del C.S. Universidad Tecnológica Nacional N° 955/2014

Woolcock and Narayan (2000). “Social Capital: Implications for Development Theory, Research, and Policy”. Final version submitted to the World Bank Research Observer To be published in Vol. 15(2), 2000

5. BIBLIOGRAFÍA

Fukuyama, F. (2001). “Social capital, civil society and development”. *Third World Quarterly*, Vol 22, No 1, pp. 7– 20.

Moreno (2010). “A review of the methodology of estimation through respondent- driven sampling for proportions of a hidden population”. *Comunicaciones en Estadística* Junio 2010, Vol. 3, No. 1

Nakagawa, Y. and Shaw, R. (2004). “Social Capital: A Missing Link to Disaster Recovery”. *International Journal of Mass Emergencies and Disasters*. March 2004, Vol. 22, No. 1, pp. 5–34

Narayan and Cassidy (2001). “A dimensional approach to measure social capital: development and validation of a social capital inventory”. *Current Sociology*. March 2001, Vol. 49(2): 59–102 SAGE Publications.

Sanchez Gonzales (2015). “Implementación del método ANP difuso en el sistema de soporte a la toma de decisiones multicriterio MultiDecision PAAT”. *Serie Científica de la Universidad de las Ciencias Informáticas* Vol. 8, No. 1, Octubre, 2015. <http://publicaciones.uci.cu/index.php/SC/article/viewFile/1660/774>

Tsai, W. and Ghoshal, S. (1998). “Social capital and value creation: The role of intrafirm networks”. *Academy of management Journal*, 41(4), 464-476.

Apache Pig en Hadoop sobre Cassandra

Susana B. Chavez¹, Adriana E. Martin², Nelson R. Rodríguez³, María A. Murazzo⁴

Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

Complejo Universitario Islas Malvinas, Av. I. de la Roza 590 (O), CP: 5402 Rivadavia,

San Juan. Tel:0264 4234129

¹schavez@iinfo.unsj.edu.ar ²arianamartinsj@gmail.com ³nelson@iinfo.unsj.edu.ar

⁴marite@unsj-cuim.edu.ar

Resumen

Los sistemas distribuidos en la web y las tecnologías informáticas distribuidas como cluster y cloud, permiten diseñar un entorno de entidades distribuidas que cooperen para resolver un problema que no puede ser resuelto individualmente. La variedad de estos sistemas pueden incluir servidores de aplicaciones, cloud privados, pequeños centros de datos y cluster para almacenamiento y búsqueda de datos.

Esto explica por qué ha crecido enormemente la habilidad de recolectar y almacenar datos en las últimas décadas, incluso hoy en día, se puede decir que este apetito por los datos no muestra signos de satisfacción.

Los científicos quieren ser capaces de almacenar más datos con el fin de construir mejores modelos matemáticos del mundo. Los vendedores quieren mejores datos para entender los deseos y hábitos de compra de sus clientes. Los analistas financieros quieren entender mejor el funcionamiento de sus mercados. Y todo el mundo quiere mantener todas sus fotografías, videos, correos electrónicos, etc.

En consecuencia, es primordial encontrar la mejor solución para el procesamiento y análisis de esta gran escala de enormes cantidades de datos.

En este sentido, un RDBMS como SQL Server o MySQL es una buena opción si el conjunto de datos de trabajo nunca va a crecer más allá de 40-50GB a lo largo de su vida útil. Incluso no necesitan ser distribuidos ya que pueden ser procesados en la memoria de una sola máquina.

Sin embargo, si se construye una aplicación que tiene un conjunto de datos que crece rápidamente y ráfagas de cargas impredecibles, será necesario optar por una solución que sacrifique cierta velocidad o consistencia en pos de poder distribuirse y así procesar el gran volumen de datos.

En los últimos años han surgido las bases de datos NoSQL que rompen una o más de las reglas de los sistemas de bases de datos relacionales. No esperan que los datos sean normalizados. En su lugar, los datos a los que accede una aplicación viven en una gran tabla, de modo que pocos o ningún *joins* son necesarios. Estos sistemas están diseñados para administrar terabytes de datos.

A esto, se suma el desarrollo de muchos sistemas alternativos de procesamiento de datos como *Apache Hadoop*. Este proyecto ha impulsado el desarrollo de lenguajes existentes y la construcción de nuevas herramientas como *Apache Pig*. Esta herramienta proporciona un

mayor nivel de abstracción para los usuarios de datos, dando acceso a la flexibilidad y potencia de Hadoop sin necesidad de tener que escribir extensas aplicaciones de procesamiento de datos en código Java de bajo nivel.

Las bases de datos NoSql que se han integrado con Pig incluyen HBase, Accumulo y Cassandra.

En este trabajo se propone realizar pruebas experimentales con Apache Pig sobre Apache Hadoop y como motor NoSql se elige Cassandra, ya que coincide muy bien con la naturaleza distribuida de Hadoop, para ejecutar consultas sobre datos que abarcan múltiples nodos.

Palabras Claves: Apache Pig, Hadoop, Nosql, Computación distribuida, Cloud Computing.

Contexto

El presente trabajo se enmarca dentro del proyecto de investigación: **Evaluación de arquitecturas distribuidas de commodity basadas en software libre**, el cual tiene como unidades ejecutoras al Departamento e Instituto de Informática de la FCEFyN de la UNSJ.

Introducción

Los sistemas de bases de datos NoSql crecieron con las principales redes sociales, como Google, Amazon, Twitter y Facebook, debido a que debieron enfrentarse a nuevos desafíos con el tratamiento de datos. Con el crecimiento de la web en tiempo real existía una necesidad de proporcionar información procesada a partir de grandes volúmenes de datos que tenían estructuras horizontales más o menos similares.

En ese sentido, las bases de datos NoSql están altamente optimizadas para las operaciones de recuperar y agregar, y normalmente no ofrecen mucho más que la funcionalidad de almacenar los registros. La pérdida de flexibilidad en tiempo de ejecución, comparado con los sistemas SQL clásicos, se ve compensada por la significativa ganancia en escalabilidad y rendimiento cuando se trata con ciertos modelos de datos [1].

Apache Cassandra es una base de datos NoSql distribuida y de código abierto, cuya principal característica es que fusiona Dynamo, de Amazon con BigTabla, de Google, siendo ambas implementaciones de código cerrado. Está basada en un modelo de almacenamiento de «clave-valor», escrita en Java. Permite almacenar grandes volúmenes de datos en forma distribuida. Su objetivo principal es la disponibilidad y la escalabilidad lineal [2][3].

“Apache Cassandra se ha convertido en una de las bases de datos NoSQL más utilizados del mundo y sirve como columna vertebral de algunas aplicaciones muy populares hoy en día”. Un ejemplo de esto, es Twitter que lo usa en su plataforma. Su objetivo principal es la escalabilidad lineal y la disponibilidad. Permite el uso de Hadoop para implementar MapReduce, ya que Hadoop puede trabajar directamente con cualquier sistema de archivos distribuido [4].

Apache Hadoop, surge como una alternativa para el procesamiento de estos datos masivos [5]. Hadoop es un proyecto de código abierto iniciado por Doug Cutting. Durante los últimos años, Yahoo! y varias otras compañías web han impulsado el desarrollo de Hadoop, basado en artículos publicados por Google que describen cómo enfrentaban al reto de almacenar y procesar las

enormes cantidades de datos que eran coleccionadas. Hadoop se instala en un grupo de máquinas y proporciona un medio para unir el almacenamiento y el procesamiento en un clúster.

El desarrollo Hadoop ha impulsado el desarrollo de herramientas y lenguajes existentes y la construcción de nuevas herramientas como Apache Pig [6].

Pig proporciona un motor para ejecutar flujos de datos en paralelo en Apache Hadoop. Incluye un lenguaje, Pig Latin, para expresar estos flujos de datos. Pig Latin incluye operadores para muchas de las operaciones de datos tradicionales (join, sort, filter, etc.), así como proporcionar a los usuarios la capacidad de desarrollar sus propias funciones para leer, procesar y escribir datos.

Pig Latin es un lenguaje de flujo de datos. Esto significa que permite a los usuarios describir cómo los datos de una o más entradas deben ser leídos, procesados y luego almacenados en una o más salidas en paralelo. Estos flujos de datos pueden ser simples flujos lineales o flujos de trabajo complejos que incluyen puntos donde se unen múltiples entradas y donde los datos se dividen en múltiples flujos para ser procesados por diferentes operadores. No hay instrucciones if o para bucles en Pig Latin. Esto se debe a que los lenguajes de programación tradicionales y orientados a objetos describen el flujo de control y el flujo de datos es un efecto secundario del programa. En cambio, Pig Latin se centra en el flujo de datos.

Pig corre en Hadoop. Utiliza el Sistema de Archivos Distribuidos Hadoop (HDFS) y el sistema de gestión de recursos de Hadoop (YARN, a partir de Hadoop 2). HDFS es un sistema de archivos distribuido que almacena archivos en todos los nodos de un clúster Hadoop. Se encarga de romper los archivos en bloques grandes y distribuirlos a través de diferentes

máquinas, incluyendo la realización de copias múltiples de cada bloque para que si una máquina falla, no se pierden datos. Presenta una interfaz similar a POSIX a los usuarios. De forma predeterminada, Pig lee los archivos de entrada de HDFS, utiliza HDFS para almacenar datos intermedios entre los trabajos de MapReduce y escribe su salida en HDFS [7].

Líneas de Investigación y desarrollo

De acuerdo a la literatura consultada, los casos de usos de Pig latino se clasifican en tres grupos: los procesos tradicionales

de extracción, transformación y carga (ETL), la investigación de los datos brutos (raw data) y el procesamiento iterativo.

El objetivo de esta línea de investigación dentro del proyecto marco, es construir modelos de predicción de comportamiento. Apache Pig se presenta como una herramienta optima para explorar todas las interacciones del usuario con un sitio web y para luego dividir a los usuarios en varios segmentos. En este caso se tomara el sitio de la UNSJ como primera experiencia.

Paso seguido, se deberá construir un modelo por cada segmento, este modelo predecirá como los miembros de ese segmento responderán a los tipos de anuncios o noticias. De esta forma, el sitio web podrá mostrar anuncios con mayor probabilidad de obtener un clic u ofrecer noticias que tengan más probabilidades de involucrar a los usuarios y que vuelvan al sitio. Por qué Pig Latin?:

- Porque puede operar en situaciones donde el esquema es desconocido, incompleto o inconsistente, además puede manejar fácilmente los datos anidados y

- Porque es un proyecto Apache open source. Esto significa que se puede descargar como fuente o binario, utilizarlo para trabajo experimental, contribuir a él y, bajo los términos de la Licencia Apache, usar sus productos y adaptarlos como mejor convenga.

Resultados Obtenidos

A pesar de que esta línea de investigación ha sido presentada para los años 2016 y 2017. Durante los últimos cinco años se trabajó en proyectos sobre Cloud Computing y en particular durante los últimos dos años sobre Cloud híbridos. La experiencia sobre los Cloud privados, junto con líneas de investigación anteriores, impulsó esta línea de investigación. El grupo ha realizado nueve publicaciones en el área durante el último año: tres trabajos en el WICC 2015-2016, un trabajo en el CACIC 2015, dos trabajos en las Jornadas de Cloud Computing, además se realizaron tres publicaciones en revistas científicas. Se han aprobado tres tesinas de grado y un trabajo de especialización.

Objetivo

En este proyecto se enfocaran las investigaciones en los sistemas de cómputo distribuidos, los cuales permiten realizar de manera más eficiente tareas de computación de alta prestaciones basadas en el paradigma de memoria distribuida. Ejemplos de Arquitecturas que soportan este tipo de sistemas distribuidos son los cluster y el cloud computing.

En particular este trabajo tiene como objetivo instalar una base de datos Nosql, en particular Cassandra, sobre un cluster montado como servicio (CaaS), para luego utilizar Apache Pig

sobre Hadoop para realizar el procesamiento de los datos.

Formación de Recursos humanos

El equipo de trabajo está compuesto por seis docentes-investigadores y cuatro alumnos.

Se están realizando cuatro tesinas de licenciatura una sobre evaluación de algoritmos de algebra lineal sobre arquitecturas diversas, otra sobre Cloud Computing Privado, otra sobre dispositivos de juegos aplicados a salud y otra sobre SOA aplicada a Cloud. Se espera realizar también una tesis de maestría sobre Metodologías de desarrollo aplicadas a SaaS, otra sobre bases de datos NoSQL y otra sobre algoritmos de Cómputo Intensivo para Big Data y su implementación en Clouds. Además aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación.

Referencias

- [1] <http://wikipedia.org/wiki/NoSQL>
- [2] <http://cassandra.apache.org>
- [3] C.Y. Kan, “**Cassandra Data Modeling and Analysis**” Copyright © 2014 - Packt Publishing.
- [4] www.siliconweeks.com
- [5] Tom White, “**Hadoop: The Definitive Guide**” Copyright © 2015 - O’Reilly
- [6] <http://pig.apache.org/>
- [7] Alan Gates y Daniel Dai, “**Programming Pig**” Copyright © 2016 - O’Reilly Media

Aplicaciones de Ontologías a Problemas Lingüísticos: Bases de Conocimiento Basadas en Texto No Estructurado

Henrión, Guillermo; Azcurra, Diego; Soria, Marcelo, Tabares, Diego

Departamento de Ingeniería
Universidad Nacional de Tres de Febrero
Valentín Gómez 4752 (1678)
Caseros, Tres de Febrero
Buenos Aires, Argentina

Facultad de Agronomía
Universidad de Buenos Aires
Av. San Martín 4453
Ciudad Autónoma
Buenos Aires – Argentina

Dpto. Desarrollo Productivo y Tecnológico
Universidad Nacional de Lanús.
29 de Septiembre 3901 (1826)
Remedios de Escalada, Lanús
Buenos Aires, Argentina.

Resumen

En proyectos anteriores de ontologías biomédicas se realizó un estudio exhaustivo de las ontologías y sus aplicaciones a las ciencias biomédicas, obteniendo además resultados sobre medidas de similitud semántica. Posteriormente se utilizó este conocimiento en problemas lingüísticos sobre textos del ámbito biomédico principalmente.

En este proyecto, y en la misma dirección, se continuará avanzando en la utilización sobre problemas lingüísticos, pero esta vez extendiendo su uso no solo a texto estático sino aplicados a distintas fuentes de conocimiento textuales, como ser RSS, redes sociales, datos abiertos, sitios específicos y datos enlazados.

Se propondrá el diseño de un motor que procese estas fuentes y las integre a una base de conocimiento unificada, utilizando ontologías como forma de representación.

Palabras clave: minería de datos, tecnologías semánticas, ontologías, representación del conocimiento, procesamiento del lenguaje natural, datos

enlazados, datos abiertos, minería de grafos.

Contexto

Este proyecto de investigación continúa la línea de trabajo en aplicaciones de Minería de datos y tecnologías semánticas en el marco de la carrera de Ingeniería en Computación de la Universidad Nacional de Tres de Febrero, en colaboración con FCEN y de la Carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús.

Introducción

Actualmente existe un creciente interés sobre cómo adquirir y estructurar conocimiento obtenido desde distintas fuentes, ya sean estas redes sociales, noticias publicadas mediante RSS e información publicada en forma de datos abiertos, y sobre cómo este conocimiento puede ser integrado de manera de poder sacar conclusiones imposibles de obtener sin la interacción de todas estas fuentes.

Para poder construir esta base de conocimiento integrada es necesario primero establecer una serie de cuestiones

sobre cómo proceder tanto al recuperar la información desde las distintas fuentes, cómo al representar esta información dentro de la base de conocimiento para que pueda ser manipulada como tal, y finalmente cómo los distintos conceptos obtenidos se integran para formar una base consistente.

Todos los conceptos deben quedar representados formalmente, en algún lenguaje de representación adecuado, motivo por el cual se incluye dentro de nuestro proyecto una línea de investigación en tecnologías de representación del conocimiento, principalmente, pero no excluyente, referidas a lógicas para la descripción.

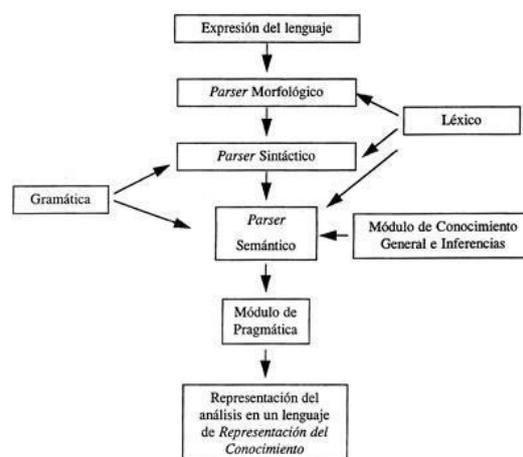
Una forma de representar el conocimiento es mediante el uso de ontologías. Una ontología es una representación formal en donde cada concepto es representado como nodos de un grafo, en donde los arcos indican las relaciones entre los conceptos. La definición formal de acuerdo a Gruber es:

“Una especificación explícita y formal de una conceptualización compartida”.

Pero para poder llegar a esta representación existe una fase de extracción de conocimiento, en donde son utilizadas técnicas de procesamiento de lenguaje natural [10]

El procesamiento del lenguaje natural (PLN) es el campo que combina las tecnologías de la ciencia computacional (como la inteligencia artificial, el aprendizaje automático o la inferencia estadística) con la lingüística aplicada, con el objetivo de hacer posible la comprensión y el procesamiento asistidos por ordenador [3].

Las etapas involucradas en el procesamiento del lenguaje natural involucran los análisis siguientes: léxico, sintáctico, semántico, pragmático y finalmente la representación de los conceptos extraídos en algún lenguaje formal de representación del conocimiento [4].



En el caso del proyecto presente las fuentes posibles a ser analizadas son las siguientes:

RSS: RSS son las siglas de Really Simple Syndication, un formato XML para syndicar o compartir contenido en la web. Se utiliza para difundir información actualizada frecuentemente a usuarios que se han suscrito a la fuente de contenidos. [9]

Datos abiertos: Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen [2].

Redes sociales: Las redes sociales son sitios de internet que permiten a las

personas conectarse con sus amigos e incluso realizar nuevas amistades, de manera virtual, y compartir contenidos, interactuar, crear comunidades sobre intereses similares: trabajo, lecturas, juegos, amistad, relaciones amorosas, relaciones comerciales, etc [7].

Datos enlazados: Los Datos Enlazados es la forma que tiene la Web Semántica de vincular los distintos datos que están distribuidos en la Web, de forma que se referencien de la misma manera que lo hacen los enlaces de las páginas web [1].

Otras fuentes: Blogs, correos electrónicos, artículos académicos, etc.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se inscribe en una línea de investigación que busca desarrollar y sistematizar el cuerpo de conocimiento de las tecnologías semánticas, principalmente en lo relacionado a representación del conocimiento, similitud semántica y procesamiento del lenguaje natural.

Así como en proyectos anteriores se estudiaron las medidas de similitud más adecuadas para establecer la similitud semántica dentro de una ontología, en el presente proyecto se estudiarán las distintas fases del procesamiento del lenguaje natural para extraer los conceptos y relaciones que constituirán la ontología que represente el conocimiento de las fuentes originales.

Resultados y Objetivos

El objetivo general de este proyecto es la aplicación de algoritmos extracción de conceptos y relaciones desde fuentes no estructuradas sobre Internet, con el fin de diseñar una base de conocimiento, que utilice ontologías como forma de representación.

En esta línea, como objetivos particulares, identificamos:

1. Proveer herramientas informáticas a investigadores de distintos ámbito (principalmente adecuados para estudios biológicos y sociales) mediante las cuales puedan organizar su información, vincular los resultados de diferentes experimentos y realizar análisis automáticos.
2. Diseño de un motor que procese fuentes de texto diversas y las integre a una base de conocimiento unificada, utilizando ontologías como forma de representación.
3. Aplicar esta tecnología en otros dominios, en donde estén disponibles grandes volúmenes de información con relaciones complejas.
4. Un objetivo transversal del proyecto es atraer e interesar a los alumnos en temas avanzados en ciencias de la computación y acercarlos a actividades de investigación.

Formación de Recursos Humanos

El equipo de trabajo está conformado por tres investigadores formados y un estudiante avanzado de la carrera de

Ingeniería en Computación. Asimismo, colaboran en el proyecto otros estudiantes de la carrera, quienes han manifestado su interés en desarrollar su trabajo de fin de carrera en la línea de investigación presentada.

Referencias

[1] Guía breve de Linked Data
<http://www.w3c.es/Divulgacion/GuiasBreves/LinkedData>

[2] ¿Qué son los datos abiertos?
<http://opendatahandbook.org/guide/es/what-is-open-data/>

[3] Procesamiento del Lenguaje Natural
<http://www.vicomtech.org/t4/e11/procesamiento-del-lenguaje-natural>

[4] Eduardo Sosa (1997)

Procesamiento del lenguaje natural: revisión del estado actual, bases teóricas y aplicaciones

[5] Gruber, T. (1995). "Toward Principles for the Design of Ontologies Used for Knowledge Sharing".

International Journal of Human-Computer Studies 43 (5-6): 907–928. doi:10.1006/ijhc.1995.1081.

[6] Chu-ren Huang "Ontology and the Lexicon - A Natural Language Processing Perspective" Cambridge University Press 2010

[7] Matthw Denny (2014)

Social Network Analysis

[8] Alexander Clark, Chris Fox, and Shalom Lappin (2010)

The Handbook of Computational Linguistics and Natural Language Processing

[9] RSS Explicada

<http://www.rss.nom.es/>

[10] John Davies, Marko Grobelnik (2009)

Semantic Knowledge Managemen

Caracterización del Aspirante a Ingresar a la Licenciatura en Ciencias de la Computación

Marisa Haderne¹, Lucía Cortés¹, Carlos García Garino^{1,2}, Carlos Catania^{1,2}, Elina Pacini^{1,2}

¹ Facultad de Ingeniería. Universidad Nacional de Cuyo.

² ITIC, Universidad Nacional de Cuyo.

hadernemarisa@yahoo.com.ar; luciacortes5519@gmail.com; cgarcia@itu.uncu.edu.ar;
ccatania@itu.uncu.edu.ar, epacini@uncu.edu.ar

RESUMEN

El problema de la deserción de los estudiantes universitarios en los primeros años de la carrera es una preocupación presente en todas las instituciones, pero el mismo puede ser analizado en una etapa previa, cuando el estudiante aun no es alumno y está transitando su ingreso a la institución seleccionada.

Entre las numerosas causas que generan el problema citado, este proyecto apunta a características y comportamientos propios del aspirante universitario. Entre otros factores que influyen en el rendimiento de los mismos pueden citarse: situación socioeconómica, condición cultural, institución en la cual cursó estudios medios, rendimiento en los exámenes de ingreso, motivación personal, comportamiento actitudinal, entre otros.

En este contexto el presente proyecto apunta a caracterizar mediante técnicas de minería de datos el comportamiento de los aspirantes universitarios de la Licenciatura en Ciencias de la Computación, a partir de datos censales y rendimiento durante el cursado de ingreso.

Luego, a partir de dicha caracterización se puede postular que los aspirantes que presenten fuertes similitudes con la población caracterizada, tendrán rendimientos similares.

Palabras clave: 1) Deserción estudiantil 2) Minería de Datos Educativa 3) Rendimiento de los estudiantes

CONTEXTO

Este trabajo se encuentra en un estado inicial y pertenece al proyecto “SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios” [11], correspondiente a los proyectos SeCTYP, convocatoria 2016, perteneciente a la Universidad Nacional de Cuyo.

1. INTRODUCCIÓN

El problema planteado en el presente trabajo es relevante para la Facultad de Ingeniería de la Universidad Nacional de Cuyo. Es fundamental conocer al aspirante de la Licenciatura en Ciencias de la Computación, su problemática académica, trayectos educativos anteriores y otros componentes sociales que pueden afectar su rendimiento al momento del ingreso.

Dicha carrera es nueva, su primera cohorte comienza en el 2017 y es muy importante para sus autoridades analizar y conocer las características de sus ingresantes desde estos primeros momentos.

De un total de 118 aspirantes, sólo 31 ingresaron (27%), 16 de los cuales presentaron equivalencias para ingreso y 15 aprobaron los exámenes por matemática y resolución de problemas correspondientes a la licenciatura. La problemática se puede observar en el 47% de alumnos que desaprobo alguna instancia de ingreso y el 26% que no realizaron ninguna actividad.



En los trabajos de García de Fanelli [9] [10], se discute el tema y en el primero de los trabajos citados se relevaron numerosos estudios que analizan los factores que influyen en el rendimiento y abandono de los estudiantes universitarios. La cuestión planteada no es solamente un problema argentino, ver por ejemplo, la carta de la Dra. Zanni-Merck acerca de la situación en Francia por ejemplo, que entre otros motivos, ha dado lugar a una colaboración entre el ICube de la Universidad de Estrasburgo, Francia y el ITIC de la UNCuyo.

El problema posee mucha relevancia local ya que en el Plan Estratégico 2021 de la UNCuyo, [15] el Objetivo Estratégico I plantea: “Contribuir al desarrollo integral de la comunidad, al bien común y a la ciudadanía plena en los ámbitos local, nacional y regional, atendiendo con pertinencia necesidades y demandas sociales, considerando los planes estratégicos provinciales y nacionales y articulando los saberes y prácticas con una clara orientación interdisciplinar, en un marco de responsabilidad institucional”.

En virtud de lo anterior, se espera que el presente proyecto contribuya a efectuar aportes concretos al plan estratégico así como también a paliar el problema del ingreso y la deserción en la Universidad Argentina; incrementando, aunque sea de manera modesta, la cantidad de egresados universitarios, disminuyendo frustraciones en el desarrollo personal de jóvenes estudiantes con dificultad para afrontar su carrera universitaria, especialmente en los primeros años y el consecuente ahorro de

recursos, siempre escasos, generalmente destinados a los primeros años de la universidad, muchas veces superpoblados de alumnos recursantes. Estas preocupaciones han motivado trabajos de difusión como resulta el ya clásico texto de Jaim [14].

Existen numerosos trabajos que discuten el tema, Garbanzo Vargas [8] agrupa a los diferentes factores que caracterizan el comportamiento de un estudiante en tres grupos: a) Determinantes personales; b) Determinantes sociales y c) Determinantes institucionales. Muchos de los factores no se pueden medir con facilidad en el momento del acceso del estudiante a la Universidad. Luego, a priori, se trabajará con factores propios de los determinantes sociales, ya que los mismos parecen, según estudios realizados, brindar un buen punto de partida.

Parece importante además contemplar otras características propias de los alumnos. Entre ellos pueden mencionarse el estilo de aprendizaje, que para el caso de los alumnos de Ingeniería puede obtenerse con relativa facilidad mediante el empleo del Test de Felder [3], disponible en forma gratuita en línea. Otro elemento importante a considerar es el de la generación Milenio, tema que ha sido abordado por la Dra. Lucía Brottier en un proyecto bienal [1] financiado por la UNCuyo ya que los alumnos ingresantes de las próximas cohortes en gran medida responden a este patrón de comportamiento. Un trabajo de interés acerca de esta generación es el estudio del empleo de redes sociales, intercambio de conocimiento y aprendizaje de los estudiantes debido a Eid y Al-Jabri [2]. Cabe señalar que estos factores fueron señalados en las conclusiones del proyecto de “Expectativas laborales de las nuevas generaciones.”[1] como indicadores a tener en cuenta para profundizar su estudio.

En cualquier caso, los factores (atributos) para caracterizar la población inicial surgirán de los estudios mencionados y de una discusión interdisciplinar entre los integrantes del equipo de trabajo, entre quienes se encuentran especialistas en informática, postgraduados en educación, profesores universitarios, etc.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Además de la relevancia general y local del problema corresponde analizar la relevancia disciplinar del tema. La Minería de Datos Educativa es una línea que ha cobrado mucha relevancia mundial en los últimos años, como se observa en el portal dedicado al tema que lleva adelante la International Educational Data Mining Society [7]. En dicho sitio se expresa textualmente: “La Minería de Datos en educación es una disciplina emergente, orientada al desarrollo de métodos para explorar el creciente volumen de datos provenientes de los sistemas de educación, y utilizar esos métodos para comprender mejor a los estudiantes y sus entornos de aprendizaje”. La citada sociedad edita un journal específico: Journal of Educational Data Mining y lleva adelante las conferencias internacionales acerca del tema, la última de las cuales tuvo lugar en Madrid en 2015.

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo principal del proyecto es caracterizar el comportamiento de los aspirantes universitarios a la licenciatura en ciencias de la computación como herramienta predictiva del ingreso. Con este propósito se plantean una serie de Objetivos Específicos:

1. *Seleccionar los atributos de Interés para caracterizar la población estudiantil*

Se tendrán en cuenta atributos que permitan detectar, definir y analizar los factores fundamentales, externos e internos, que influyen en el rendimiento académico, desgranamiento e igualdad de oportunidades para el ingreso de los aspirantes a la carrera. Entre otros factores de interés para caracterizar la población estudiantil se pueden mencionar la situación socioeconómica; condición cultural, institución en la cual cursó estudios medios; rendimiento académico en trayectos anteriores, motivación personal, comportamiento actitudinal, etc. Estará a cargo de un equipo multidisciplinar conformado por

Ingenieros, Informáticos y especialistas en Educación. Se tomarán en cuenta estudios previos como los realizados por miembros del grupo de trabajo del proyecto Fernández [5], Raichman [16], Haderne [12] [13], Brottier [1].

2. *Obtener una base de datos inicial con información sobre la población estudiantil.* Los atributos de interés de la población estudiantil se encuentran distribuidos en diferentes sistemas de información, como el Sistema SIU Guaraní, que muchas veces se complementan con registros ad-hoc propios de cada unidad académica.

Esta información se tendrá que integrar debidamente en una base de datos que incluirá además información sobre el rendimiento académico de la población. La base de datos inicial resulta de vital importancia para la adecuada creación y verificación de los distintos perfiles de alumnos a través de las técnicas de aprendizaje de máquinas.

3. *Elaborar modelos descriptivos de alumnos en base a técnicas de minería de datos y aprendizaje automático.*

A partir del análisis de los atributos disponibles en la base de datos, es posible construir diferentes modelos de aspirante. Estos modelos tendrán como objetivo describir el comportamiento de los estudiantes a lo largo de su ingreso. En particular, los modelos permitirán reconocer diferentes grupos a partir de similitudes en los atributos de interés, para luego correlacionar con el rendimiento y comportamiento observado durante el cursado.



4. *Elaborar modelos predictivos acerca del comportamiento y rendimiento esperado de una nueva población de alumnos.*

Conociendo el rendimiento histórico de los diferentes grupos de aspirantes es posible, mediante la utilización de técnicas de aprendizaje automático, construir modelos que permitan la predicción del rendimiento académico de una nueva población de alumnos. La elaboración de estos modelos se focalizará en la realización de una detección temprana de posibles fallas en la población de

estudiantes, para así facilitar la aplicación de medidas para contrarrestar los resultados negativos (aspirantes que desapruban o no realizan ninguna actividad, 73% durante el 2016).

Así, la principal contribución que se espera obtener de la investigación propuesta en el corto plazo es la creación de una base de datos de aspirantes sobre la cual se construirán diversos modelos descriptivos y predictivos aplicados al comportamiento y rendimiento de los estudiantes en su estado inicial, la etapa de ingreso, y luego poder acompañarlos a fin de lidiar con la problemática de la deserción en los primeros años del ciclo. Para ello se está trabajando con el proyecto “SIPREU: Un sistema inteligente para caracterizar y predecir el rendimiento de estudiantes universitarios” que engloba al presente trabajo.

Este proyecto es bianual y se inició en 2016, con lo cual aún no tenemos resultados.

Resultados esperados

Se espera que muchos de los análisis y estudios, plasmados con las herramientas correspondientes, puedan extenderse con relativa simplicidad a otras Unidades Académicas de la UNCuyo, así como a otras universidades, como resulta por ejemplo el módulo de extracción y pre procesamiento de datos del Sistema Guaraní.

Al finalizar el trabajo se podrá obtener una caracterización del comportamiento de los aspirantes universitarios y utilizar la misma como herramienta predictiva durante los primeros años, y poder colaborar con los actores principales de la Facultad que trabajan junto con los alumnos para poder disminuir la actual deserción.

Se espera poder completar un diagnóstico de la problemática de rendimiento académico de los aspirantes de la Facultad, así como el diseño de herramientas que faciliten la gestión de datos académicos y permitan la detección temprana de alumnos de alto riesgo. Asimismo, este proyecto pretende contribuir a la formación de recursos humanos y responde a la necesidad de incentivar en los docentes del área a la realización de trabajos de

investigación y su difusión entre los actuales o potenciales investigadores.

Es importante destacar que el presente trabajo está en concordancia con el Plan Estratégico 2021 [15], de la UNCuyo en cuanto a los objetivos de:

- Fortalecimiento de las políticas de ingreso, permanencia y egreso de los estudiantes.
- Fortalecimiento y diversificación de la modalidad de educación a distancia y promoción del uso de las tecnologías de la información y la comunicación en los procesos de enseñanza y aprendizaje, tanto en los ámbitos educativos presenciales como virtuales.
- Uso intensivo de las tecnologías de la información y la comunicación y del Sistema Informático Universitario, para la gestión institucional de la Universidad en todas sus funciones.

4. FORMACIÓN DE RECURSOS HUMANOS

Parte del equipo de trabajo del presente proyecto ha trabajado en los últimos años en diversas áreas relacionadas con la caracterización de los alumnos en la Universidad Nacional de Cuyo. Se contará con su experiencia a la hora de definir los atributos de interés para la construcción de los modelos, como así también la evaluación de los resultados generados por estos.

Es importante señalar que el equipo de trabajo reúne especialistas de diferentes campos: ingeniería, informática y educación. También hay una conjunción de especialistas con experiencia en investigación que cuentan con publicaciones indexadas y otros investigadores menos habituados a publicar, pero que cuentan con títulos de postgrado (maestrías y doctorados).

El equipo de trabajo está conformado por investigadores de las Facultades de Ingeniería, Ciencias Exactas y Naturales y Ciencias Económicas de la UNCuyo, así como integrantes del ITIC (Instituto de Investigación y Desarrollo para las Tecnologías de la Información y las Comunicaciones) y del Instituto Tecnológico Universitario (ITU).

5. BIBLIOGRAFIA

- [1] L. Brottier: Proyecto 06/B005 Expectativas laborales de las nuevas generaciones. Proyecto bienal, Universidad Nacional de Cuyo, 2013-2015.
- [2] M. Eid and I. Al-Jabri. Social Networking, knowledge sharing, and student learning: The case of university students, pp. 14-27, vol. 99, Computers and Education, 2016.
- [3] R. Felder and L. Silverman: Learning and Teaching Styles In Engineering Education, pp. 674–681, Vol. 78, No. 7, Engineering Education, 1988.
- [4] Felder: Test de Felder.
<http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSpa.html>
- [5] D. Fernández . Proyecto 06/B188 Una mirada introspectiva de los indicadores académicos que influyen tanto en el desempeño y el desgranamiento de los estudiantes como en la duración real de la carrera. Proyecto bienal, Universidad Nacional de Cuyo, 2009-2011
- [6] D. Fernández. Proyecto 06B/246 Detección, identificación y análisis de los factores fundamentales, externos e internos, que influyen y definen el ingreso de los aspirantes a Ingeniería y Arquitectura de la Universidad Nacional de Cuyo. Proyecto bienal, Universidad Nacional de Cuyo, 2011-13 [7]
 Educational Data Mining Web Site:
<http://www.educationdatamining.org/>
- [8] G. M. Garbanzo Vargas. Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública: pp. 43-63, vol. 31, núm. 1, 2007, Educación. Universidad de Costa Rica. San Pedro, Montes de Oca. Disponible en:
<http://www.redalyc.org/articulo.oa?id=44031103>
- [9] García de Fanelli. Rendimiento académico y abandono universitario: Modelos, resultados y alcances de la producción académica en la Argentina, Revista Argentina de Educación Superior (RAES), Año 6, Número 8, junio 2014, ISSN 1852-8171.
- [10] A.García de Fanelli. La cuestión de la graduación en las universidades nacionales de la Argentina. Indicadores y políticas públicas a comienzos del siglo XXI. Dossier. Propuesta Educativa, 17-31, Año 24, No 43, Vol. 11, Junio 2015.
- [11] C.G.Garino, M.Haderne, C.Catania, E. Pacini. SIPREU: An Intelligent System for characterization and Prediction of University Students performance. ADNTIIC 2016: Advances in New Technologies, Interactive Interfaces and Communicability. Córdoba.
- [12] M. Haderne. Uso de Tecnologías de la Información para detectar posibles deserciones universitarias. VII Congreso de Tecnología en Educación y Educación en Tecnología, TE & ET 2012
- [13] M. Haderne y A. Marin. Hacia una educación de calidad: uso de minería de datos para la detección precoz de factores de riesgo académico, pp 90-102, Vol. 1. Tecnologías y aprendizaje. Avances en Iberoamérica. 2013
- [14] G. Jaim Etcheverry. La tragedia educativa, Fondo de Cultura Económica, Octubre de 2000, ISBN 9789505573219.
- [15] Universidad Nacional de Cuyo: Plan estratégico 2021. Disponible en:
<http://www.uncu.edu.ar/planificacion/upload/plan-estrategico-1.pdf>. 2014
- [16] S. Raichman. 06/B157 Una propuesta de Educación a Distancia como apoyo a la modalidad presencial de la asignatura Geometría Analítica en Carreras de Ingeniería. Proyecto Bienal 2007 – 2009.

Contribuciones a las Bases de Datos no Convencionales

Jorge Arroyuelo, Maria E. Di Genaro, Susana Esquivel, Alejandro Grosso, Verónica Ludueña,
Cintia Martínez, Nora Reyes

Dpto. de Informática, Fac. de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis

{bjarroyu, mdigena, esquivel, agrosso, vlud, nreyes}@unsl.edu.ar, cintiavmartinez@hotmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México

elchavez@cicese.mx

Karina Figueroa

Fac. de Cs. Físico-Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, México

karina@computo.fismat.umich.mx

Rodrigo Paredes

Dpto. de Cs. de la Computación, Fac. de Ingeniería, Universidad de Talca, Chile

raparedede@utalca.cl

Resumen

El advenimiento de las ciencias de la computación a todos los ámbitos de la vida moderna, ha exigido el desarrollo de aplicaciones que satisfagan los requerimientos de distintos tipos de usuarios, desde campos muy dispares, adaptándose a todo tipo de exigencias para lograr un alcance masivo. Claramente, esto implica lograr manipular eficientemente datos no convencionales muy disímiles como: huellas digitales, imágenes, audio, secuencias de ADN, texto, video, etc. Como las soluciones tradicionales no suelen hacer frente a tales requerimientos, es necesario utilizar depósitos especializados y búsquedas no exactas sobre estos tipos de datos.

Además de proveer una respuesta rápida y adecuada a dichas demandas, es necesario un uso eficiente del espacio disponible, y al considerar bases de datos masivas, las estructuras en particular serán *estructuras de datos con I/O eficiente*. Las *Bases de Datos Métricas* son uno de los modelos generales en los cuales se pueden utilizar estructuras de datos especializadas que contemplen estos aspectos. Los lenguajes de consulta no siempre poseen el poder expresivo necesario para reflejar las consultas consideradas de interés. Así, nuestra investigación pretende contribuir a consolidar este nuevo modelo de bases de datos desde varias perspectivas.

Palabras Claves: bases de datos no convencionales, expresividad, índices, lenguajes de consulta.

Contexto

El desarrollo del presente trabajo se enmarca en la línea de investigación *Bases de Datos no Convencionales* del Proyecto Consolidado 30314, *Tecnologías Avanzadas de Bases de Datos*, de la Universidad Nacional de San Luis La investigación que se realiza en

este proyecto, abarca aspectos relacionados con lograr que las bases de datos no convencionales, destinadas a manipular datos no estructurados, alcancen la madurez de las bases de datos tradicionales. Esto incluye la expresividad de los lenguajes de consulta, los operadores necesarios para responder preguntas de interés, y el análisis de aspectos teóricos, empíricos y aplicativos de los mismos; contribuyendo así a distintos campos de aplicación: sistemas de información geográfica, computación móvil, robótica, visión artificial, motores de búsqueda en internet, diseño asistido por computadora, etc.

Nuestras investigaciones se realizan en la colaboración con investigadores de otros grupos de: Universidad de Talca (Chile), Universidad Michoacana de San Nicolás de Hidalgo (México) y Centro de Investigación Científica y de Educación Superior de Ensenada (México).

Introducción

La evolución de la tecnología ha permitido que la computación se introduzca en los diferentes aspectos de la vida, tanto el social, como el laboral, productivo, recreativo, de la salud, etc. Esto se ha logrado gracias al desarrollo de aplicaciones capaces de adaptarse tanto a estos nuevos entornos como a los diversos usuarios de las mismas. Para ello, las bases de datos han debido evolucionar hasta ser capaces de administrar todo tipo de datos y responder consultas sobre los mismos de una manera totalmente diferente a la tradicional, muchas veces más intuitiva. Estos

avances contribuyen en aplicaciones como: comparación de huellas digitales, bases de datos médicas, reconocimiento de voz, reconocimiento facial, reconocimiento de imágenes, recuperación de texto, biología computacional, minería de datos, clasificación y aprendizaje automático, etc.

A pesar de su diversidad, todas estas aplicaciones tienen características comunes, que pueden englobarse en el modelo de *espacios métricos*. Formalmente un espacio métrico consiste de un universo de objetos U y una función de distancia definida entre ellos $d : U \times U \rightarrow \mathbb{R}^+$ que mide la disimilitud entre los objetos. Este escenario es propicio para resolver demandas tales como, introducir una imagen en un buscador, esperando obtener aquellas que sean similares a la muestra. Por ello, las *búsquedas por similitud*, provistas por este modelo, son más naturales sobre estos tipos de datos.

Los *Métodos de Acceso Métricos* (MAMs) permiten responder eficientemente a este tipo de búsquedas, evitando la examinación secuencial de los datos. Sin embargo, es esencial su optimización, ya que la mayoría de estos métodos no están diseñados para soportar conjuntos masivos de datos, ni operaciones de búsqueda complejas, y además no admiten dinamismo. El trabajo con bases de datos masivas, o con objetos muy grandes, da lugar también a investigar sobre diseñar MAMs más eficientes para memorias jerárquicas, al cambiar el modelo de costos.

Otras áreas de investigación exploradas buscan incrementar la expresividad de los lenguajes de consulta para expresar consultas más precisas y caracterizar nuevas arquitecturas que permitan reducir el flujo de bits entre el procesador y la memoria en relación a la cantidad de datos utilizados por cada programa, para mejorar el desempeño en administradores de bases de datos (DBMS) a bajo nivel.

Líneas de Investigación y Desarrollo

Expresividad de los Lenguajes

Si se piensa en una base de datos simplemente como una estructura finita, se pueden utilizar las lógicas para expresar consultas sobre éstas. El empleo de lógicas para expresar consultas (o problemas) da origen a la complejidad descriptiva, que clasifica a los problemas según el uso de recursos lógicos tales como el número de variables, cuantificadores, operadores, etc. Existe una relación estrecha entre estos dos tipos de complejidades para clases que se identifican con la computación factible, pero el dominio de las estructuras debe estar ordenado. En ese ca-

so la clase de complejidad P es capturada por FO (*First-Order Logic*) extendida con un operador de punto fijo. Aún así, estas lógicas todavía resultan incompletas, ya que ninguna caracterización lógica de computación factible es conocida para estructuras cuyo dominio no está ordenado.

A. Dawar demuestra en [9] que ninguna extensión de la lógica de punto fijo, con un número finito de cuantificadores generalizados, captura la clase de complejidad P y en [10] que ciertos problemas NP completos sobre inequivalencia de autómatas finitos son expresables en el fragmento existencial de la lógica SO^ω mientras que el problema NP completo 3-coloreabilidad no lo es. Es importante utilizar diferentes lógicas para separar problemas que en complejidad clásica son vistos como similares.

En esta línea de investigación se continúa con la línea estudiada por Dawar en SO^ω , la cual plantea una restricción semántica a la SO , donde la valoración de las variables relacionales para los cuantificadores de segundo orden son cerrados bajo \exists_k . Se define una nueva lógica de tercer orden (TO), denominada TO^ω . Ésta intenta caracterizar y estudiar clases de complejidad relacionales (temporales) de lógicas de orden superior. Una relación se dice *cerrada bajo \exists_k* si todas las tuplas del dominio sobre el que trabaja, que satisfacen las mismas formulas de FO^k , están en la relación. Se definió una variación de una máquina relacional no determinística, denotada como 3-NRM, donde se permiten relaciones de tercer orden en el *relational store*; esto permite asociar TO^ω a una clase de complejidad temporal. Esa clase de complejidad fue llamada $NEXPTIME_{3,r}$, como la clase de máquinas 3-NRM que trabajan en tiempo exponencial de acuerdo al tamaño de la entrada. La clase $NEXPTIME_{3,r}$ es exactamente caracterizada por el fragmento existencial de TO^ω [1].

Arquitecturas de Procesadores Orientadas a Bases de Datos

La arquitectura del procesador es la funcionalidad que se le provee al programador en lenguaje de máquina, modos de direccionamiento, operaciones, interrupciones y entrada-salida [2]. En ella se distinguen: la organización básica del flujo de datos y el control (*implementación*) y la estructura física que comprende la implementación (*realización*). El lenguaje de máquina actual no es ni un lenguaje de aplicación ni un lenguaje de hardware, sino algo intermedio. Entonces, ¿por qué no interpretar directamente un lenguaje de alto nivel en lugar de compilar a un lenguaje intermedio? o ¿por qué no darle acceso

directo al programador, o compilador, al hardware en lugar de restringirlo al lenguaje de máquina?

Se puede elegir la estrategia de “impulso hacia arriba”; es decir subir el nivel, para mejorar el desempeño de la máquina, además de facilitar el uso del lenguaje de máquina. Un aspecto a considerar en este caso es el tráfico de bits y la forma usual de reducirlo es tener una arquitectura que haga lo más posible con cada búsqueda de instrucción, abandonando la arquitectura de bajo nivel y yendo tan alto como el software lo permita. El otro aspecto es explotar la concurrencia, porque si una implementación conoce más sobre lo que debe ser hecho entonces es posible que a menudo realice varias acciones simultáneamente. El implementador posee varias técnicas para aumentar la concurrencia: paralelizar, segmentar (pipelining), adelantar, poner a un lado (cache look-aside), adivinar y corregir (control and data prediction). La otra estrategia es considerar el “impulso hacia abajo”. Aún si todas las aplicaciones fueran escritas en lenguaje de alto nivel, hay razones para definir una arquitectura de computadora de nivel más bajo, pues existe conflicto de intereses entre usuario e implementador: el usuario desea expresar en forma simple y breve, haciendo uso del contexto, y el implementador desea que cada instrucción sea interpretada independientemente del resto.

Por lo tanto, es importante definir una arquitectura cuando se construye una computadora. En la actualidad la investigación sobre arquitecturas de procesadores ha sido desplazada por la de implementación de procesadores. La mayoría de los trabajos de investigación se dedican a mejorar técnicas de predicción (tanto de control como de datos), técnicas para sincronizar y comunicar procesadores (núcleos) mediante mensajes y/o memoria compartida. Muchas de estas técnicas de implementación surgieron en los años 60 y hoy se han incorporado a los diseños de microprocesadores actuales. Sin embargo, estas técnicas de implementación se podrían aplicar a todo tipo de arquitectura, desde una arquitectura

RISC¹ (que intenta acercar el lenguaje de máquina al hardware del procesador) a una arquitectura que se aleje del hardware e intente disminuir el tráfico de bits entre procesador y memoria. El objetivo en esta área es plantear nuevas arquitecturas que minimicen el tráfico de bits entre el procesador y la memoria. Se está construyendo un simulador del set de instrucciones AMD-64 o x86-64, como “benchmark”, para evaluar el tráfico de bits, como Specint y Specfp para

la arquitectura x86. Luego, se evaluará el tráfico de bits para la arquitectura propuesta sobre los mismos benchmarks, lo que implica construir tanto el simulador de la arquitectura como el compilador C para la misma. Finalmente, se pretende aprovechar al conocimiento adquirido para, desde bajo nivel, mejorar el desempeño de los DBMSs.

Bases de Datos Métricas

Las bases de datos no convencionales, que gestionan imágenes, videos, texto libre, secuencias de ADN, audio, etc., pueden modelarse con espacios métricos generales. Debido a lo costoso que resultan los cálculos de distancia, el número de cálculos realizados al crear el índice o al realizar búsquedas es usado como medida de complejidad. Por ello, el objetivo aquí es optimizar los MAMs, necesarios al momento de responder las diversas consultas a una base de datos, analizando aquellos que han mostrado buen desempeño en las búsquedas para reducir su complejidad considerando, cuando sea necesario, la jerarquía de memorias. En general, dada una base de datos $X \subseteq U$ y una consulta $q \in U$, las consultas por similitud son de dos tipos: por *rango* o de *k-vecinos más cercanos* (*k*-NN).

Métodos de Acceso Métricos

Como se dijo anteriormente, una de las optimizaciones necesarias a los MAM's es el dinamismo. Por ejemplo, considerando el *Árbol de Aproximación Espacial (SAT)*, un índice con muy buen desempeño en espacios de mediana a alta dimensión, pero totalmente estático, se desarrolló el *Árbol de Aproximación Espacial Dinámico (DSAT)* [12] que permite realizar inserciones y eliminaciones, conservando muy buen desempeño en las búsquedas, pero que agrega un parámetro a sintonizar. El *Árbol de Aproximación Espacial Distal (DiSAT)* [6], una variante también estática del *SAT* y sin parámetros, logra optimizar las búsquedas respecto de ambos (*SAT* y *DSAT*). Por ello, se ha propuesto la *Foresta de Aproximación Espacial Distal (DiSAF)* [4], que es dinámica, para memoria principal y que para lograr mejorar al máximo su desempeño, aplica la técnica de dinamización de Bentley y Saxe al *DiSAT* y aprovecha el profundo conocimiento que se tiene sobre la aproximación espacial.

Sin embargo, muchas veces los índices no caben en memoria principal, ya sea porque administran una base de datos masiva, o porque los objetos de la misma son muy grandes. Entonces surge la necesidad de diseñar índices para memoria secundaria. Muchos de estos índices se basan en “agru-

¹ Acrónimo del inglés “Reduced Instruction Set Computer”.

par elementos”; y para analizar cuán buenos son los agrupamientos que logran, se pueden utilizar estrategias de optimización basadas en heurísticas bioinspiradas. Teniendo esto en consideración, se han diseñado dos nuevos índices basados en la *Lista de Clusters(LC)* [7] que son totalmente dinámicos, es decir, admiten inserciones y eliminaciones de elementos y están especialmente diseñados para trabajar sobre grandes volúmenes de datos [13]. La *Lista de Clusters Dinámica (DLC)*, tiene buen desempeño en espacios de alta dimensión, con buena ocupación de página y operaciones eficientes tanto en cálculos de distancia como en operaciones de I/O. Sin embargo, las búsquedas en ella deben recorrer completamente la lista de centros de los clusters, elevando los costos. El *Conjunto Dinámico de Clusters (DSC)*, también mantiene los clusters en memoria secundaria, pero organiza los centros de clusters en un *DSAT* en memoria principal, permitiendo que las búsquedas realicen menos cálculos de distancia y accedan a menos páginas/clusters. La información de ese *DSAT* también se aprovecha en las inserciones, mejorando los costos de las operaciones en cálculos de distancia y manteniendo los bajos costos de acceso a disco. Ambos, *DLC* y *DSC*, han demostrado tener una razonable utilización de páginas de disco y son competitivas respecto a las alternativas representativas del estado del arte.

Algunas aplicaciones requieren que las respuestas sean aún más rápidas, aunque sea a costa de algunos errores: se intercambia precisión (devolviendo sólo algunos objetos relevantes) por velocidad en la respuesta. Este tipo de búsquedas se denominan *aproximadas*. Para conjuntos de datos masivos, las búsquedas por similitud aproximadas permiten obtener un buen balance entre el costo de las búsquedas y la calidad de la respuesta obtenida. El *Algoritmo Basado en Permutaciones (PBA)* [3], es uno de los mejores representantes de este tipo de consultas, logrando una respuesta de alta calidad a un bajo costo. Por ello, se ha diseñado la *Lista Dinámica de Permutaciones Agrupadas (DLCP)* [11] (combina *LC* con *PBA*), que además es dinámica y para memoria secundaria. Este índice agrupa por distancia entre las permutaciones de los objetos, en lugar de por distancia entre objetos y se le puede indicar cuántos cálculos de distancia y/o operaciones de I/O utilizar, para obtener una respuesta rápida, aunque menos precisa.

All-k-NN Aproximados

Entre las aplicaciones que pueden modelizarse con espacios métricos, se encuentran algunas como

la predicción de funciones, la clasificación y aprendizaje automático, la cuantificación y compresión de imágenes; que utilizan las búsquedas por similitud *k-NN*. Dado un elemento $u \in U$ y sea $X \subseteq U$ la base de datos, $k\text{-NN}(u)$ obtiene los k elementos en $X - \{u\}$ que tengan la menor distancia d a u . Una variante menos analizada de esta primitiva, denominada *All-k-NN*, es la búsqueda de los k -vecinos más cercanos de todos los elementos de X ; es decir, obtiene los $k\text{-NN}(u_i)$ para cada $u_i \in X$. Si $|X| = n$, se pretende resolver este problema realizando menos de n^2 cálculos de distancia, preprocesando los datos para reducir el número de cálculos de distancia (la complejidad en este modelo).

Se han propuesto soluciones a este problema para espacios métricos generales [14], basadas en la construcción del *Grafo de los k-vecinos más cercanos (kNNG)*. El *kNNG* indexa un espacio métrico y luego se emplea en la resolución de las consultas por similitud. Su desempeño en las búsquedas supera al de las técnicas clásicas basadas en pivotes. La solución habitual para obtener el *kNNG* suele construir un índice y luego realizar búsquedas de los k vecinos más cercanos, para todos los elementos de la base de datos. Una aproximación al *kNNG* puede ser el *Grafo de vecinos cercanos (knNG)* [5]. Un caso particular se obtiene cuando $k = 1$ y así el *1nNG* sería el grafo que conecta a cada elemento con un elemento cercano, que puede ser, o no, su vecino más cercano.

Las *búsquedas por similitud aproximadas* se requieren en muchas aplicaciones que priorizan la velocidad sobre la precisión [15, 7, 16, 8]. También resultan útiles en espacios métricos de alta dimensión donde, resolver consultas requiere revisar casi todo el conjunto de datos sin importar la estrategia utilizada. Lo mismo sucede cuando la función de distancia es demasiado costosa de calcular, o si se tiene una base de datos masiva. En estos casos además, puede surgir otro problema: el costo de la construcción de un índice, para luego obtener los vecinos cercanos, puede resultar excesivo.

Considerando estas situaciones y que gracias al conocimiento profundo que se tiene del *DiSAT* se sabe que la información que se obtiene durante la construcción del índice puede aprovecharse para la obtención del *1nNG*, se ha propuesto en [5] un enfoque novedoso al problema: cada objeto es relacionado con el elemento más cercano de la base de datos con el que se comparó durante la construcción. Más aún, la construcción del *DiSAT* retorna una aproximación del *1-NNG*, la cual puede mejorarse median-

te reconstrucciones adicionales. Es posible obtener así un adecuado 1nNG, sin realizar búsquedas y utilizando una fracción del costo de la solución habitual. La nueva propuesta permite recuperar el 1nNG con bajo costo, una muy buena precisión y un error bajo, logrando un buen compromiso calidad/tiempo.

Resultados y Objetivos

Los estudios realizados sobre el modelo de espacios métricos permitirán mejorar el desempeño de los MAMs analizados y estudiar la aplicación de los resultados obtenidos a otros [4, 5, 6, 13, 11].

Se profundizará el estudio del diseño de estructuras de datos, buscando incrementar su eficiencia en espacio y en tiempo: que se adapten mejor al nivel de la jerarquía de memorias donde se almacenarán y a las características de los datos a ser indexados. Se espera brindar nuevas herramientas de administración eficiente para las bases de datos métricas, que permitan que su desarrollo se acerque al que tienen los modelos tradicionales de base de datos.

Se continuará analizando la expresividad de las distintas lógicas, para lograr caracterizar la clase de las consultas computables sobre bases de datos no convencionales. Además, se espera mejorar el desempeño de las operaciones de bajo nivel que realizan los DBMS, mediante la propuesta de una nueva arquitectura del procesador.

Actividades de Formación

Formación de investigadores dentro de la línea:

Doctorado en Cs. de la Computación: una tesis sobre expresividad de la lógica como lenguaje de consulta.

Maestría en Cs. de la Computación: una tesis sobre búsqueda por similitud aproximada y otra sobre un índice dinámico eficiente.

Maestría en Informática: una tesis, de la Universidad Nacional de San Juan, sobre un índice dinámico para búsquedas por similitud aproximadas en memoria secundaria.

Referencias

- [1] J. Arroyuelo and J. M. Turull Torres. The existential fragment of third order logic and third order relational machines. In *Proc. del XX CACIC*, 324–333, 2014.
- [2] G. Blaauw and F. Brooks, Jr. *Computer Architecture: Concepts and Evolution*. Addison-Wesley Longman Pub. Co., 1st edition, 1997.
- [3] E. Chávez, K. Figueroa, and G. Navarro. Effective proximity retrieval by ordering permutations. *Pattern Analysis and Machine Intel., IEEE Trans. on*, 30(9):1647–1658, 2008.
- [4] E. Chávez, M. Di Genaro, N. Reyes, and P. Roggero. Distal dynamic spatial approximation forest. In *XXII CACIC*, 804–813, 2016.
- [5] E. Chávez, V. Ludueña, N. Reyes, and F. Kasián. Approximate nearest neighbor graph via index construction. In *Proc. del CACIC*, 824–833, 2016.
- [6] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal SAT. *Information Systems*, 59:15 – 47, 2016.
- [7] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM Comp. Surveys*, 33(3):273–321, 2001.
- [8] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.
- [9] A. Dawar. *Feasible Computation Through Model Theory*. PhD thesis, University of Pennsylvania, 1993. UMI Order No. GAX93-21378.
- [10] A. Dawar. A restricted second order logic for finite structures. *Information and Computation*, 143(2):154 – 174, 1998.
- [11] K. Figueroa, C. Martínez, R. Paredes, N. Reyes, and P. Roggero. Dynamic list of clustered permutations on disk. *Computer Science and Technology*, 201–211, 2016.
- [12] G. Navarro and N. Reyes. Dynamic spatial approximation trees. *Journal of Exp. Algorithms*, 12:1–68, 2008.
- [13] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.
- [14] R. Paredes, E. Chávez, K. Figueroa, and G. Navarro. Practical construction of k -nearest neighbor graphs in metric spaces. In *Proc. 5th WEA*, LNCS 4007, 85–97, 2006.
- [15] H. Samet. *Foundations of Multidimensional and Metric Data Structures*. Morgan Kaufmann Publishers Inc., 2006.
- [16] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach*. Springer-Verlag New York, Inc., 2005.

Data Viz en Bibliotecas Universitarias

Prog. Luis Olguin, Mag. Alejandra Malberti, Mag. Raúl Klenzi
 Instituto de Informática – Departamento de Informática
 Facultad de Ciencias Exactas, Físicas y Naturales
 Universidad Nacional de San Juan

lolguin@iinfo.unsj.edu.ar; amalberti@gmail.com; rauloscarklenzi@gmail.com

Resumen

Las bibliotecas universitarias generan gran cantidad de datos a partir de sus sistemas de gestión, redes sociales, etc. Sin embargo, la denominada *superabundancia de información* hace que gran parte pase desapercibida y al no tratarla con las herramientas apropiadas, perdamos la oportunidad de encontrar valiosa información oculta en ellos.

En el presente trabajo abordamos la temática de *Data Viz* (Data Visualization) para datos masivos extraídos de una biblioteca universitaria. Entendiendo la visualización como la representación gráfica de información con el doble objetivo: de dar sentido a las decisiones de la biblioteca y ser una herramienta de comunicación con los usuarios finales.

Para ello proponemos el uso de algunas herramientas informáticas para trabajar los datos generados por una biblioteca universitaria y obtener *nuevas vistas gráficas* de la información en ellos contenida

Palabras clave: Data Science, Data Viz, Visualization, Data Analysis, Biblioteca Universitaria.

Contexto

Este trabajo se enmarca en las actividades establecidas en el proyecto “**La Ciencia de Datos en grandes colecciones de datos**” (CICITCA E/1014) **ejecutado** por el Instituto de Informática y el Departamento de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan, sujeto a evaluación externa y contenido en el marco del Laboratorio de sistemas inteligentes para la búsqueda de conocimiento en datos masivos.

Introducción

“Existe una cantidad inimaginable de información y datos que constantemente se recogen y se procesan con mayor velocidad. Estamos en el tiempo de los datos masivos pero cuando éstos son demasiados o muy variados hasta el propio analista experto puede perderse en esa lectura, y es en ese momento cuando se hace imprescindible que la visualización de datos sea capaz de mostrar un relato que permita comprender y/o interactuar con los usuarios, clientes o público en general” (Analyze, 2015).

En el trabajo de Olguin, Malberti, Klenzi (2016) se aborda un nuevo enfoque al tratamiento de los datos obtenidos del área de circulación para *modelarlos* con herramientas del área de las *redes sociales* y obtener novedosas *nuevas vistas de los datos transaccionales de una biblioteca universitaria*.

Finch, Jannette (2016) expresa que “la recopilación y visualización de datos en representaciones visuales ayuda a informar al cerebro de forma más rápida y efectiva que la lectura de líneas textuales de información”. Tufte, Edward (2001) enuncia conceptos similares en referencia a la importancia de la visualización, asegura que *los gráficos revelan datos* indicando que *la manera más eficaz de describir, explorar y resumir un conjunto de números es mirar imágenes de esos números*.

En el ambiente de bibliotecas universitarias, donde se centra nuestro trabajo, es posible obtener *miradas más reveladoras* que las simples gráficas generadas por medio de planillas de cálculo que acostumbran nuestros bibliotecarios a explotar.

“La **visualización de datos** es el proceso de representación de datos, en formato gráfico, de una manera clara y eficaz. Ha surgido como una herramienta poderosa y ampliamente aplicable para analizar e interpretar datos grandes y complejos. Se ha convertido en un medio rápido y fácil de transmitir conceptos en un formato universal” (Sadiku, 2016).

Fry, Ben (2007) propone como proceso necesario para la correcta visualización de datos los siguientes pasos:

- **Adquirir:** Recuperar los datos de algún medio magnético.
- **Analizar:** La cantidad de datos que se pueden recopilar y analizar es inmensa. Es necesario poner los datos que se adquieren en una *estructura*. Esta estructura hará que sea más fácil saber transmitir a otros qué datos se tienen (etiquetas).
- **Filtrar:** Eliminar todos los datos que *no son de interés* para la visualización en curso.
- **Minar:** Colocar los datos en un contexto matemático, aplicando técnicas estadísticas o de minería de datos.

- **Representar:** Elegir el mejor modelo visual gráfico para representar los datos.
- **Refinar:** Mejorar la representación visual aplicando algunas herramientas como hojas de estilo (CSS3), gráficos basados en vectores escalables (SVG), etc.
- **Interactuar:** Agregar métodos para controlar las características a ser visibles en la representación visual.

En Finch, Jannette (2016) y Cox, Brian (2012) se generaron visualizaciones de datos para comparar secciones de libros en la biblioteca vs. los gastos en las respectivas áreas, lo que sirvió de punto de partida para discusiones de presupuestos a aplicar. Con los mismos datos se generaron en este trabajo visualizaciones que permitieron detectar obras para expurgo, relaciones entre obras y usuarios, etc.

El proyecto Seattle Central Library (2014) es una interesante muestra de que la exposición de datos en formatos gráficos no tradicionales y dinámicos le brinda una *nueva vista* a los usuarios de la biblioteca. Mediante el uso de pantallas LED ubicadas en el área de circulación de la biblioteca, se ofrecen vistas gráficas de la evolución de los préstamos, áreas de más interés (Dewey), listas de palabras claves de las obras solicitadas, etc. Ésta abundante información, que se procesa cada 60 minutos, es un atractivo entre los servicios que ofrece esta biblioteca.

Mediante la aplicación de estas *vistas de los datos* se busca que la biblioteca las utilice para la toma de decisiones y como elemento orientativo para los usuarios.

El desafío es determinar cómo se aplicará la visualización, pasando del solo hecho de representar datos de una *manera distinta*, a guiarnos en cómo seleccionar los datos a desplegar para que representen información.

“Si la *ciencia de los datos* es la capacidad de extraer conocimiento e ideas de conjuntos de datos grandes y complejos, entonces la *visualización efectiva de datos* es la capacidad de transmitir los conocimientos y las ideas gráficamente con precisión y significado” (Gutsche, Betha (2016).

La manera más tradicional de utilizar *Data Viz* es sin duda el gráfico denominado básico, que incluye los tradicionales de líneas, barras, torta, a los que se le agregan indicadores numéricos que en conjunto permiten a quien lo examina, entender, obtener conocimiento y tomar decisiones en menos tiempo que el utilizado si se trata de excavar buscando información en listas numéricas.

Un paso más adelante en esta disciplina lo dan los generadores de gráficos de burbujas, líneas de tiempo, tree maps, etc. En estos casos se pasa de la vista general al zoom y posteriormente al filtro, con el objetivo de obtener el detalle de información contenido en alguno(s) de los atributos que componen la gráfica.

En el trabajo aquí expuesto, se implementarán herramientas open source/freeware para el tratamiento de los datos a visualizar:

D3.js, librería de JavaScript que permite, a partir de datos numéricos, presentar gráficos de mediana complejidad los cuales pueden ser mostrados en formatos web y mobile-Gráfico 1.

Tableau Public, herramienta gratuita de visualización de datos que permite combinar de manera natural el diseño y los datos. A partir de simples planillas de Excel, se puede llegar a visualizaciones interactivas de alto impacto gráfico- Gráfico 2.

Google Fusion Tables, un servicio web de Google para la gestión de datos.

Proporciona un medio para visualizar los datos con gráficos circulares, gráficos de barras, diagramas de dispersión y líneas de tiempo; así como mapas geográficos basados en Google Maps- Gráfico 3.

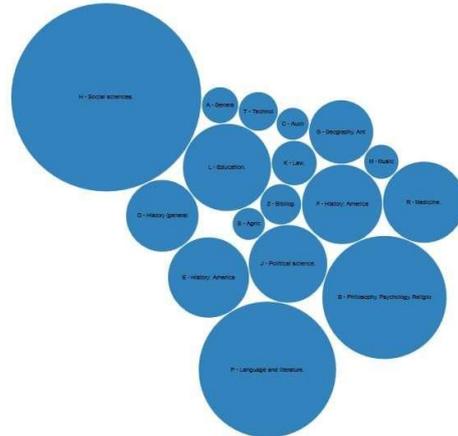


Gráfico 1 -Despliegue en gráfico de burbujas (D3.js) de una búsqueda agrupada por Clasificación

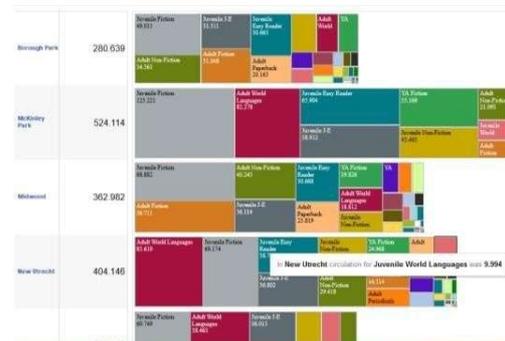


Gráfico 2 - Total de préstamos por área, usando Tableau Public

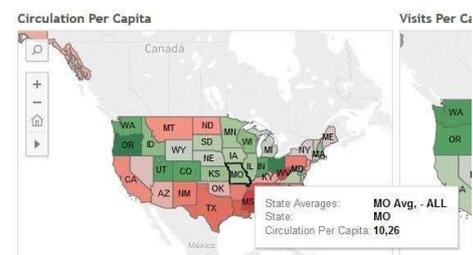


Gráfico 3 - Total de préstamos por región usando Fusion Tables y Tableau Public

Líneas de Investigación, Desarrollo e Innovación

La posibilidad de almacenar enormes volúmenes de datos exige el estudio de nuevas formas de representarlos para *extraer información valiosa*.

En este trabajo se propone el estudio de herramientas y técnicas que permitan analizar grandes volúmenes de datos y generar representaciones visuales que posibiliten la toma de decisiones tanto a los bibliotecarios como a los usuarios finales de una biblioteca universitaria.

La línea principal de investigación, que se ajusta a una investigación en informática aplicada, consiste en el empleo de soluciones, preponderantemente open source o freeware, para la toma de decisiones en el ámbito de una biblioteca universitaria.

Resultados y Objetivos

Se propone la aplicación de herramientas informáticas que permitan el tratamiento de grandes volúmenes de datos y su posterior vista gráfica (*Data Viz*) como otra forma de *descubrir información* oculta en los datos de una biblioteca universitaria.

Se espera que al implementar estas herramientas informáticas en el ámbito de la biblioteca se pueda establecer, mediante el análisis comparativo, cuales son las más apropiadas desde un enfoque de HCI (interacción humano computador), entendiendo que los usuarios finales deben ser la referencia para el diseño de la interfaz de un sistema y este debe adecuarse a la forma de interactuar de ellos.

Formación de Recursos Humanos

En esta instancia se está dirigiendo una Tesis de Maestría en Informática, orientada a la aplicación de herramientas de Ciencia de Datos y Redes Sociales en el contexto de bibliotecas universitarias.

De manera continua se transfieren los conocimientos adquiridos a las cátedras de las carreras del Dpto. de Informática de la FCEF y al equipo de trabajo que forma parte del **Laboratorio de Sistemas Inteligentes para la Búsqueda de Conocimiento en Datos Masivos** (Instituto de Informática, Resol. 02-2015), cuyo objetivo es la extracción de conocimiento en grandes bases de datos mediante la utilización de algoritmos de minería de datos y aprendizaje de máquina soportado por arquitecturas secuenciales y paralelo-distribuidas.

Referencias

Analyze (2015). Data Visualization. *Recuperado de* <http://www.analyze.cl/data-visualization/>

Cox, Brian (2012). *Discovering the impact of library use*. Educause Review Online. *Recuperado de* <http://er.educause.edu/articles/2012/7/discovering-the-impact-of-library-use-and-student-performance>

Finch, Jannette (2016). Using Data Visualization to Examine an Academic Library Collection. *College & Research Libraries*, Noviembre, 165-178.

Fry, Ben (2007). *Visualizing Data: Exploring and Explaining Data with the Processing Environment*. O'Reilly Media, Inc

Gutsche, Betha (2016). *Data Data Everywhere – Finding Focus in a Sea of Visualizations*. *Recuperado de* http://www.wla.org/assets/Alki/Alki_Mar2016_v32-1.pdf

Olguin, Malberti, Klenzi (2016). *Red social de co-prestamo en bibliotecas universitarias*. Workshop de Investigadores en Ciencia de la Computación. Entre Ríos.

Sadiku, Matthew (2016). *Data Visualization brief*. International Journal of Engineering Research And Advanced Technology(IJERAT).Volume. 02 Issue.12. p. 11-16.

Seattle Central Library (2014). *Making visible the invisible*. Recuperado de <http://www.mat.ucsb.edu/~g.legrady/gWeb/Projects/spl/spl.html>

Tufte, Edward (2001). *The Visual Display of Quantitative Information* (2^a. ed.). Cheshire, Conn.:Graphics Press.

Desarrollo de Capacidades Científico-Tecnológicas para la Gestión de Datos Masivos

Fernando Emmanuel Frati¹, Jose Texier¹, Daniel Robins^{1,2}, Fernanda Carmona¹, Alberto Riba¹, Javier Ruitti¹, Jonatan Alvarez^{1,2}, Cristian Rios¹, Lucas Loto²

¹ Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina

² KUNAN, Córdoba, Argentina

{fegrati, jtexier, drobins, fbcarmona, ariba, jruiitti, jalvarez}@undec.edu.ar,
riosbourne555@gmail.com, lucas.loto@kunan.com.ar

Resumen

El incremento en la capacidad de almacenamiento y procesamiento de los equipos de cómputo, sumado a conexiones a Internet cada vez más veloces, permiten diseñar soluciones de datos masivos -antes restringidas a las ciencias tradicionales- para problemas de diferentes áreas de la sociedad. Actualmente, es posible poner a la par ciencia y sociedad como grandes generadores de datos masivos y, en consecuencia, es necesario reconocer la oportunidad estratégica de formar recursos humanos en ésta área de conocimiento. Sin embargo, al igual que toda tecnología emergente, el tema de gestión de grandes datos demora en llegar a la currícula de las carreras de grado. Mientras tanto, se genera una brecha entre lo que la ciencia/sociedad/industria requiere y lo que la instituciones educativas están en condiciones de ofrecer. La forma en que la comunidad académica lidia con esta brecha es fomentando la investigación y desarrollo en esos temas. Esta línea de I/D/I corresponde al diseño, desarrollo e implementación de proyectos que fortalecerán la capacidad científico-tecnológica necesaria para abor-

dar problemas en el campo de la gestión y aprovechamiento de datos masivos (Big Data) que sean de interés para el desarrollo regional y nacional, sobre la base del recurso humano presente en la Universidad Nacional de Chilecito (UNDeC).

Palabras clave: *Big Data, Bases de datos analíticas-columnares, información no estructurada (NoSQL), visualización de grandes volúmenes de datos*

Contexto

El equipo de trabajo ha presentado un proyecto en esta línea en la convocatoria "Proyectos de Investigación Científica y Tecnológica 2016, Plan Argentina Innovadora 2020" en la categoría Equipo de Reciente Formación (PICT-2016-4293). Además, durante el año en curso se adquirirá equipamiento tecnológico solicitado para dar soporte a esta línea con fondos del PROMINF.

Dos de los miembros dirigen proyectos vinculados a esta línea de trabajo, aprobados por la UNDeC en la convocatoria 2013-2014 del programa "Financiamiento para el Estímulo y Desarrollo de la Investiga-

ción Científica y Tecnológica”. Es importante destacar que el estudiante involucrado en el proyecto ha sido beneficiado con una beca de Estímulo a las Vocaciones Científicas, convocatoria EVC 2016(CIN).

Introducción

La caracterización típica de un problema de Big Data es si cumple con un gran volumen, con variedad de fuentes (tanto estructuradas o no), requiere velocidad y frecuencia de las actualizaciones y con veracidad de la información [1]. A continuación, se presentan posibles orígenes de problemas de datos masivos presentes en la comunidad regional pero también a nivel nacional.

Actividades agroindustriales. El clima de la región se caracteriza por la extrema aridez, con grandes amplitudes térmicas, escasas lluvias anuales concentradas en época estival; fuerte insolación anual, frecuentes vientos desecantes y baja humedad atmosférica. Pese al marcado déficit hídrico típico de la región de los valles áridos, lleva adelante una intensa actividad agrícola industrial. La mayor concentración de cultivos en el subsector de fruticultura superando las 20000 hectáreas lo tienen el olivo, la vid y el nogal, los cuales se comercializan a nivel local, regional, nacional e internacional. El proceso de industrialización de algunos cultivos como el de la vid se lleva adelante en 15 bodegas que se distribuyen en la ciudad y distritos del Departamento Chilecito, mientras que la fabricación de aceite de oliva se encuentra en pleno crecimiento. El sector mantiene un estrecho vínculo con la universidad, dispuesto a colaborar en trabajos de investigación y transferencia tecnológica. En este sentido, existe un gran potencial de análisis de grandes volúmenes de datos (control de plagas, monitorización del crecimiento de cultivos, gestión de datasets agrometeorológicos, etc.).

Laboratorio de Altura y Laboratorio de Alta Complejidad. La UNDeC cuenta con dos importantes laboratorios que ofrecen servicios a la comunidad y que representan un enorme potencial de trabajo con grandes volúmenes de datos y procesos complejos con requerimientos de tiempo real. El *Laboratorio de Altura* es el primero de América de esta clase. Se encuentra a 5200 metros sobre el nivel del mar, es de fácil acceso y posee excelentes condiciones atmosféricas. Este laboratorio permite realizar mediciones imposibles de hacer a nivel del mar en campos como medicina, biología, astronomía, física, etc. Por otro lado el *Laboratorio de Alta Complejidad* presta los siguientes servicios a la comunidad: análisis de suelos, análisis de aguas para riego, análisis de aguas para consumo, análisis microbiológico de agua, análisis de efluentes, entre otros. Cuenta con una gran cantidad de instrumental de laboratorio y de campo, y actualmente está en proceso la adquisición de un secuenciador de ADN, un secuenciador genómico y un microscopio electrónico de barrido de alta resolución, lo que permitirá ampliar los servicios ofrecidos.

Servicios de información con valor agregado. Es posible generar información con valor agregado para la comunidad a partir de un gran volumen de datos que está disponible públicamente en Internet. Una expresión importante del fenómeno conocido como sociedades de la información y el conocimiento [2, 3] son las redes sociales Twitter, Facebook, Instagram, Snapchat, Whatsapp, entre otras [4, 5]. Los datos que circulan a través de ellas pueden ser transformados en información de relevancia y utilidad sobre tendencias de consumo, pensamientos políticos, ideologías, preferencias y costumbres [6, 7, 8]. Por ejemplo, es posible predecir tendencias en elecciones a partir de un análisis de opinión de los

datos disponibles en estas redes [9, 10, 11]. Otro enfoque en pleno crecimiento consiste en relacionar datos abiertos [12] difundidos por organizaciones públicas o privadas en búsqueda de ofrecer nuevos servicios a la comunidad. En Argentina varios organismos gubernamentales comenzaron a promover el uso de Datos Abiertos, con el objetivo de facilitar información a los ciudadanos para su consulta y libre uso [13, 14]. Sin embargo, también aparecen expresiones en el sector privado que utilizan estos recursos para ofrecer mejores servicios. Como ejemplo se puede mencionar el proyecto del diario La Nación Data [15] que utiliza datos abiertos para generar noticias que han tenido gran impacto en la comunidad, o emprendimientos como Properati Data [16], una inmobiliaria con un modelo de negocios basado en el cruce de información con datasets públicos en búsqueda de atraer la atención de clientes.

Líneas de Investigación, Desarrollo e Innovación

Bases de datos, Minería de datos, bases de datos columnares; visualización de grandes volúmenes de datos; sistemas de procesamiento estadístico; Cloud Computing, Arquitecturas paralelas; Análisis social web; Simulación; Internet de las cosas; Agromática, Vitivinicultura, Genética; Repositorios institucionales y bibliotecas digitales, Análisis semántico de la información

Resultados y Objetivos

Desde el año 2008 se coopera con empresas del sector agrícola dedicadas especialmente al cultivo de olivo, lo que dio lugar a los primeros proyectos de investigación con financiamiento interno de la UNdeC para su desarrollo. Estos proyectos trabajaron sobre monitorización, planificación y automatización del riego de los cultivos, y su ejecución derivó en desarrollos de software a medida que continúan siendo utili-

zados. Las relaciones de cooperación entre las empresas del sector y los miembros del grupo representan una potencial fuente de problemas de datos masivos. Es de esperar que como parte de la ejecución de esta línea sea posible llevar esos desarrollos a gran escala.

Uno de los miembros está desarrollando su tesis de maestría en el tema “Mejora de la precisión posicional utilizando receptores GPS de bajo costo”. Es un tema de gran impacto para la región debido a que el único tipo de agricultura sustentable es la de precisión, y esta sólo es posible con la tecnología adecuada. Posee financiamiento propio y está indirectamente relacionado con esta línea a través de los servicios que pueden ser ofrecidos al sector agrícola.

En el contexto de esta línea, se está desarrollando un trabajo final de grado en el tema “Ambiente colaborativo para mejorar las prácticas de viticultores independientes en la zona de Pituil”. Se espera organizar y estructurar la captura y almacenamiento de la información que circula por la red para ser puesta al servicio de la comunidad científica y académica a través de datasets públicos. Se está trabajando con investigadores, profesionales y productores del sector, lo que sienta las bases para futuros trabajos en el área. Se prevé su finalización para julio de 2017.

Otro de los miembros está desarrollando su tesis de maestría con el tema “Métricas de calidad de los datos obtenidos en un Sistema de Integración de Datos”. El trabajo destaca el efecto negativo que pueden tener los distintos tipos de inconsistencias para la integración de datos y propone definir un conjunto de métricas basadas en distintos algoritmos de detección de inconsistencias para predecir la calidad de la información integrada. Estas métricas serán luego utilizadas en un framework para la publicación de datasets públicos. Se espera con es-

te trabajo avanzar en una infraestructura para dar soporte a otros proyectos vinculados a la administración pública para la publicación de datos abiertos (universidad, municipalidad, poder judicial, consejo deliberante, etc.). Está planificada la finalización de la tesis para marzo de 2018.

Se está trabajando con la empresa cordobesa Kunan S.A., sobre la idea de predecir comportamientos sociales a partir de análisis de sentimiento en redes sociales. Como resultado de esta colaboración se presentó un artículo sobre el Balotaje Argentina 2015 en las IV Jornadas de Cloud Computing & Big Data 2016. Para el trabajo se adquirieron los comentarios vertidos voluntariamente en la red social twitter por los usuarios referidos al balotaje presidencial con el agente Apache Flume de Hadoop, y se utilizó el motor de base de datos Vertica con su componente Pulse, y el software de visualización Tableau. Para el análisis de correlación de los datos se utilizó Stata.

En esta misma línea a finales de 2016 dos docentes en colaboración con Kunan S.A. utilizaron las técnicas de análisis de sentimiento en tweets para determinar el interés energético de sus usuarios a nivel mundial, catalogadas por tipo de energías. Este trabajo fue realizado para plantas generadoras de energía, y permitió conocer rápidamente la opinión de la población sobre el tema, determinar competidores y en función de ello proyectar inversiones en zonas con mayor oportunidad de aceptación. Igual que en el caso anterior, se emplearon Vertica, Tableau y Stata. Los resultados finales aún están pendientes de publicación.

Dos miembros del equipo ofrecieron un curso de Vertica durante las “VIII Jornadas de Informática y Comunicaciones 2016”. A partir de este año, estos miembros comenzarán a dar una asignatura sobre Big Data en las carreras ofrecidas por la UNdeC.

Objetivos

Crear dentro del marco de la UNdeC la capacidad científico-tecnológica necesaria para abordar problemas en el campo de la gestión y aprovechamiento de datos masivos (Big Data) que sean de interés para el desarrollo regional y nacional.

- Definir los requerimientos de una plataforma de experimentación, desarrollo y producción de soluciones a problemas de datos masivos.
- Desarrollar una infraestructura acorde a los requerimientos anteriores.
- Estudiar técnicas y herramientas para la gestión y aprovechamiento de datos masivos.
- Difundir a nivel regional y nacional la potencialidad de trabajo del equipo.
- Explorar oportunidades de colaboración con la comunidad académica, productiva y en general que deriven en problemas de datos masivos.
- Canalizar esas oportunidades a través de trabajos de finalización de grado y tesis de postgrado.
- Promover el diseño y desarrollo de algoritmos paralelizados orientados a la optimización de problemas de cómputo con grandes volúmenes de datos.
- Consolidar un grupo de investigación multidisciplinario en la UNdeC.

Formación de Recursos Humanos

El equipo de trabajo está formado por: Dos doctores especializados en repositorios institucionales, bibliotecas digitales, desarrollo de software, cómputo paralelo y tecnología grid. Tres estudiantes de maestría en informática en su etapa final, dos de los cuales trabajan en temas relacionados con esta línea. Un estudiante de grado, el cual

presentará su trabajo final en julio en un tema directamente relacionado a esta línea. Tres de los miembros están categorizados en el programa de incentivos. Además, el grupo mantiene vínculos de colaboración con los miembros del III-LIDI (UNLP) y con los miembros del grupo ARTECS de la Universidad Complutense de Madrid.

Referencias

- [1] M. Tascón, “Introducción: Big Data. Pasado, presente y futuro,” *Telos: Cuadernos de comunicación e innovación*, no. 95, pp. 47–50, 2013.
- [2] I. Murua Anzola, M. L. Cacheiro González, and D. J. Gallego Gil, “Las cibercomunidades de aprendizaje (cca) en la formación del profesorado,” *RED.*, vol. XIII, no. 43, p. 29, 2014.
- [3] M. Meirinhos and A. Osório, “Las comunidades virtuales de aprendizaje: el papel central de la colaboración,” *Pixel-Bit. Revista de Medios y Educación*, no. 35, pp. 45–60, 2009.
- [4] A. G. Sans, “Las Redes Sociales como Herramientas para el Aprendizaje Colaborativo: Una Experiencia con Facebook,” *Re-Presentaciones: Periodismo, Comunicación y Sociedad*, no. 5, pp. 48–63, 2009.
- [5] R. V. Argüelles, “Las redes sociales y su aplicación en la educación,” *Revista Digital Universitaria*, vol. 14, no. 4, pp. 1–14, 2013.
- [6] M. Gil Mediavilla, V. Ausín Villaverde, and F. Lezcano Barbero, “Redes sociales educativas como introducción a los entornos personales de aprendizaje (PLE’s),” *EduSer-Revista de educação*, vol. 4, no. 1, pp. 17–29, 2012.
- [7] J. Lorca and L. Pujol, “Redes sociales: descripción del fenómeno, situación actual y perspectivas,” *Revista eSalud. com-Fesalud. Fundación para la eSalud*, vol. 4, no. 15, pp. 1–15, 2008.
- [8] V. Miguel, M. Fernández, E. V. de Enseñanza Aprendizaje, and P. de Gestión, “Redes Sociales y Construcción del Conocimiento,” *AB Martínez y N. Hernández “Comunidades Virtuales de Aprendizaje”*. Caracas, Venezuela. Consejo de Desarrollo Científico y Humanístico, Universidad Central de Venezuela, 2013.
- [9] L. Deltell Escolar, “Predicción de tendencia política por Twitter: Elecciones Andaluzas 2012,” *Ambitos: Revista internacional de comunicación*, no. 22, pp. 91–100, 2013.
- [10] A. Ceron, L. Curini, and S. M. Iacus, “Using Sentiment Analysis to Monitor Electoral Campaigns: Method Matters— Evidence From the United States and Italy,” *Social Science Computer Review*, vol. 33, no. 1, pp. 3–20, Feb. 2015.
- [11] S. Unankard, X. Li, M. Sharaf, J. Zhong, and X. Li, *Predicting Elections from Social Networks Based on Sub-event Detection and Sentiment Analysis*, ser. Lecture Notes in Computer Science, B. Benatallah, A. Bestavros, Y. Manolopoulos, A. Vakali, and Y. Zhang, Eds. Springer International Publishing, Oct. 2014, no. 8787.
- [12] “El manual de Open Data.” [Online]. Available: <http://opendatahandbook.org/guide/es/>
- [13] S. Fumega, “Opening Cities: Open Data in Buenos Aires, Montevideo and Sao Paulo,” Exploring the Emerging Impacts of Open Data in Developing Countries (ODDC), Tech. Rep., Apr. 2014.
- [14] “Datos Argentina.” [Online]. Available: <http://datos.gob.ar>
- [15] “LA NACION Data - LA NACION.” [Online]. Available: <http://www.lanacion.com.ar/data>
- [16] “Properati Data.” [Online]. Available: <http://www.properati.com.ar/data>

Detección de Ataques DoS con Herramientas de Minería de Datos

Klenzi, Raúl; López, Marcelo

Instituto de Informática / Departamento Informática / Facultad de Ciencias Exactas Físicas y Naturales / Universidad Nacional de San Juan

Domicilio: Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas",
Rivadavia, San Juan, CPA: J5402DCS, 0264-260353 0264-4260355
{rauloscarklenzi; marcelo.sanjuan.ar;}@gmail.com

Resumen

En el marco de proyectos “La ciencia de los datos en grandes colecciones de datos” y “Evaluación de arquitecturas distribuidas de commodity basadas en software libre” contenidos en el “Laboratorio de Sistemas Inteligentes para la búsqueda de Conocimiento en Datos Masivos”, integrado por docentes investigadores del Departamento e Instituto de Informática (DI-Idel) de la Facultad de Ciencias Exactas Físicas y Naturales FCEF, se trabaja en modelar y desde allí mitigar, ataques a un servidor de red por denegación de Servicios (Denial of Services)-DoS- mediante el análisis offline de un flujo de datos simulados y la utilización de algoritmos y herramientas correspondientes a Data Stream Mining (Minería de datos -MD- en flujos de datos continuos). La aplicación utiliza módulos y algoritmos específicos de las herramientas de software libre RapidMiner (RM) 5.3.015 y KNIME 3.3

Palabras clave: Extracción de Conocimiento, Análisis off line, DNS, software libre

Contexto

En el ámbito del Idel, por ordenanza 02/2015-CD-FCEF, se conformó el “Laboratorio de Sistemas Inteligentes para la extracción de Conocimiento en Datos Masivos”, aquí se cubren diferentes áreas que abarcan conocimiento del hardware, software y

fundamentalmente datos que por su magnitud, necesitan de herramientas específicas.

En los proyectos insertos en el laboratorio, se desarrollan aplicaciones del área temática: Descubrimiento de Conocimiento en Bases de Datos (Knowledge Discovery Data -KDD-) que es un análisis automático exploratorio y modelado de grandes depósitos de datos e involucra inteligencia artificial, aprendizaje automático (Machine Learning -ML-), estadística, sistemas de gestión de base de datos, técnicas de visualización de datos y medios que apoyan toma de decisiones.

Data Stream Mining es el proceso de extraer conocimiento en estructuras de datos continuas y con rápidas transiciones. Un data stream es una secuencia ordenada de instancias que herramientas de extracción de conocimiento en datos pueden leer utilizando capacidades de cómputo y almacenamiento limitadas. Ejemplos de data streams incluye análisis de tráfico en redes de computadoras, comunicaciones telefónicas, transacciones ATM, búsquedas web y/o datos relevados desde sensores. El Data stream mining se considera como un subcampo del ML, KDD y MD.

Desde la medición del flujo de bits que ingresan al servidor se propone una primera instancia de modelación de trazas de red que caracterizan a un ataque DoS o DDoS (Distributed DoS), el cual es un ataque a un sistema de computadoras que causa que un servicio o recurso sea inaccesible a usuarios legítimos. Aquí se analiza un flujo de datos simulado, utilizando herramientas de software libre de ML Rapidminer 5.3.15 y KNIME 3.3.

Introducción

En el ámbito de redes de comunicaciones, un ataque o intrusión se define como un evento en la red que aprovecha cualquier tipo de vulnerabilidad de un sistema informático para causar daño sin consentimiento del usuario de dicha red, afectando la confidencialidad, integridad, disponibilidad o no repudio y pueden presentar los siguientes signos verificables: *interrupción* (el recurso se vuelve no disponible), *intercepción* (“alguien” no autorizado consigue acceso a un recurso) y *modificación* (además de la intercepción es capaz de manipular los datos) [1].

Según cómo afecte el tipo de ataque a la red, se clasifica en ataque pasivo o activo. En un *ataque pasivo*, el intruso monitoriza el tráfico en la red y hace uso de la información capturada; se centra en la intercepción de datos y análisis de tráfico, para obtener información de la propia red, siendo muy difícil de detectar por no alterar los datos interceptados [2].

En un *ataque activo* el intruso interfiere con el tráfico que fluye por la red, explotando sus vulnerabilidades o de una víctima particular. Según su objetivo, estos ataques se clasifican en cuatro categorías: *Suplantación de identidad*, *Reactuación*, *Modificación de mensajes* y *Degradación fraudulenta del servicio*. Este último, conocido como DoS provoca una denegación de los recursos informáticos y de comunicación de un elemento de la red [3].

Un ataque de DoS tiene como objetivo atacar una infraestructura de red, causando que sus servicios sean inaccesibles a usuarios que acceden de una forma legítima, normalmente ocasiona pérdida total de conectividad a la red debido al excesivo consumo del ancho de banda de la víctima y se implementa a través de la saturación intencional de los puertos del host atacado con un flujo constante de información, sobrecargando los recursos de los servidores y la capacidad de responder a las peticiones realizadas por los usuarios originales.

Existen varios métodos mediante los cuales se pueden realizar DoS a los servicios [4]: *Spoofed* (paquetes con una dirección de origen falsificada), *Malformed* (paquetes con bits o flags encendidos en forma anormal), *Floods* (paquetes conformados de manera legítima en gran cantidad), *Null* (paquetes sin contenido), *Protocol* (paquetes con protocolos ilegítimos), *Fragmented* (paquetes fragmentados los cuales nunca se completarán) y *BruteForce* (paquetes que exceden el umbral definido de „flowrates“) [5].

En ataques de tipo (flooding) se centra en consumir recursos disponibles para el servicio o recursos de todo tipo que existan en el camino como routers, firewalls, etc., mediante la inyección de grandes volúmenes de tráfico.

Las estrategias de inundación, se clasifican como de baja o alta tasa. El flooding de tasa baja explota vulnerabilidades de los protocolos de red y permite que el tráfico inyectado adopte patrones periódicos de volumen fluctuante en el tiempo, en tanto el flooding de tasa alta consiste en la emisión de grandes cantidades de tráfico de manera constante y uniforme [6].

Los métodos más populares de ataque siguen siendo SYN-DDoS, TCP-DDoS y HTTP-DDoS y UDP-DDoS, y pueden detectarse analizando la variación temporal de señales del protocolo de comunicación. En el SYN Flood, un usuario realiza un número especialmente alto de inicios de conexión que nunca son finalizados evidenciándose por permanecer activa la señal syn del protocolo un tiempo prolongado.

Sin herramientas que permitan un adecuado monitoreo del tráfico de red en busca de anomalías, puede presentarse una paulatina degradación de los servicios ofrecidos.

El éxito de un método para la detección y mitigación de ataques DoS, depende en parte de factores como el consumo de recursos, tiempo de respuesta, complejidad, además de garantizar que su implementación no ocasionará interrupciones en los servicios provistos.

Las actuales herramientas para combatir este problema supone contar con grandes recursos de hardware y software, ya que la gran mayoría de estos sistemas afectan recursos valiosos como procesador, memoria física y ancho de banda, requiriéndose en casi todos ellos, horas, semanas o incluso meses de análisis previos antes de mitigar realmente un ataque. Algunas herramientas para combatir ataques DoS, sólo realizan investigación forense que generan estrategias de prevención en el futuro, es decir, no son reactivas [8].

El aporte del trabajo a la mitigación de ataques DoS, se orienta al estudio de la efectividad de su detección mediante el Análisis de Datos (AD) utilizando técnicas de MD.

Las plataformas de AD para Ciberseguridad se agrupan en tres categorías (Figura 1): (1) de propósito experimental, (2) de propósito específico y (3) de propósito general. La primera agrupa aquellas plataformas dirigidas al desarrollo y prueba de algoritmos que posteriormente se aplican al escenario real de la Ciberseguridad. El segundo grupo contiene plataformas implementadas en escenarios reales de Ciberseguridad aunque realizan un número específico y limitado de tareas. Por último, las plataformas de propósito general tienen en común con las de propósito específico que pueden insertarse en escenarios reales, sólo que estas últimas, se pueden adecuar a cualquier tarea de Ciberseguridad.



Fig. 1. Taxonomía de las plataformas de AD orientadas a la Ciberseguridad.

Ejemplos de plataformas experimentales de AD son Clementine, Weka, RapidMiner, KNIME y Orange. Existe una gran cantidad de trabajos vinculados con la aplicación del AD en

Ciberseguridad que solo se centran en el mejoramiento de los algoritmos de MD[9].

Aquí se propone un análisis de tráfico de red off_line basado en estrategias derivadas del ML, de MD y particularizando en el uso de herramientas de software libre que permitendese desde la generación de WF con módulos específicos, la aplicación comparativa de diferentes algoritmos y formas de visualización de extracción de conocimiento en datos.

La Figura 2 presenta la encuesta del sitio KDnuggets.com respecto a las herramientas más utilizadas en ML, KDD y MD.

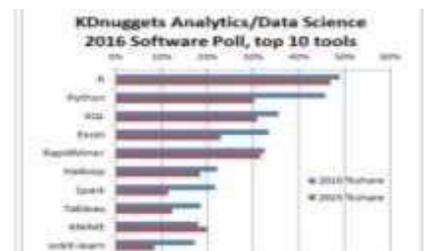


Fig. 2: Herramientas de software de ML más utilizadas en 2016 según kdnuggets.com

Desde la Fig. 2 y por ser herramientas de software libre, se seleccionan RapidMiner, y KNIME. Estas herramientas disponen módulos específicos que procesan flujos de datos y gran riqueza de formas de visualización.

- **RapidMiner:** Herramienta de software libre licencia AGPL RapidMiner 5.3.015, desarrollado en java en Dortmund Alemania.
- **KNIME** - KonstanzInformationMiner- en versión KNIME ANALYTICS 3.3 bajo licencia GNU 3, es una plataforma de ML construida en Eclipse y programada en Java, desarrollada en el departamento de MD y bioinformática de la Universidad de Constanza, Alemania.

Por la disponibilidad de grandes cantidades de datos surge el área de estudio del KDD, definida como “un proceso no trivial de identificación de patrones válidos, novedosos, potencialmente útiles y comprensibles a partir de los datos”[10].

Figura 3: Tabla con fracción de datos a analizar

Desde el análisis del data streamming que ingresa al servidor, simulado en el laboratorio y presentado en la Fig. 3, se propone una primera instancia de modelación de tramas que caracterizan a un ataque DoS reconociendo en ciertos lapsos de tiempo el estado de un conjunto de señales como SYN, LEN, etc, que forman parte del atributo Info y permiten reconocer la existencia de un posible DoS.

Dado que el análisis de datos en toda estructura de tablas se realiza por filas, es necesario a los efectos de procesar información temporal (cada registro asociado a un evento temporal expreso), realizar una transformación mediante el uso del operador windowing (sliding window technique) [12] que permite establecer una ventana de análisis temporal ajustada por el usuario y así tratar de detectar patrones correspondientes a un ataque DoS, como se aprecia en la Figura 4.

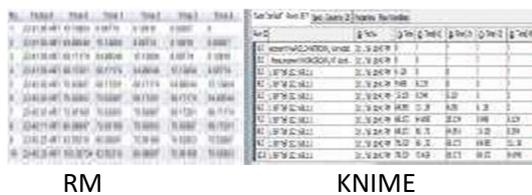


Figura 4: Sliding Windows, para el análisis de flujo de datos en las herramientas utilizadas

Para la tarea nos valemos de las potencialidades visuales de cada herramienta que permitirá de manera rápida y amigable reconocer, en ese análisis fuera de línea, las características asociadas a un ataque DoS y así tratar de mitigar posteriores intromisiones.

Líneas de Investigación, Desarrollo e Innovación

En el marco del laboratorio se llevan adelante diferentes trabajos de investigación aplicada, caracterizada por el tipo de datos observados y analizados. En este trabajo se ha realizado una aplicación de Data streamming siendo el objetivo extender la aplicación a otros tipos de datos anómalos que puedan reconocerse en diferentes trazas de red y generar la mejor forma de presentar conclusiones desde las potencialidades de visualización que poseen las herramientas.

Así mismo según la variedad y tipología de datos se está trabajando en análisis de series temporales, aplicaciones georeferenciadas, y reconocimiento de perfiles de usuarios, intentando llevarlas a plataformas paralelas en cluster de computadoras.

Resultados y Objetivos

Desde al análisis en diferentes ventanas temporales de las señales que conforman los protocolos de comunicación se han logrado detectar las diferentes variantes de ataques DoS. La Figura 5 muestra el reconocimiento de dos tipos de ataques DoS. El objetivo próximo es extender la aplicación al reconocimiento de otros Data Stream que caracterizan datos anómalos existentes en bases de datos del sitio www.Kaggle.com

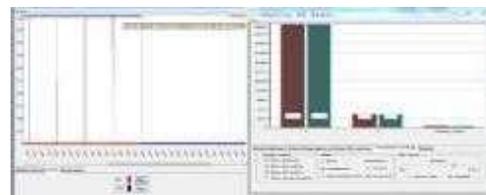


Figura 5: Histograma de Visualización de ataques SYN y UDP Flood

Al utilizar las herramientas de ML se ha observado que KNIME genera archivos de simulación de mayor tamaño que los que genera

RM, guardando las visualizaciones de salida, conjuntamente con la estructura del WF. Así, al momento de cargar un WF, se pueden observar las salidas de los diferentes módulos, en tanto en RM cada vez que se carga un WF se debe ejecutar y esperar las visualizaciones de salida hasta que la ejecución concluya.

Formación de Recursos Humanos

En el último año desde la actividad de ambos proyectos de investigación y del área de contención generada por el laboratorio se ha generado una gran actividad que tuvo su prólogo inicial en el trabajo publicado en el WICC 2016 realizado en Concordia con el trabajo “Confluencia de áreas de trabajo en un laboratorio de sistemas inteligentes”. Allí se planteaba la posibilidad de llevar adelante dos trabajos finales de grado en Licenciatura en Ciencias de la Computación que destacaba la conjunción de dos áreas de trabajo bien diferenciadas y que hoy están en el proceso de redacción de informe final, como así también la posibilidad cierta de incorporar alumnos becarios y adscriptos. Así es como en el marco del proyecto “Ciencia de los Datos en grandes colecciones de datos” se cuenta con dos alumnos de grado adscriptos al proyecto, se están dirigiendo otros tres trabajos finales de grado y dos trabajos de maestría, en tanto integrantes del proyecto son doctorandos del Doctorado en Ciencias de la Informática que se dicta en la FCFN_UNSJ. De manera similar se cumplen estas tareas en el otro proyecto de investigación mencionado en el trabajo.

Referencias

- P. Deepa Lakshmi, J. S. Praveen, V. Venkatraman, and N. Manoharan Director-Research, “A Review On Data Security In Distributed System,” *Int. J. Comput. Eng. Technol.* N. A. Rev. Data Secur. Distrib. Syst. Int. J. Comput. Eng. Technol., vol. 6, no. 610, pp. 13–16, 2015.
- [2] H. Sandberg, S. Amin, and K. H. Johansson, “Cyberphysical Security in Networked Control Systems: An Introduction to the Issue,” *IEEE Control Syst.*, vol. 35, no. 1, pp. 20–23, Feb. 2015.
- [3] N. Paulauskas and E. Garsva, “Computer System Attack Classification,” *Elektron.irElektrotechnika*, vol. 66, no. 2, pp. 84–87, 2015.
- [4] G. Kumar, “Denial of service attacks – an updated perspective,” *Syst. Sci. Control Eng.*, vol. 4, no. 1, pp. 285–294, Jan. 2016.
- [5] P. Gasti, G. Tsudik, E. Uzun, and L. Zhang, “DoS and DDoS in Named Data Networking,” in *2013 22nd International Conference on Computer Communication and Networks (ICCCN)*, 2013, pp. 1–7.
- [6] N. Sharma, M. Alam, and M. Singh, “Denial of Service: Techniques of Attacks and Mitigation,” *J. Comput. Sci. Eng. Softw. Test.*, vol. 1, no. 2, 2015.
- [7] H. Wang, Q. Jia, D. Fleck, W. Powell, F. Li, and A. Stavrou, “A moving target DDoS defense mechanism,” *Comput. Commun.*, vol. 46, pp. 10–21, 2014.
- [8] K. Singh, N. Kaur, and D. Nehra, “A Comparative Analysis of Various Deployment Based DDoS Defense Schemes,” *Springer Berlin Heidelberg*, 2013, pp. 606–616.
- [9] A. F. Rivas and O. A. P. García, “Estado del Arte de las Plataformas de Análisis de Datos en la Ciberseguridad,” 2016.
- [10] D. T. Larose and C. D. Larose, “Data Mining and Predictive Analytics-Wiley,” 2015.
- [11] J.-P. Mens, *Alternative DNS Servers*. 2009.
- [12] H. Ryang and U. Yun, “High utility pattern mining over data streams with sliding window technique.”

Determinación del Rendimiento Académico Universitario

Myriam Herrera¹, María Inés Lund², Susana Beatriz Ruiz¹, Lilian Adriana Mallea³,
María Gema Romagnano², Estela Liliana Torres¹

¹Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales,
Universidad Nacional de San Juan

²Instituto de Informática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad
Nacional de San Juan

³Departamento de Matemática, Facultad de Filosofía, Humanidades y Artes, Universidad
Nacional de San Juan

mherrera, mlund, mromagnano@{iinfo.unsj.edu.ar}

RESUMEN

El estudio del rendimiento académico constituye actualmente uno de los temas “estrella” en la investigación educativa. En una sociedad de la información como la actual, uno de los desafíos de la educación es transformar esa gran cantidad de información disponible en conocimiento personal para desenvolverse con eficacia en la vida. Además, según Escudero Escorza [1] la calidad de la enseñanza, la cual se manifiesta a través del Rendimiento Académico, es responsabilidad de los universitarios y la sociedad en general. Hasta el momento en la UNSJ se estima el rendimiento académico tan sólo por una calificación numérica de los alumnos (notas, cantidad de materias rendidas, aprobadas, boletas, etc). En este proyecto se determinarán indicadores directos de la calidad de la enseñanza que permitirán medir el rendimiento académico de los estudiantes, en base a otra información.

Esto se obtendrá mediante la aplicación de las técnicas de conglomerados y el análisis discriminante, las cuales algunos autores ubican entre las más potentes para aplicar en investigaciones sociales; permiten clasificar sujetos u objetos a partir de características similares.

Palabras clave: Clasificación, Rendimiento, Calidad Universitaria

Este proyecto se encuentra a la mitad de su ejecución es de carácter bi-anual (2016-2017) y financiado por la UNSJ. Se enmarca en las líneas de investigación de los Gabinetes Estadística e Ingeniería de Software del Instituto de Informática de la FCFN de la UNSJ.

Además se vincula a cátedras de las carreras de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, que se dictan en la Institución.

1. INTRODUCCION

El análisis del rendimiento académico, en torno a la problemática de valorar la calidad educativa, se obtendrá a partir de datos que se convierten en un recurso crítico.

El análisis de conglomerados y el análisis discriminante, son técnicas que algunos autores ubican entre las más potentes para aplicar en investigaciones sociales, permiten clasificar sujetos u objetos a partir de características similares [2].

Estas dos técnicas se pueden diferenciar por la manera de extraer conocimiento útil escondido en esos datos. El Análisis Discriminante cuenta con grupos de datos conocidos, con observaciones de unidades de pertenencia desconocida inicialmente y tiene que ser determinada a través del análisis de los datos. Este tipo de problemas de clasificación es referido como reconocimiento de patrones asistido o aprendizaje

CONTEXTO

supervisado; en terminología estadística cae bajo el título de Análisis Discriminante [3], [4], [5].

Por otro lado, hay problemas de clasificación donde los grupos son desconocidos a priori y el principal propósito del análisis es determinar los grupos a partir de los propios datos, de modo que las unidades dentro del mismo grupo sean, en algún sentido, más similares u homogéneas que aquellas que pertenecen a grupos diferentes. Este tipo de problema de clasificación es referido como reconocimiento de patrón no supervisado o conocimiento sin guía, y, en terminología estadística cae bajo el título de Análisis de Conglomerados [6], [7].

El análisis discriminante es la prueba estadística apropiada para seleccionar qué variables independientes o predictivas permiten diferenciar grupos y cuántas de estas variables son necesarias para alcanzar la mejor clasificación posible. Además permite cuantificar su poder de discriminación en la relación de pertenencia de un sujeto u objeto a un grupo u otro. Por ello esta técnica es considerada, además de una prueba de clasificación, una prueba de dependencia. De hecho, su propósito es similar al análisis de regresión logística; la diferencia radica en que solo admite variables cuantitativas. Mediante un análisis discriminante se puede establecer el poder explicativo y discriminatorio de las características que diferencian a los alumnos según su rendimiento. Además del rendimiento se tendrán en cuenta en el estudio una serie de variables independientes como, por ejemplo, variables de carácter socioeconómico, variables académicas referentes a la preparación en el nivel secundario y variables actitudinales en relación con la variable dependiente que clasifica a los sujetos según el rendimiento obtenido [8], [9].

Según las características analizadas a través de la descripción del grado de relación existente entre el conjunto de variables, se puede encontrar la frontera que separa los grupos. Se espera como resultado final obtener una regla de clasificación que permita pronosticar la adscripción al grupo de

rendimiento establecido para nuevos estudiantes.

Actualmente en el proyecto se están aplicando ambas técnicas y también una combinación de ellas. Además se está investigando para encontrar otra técnica para analizar lo que llamamos rendimiento académico universitario o bien generar una nueva. Se puede afirmar que, en general, un indicador directo de la calidad de la enseñanza es el rendimiento académico, medido a través del nivel alcanzado por los estudiantes. Para ello se determinaron las principales variables que influyen en el rendimiento como así también tipologías básicas de grupos, obtenidos de los alumnos de Licenciatura en Sistemas de Información y Ciencias de Computación de la Facultad de Ciencias Exactas como de los alumnos del Dpto. de Matemática de la Facultad de Filosofía de la UNSJ.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En muchas de las investigaciones, independientemente del área de conocimiento, es habitual tener la necesidad de identificar cuáles son las características que diferencian grupos de sujetos u objetos respecto de otros, para así poder realizar predicciones futuras. Tanto el análisis de conglomerados como el análisis discriminante son técnicas que nos permiten clasificar sujetos u objetos a partir de características similares. La diferencia fundamental entre ambas pruebas es el momento del establecimiento de los grupos. En el análisis discriminante (AD) el investigador conoce a priori a qué grupo pertenece cada sujeto u objeto; en cambio, en el análisis de conglomerados los grupos o clúster se determinan y configuran a posteriori, es decir, una vez estudiadas y analizadas las agrupaciones.

En el presente proyecto utilizaremos estas técnicas en el ámbito educativo como es el estudio del rendimiento estudiantil y la identificación de las variables que mejor lo predicen, usando no sólo las calificaciones numéricas sino también características sociales del alumno.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El equipo de investigación ha presentado el proceso de investigación realizado en el marco de las “Charlas de los Viernes”, un evento mensual que se realiza en el Instituto de Informática, donde se convoca a todos los investigadores del área a participar y escuchar los avances de los proyectos de investigación. Se ha elaborado la encuesta para aplicar a los alumnos de la UNSJ, en ambas carreras de la FCEFyN (<https://exactas.unsj.edu.ar/alumnos/>) Los resultados parciales obtenidos de esta encuesta han sido presentados en el XX Encuentro Nacional y XII Internacional sobre Educación Matemática en Carreras de Ingeniería - EMCI 2017, actualmente en proceso de revisión por parte de pares evaluadores.

Al momento el grupo de investigación está realizando las actuaciones correspondientes para que esta encuesta sea respondida para los alumnos de todas las carreras de ambas facultades, ya que hasta ahora se ha trabajado con alumnos de las carreras del Departamento de Informática de la FCEFyN y Departamento de Matemática de la FFHA.

Trabajo realizado:

Actividad N°1: Se estudiaron y analizaron las técnicas de agrupamiento o clasificación de objetos o sujetos teniendo en cuenta el marco teórico considerado, como así también trabajos relacionados.

Actividad N° 2: Se confeccionó una encuesta para relevar las variables influyentes. La misma fue elaborada con la herramienta web de encuestas online **EncuestaFácil.com**, (<https://www.encuestafacil.com>). Estas variables fueron agrupadas en varias secciones (Fig. 1).

Sección A: Localización del encuestado en la UNSJ

Sección B: Datos demográficos y socio-económicos

Sección C: Estudios y actividades en la UNSJ

Sección D: Conocimiento y uso de instalaciones y servicios en la UNSJ

Sección E: Estado de Salud reciente

Sección F: Antecedentes de problemas de salud

Sección G: Hábitos relacionados con la salud

Sección H: Percepciones y estado anímico

Sección I: Hábitos cotidianos

Actividad N° 3: Poner en práctica la encuesta con alumnos de la FCEFyN y FFHyA de la UNSJ. La encuesta ya se encuentra habilitada, y algunas respuestas se han obtenido, con alumnos correspondientes a los primeros años de las carreras. La información se vuelca en una base de datos para ser procesada. En la Fig. 2 se observa parcialmente las tablas de datos. Para el procesamiento de datos se ha trabajado con el software SPAD (Système Protable pur l'Analyses des Données) es un paquete estadístico general. Este paquete implementa métodos descriptivos y multidimensionales aplicables a grandes matrices de datos cuantitativos, cualitativos o textuales. [10].

4. FORMACION DE RECURSOS HUMANOS

El equipo de investigación está formado por docentes investigadores de dos facultades de la UNSJ. Se espera sumar alumnos tesis de grado y posgrado (maestría y doctorado), interesados en estas líneas de investigación.

2.- SECCIÓN A

LOCALIZACIÓN DEL ENCUESTADO EN LA UNSJ

***1. ¿En qué dependencia de la UNSJ estás cursando la carrera por la cuál te convocamos para contestar esta encuesta?**



- Facultad de Filosofía, Humanidades y Artes (FFHA)



- Facultad de Ciencias Exactas, Físicas y Naturales (FCEFYN)

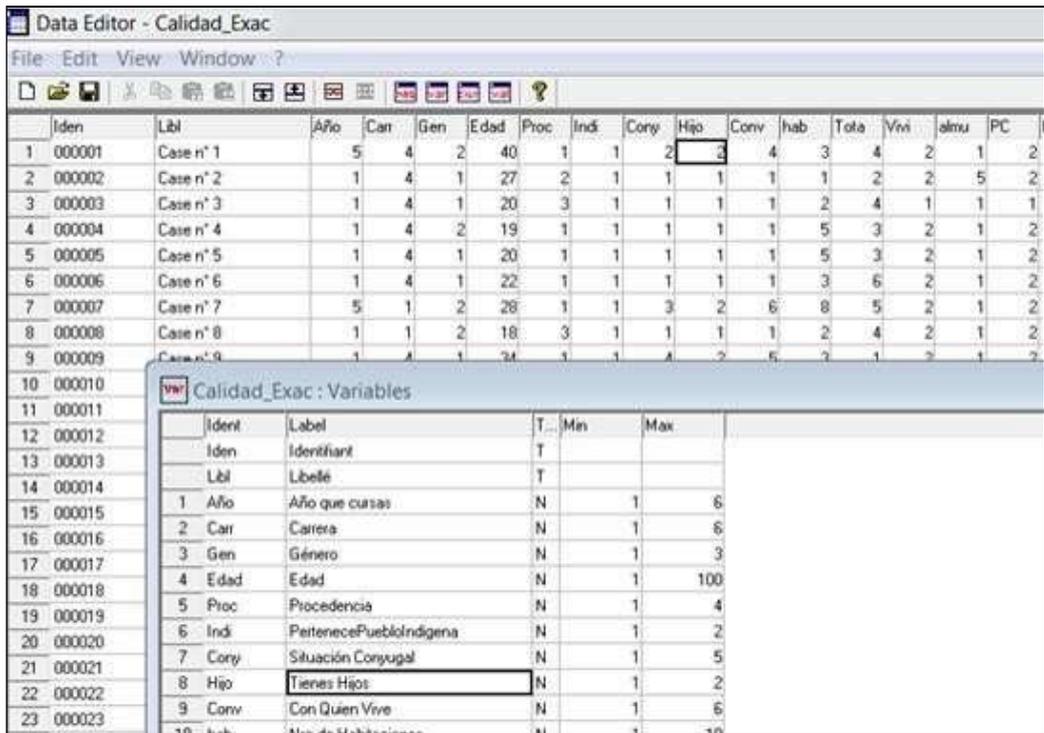
***2. ¿De qué año son la mayoría de las materias que estás cursado actualmente?**
Por favor seleccione alguna de las siguientes respuestas:

- Primero
- Segundo
- Tercer
- Cuarto
- Quinto
- Sexto

***3. Indica la carrera que cursas actualmente en tu facultad**

- TECNICATURA EN PROGRAMACIÓN WEB
- LICENCIATURA EN SISTEMAS DE INFORMACIÓN - TÍTULO INTERMEDIO TÉCNICO UNIVERSITARIO EN SISTEMAS DE INFORMACIÓN
- LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN - TÍTULO INTERMEDIO TÉCNICO UNIVERSITARIO EN PROGRAMACIÓN

Fig. 1. Imagen parcial de la encuesta en la web.



Data Editor - Calidad_Exac

File Edit View Window ?

Iden	Libl	Año	Carr	Gen	Edad	Proc	Indí	Cony	Hijo	Conv	hab	Tota	Vivi	almu	PC	
1	000001	Case n° 1	5	4	2	40	1	1	2	3	4	3	4	2	1	2
2	000002	Case n° 2	1	4	1	27	2	1	1	1	1	1	2	2	5	2
3	000003	Case n° 3	1	4	1	20	3	1	1	1	1	2	4	1	1	1
4	000004	Case n° 4	1	4	2	19	1	1	1	1	1	5	3	2	1	2
5	000005	Case n° 5	1	4	1	20	1	1	1	1	1	5	3	2	1	2
6	000006	Case n° 6	1	4	1	22	1	1	1	1	1	3	6	2	1	2
7	000007	Case n° 7	5	1	2	28	1	1	3	2	6	8	5	2	1	2
8	000008	Case n° 8	1	1	2	18	3	1	1	1	1	2	4	2	1	2
9	000009	Case n° 9	1	4	1	24	1	1	4	2	5	2	1	2	1	2
10	000010															
11	000011															
12	000012															
13	000013															
14	000014															
15	000015															
16	000016															
17	000017															
18	000018															
19	000019															
20	000020															
21	000021															
22	000022															
23	000023															

Calidad_Exac : Variables

Ident	Label	T...	Min	Max
Iden	Identifiant	T		
Libl	Libellé	T		
1	Año que cursas	N	1	6
2	Carrera	N	1	6
3	Genéro	N	1	3
4	Edad	N	1	100
5	Procedencia	N	1	4
6	PertenecePuebloIndigena	N	1	2
7	Situación Conyugal	N	1	5
8	Tienes Hijos	N	1	2
9	Con Quien Vive	N	1	6
10	...	N	1	10

Fig. 2. Tabla parcial de datos de encuesta

5. REFERENCIAS

- [1] T. Escudero Escorza, “La evaluación y mejora de la enseñanza en la Universidad: otra perspectiva,” in *Revista de Investigación Educativa*, vol. 18, no. 2, 2000, pp. 405–416.
- [2] M. Ato García, J. J. López García, and C. García Caro, *Análisis estadístico para datos categóricos*. Síntesis, 1996.
- [3] Benzecri, *Correspondence Analysis Handbook*. CRC Press, 1992.
- [4] I. González López, “Realización de un Análisis discriminante explicativo del rendimiento académico en la Universidad,” in *Revista de Investigación Educativa*, vol. 22, no. 1, 2004, pp. 43–59.
- [5] M. M. Torrado-Fonseca and V. Berlanga-Silvente, “Revista d’innovació i recerca en educació,” in *Revista d’Innovació i Recerca en Educació*, vol. 6, no. 2, Universitat de Barcelona, 2013, pp. 150–166.
- [6] E. Diday, “Análisis de Datos Simbólicos,” *Rev. IRICE*, vol. 11, 1997.
- [7] L. Lebart, A. Morineau, and M. Piron, *Statistique exploratoire multidimensionnelle*. DUNOD, 1995.
- [8] M. de M. Diaz, P. A. Urquijo, J. M. Arias Blanco, T. Escudero Escorza, S. Rodriguez Espinar, and J. Vidal García, “Evaluación del rendimiento en la enseñanza superior. Comparación de resultados entre alumnos procedentes de la LOGSE y del COU,” in *Revista de Investigación Educativa*, vol. 20, no. 2, M. de Miguel, P. A. Urquijo, J. M. A. Blanco, T. E. Escorza, S. R. Espinar, and J. V. García, Eds. 2002, pp. 357–383.
- [9] G. M. Garbanzo Vargas, “Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública,” *Rev. Educ.*, vol. 31, no. 1, pp. 43–63, 2007.
- [10] “Spad - Software Informer. SPAD, a data analytics software, uses company data to anticipate risks.” [Online]. Available: <http://spad.software.informer.com/>. [Accessed: 13-Mar-2017].

El Proceso de Extracción de Conocimiento en la Determinación del Perfil del Autor y la Atribución de Autoría

Mercado V.¹, Villagra A.², Errecalde M.³

¹⁻²Laboratorio de Tecnologías Emergentes (LabTEm), Instituto de Tecnología Aplicada (ITA)
Unidad Académica Caleta Olivia, Universidad Nacional de la Patagonia Austral.
Santa Cruz - Argentina.

³Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC),
Departamento de Informática - Universidad Nacional de San Luis. San Luis - Argentina.

{vmercado,avillagra}@uaco.unpa.edu.ar, merreca@unsl.edu.ar

RESUMEN

En el presente trabajo se describen, brevemente, las tareas de investigación y desarrollo que se están llevando a cabo en forma conjunta en el área de análisis de autor de documentos entre el LIDIC de la UNSL y el LabTEm de la UNPA. En particular, se ha tomado como caso de estudio primario los documentos de periodistas con diversas orientaciones políticas (oficialista vs opositor) con el objetivo de realizar con los mismos el Análisis de Autor y la Determinación/Caracterización del perfil del autor. Ambos tipos de tareas, han ganado creciente interés en la comunidad científica internacional y en empresas dedicadas al análisis de la información en la Web, por lo que la línea de investigación propuesta permitiría la formación de recursos humanos en temáticas relevantes a corto y mediano plazo tanto en el ámbito académico / científico como en la industria.

Palabras claves: Minería de Textos, Análisis de Autoría, Determinación del perfil del Autor, orientación política en Artículos Periodísticos. El proceso KDD.

CONTEXTO

Esta línea de trabajo se enmarca en los trabajos conjuntos que desde hace varios años llevan a cabo investigadores del LabTEm de la UNPA y el LIDIC de la UNSL. En particular, las tareas de investigación desarrolladas tienden a consolidar trabajos previos conjuntos relacionados a la Minería de Textos y la Web [Taquias et al., 2014], y complementarlos con los desarrollos que en el LIDIC se están llevando a cabo en las áreas específicas de análisis de autoría y determinación del perfil del autor [Funez et al., 2013., Villegas et al., 2014].

En este contexto, ambos laboratorios no sólo disponen de financiación obtenida de proyectos de investigación consolidados, sino que además se mantienen relaciones fluidas de investigación con centros de excelencia mundial especializados en estos temas como el Laboratorio de Tecnologías del Lenguaje del INAOE (Puebla, México) y el *Artificial Intelligence Laboratory-DICSE* de la *University of the Aegean* (Karlovasi, Grecia). En particular, una integrante del LabTEm desarrollará su trabajo de Maestría en esta temática, mientras que en el LIDIC un becario de

doctorado y uno post-doctoral de CONICET trabajarán en la temática específica de determinación del perfil del autor, y colaborarán en aquellos temas que se solapen con la presente investigación.

1. INTRODUCCIÓN

A partir de la disponibilidad de volúmenes inmensos de información en la Web, se reconoce cada día más el rol de la Minería de Datos (MD) como una herramienta fundamental para hacer un uso adecuado y ventajoso de esta información. Esta tendencia crece día a día y se plantean nuevos escenarios relevantes como es el caso de *Big Data*, donde el contexto en el cual deben ser aplicados los métodos de MD es sumamente desafiante. En particular, un área que comienza a ganar creciente interés es la *determinación del perfil del autor* (DPA), es decir, aquella que identifica patrones compartidos por un *grupo de gente* y que aborda problemas de clasificación de los usuarios de la Web de acuerdo a la edad, género, orientación política, etc. La DPA, un sub-campo del área más general conocida como *análisis de autoría* (AA), es un tema muy importante de investigación por sus potenciales (y actuales) aplicaciones en problemas de seguridad nacional e inteligencia, lingüística forense, análisis de mercados e identificación de rasgos de personalidad, entre otros. Otro sub-campo de la AA muy estudiado, denominado *atribución de autoría* (ATA), consiste en la atribución de un texto de autoría desconocida a uno de un conjunto de autores potenciales.

Si bien la MD, la DPA y la ATA son áreas de investigación científica muy activas, cuando se aplican a problemas concretos de la vida real se las debe considerar en el contexto más general del

proceso de extracción de conocimiento, que involucra varias etapas y herramientas para la recopilación de información, pre-procesamiento y extracción de características, análisis y visualización. El problema es que, usualmente, estas herramientas están dispersas, escritas en lenguajes y plataformas diferentes y, en muchos casos, como en el análisis de información textual, no están disponibles para el idioma español.

En este contexto, esta línea de investigación se propone el abordaje de dos tareas de AA, una de ATA y otra de DPA, como lo son la atribución de autoría y la determinación de la orientación política en documentos periodísticos, en el contexto de un proceso completo de extracción de conocimiento.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describen las líneas de investigación que se llevan a cabo en el proyecto:

El *análisis de autoría* (AA) [Stamatatos, 2009] es un área de investigación que ha ganado interés creciente en los últimos años principalmente por sus potenciales (y actuales) aplicaciones en problemas de seguridad nacional e inteligencia, lingüística forense, análisis de mercados e identificación de rasgos de personalidad, entre otros. El AA se enfoca en la clasificación automática de textos basándose fundamentalmente en las elecciones estilísticas de los autores de los documentos, e incluye distintas tareas de análisis como, por ejemplo: a) la *atribución de autoría*, b) la *verificación de autor*, c) la *detección de plagios*, d) la *determinación del perfil del autor* y e) la *detección de inconsistencias estilísticas*.

Los enfoques predominantes en esta área están basados en el aprendizaje automático/de máquina supervisado. En pocas palabras, estos enfoques derivan, a partir de un conjunto de datos etiquetados (conjunto de entrenamiento) y un proceso inductivo de aprendizaje/entrenamiento, un clasificador que puede generalizar sus predicciones a otros datos no observados previamente. La representación clásica de los textos/documentos en estos casos, incluye tanto atributos basados en el contenido (palabras) como en el estilo de escritura de los autores.

A partir de la disponibilidad de volúmenes inmensos de información en la Web, se reconoce cada día más el rol de la AA como una herramienta fundamental para hacer un uso adecuado y ventajoso de esta información, lo que ha quedado plasmado en un incremento de Workshops y Competencias específicos de esta temática. En particular, un área que comienza a ganar creciente interés es la determinación del perfil del autor, es decir, aquella que identifica patrones compartidos por un grupo de gente y que aborda problemas de clasificación de acuerdo a la edad y género [Peersman et al., 2011, Schler et al., 2006, Argamon et al., 2009], nacionalidad, personalidad [Celli et al., 2014, Mairesse et al., 2007], orientación política [Abooraig et al., 2014, Conover et al., 2011, Malouf & Mullen, 2007], etc.

Más allá de la relevancia y ventajas que pueden tener este tipo de tareas existe, actualmente, un desarrollo limitado en nuestro país de trabajos y grupos de investigación especializados en la problemática del AA. En este contexto, en esta línea de investigación nos enfocaremos en dos áreas claves de la AA como lo son la *determinación del perfil del autor* (DPA), y la *atribución de autoría* (ATA).

Respecto a la DPA, también conocida como *caracterización del autor* (en inglés *author profiling*), incluye actividades como la determinación automática de la edad, género, rasgos de personalidad y orientación política, entre otras. En nuestro caso, nos concentraremos en la orientación política (pro-gobierno vs opositor) de documentos periodísticos de acceso público, como libros de investigación periodística, blogs periodísticos, artículos en revistas y diarios on-line, etc.

Respecto a la ATA, analizaremos las particularidades que surgen para la identificación automática de autores, en aquellos contextos en donde los mismos tienen igual o diferente orientación política. En estos casos, se analizará cuáles son las *features* (estilográficas o de contenido) que son más relevantes para discriminar los distintos autores que pertenecen al mismo (o diferente) espectro político.

Por otra parte, a diferencia de los estudios de laboratorio, donde es usual disponer de datos recolectados y procesados *a priori*, listos para ser analizados, el proceso de extracción de conocimiento (en inglés *KDD*, por *Knowledge Discovery in Data*) [Kurgan & Musilek, 2006, Fayyad et al., 1996] involucrado en problemas prácticos concretos requiere de varias etapas y herramientas para la recopilación de información, pre-procesamiento y extracción de características, análisis y visualización. El problema es que, usualmente, estas herramientas están dispersas, escritas en lenguajes y plataformas diferentes y, en muchos casos, como en el análisis de información textual, no están disponibles para el idioma español. Si bien existen hoy en día nuevas herramientas y plataformas como

KNIME¹ y RapidMiner² que se suponen asisten al usuario en identificar e integrar estas etapas y herramientas, no siempre es claro cómo compatibilizan estas plataformas aspectos como la claridad, flexibilidad, facilidad de uso y extensión, entre otros. Por lo tanto, realizar una experiencia concreta sobre uno o varios problemas particulares (como la DPA y la ATA) utilizando una plataforma de este tipo, permitirá ganar experiencia que podrá servir no sólo en problemas de Minería de Textos y de la Web sino en otras tareas de análisis futuros que involucren otros datos arbitrarios como, por ejemplo, imágenes, videos, sonido, datos de redes de sensores, etc.

3. RESULTADOS OBTENIDOS / ESPERADOS

En cuanto a los resultados se pretende lograr un sistema integrado de *atribución de autoría* con periodistas de la Argentina y de *determinación de la orientación política en documentos periodísticos*, que también soporte el *descubrimiento de tópicos* en estos documentos. Este resultado quedará plasmado en distintos *workflows* del tipo de los soportados por KNIME, en los cuales los distintos “nodos” que componen las tareas quedan explícitamente expresados, facilitándose su uso y modificación por parte de aquellos usuarios no familiarizados con este tipo de tareas. Estos *workflows* contarán además con nodos dedicados a la evaluación y clara visualización de los resultados obtenidos.

El sistema anterior será utilizado en trabajos experimentales realizándose comparaciones con enfoques similares representativos del estado del arte en el

¹ <https://www.knime.org/>

² <https://rapidminer.com/>

área. Un objetivo adicional a largo plazo es que la experiencia obtenida con estas tareas sirva para abordar otros procesos de KDD que involucren otros tipos de datos (imágenes, videos, etc.) lográndose así consolidar en la UNPA un equipo de trabajo especializado en temáticas de gran relevancia nacional e internacional.

4. FORMACIÓN DE RECURSOS HUMANOS

Un integrante de este proyecto de investigación cuenta con una beca Post-Doctoral de CONICET enfocada en aspectos psicológicos relacionados a las tareas de DPA.

Un integrante de este proyecto de investigación está desarrollando su Tesis de Doctorado sobre DPA multimodal con una beca Doctoral de CONICET.

Un integrante está desarrollando su tesis de Maestría orientada en esta línea de investigación.

5. REFERENCIAS

- [1] **Abooraig R.**, Alwajeeh A., Al-Ayyoub M., and Hmeidi I.(2014). On the automatic categorization of arabic articles based on their political orientation. In *Proc. of the Third International Conference on Informatics Engineering and Information Science (ICIEIS2014)*.
- [2] **Argamon S.**, Dhawle S., Koppel M. and Pennebaker J. (2005). *Lexical Predictors of Personality Type*. Joint Annual Meeting of the Interface and the Classification Society of North America.
- [3] **Celli F.**, Lepri B., Biel J.-I., Gatica-Perez D., Riccardi G., and Pianesi F.(2014). *The workshop on computational personality recognition 2014*. In *Proceedings of the ACM International*

Conference on Multimedia, MM '14, pages 1245-1246, New York, NY, USA. ACM.

[4] **Conover M.**, Goncalves B., Ratkiewicz J., Flammini A., and Menczer F. (2011). *Predicting the political alignment of twitter users*. In Proceedings of 3rd IEEE Conference on Social Computing (SocialCom).

[5] **Kurgan L. A.** and Musilek P. (2006). *A survey of Knowledge Discovery and Data Mining process models*. Knowledge Engineering Review. 21, 1 (March 2006), 1-24.

[6] **Fayyad U. M.**, Piatetsky-Shapiro, G. and Smyth, P. (1996). *From data mining to knowledge discovery: an overview*. In Advances in knowledge discovery and data mining, Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy (Eds.). American Association for Artificial Intelligence, Menlo Park, CA, USA 1-34.

[7] **Funez D. G.**, Cagnina L., and Errecalde M. L. (2013). *Determinación de género y edad en blogs en español mediante enfoques basados en perfil*. In Anales del XIX Congreso Argentino de Ciencias de la Computación (CACIC 2013), pages 1003-1012.

[8] **Mairesse F.**, Walker M. A., Mehl M. R. and Moore R. K. (2007). *Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text*. In JAIR, 30, 457-500.

[9] **Malouf R.** and Mullen T. (2007) *Graph-based user classification for informal online political discourse*.

[10] **Peersman C.**, Daelemans, W. and Van Vaerenbergh L. (2011). *Predicting age and gender in online social networks*. In Proceedings of the 3rd international workshop on Search and mining user-

generated contents, SMUC '11, pages 37-44, New York, NY, USA. ACM.

[11] **Schler J.**, Koppel M., Argamon S., and Pennebaker, J. W. (2006). Effects of age and gender on blogging. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 199-205, 2006.

[12] **Taquias D.**, Villagra A., and Errecalde M. L.. *Detección de plagios con adversarios*. In Anales del XVI Workshop de Investigadores en Ciencias de la Computación (WICC 2014), pages 233, 237, 2014.

[13] **Villegas M. P.**, Garcíarena-Ucelay M. J., Errecalde M. L. and Cagnina L. C. (2014). *A Spanish Text Corpus for the Author Profiling Task*, XX Congreso Argentino de Ciencias de la Computación, Buenos Aires, Argentina.

Evaluación de Técnicas de Clasificación para Predecir el Rendimiento Académico de Ingresantes a la Universidad en Temáticas de Matemática

Maria Paula DIESER⁽¹⁾, Lorena Verónica CAVERO⁽¹⁾, María Cristina MARTÍN⁽¹⁾⁽²⁾,
Erica SCHLAPS⁽¹⁾, Diamela TITIONIK⁽¹⁾, Laura WAGNER⁽¹⁾

⁽¹⁾ Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa

⁽²⁾ Departamento de Matemática, Universidad Nacional del Sur

{pauladieser, cavero, maritamartin}@exactas.unlpam.edu.ar

Resumen

En el proceso de inscripción a las carreras de grado de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa, y en el desarrollo de las actividades del Programa de Ambientación a la Vida Universitaria de la Institución, se recolectan múltiples datos aportados por los aspirantes a través de los sistemas de gestión. Éstos constituyen una importante fuente de información, en tanto se extraiga conocimiento para el análisis de la realidad de los estudiantes y los contextos en los que ellos aprenden, y para el diseño de eventuales planes de acción. Es una realidad la constante preocupación de la comunidad institucional por los elevados índices de deserción, o retrasos en alcanzar su título de grado, por dificultades en asignaturas vinculadas con la matemática.

La línea de investigación presentada propone procesar los datos recolectados a través de los sistemas de gestión durante el ingreso, y resultados del seguimiento de la actividad académica en asignaturas de matemática, para obtener posibles patrones entre los estudiantes que alcancen idénticos logros. Los modelos resultantes permitirán predecir el rendimiento académico en el área y determinar factores que lo afectan para implementar políticas de retención adecuadas.

Palabras clave: clasificación, predicción, minería de datos, rendimiento académico

Contexto

Desde 2014 se vienen realizando tareas de investigación, en el ámbito de la Facultad de Ciencias Exactas y Naturales (FCEyN) de la Universidad Nacional de La Pampa (UNLPam), relacionadas con el estudio y aplicación de métodos multivariados de discriminación y de clasificación, con el propósito de establecer similitudes y diferencias, y analizar las estimaciones que se obtienen con ellos al aplicarlos efectivamente en el Análisis de Datos Multivariados. El Proyecto, acreditado y financiado por la Institución mencionada, ha contado también con la participación de estudiantes de postgrado de la Universidad Nacional de Asunción (UNA). Entre los métodos estudiados en el marco del Proyecto, se encuentran algunos que podrían entenderse como clásicos y de una esencia más estadística (discriminación debida a Fisher), y otros propios del *Data Mining* (Árboles y Reglas de Clasificación, Redes Neuronales, y el Análisis de *Clusters*). Se ha desarrollado la teoría sobre estas técnicas, y aplicado a diferentes conjuntos de datos a fin de analizar su sensibilidad y fiabilidad, realizando prácticas con el lenguaje de programación R. De las investigaciones realizadas, surge el campo de la educación como un terreno propicio para las aplicaciones de *Data Mining*, dada la multiplicidad de fuentes de datos y los diversos grupos de interés implicados.

Asimismo, el área educativa ofrece la posibilidad de aplicar elementos de la Teoría de Respuesta al Ítem para el análisis de las respuestas en cuestionarios, y el Análisis de Supervivencia, para extraer conclusiones del tiempo requerido para la aprobación de espacios curriculares o la graduación.

1. Introducción

La comunidad universitaria en su conjunto se plantea y propone la mejora continua de la calidad de los procesos educativos que se desarrollan en sus instituciones y de los servicios que ofrecen. La FCEyN de la UNLPam no es ajena a esta realidad. El equipo de gestión, cuerpo docente y agrupaciones estudiantiles, a través de la Comisión *ad hoc* de Ingreso y Permanencia (CIP), han diagnosticado altos niveles de deserción y desgranamiento en los primeros años de estudio, en muchos casos asociados a los bajos rendimientos en asignaturas vinculadas con la matemática. No obstante, los diagnósticos realizados carecen de la sistematización necesaria que permita revelar a tiempo el abandono de estudiantes en diferentes tramos de las carreras elegidas.

Entre las políticas de gestión impulsadas por la CIP, se organiza en cada ciclo lectivo, previo al inicio de las cursadas regulares, una serie de acciones en el marco del Programa de Ambientación a la Vida Universitaria (PAVU) que incluye charlas, talleres y actividades recreativas destinadas a los aspirantes. Entre éstas, desde 2015, se desarrolla el Taller “Introducción a la Matemática” cuyo propósito es recuperar los conocimientos de matemática elemental que poseen los ingresantes y que son requeridos para el cursado de las asignaturas del área, contempladas en la oferta académica de grado de la FCEyN (UNLPam).

El Taller está a cargo de docentes y auxiliares docentes del Departamento de Matemática de la institución y cuenta con la colaboración de estudiantes avanzados de Profesorado y Licenciatura en Matemática en las tareas de tutorías. Es de carácter semipresencial, durante las tres semanas previas al inicio del primer período de clases. Los encuentros presenciales se distribuyen en 8 encuentros de 2 horas reloj

cada uno, y están acompañados por actividades diversas implementadas sobre el curso *online* del Taller, desarrollado en el entorno virtual de enseñanza y aprendizaje *Moodle*. Este curso se estructuró en tres temas correspondientes a los tres bloques temáticos considerados para el tratamiento de los contenidos (Números, Álgebra, y Funciones). Se pusieron a disposición de los estudiantes, los materiales diseñados *ad hoc* (apuntes teóricos y trabajos prácticos), así como diferentes foros destinados a la comunicación de las novedades del Taller, el establecimiento de lazos sociales, y la evacuación de dudas y consultas. A partir de 2016, finalizado el Taller, se propuso a los estudiantes completar una autoevaluación de los contenidos trabajados y un cuestionario diseñado con preguntas cerradas vinculadas con diversos aspectos de índole demográfica, social, emocional, y escolaridad previa que se supone pueden afectar el rendimiento del estudiante. Estos datos, junto con los referidos a la asistencia al Taller y la participación en las actividades del curso *online* asociado, son considerados para analizar la influencia de las variables involucradas en el rendimiento académico de los estudiantes en asignaturas de matemática cursadas en el primer año de sus carreras. La finalidad es obtener información útil para la identificación temprana de estudiantes en riesgo, y el establecimiento de una política de apoyo académico adecuada para atender la situación y, eventualmente, disminuir los índices de fracaso y abandono. Este tipo de estudios en el campo de la educación corresponde a aplicaciones de una rama particular del *Data Mining* (DM) conocida como Minería de Datos Educativos (EDM, por sus siglas en inglés). Este nuevo espacio de investigación interdisciplinario se ocupa del desarrollo y utilización de métodos para explorar los datos que se dan en el ámbito educativo, así como también para entender mejor a los estudiantes y los contextos en que ellos aprenden (Romero & Ventura, 2010). Romero et al. (2010) definen la EDM como el desarrollo, investigación y aplicación de métodos computacionales para detectar patrones en grandes conjuntos de datos

educativos que, de otro modo, serían difíciles o imposibles de analizar debido a su volumen. Revisiones de investigaciones realizadas en EDM dan cuenta de los objetivos perseguidos y las diversas aplicaciones posibles en el área (Romero & Ventura, 2007, 2010; Baker & Yacef, 2009). En particular, Romero & Ventura (2010) elaboran una taxonomía de las áreas de aplicación de EDM, entre las que se menciona la predicción del desempeño de estudiantes.

Sin embargo, el estudio del rendimiento académico de los estudiantes y el abandono escolar no es de interés reciente, y siempre ha estado relacionado con factores sociales, económicos y psicológicos. Varios estudios han abordado estos temas usando distintas metodologías: análisis discriminante, reglas de asociación, modelos de regresión logística y de imputación múltiple, análisis de la varianza, árboles de decisión, redes neuronales, redes bayesianas, entre otros (Streeter & Franklin, 1991; Ma et al., 2000; Wayman, 2001; Pursley, 2002; Minaei-Bidgoli et al., 2003; Kotsiantis et al., 2004; Pardos et al., 2006; Cortez & Silva, 2008; Márquez Vera et al., 2012).

Por otra parte, la Teoría de Respuesta al Ítem (TRI) y el Análisis de Supervivencia (AS) son áreas de investigación estadística que podrían ofrecer técnicas adecuadas para el análisis de datos educativos. En particular, la TRI ofrece estimaciones del rasgo latente de individuos medidos mediante un test o cuestionario (Hidalgo Flores, 2007). Su utilidad en el campo educativo radica en determinar si un estudiante consigue responder correctamente a cada una de las preguntas que componen el cuestionario y en atender al puntaje bruto obtenido en la prueba (Debera & Nalbarte, 2006). Por su parte, el AS permite modelizar el tiempo que se tarda en que ocurra un determinado suceso y su dependencia con otras posibles variables explicativas. En el ámbito de la educación, aporta técnicas apropiadas para analizar el tiempo requerido para graduarse o alcanzar determinado objetivo, así como su relación con predictores

sociodemográficos y de aptitud académica, entre otros (Gallardo Allen et al., 2016).

2. Líneas de Investigación y Desarrollo

La línea de investigación aquí presentada surge de un Proyecto más amplio cuyo objetivo general es investigar técnicas de discriminación y clasificación multivariadas, con el propósito de establecer similitudes o diferencias y analizar la eficiencia de las mismas al aplicarlas, efectivamente, en el análisis de datos multivariados. En esta línea se pretende aplicar técnicas de DM para predecir el rendimiento académico de estudiantes ingresantes a la FCEyN (UNLPam) en asignaturas vinculadas con la matemática, empleando diversos métodos de clasificación (reglas y árboles de clasificación, y redes neuronales), en combinación con otros propios de la TRI y el AS. Las técnicas seleccionadas serán aplicadas sobre una base de datos construida a partir de distintas fuentes: cuestionarios con preguntas cerradas vinculadas con diversos aspectos de índole demográfica, social, emocional, y escolaridad previa que se supone pueden afectar el rendimiento del estudiante; resultados obtenidos del Taller "Introducción a la Matemática" desarrollado durante 2016 (asistencia, participación, y autoevaluación); e informes finales incluyendo las calificaciones obtenidas por los estudiantes en las asignaturas de matemática cursadas en el primer año de las carreras respectivas. Los resultados obtenidos serán comparados, y los mejores modelos resultantes podrían ser de utilidad en la identificación temprana de estudiantes en riesgo, y el establecimiento de una política de apoyo académico adecuada para atender la situación y, eventualmente, disminuir los índices de fracaso y abandono.

3. Resultados Obtenidos y Esperados

Hasta el momento se han llevado a cabo la fase de integración y recopilación de datos (determinando las fuentes de información descriptas anteriormente consideradas de utilidad para conformar una base de datos unificada); y la fase de limpieza y transformación como parte del

preprocesamiento de los datos (detectando la existencia de ciertos “problemas” en los datos, analizando y decidiendo formas adecuadas para su tratamiento). Como resultado de estos procesos se ha obtenido una vista minable de los datos recopilados. Actualmente se están realizando las tareas vinculadas con la selección de variables relevantes para el objetivo de estudio, y la transformación o combinación de estas últimas. Esto permitirá reducir la dimensionalidad del problema de manera adecuada para simplificar el trabajo en las fases posteriores, a saber, (a) selección del método que produzca los patrones y modelos más expresivos; (b) evaluación de los patrones obtenidos a partir de un análisis e interpretación del conocimiento obtenido; y (c) comparación de los modelos obtenidos con los que pudieran surgir de la aplicación de otras técnicas de clasificación clásicas.

Se espera que los resultados alcanzados contribuyan a la identificación temprana de estudiantes en riesgo, y al establecimiento de estrategias académicas adecuadas para atender la situación y, eventualmente, disminuir los índices de fracaso y abandono.

Esta línea de investigación podría dar origen a un nuevo Proyecto más amplio en el que se consideren los datos registrados en el sistema de gestión de información estudiantil (SIU Guarani) de la FCEyN (UNLPam) a fin de vincular la información referida al rendimiento académico de los estudiantes con aquellos de índole socioeconómica, familiar, escolaridad previa, entre otros.

4. Formación de Recursos Humanos

En el área del Proyecto de Investigación, bajo la Dirección de la Dra. Martín, se han formado dos de los integrantes como graduados de la Maestría en Estadística y Metodología de la Investigación Científica Básica y Aplicada de la FCEyN (UNA) y, otros dos como egresadas de la Licenciatura en Matemática de la FCEyN (UNLPam), en 2014 y 2016, respectivamente. En la línea aquí presentada (DM), de un total de once integrantes, trabajan dos con formación de base matemática, computacional y pedagógica, quienes finalizaron el cursado

de la Maestría en Tecnología Informática Aplicada en Educación de la Facultad de Informática (UNLP) y se encuentran en proceso de elaboración del proyecto de tesis. Además, tres de las integrantes han comenzado sus estudios de Doctorado en Estadística en la Universidad Nacional de Rosario, y una de ellas proyecta realizar su trabajo de Tesis Doctoral en AS, línea que, como ya se manifestara, se plantea aplicar para el estudio de la permanencia de los estudiantes universitarios. Respecto de la otra línea de investigación (TRI) a estudiar, a la brevedad, se solicitará al Consejo Directivo de la FCEyN (UNLPam) la incorporación de una nueva integrante del proyecto, quien actualmente se desempeña como auxiliar docente en la Institución y ha obtenido una beca del Programa Becas de Investigación y Postgrado / Subprograma Becas de Postgrado para iniciar Doctorados y Maestría de la UNLPam, por el período 2017-2018. El plan propuesto, también bajo la dirección de la Dra. Martín, es “La Teoría de Respuesta al Ítem aplicada a prueba diagnóstica de ingreso universitario”, y busca la obtención del grado de *master* en la Maestría en Estadística Aplicada de la Universidad Nacional de Córdoba. Finalmente, cabe señalar que, se prevé sumar estudiantes de grado interesados en esta línea de investigación.

5. Bibliografía

Baker, R. S. J. D. & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1(1):3–16.

Cortez, P. & Silva, A. (2008). Using data mining to predict secondary school student performance. En Brito, A. and Teixeira, J. (Eds.), *Proceedings of 5th Future Business Technology Conference*, pp. 5–12, Porto, Portugal. EUROISIS.

Debera, L. & Nalbarte, L. (2006). *Pruebas diagnósticas: una aplicación a la teoría de respuesta al ítem, aproximación clásica y*

bayesiana. Instituto de Estadística. F.C.E. y Administración, Universidad de la República.

Gallardo Allen, E, Molina Delgado, M. & Cordero Cantillo, R. (2016). Aplicación del Análisis de Supervivencia al Estudio del Tiempo Requerido para Graduarse en Educación Superior: El Caso de la Universidad de Costa Rica. *Páginas de Educación*, 9(1):61–87.

Hidalgo Flores, R. (2007). *Teoría de respuesta al ítem: una aplicación educativa*. Facultad de Ingeniería, Universidad Autónoma de Querétaro, México.

Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2004). Predicting student's performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, 18(5):411–426.

Ma, Y., Liu, B., Wong, C. K., Yu, P. S., & Lee, S. M. (2000). Targeting the right students using data mining. En *Proceedings of 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 457–464, Boston, USA.

Márquez Vera, C., Romero Morales, C., & Ventura Soto, S. (2012). Predicción del Fracaso Escolar Mediante Técnicas de Minería de Datos. *IEEE-RITA*, 7(3):109–117.

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). Predicting student performance: an application of data mining methods with an educational web-based system. En *Proceedings of 33rd Annual Frontiers in Education, FIE 2003*, pp. 13–18, Colorado, USA.

Pardos, Z. A., Heffernan, N. T., Anderson, B., and Heffernan, C. L. (2006). Using fine-grained skill models to fit student performance with bayesian networks. En *Proceedings of the Workshop in Educational Data Mining held at the 8th International Conference on Intelligent Tutoring Systems*, Taiwan.

Pursley, M. (2002). *Changes in Personal Characteristics of Mexican-American High*

School Graduates and Dropouts During the Transition from Junior High to High School. Texas Tech University.

Romero, C. & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Syst. Appl.*, 33(1):135–146.

Romero, C. & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 40(6):601–618.

Romero, C., Ventura, S., Pechenizky, M., & Baker, R. (2010). *Handbook of Educational Data Mining*. Chapman and Hall CRC Press, Taylor & Francis Group, Boca Raton.

Streeter, C. L. & Franklin, C. (1991). Psychological and family differences between middle class and low income dropouts: A discriminant analysis. *The High School Journal*, 74(4):211–219.

Wayman, J. C. (2001). Factors influencing GED and diploma attainment of high school dropouts. *Education Policy Analysis Archives*, 9(4):1–19.

Extensiones Biométricas para Bases de Datos Objeto-Relacionales

Ernesto Miranda, Silvia Ruiz, Juan José Aguirre, Mauro Herlein, Graciela Etchart, Carlos Alvez

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

emiranda@fcad.uner.edu.ar, sruiz@fcad.uner.edu.ar, juaagu@fcad.uner.edu.ar, herlein.mauro@gmail.com,
getchart@fcad.uner.edu.ar, caralv@fcad.uner.edu.ar

Resumen

La autenticación de personas basadas en rasgos biométricos se ha vuelto muy popular en los últimos años como consecuencia de la baja en los costos de los sensores requeridos, su inclusión en dispositivos de consumo masivo y el surgimiento de vulnerabilidades debido al uso de múltiples claves de acceso a diferentes sitios que requieren cierto nivel de seguridad como ser cuentas de correo, sitios de banca electrónica, sistemas corporativos, etc. De todos los rasgos utilizados en biometría, el iris es uno de los más estables a lo largo de la vida de una persona y también uno de los más difíciles de falsificar. El citado incremento tiene especial efecto en el requerimiento de desarrollo de aplicaciones biométricas para las cuales deben preverse estructuras de datos adecuadas para el almacenamiento de información, la cual posee cierto grado de complejidad. Las bases de datos relacionales presentan deficiencias al tratar con estructuras complejas de información, es por eso que se opta por utilizar un modelo objeto-relacional. El presente trabajo tiene como objetivo el desarrollo en forma de extensión para bases de datos objeto-relacionales de modelos de datos que representen el rasgo biométrico de iris.

Palabras clave: base de datos objeto-relacional, biometría, iris

Contexto

Este artículo constituye un trabajo de Tesis de Maestría, que se lleva adelante en la Universidad Nacional de Entre Ríos - Facultad de Ciencias de la Administración, en el ámbito de la Maestría en Sistemas de Información acreditada por CONEAU según Res. No 1276/12. El mismo se enmarca en el Proyecto de Investigación PID UNER 7044 "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales" desarrollado en la misma unidad académica, cuyo objetivo general es desarrollar sistemas para la identificación de personas mediante el reconocimiento de iris y voz que permitan la interoperabilidad entre organismos acorde a normas internacionales.

Introducción

La autenticación basada en biometría hace referencia al proceso de establecer una identidad de un individuo en base a características físicas y de comportamiento tales como su rostro, huellas dactilares, geometría de la mano, iris, voz, firma, entre otros [1]. Los sistemas biométricos ofrecen diversas ventajas respecto de los sistemas tradicionales de autenticación, los cuales se

basan principalmente en contraseñas, dado que los rasgos biométricos no pueden perderse u olvidarse y resultan difíciles de copiar y distribuir.

Como se ha mencionado, existen varios rasgos biométricos, cada uno de los cuales poseen fortalezas y debilidades y su elección dependerá de la aplicación que se le quiera dar. La idoneidad de un rasgo biométrico para una aplicación será determinada por los requerimientos de esa aplicación y las propiedades del rasgo biométrico.

En el caso del iris se trata de la región anular del ojo rodeada por la pupila y la esclerótica. La estructura visual del iris se forma durante el desarrollo del feto y se estabiliza en los primeros dos años de vida. Su compleja estructura porta información sumamente útil para el reconocimiento de personas. Se cree que cada iris es distintivo e incluso los iris de gemelos idénticos son diferentes. Es muy difícil cambiar su estructura quirúrgicamente e incluso la habilidad de detectar iris artificiales ha sido demostrada.

Dada su aplicación en seguridad, ya sea para procesos de autenticación o de identificación, para controles de acceso a sistemas como a lugares físicos la eficiencia en la recuperación y comparación de datos resulta crítica. Dado que las estructuras de metadatos requeridas para la representación de rasgos biométricos son no atómicas, deben ser separadas entre múltiples tablas en bases de datos relacionales, lo que hace menos eficiente su tratamiento [2]. Además se debe proveer del procesamiento de información realizado en un lenguaje separado, debiendo transferir la información del motor de base de datos a la plataforma donde se realice el procesamiento.

Estos inconvenientes pueden ser superados utilizando tecnología de bases de datos objeto-relacionales las cuales proveen soluciones a las limitaciones de las bases de datos relacionales mediante las siguientes posibilidades [3]:

- Definición de tipos de datos: que involucren estructuras complejas como colecciones, arreglos, objetos largos.
- Definición e implementación de comportamientos: se pueden definir métodos que manipulen la información de los tipos definidos, pudiendo implementar, por ejemplo, la comparación de dos rasgos en la base de datos directamente, lo cual mejora la eficiencia.

Teniendo en cuenta lo antes expuesto, el aporte esperado del presente trabajo de tesis consistirá en el desarrollo de modelos de datos que permitan el almacenamiento, recuperación, comparación, etc. de rasgos biométricos de iris sobre bases de datos objeto-relacionales en forma de extensión.

Las extensiones en bases de datos son librerías agregadas al sistema gestor de base de datos (SGBD) que tienen como objetivo agregar funcionalidades específicas desarrolladas por usuarios. Resulta importante destacar que existen diversos sistemas gestores de bases de datos objeto-relacionales (SGBDOR) con soporte de extensiones de diverso tipo y función, pero que no se han encontrado extensiones desarrolladas que den soporte a problemáticas en el campo de la biometría.

Líneas de Investigación y Desarrollo

En el presente plan de tesis confluyen varias líneas de investigación entre las cuales encontramos biometría, bases de datos objeto-relacionales y procesamiento de imágenes.

La creciente preocupación por la seguridad en las organizaciones, en todos los niveles, ha abierto varias líneas de investigación, entre ellas, las relacionadas con biometría. La biometría es una tecnología en pleno desarrollo, tanto en el ámbito de la vida cotidiana como en la investigación [4].

Los SGBDOR surgieron de las investigaciones llevadas a cabo a principios de los años 90 que buscaban extender las bases de datos relacionales agregando soporte a objetos. Los primeros productos comerciales surgieron a mediados de los 90 y muchas de las ideas de la orientación a objetos fueron incorporadas al estándar SQL:1999 [5].

Los SGBDOR impulsaron el desarrollo de extensiones para bases de datos que dieron soporte a diversos problemas del mundo real, los cuales presentaban una cierta complejidad al ser tratados en bases de datos tradicionales. A pesar de esto, ninguna de ellas da soporte al almacenamiento de metadatos biométricos.

En cuanto a los modelos de datos utilizados existe el estándar ANSI / NIST ITL 1-2011 [6] el cual define como asegurar la interoperabilidad de la información biométrica entre diferentes sistemas. Este estándar define el contenido, formato y unidades de medida para el intercambio electrónico de datos biométricos de diferentes rasgos. Cada uno de los rasgos es soportado por un registro en especial dentro del estándar, siendo el registro tipo 17 el que se encarga de representar al iris.

En [2] se presenta un modelo realizado en un diagrama de clases UML que representa al registro tipo 17 del estándar ANSI / NIST ITL 1-2011 el cual podría representar un punto de partida para este trabajo.

Resultados Obtenidos/Esperados

El objetivo general de este trabajo consiste en el desarrollo e implementación de modelos de datos que permitan el almacenamiento, recuperación y comparación de rasgos biométricos de iris sobre bases de datos objeto-relacionales en forma de extensión.

Los objetivos específicos son:

- Diseñar modelos de datos que den soporte al almacenamiento de rasgos de iris.
- Implementar una extensión para SGBDOR con dichos modelos.
- Implementar métodos en dicha extensión con las operaciones básicas a realizar con los objetos que se almacenen en la base de datos.
- Verificar su funcionamiento y rendimiento.

El trabajo de esta tesis intenta beneficiar a todos aquellos que se encuentren trabajando en el desarrollo aplicaciones en el campo de la biometría, facilitando la implementación de las estructuras necesarias para la representación de rasgos de iris. Esperamos que este desarrollo sirva como base para la incorporación de otros rasgos biométricos a la extensión.

Formación de Recursos Humanos

La estructura del equipo de Investigación está formada por un Director, una Co Directora, cuatro docentes, un becario de Iniciación a la Investigación y un integrante interno alumno de la carrera Licenciatura en Sistemas.

En la actualidad cuatro planes de tesis se desarrollan en relación con el proyecto de investigación, dos dentro del área de Bases de Datos y dos dentro del área de Biometría de Voz, en el marco de la Maestría en Sistemas de Información que se dicta en la Universidad Nacional de Entre Ríos. Dichas tesis están siendo dirigidas por el Director del Proyecto de Investigación.

Bibliografía

1. Anil K. Jain, Arun Ross y Sharath Pankanti: Biometrics: A Tool for Information Security. Ieee Transactions On Information Forensics And Security, Vol. 1, No. 2, Junio 2006

2. Alvez Carlos, Etchart Graciela, Ruiz Silvia, Miranda Ernesto: Iris Information Management in Object-Relational Databases. XXI Congreso Argentino de Ciencias de la Computación, Octubre 2015.
3. Carlos E. Alvez, Aldo R. Vecchiatti: Combining Semantic and Content Based Image Retrieval in ORDBMS. Knowledge-Based and Intelligent Information and Engineering Systems Lecture Notes in Computer Science, 2010, Volume 6277/2010, pp. 44-53. Springer-Verlag Berlin Heidelberg (2010).
4. Alvez Carlos, Etchart Graciela, Ruiz Silvia, Miranda Ernesto: Representación e Interoperabilidad de Imágenes Biométricas. Workshop de Investigadores en Ciencias de la Computación. WICC 2015, Abril 2015
5. ISO/IEC 9075-1:1999 Information technology – Database languages – SQL – Part 1: Framework.
6. ANSI/NIST-ITL 1-2011. Update: 2013. Information Technology: American National Standard for Information Systems Data Format for the Interchange of Fingerprint, Facial and Other Biometric Information. (2013).

Formulación de Esquemas de Almacenamiento de Datos Médicos para aplicar Minería de Datos en el Diagnóstico de Enfermedades

Ana Lía Carabio¹, Elizabeth Silva Layes¹, Marcelo A. Falappa²

¹Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos
Monseñor Tavella 1424 – Concordia, Entre Ríos (3200) - Tel.: +54(0345)4231406
anacar@fcad.uner.edu.ar, elizabeth.silva@gmail.com

²Departamento de Ciencias e Ingeniería de la Computación - Universidad Nacional del Sur
San Andrés 800 – Campus de Palihue - Bahía Blanca (8000) - Tel.: +54(0291)4595135
mfalappa@cs.uns.edu.ar

Resumen

El sector salud administra grandes volúmenes de datos, centrándose la mayoría de las tomas de decisiones en el área clínica. Por tal motivo, contar con información útil, inmediata y efectiva es sumamente relevante en éste ámbito.

En este sentido, la minería de datos es una herramienta que permite encontrar patrones de comportamiento de utilidad para la toma de decisiones clínicas, como lo son la realización de estudios epidemiológicos, cálculo de expectativas de vida, identificación de terapias médicas satisfactorias para diferentes enfermedades, entre otros.

El objetivo del presente trabajo es construir un esquema que, a partir de la extracción de datos médicos relevantes de las Historias Clínicas Electrónicas (HCE), permita almacenarlos de manera eficiente en una base de datos NoSQL, como *HBase*, con la finalidad de aplicar técnicas de minería de datos.

Palabras clave: Historia Clínica

Electrónica, Minería de Datos, Big Data, Bases de Datos NoSQL, HBase, Hadoop.

Contexto

Este trabajo se desarrolla dentro del Proyecto de Investigación y Desarrollo PID 7042 “Estudio Comparativo y Análisis de Rendimiento de los Lenguajes de Manipulación de Datos en Bases de Datos Orientadas a Objetos y Bases de Datos Objeto-Relacionales”[1], cuyo período de ejecución será desde noviembre de 2014 a noviembre de 2017, en el marco de un Acuerdo de Colaboración Académico-Científico entre la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos (UNER) y el Instituto de Ciencias e Ingeniería de la Computación (ICIC) del Departamento de Ciencias e Ingeniería de la Computación (DCIC) de la Universidad Nacional del Sur (UNS).

Uno de los objetivos del proyecto apunta a establecer comparaciones en el rendimiento de sistemas desarrollados en

lenguajes orientados a objetos que interactúan con diversos modelos de bases de datos.

Además, este proyecto prioriza la formación de recursos humanos para investigación en la Facultad de Ciencias de la Administración de la UNER, especializados en la línea de investigación denominada “Ingeniería de Software y Lenguajes de Programación” establecida por Res. 25/11 del Consejo Directivo.

Introducción

En la actualidad, el procesamiento de grandes volúmenes de datos (*Big Data*) para la toma de decisiones ha dejado de ser privativo de organizaciones comerciales y de negocios, y se ha inmiscuido en otros ámbitos, con actividades variadas y diversos intereses. Entre ellos, el sector salud es uno de los sectores que más se ha visto beneficiado con la utilización de herramientas de análisis de datos.

La vasta acumulación de datos clínicos existentes en los *Electronic Health Records* (EHR's) [1] (diagnósticos, tratamientos indicados, paraclínicas, medicamentos suministrados, procedimientos realizados, etc.) presentes en la mayoría de las instituciones sanitarias, brinda una oportunidad inmejorable para: la realización de estudios epidemiológicos, el cálculo de expectativas de vida, la identificación de terapias médicas satisfactorias para diferentes enfermedades, etc [2].

El sector sanitario, en su totalidad, es uno de los que administra los mayores volúmenes de datos, centrándose la

mayoría de las tomas de decisiones en el área clínica. Sin duda, los Sistemas de Soporte a Decisiones Clínicas (CDSS), además de apoyar al médico en la toma de decisiones vinculadas a diagnósticos, protocolos clínicos que se deben activar ante un diagnóstico, medicación y/o procedimientos a prescribirse a un paciente, también deben verse como un soporte para la prevención de enfermedades [3]. Por esta razón, contar con información útil, inmediata y efectiva es sumamente relevante en éste ámbito.

En este sentido, la minería de datos es una herramienta que permite encontrar patrones de comportamiento de utilidad para la toma de decisiones vinculadas a este último punto.

La minería de datos se relaciona de manera estrecha con la estadística, utilizando técnicas de muestreo y visualización de datos, y depuración y cálculo de indicadores, entre otros. Según Hand et al. [4], “la minería de datos es el análisis de grandes conjuntos de datos observacionales para encontrar relaciones insospechadas, y para resumir los datos en nuevas formas, comprensibles y útiles para el titular de los datos”.

Entre los obstáculos que puede encontrar la aplicación de minería de datos en la medicina, podemos mencionar la voluminosidad y heterogeneidad de los datos médicos, la complejidad de su representación, la posible incompletitud de los mismos, entre otros. Esto fuerza a las instituciones sanitarias a realizar grandes inversiones en tiempo y dinero para poder procesar esta información adecuadamente [5].

En lo que se refiere al tratamiento de la

voluminosidad y heterogeneidad de los datos, el mismo se ha visto mejorado por la aparición de las bases de datos NoSQL. En particular, las del tipo orientadas a columnas (*Column-Oriented Databases*), adecuadas para aplicarlas en minería de datos y aplicaciones analíticas, por su forma de almacenamiento, la compresión eficiente de los datos y por su diseño, que permite cargar y analizar grandes volúmenes de datos [6, 7, 8, 9].

Considerando lo enunciado en [6] que, “...las bases de datos orientadas a columnas son adecuadas para aplicaciones analíticas y de minería de datos, donde el método de almacenamiento es ideal para las operaciones comunes que se realizan en los datos...”, se utilizará *HBase* sobre *Hadoop* para almacenar los datos relevantes que se pretenden obtener de las HCE [10, 11].

El proyecto *Apache Hadoop*, de la *Apache Software Foundation*, es uno de los enfoques existentes para el análisis de datos no estructurados. *Hadoop* es un *framework open source* que permite el procesamiento distribuido de grandes conjuntos de datos a través de *clusters* de computadoras, ofreciendo escalabilidad y confiabilidad [12].

HBase [13], que también forma parte del proyecto *Apache Hadoop*, es un sistema de gestión de bases de datos orientado a columnas que se ejecuta en la parte superior del HDFS [12], que no admite un

lenguaje de consulta como SQL (*Structured Query Language*), y que se utiliza con frecuencia para analizar grandes conjuntos de datos. Es una base de datos distribuida, no relacional, *open-*

source modelada a partir de *Google Big Table*¹, y que soporta scripts escritos en Java [12].

Líneas de Investigación, Desarrollo e Innovación

En la actualidad han cobrado importancia las bases de datos no puramente relacionales, caracterizadas, principalmente, por su almacenamiento distribuido y su fácil escalabilidad. En esta línea, se buscará analizar el comportamiento de una base de datos del tipo NoSQL como lo es *HBase*, al aplicar herramientas de minería de datos, con la finalidad de evaluar su rendimiento ante la necesidad de analizar grandes volúmenes de datos.

Entendiendo que uno de los campos de aplicación fértiles de la minería de datos es el campo de la bioingeniería, nos proponemos integrar el conocimiento que formula la minería de datos a la HCE como apoyo en la toma de decisiones clínicas, implementando un esquema que permita obtener y generar un repositorio con datos relevantes a fin de agilizar la obtención de resultados.

Resultados y Objetivos

Dada la importancia que en la actualidad está presentando el manejo de grandes volúmenes de datos, y la importancia que están adquiriendo las bases de datos NoSQL, se hace necesario integrar la

¹ *Bigtable* es un sistema de almacenamiento distribuido para gestionar datos estructurados, diseñado para escalar a un tamaño muy grande utilizado por *Google* [14].

utilización de este tipo de BD para manejar la variedad y complejidad de los datos médicos al aplicar herramientas de análisis. Para ello se prevé:

- Instalar y configurar *Hadoop* y *HBase*. En una primera etapa, se instalará y configurará *Hadoop* de una forma pseudo-distribuida. Luego se instalará y configurará *HBase* sobre la misma instalación.
- Obtener datos relevantes desde una base de datos relacional que contiene las HCE. Para ello se diseñará y desarrollará una interface en lenguaje *Java* [15] que permita extraer los datos relevantes vinculados al área de interés de análisis y volcarlos a la base de datos generada en *HBase*.
- Evaluar el correcto funcionamiento del proceso desarrollado. Para realizar las pruebas se trabajará con los datos obtenidos y referidos en [9], a fin de probar la correctitud del mismo.
- Aplicar técnicas de minería de datos sobre la base de datos NoSQL generada, con el objetivo de analizar el tema de interés clínico. Se utilizará el software *WEKA*² y los resultados se cotejarán con los resultados obtenidos en [9], para verificar la exactitud de los datos manipulados.

Finalmente, se pretende incorporar este esquema a un componente desarrollado en *Java* que puede prestar servicios e integrarse a una HCE para brindar apoyo en el diagnóstico médico.

Formación de Recursos Humanos

Como parte del actual proyecto de investigación se espera que uno de los docentes investigadores, y que es autor de este artículo, complete su Tesis de Magister en Redes de Datos en la Facultad de Informática de la Universidad Nacional de La Plata. También, se espera que otro de los autores de este proyecto complete su Tesis de Magister en Sistemas de Información en la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos. Finalmente, se buscará formar nuevas sublíneas de investigación relacionadas a este proyecto, así como también la formación de nuevos alumnos en los posgrados dictados en el ámbito de la Universidad Nacional de Entre Ríos y de la Universidad Nacional del Sur.

Referencias

- [1] Balas, E. A., Vernon, M., Magrabi, F., Gordon, L. T. & Sexton, J. (2015). *Big Data Clinical Research: Validity, Ethics, and Regulation*. In **MEDINFO 2015: EHealth-enabled Health: Proceedings of the 15th World Congress on Health and Biomedical Informatics**, Vol. 216. IOS Press, 448.
- [2] Molina, J. & García, J. (2006). *Técnicas de análisis de datos: Aplicaciones prácticas utilizando Microsoft Excel y Weka*. Universidad Carlos III de Madrid España.

² <http://www.cs.waikato.ac.nz/ml/weka/index.html>

- [3] Silva Layes, M. E., Falappa, M. A., & Simari, G. R. (2). *Sistemas de Soporte a las Decisiones Clínicas*. http://sedici.unlp.edu.ar/bitstream/handle/10915/19976/Documento_completo.pdf?sequence=1.
- [4] Hand, D. J. & Mannila, H., Smyth, P. (2001). *Principles of Data Mining*. MIT press ISBN: 026208290x.
- [5] Milovic, B. & Milovic, M. (2012). *Prediction and Decision Making in Health Care using Data Mining*. **International Journal of Public Health Science (IJPHS)**, Vol. 1, N° 2, 69-78 ISSN: 2252-8806.
- [6] Nayak, A., Poriya, A. & Poojary, D. (2013). *Type of NOSQL databases and its comparison with relational databases*. **International Journal of Applied Information Systems**, Vol. 5, N° 4, 16-19.
- [7] Mehta, R. G., Mistry, N. J. & Raghuvanshi, M. (2013). *Impact of Column-oriented Databases on Data Mining Algorithms*. **International Journal of Advanced Research in Computer and Communication Engineering**.
- [8] Carabio, A. L. R., Benedetto, M. G. & Falappa, M. A. (2016). *Comportamiento de Bases de Datos No Relacionales en Entornos Distribuidos*. In **XVIII Workshop de Investigadores en Ciencias de la Computación, WICC'2016**.
- [9] Carabio, A. L. R., Silva Layes, M. E., Frola, F., & Falappa, M. A. (2016). *Bioingeniería aplicada en el diagnóstico de enfermedades*. In **VII Congreso Argentino de Informática en Salud (CAIS 2016)-JAIHO 45**.
- [10] Das, T. K., & Kumar, P. M. (2013). *Big Data Analytics: A framework for unstructured data analysis*. **International Journal of Engineering and Technology (IJET)**, 5(1), 153-156.
- [11] Song, H., Li, L., & Fan, Y (2014). *Applied research on data mining platform for weather forecast based on cloud storage*. **Computer Modelling & New Technologies**, 18(12C) 1226-1230.
- [12] The *Apache Hadoop Project*, <http://hadoop.apache.org/> (2016).
- [13] *Apache HBase Project*, <http://hbase.apache.org/> (2016).
- [14] Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., ... & Gruber, R. E. (2006). *Bigtable: A Distributed Storage System for Structured Data*. To appear in **OSDI**, 1.
- [15] *Java Platform*, Standard Edition (Java SE) 8, <http://docs.oracle.com/javase/8/index.html> (2016).

Geometría Computacional y Bases de Datos

Susana Esquivel, Edilma Olinda Gagliardi, Maria Gisela Dorzán, Maria Teresa Taranilla,
Pablo Rafael Palmero y Carlos Andrés Casanova

Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{esquivel, oli, mgdorzan, tarani, prpalmero}@unsl.edu.ar

Gregorio Hernández Peñalver

Departamento de Matemática Aplicada
Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es

Resumen

Consideramos la línea de investigación denominada *Geometría Computacional y Bases de Datos* del proyecto *Tecnologías Avanzadas de Bases de Datos*, orientada a vincular las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas. El objetivo general consiste en utilizar métodos y herramientas de estas disciplinas para investigación de base en la resolución de problemas NP orientados a optimización, y también para la resolución de problemas en diversos dominios de aplicación.

Palabras clave: Bases de Datos, Geometría Computacional, Metaheurísticas, Bases de Datos Espaciales y Espacio-Temporales.

Contexto

El Proyecto Tecnologías Avanzadas de Bases de Datos desarrolla actividades vinculadas al tratamiento de objetos de diversos tipos, estructurados y no estructurados que son de utilidad en diversos campos de aplicación, tales como sistemas de información geográfica, computación gráfica, computación móvil, robótica, diseño asistido por computadora, motores de búsqueda en internet, entre otras, y que se relacionan en tales bases de datos.

En el proyecto existen tres líneas de investigación, orientadas al desarrollo de nuevos modelos para administrar y recuperar información almacenada en repositorios de datos no estructurados, donde los escenarios de exploración requieren modelos tales como las bases de datos de texto, bases de datos espaciales, espacio-temporales, bases de datos de imágenes, bases de datos de sonidos, espacios métricos, entre otros.

Así, surge el estudio de modelos como las bases de datos espaciales y bases de datos espacio-temporales. También, la necesidad de construir y manipular diferentes objetos y estructuras geométricas útiles en diversas áreas de aplicación.

En particular, las estructuras geométricas que se estudian deben cumplir con propiedades deseables, y algunos de los problemas relacionados con la optimización de las mismas son problemas NP-duros, por tanto en la búsqueda de soluciones aproximadas se aplican metaheurísticas.

Por otro lado, con el avance de las tecnologías, situaciones del mundo real ameritan el uso de las mismas, en pos de una mejora sustancial en cuestiones de calidad, gestión, economía, etc. A nivel mundial, las tecnologías de la información y la comunicación (TIC) continúan introduciendo cambios políticos,

económicos, sociales y culturales, entre muchos otros, convirtiéndose en procesos clave para cualquiera que quiera accionar sobre el presente y proyectar hacia el futuro.

Por lo expuesto, en la línea de investigación *Geometría Computacional y Bases de Datos* se vinculan temáticas que surgen de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, a fin de tratar con la optimización en problemas NP, o bien de poder trabajar en dominios de aplicación con soluciones integradas que hagan uso de técnicas y herramientas de estas disciplinas.

El trabajo de investigación se desarrolla en forma conjunta con investigadores afines de proyectos de esta Universidad, o de alianzas y convenios entre organizaciones nacionales y provinciales con presencia en San Luis, como así también de universidades extranjeras (Universidad Politécnica de Madrid - España, Universidad Veracruzana - México, Universidad del Bío Bío- Chile, entre otras) mediante convenios de cooperación interinstitucional.

En este trabajo, se exponen los tópicos en estudio, junto con las propuestas más recientes de interés.

1. Introducción

La optimización es una línea de investigación en Ciencias de la Computación, donde se procura encontrar la mejor solución posible a un problema dentro de un período de tiempo limitado. En el caso particular de problemas de optimización combinatoria, los hay NP-duros y los hay polinómicos; y no se puede garantizar encontrar la mejor solución en un tiempo razonable, para todas las instancias del problema.

En Geometría Computacional, la optimización de configuraciones geométricas respecto de ciertos criterios de calidad, pertenecen a esta clase de problemas, y pueden resolverse utilizando métodos de aproximación, tales como las técnicas metaheurísticas [BCKO08].

Una metaheurística es un proceso de generación iterativo que guía la búsqueda de soluciones combinando inteligentemente diferentes conceptos de campos diversos como inteligencia artificial, evolución biológica, inteligencia colectiva, sistemas inmunes, entre otros [MF04].

Entre los objetivos de la línea está proponer soluciones aproximadas para problemas geométricos para los cuales aún no se han encontrado algoritmos eficientes que los solucionen debido a su complejidad, mediante la aplicación de técnicas metaheurísticas.

Las propuestas de estudio están orientadas a la optimización de diferentes configuraciones geométricas, tales como triangulaciones y pseudotriangulaciones, poligonizaciones, entre otras. Los criterios de calidad considerados son peso, dilación, número de apuñalamiento mínimo, número de guardias en problemas de vigilancia, área, perímetro, entre otros. Estos criterios inducen a buscar soluciones óptimas respecto de ellos, mediante la aplicación de técnicas metaheurísticas.

Algunos de los problemas de optimización estudiados son la Triangulación de Peso Mínimo (*Minimum Weight Triangulation, MWT*) y la Pseudo-Triangulación de Peso Mínimo (*Minimum Weight Pseudo-Triangulation, MWPT*), problemas de carácter NP-duro [MR06], [GL07]. La Triangulación de Dilación Mínima (*Minimum Dilation Triangulation, MDT*) es otro problema estudiado, donde la dilación mide la calidad de conexión entre puntos de la triangulación. Con respecto a poligonizaciones, se estudió el problema de obtener poligonizaciones con mínima área para un conjunto de puntos en el plano [TGH11].

Otro tópico de interés actualmente en análisis, refiere a problemas de vigilancia. Los problemas de vigilancia pueden interpretarse, muchas veces, como problemas de iluminación o vigilancia. Se han planteado numerosas variantes del problema, cuestionándose ¿qué se vigila? y ¿cómo se vigila? De esta forma, se trata con diversos objetos geométricos a

vigilar y con diversas formas de vigilancia. En particular, nos interesan como objeto geométrico de estudio las triangulaciones planas las cuales constituyen un entorno geométrico adecuado, en el que también tienen sentido las preguntas sobre vigilancia. Un guardia situado en un vértice (ó arista) de la triangulación vigila todos los triángulos incidentes al vértice (ó arista). Por tanto, dada una triangulación T , podemos considerar las siguientes preguntas ¿cuántos guardias, ubicados en vértices (ó aristas), se necesitan para vigilar todos los triángulos de T ?, y ¿dónde se deben ubicar? es decir, cuáles son los vértices (ó aristas) seleccionados para ubicar los guardias.

Además de abordar el estudio de estos problemas mencionados, se busca poder hacer uso de soluciones en dominios de aplicación, naturales para su desarrollo.

La utilización de estas configuraciones geométricas optimizadas respecto de algún criterio de calidad, resultan como soporte de estrategias en la resolución de problemas vinculados con bases de datos espaciales y espacio-temporales. En este contexto, proponemos el estudio y el desarrollo de herramientas para la visualización de aplicaciones vinculadas a las bases de datos mencionadas.

2. Línea de investigación

En la línea de investigación se estudia el diseño y desarrollo de índices espacio- temporales, aplicables a diversos escenarios de movimiento (redes, espacios libres de obstáculos, etc.), considerando la geometría como una disciplina marco en la cual se formalizan aspectos propios de los problemas involucrados. En este contexto, se propone el estudio de optimización de estructuras geométricas que están relacionadas con las bases de datos ya mencionadas y al desarrollo de herramientas para la visualización de estructuras geométricas y aplicaciones vinculadas con bases de datos espacio- temporales.

Como objetivos específicos de estudio en la línea de investigación se enumeran los siguientes:

a) Indexación espacio-temporal sobre objetos en movimiento para diversos escenarios. Desarrollo de las estructuras de almacenamiento, los algoritmos de consulta y la evaluación experimental, mostrando el desempeño de los distintos índices en aplicaciones de diferentes magnitudes respecto de la población de objetos en movimiento. Desarrollo de aplicaciones con herramientas de Geometría Computacional y Bases de Datos Espacio-Temporales.

b) Estudio de configuraciones geométricas generales de puntos en el plano considerando medidas de calidad mínimas o máximas, aplicando técnicas metaheurísticas y diversas estrategias algorítmicas.

c) Estudio de problemas de vigilancia considerando variaciones en el tipo de guardia, su posible ubicación y la forma de vigilancia.

d) Diseño y desarrollo de herramientas para la generación, visualización y manipulación de diferentes configuraciones geométricas de conjuntos de puntos en el plano.

3. Resultados Obtenidos /Esperados

Para la resolución de problemas de optimización de Triangulaciones y Pseudo-triangulaciones de Peso Mínimo se aplicaron las técnicas metaheurísticas: Optimización basada en Colonia de Hormigas (Ant Colony Optimization, ACO) y Recocido Simulado (Simulated Annealing, SA), técnicas determinísticas Voraces (Greedy) y Triangulación de Delaunay. Se llevó a cabo el estudio, adecuación y evaluación experimental de las técnicas metaheurísticas mencionadas para la búsqueda de triangulaciones y pseudo-triangulaciones que cumplan la propiedad. Se diseñaron generadores de instancias de problema para ser utilizados en la evaluación experimental.

Los resultados obtenidos para los problemas MWT y MWPT utilizando la técnica

metaheurística ACO fueron publicados en [DGLH11b] [GDLH11] [DGLH12]. Se llevó a cabo una evaluación experimental y análisis de los resultados obtenidos con la técnica Recocido Simulado [DGLH11a] [DGLH11c].

Para el problema Triangulación de Dilación Mínima (Minimum Dilation Triangulation - MDT) todavía no se conoce un algoritmo que lo resuelva en tiempo polinomial y tampoco se ha demostrado que sea NP-duro. Las técnicas aplicadas para este problema fueron: Greedy, Local Search, Iterated Local Search, Simulated Annealing y Random Local Search. Para cada estrategia se propuso un conjunto de operadores adecuados. Debido a la complejidad que implica la puesta a punto de los parámetros de técnicas metaheurísticas, se utilizó Optimización de Parámetros Secuencial (Sequential Parameter Optimization - SPO) para el ajuste de los parámetros requeridos por Simulated Annealing. Se realizó un análisis experimental en el cual se compararon dichos algoritmos con otras técnicas, como por ejemplo, Delaunay. Se crearon las instancias de prueba, ya que para estos problemas no se encontraron disponibles ningún tipo benchmark con el cual comparar nuestros resultados. Las conclusiones fueron afirmadas desarrollando un estudio estadístico aplicando diferentes test estadísticos y métodos de visualización [DLMH14]. Los resultados del tratamiento de estos problemas han dado lugar a tesis de doctorado.

Por otra parte, se desarrolló una herramienta para la generación y visualización de triangulaciones pseudo-triangulaciones y poligonizaciones de conjuntos de puntos en el plano [PDG13].

Además, se implementó una aplicación en el ámbito de la Salud para el seguimiento de focos epidémicos utilizando base de datos espaciotemporales y herramientas de Geometría Computacional [GPDGT14].

El desarrollo de las herramientas antes mencionadas se plasmó en trabajos finales de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas.

Como trabajo futuro, considerando los problemas sobre configuraciones geométricas, se pretende continuar con el estudio de los problemas de optimización aplicando otras técnicas metaheurísticas, adecuadas para su resolución.

Se comenzará con el estudio de problemas relacionados con otras configuraciones geométricas, tales como cuadrangulaciones, pseudocuadrangulaciones. Se llevarán a cabo, la evaluación de las distintas técnicas para determinar su funcionalidad e impacto en la comunidad científica y el análisis que incluirá el correspondiente tratamiento estadístico, y comparativo con otros algoritmos de tipo aproximado o de tipo exacto.

Por otra parte, con el fin de promover una plataforma de investigación, intercambio y desarrollo de tecnologías de la información y la comunicación (TIC), se integró el Proyecto Campo Conectado. Las organizaciones socias se proponen realizar acciones vinculadas a la promoción de las TIC en el ámbito de la producción agropecuaria con un horizonte inicial a dos años (2017-2018). El objetivo principal es aportar a la cooperación científico tecnológica y a las prácticas sociales, productivas y comerciales de los actores de la producción agropecuaria del semiárido central argentino.

En el marco de Campo Conectado, se propone accionar en la gestación de herramientas y desarrollos tecnológicos aplicados a la gestión de la producción agropecuaria en sistemas reales de producción.

4. Formación de Recursos Humanos

La formación del grupo de trabajo en la Universidad Nacional de San Luis, se consolida con actividades de cooperación mutua e intercambio recíproco de información científica, tecnología y desarrollo de nuevos conocimientos con investigadores locales y de otras universidades.

Entre las actividades más destacadas, se mencionan:

- i) Formación de recursos humanos plasmada en tesis doctorales, tesis de maestría y Licenciados en Ciencias de la Computación.
- ii) Realización y dirección de pasantías de investigación con docentes de otras universidades.
- iii) Actividades de divulgación científica, conferencias y publicaciones en congresos y revistas en el ámbito nacional e internacional.
- iv) Actividades de formación académica, a

través de la realización y dictado de cursos de posgrado y de especialización.

- v) Integración del proyecto Campo Conectado. Como objetivo se propone continuar con las actividades integradoras relacionadas al presente proyecto, proponiendo actividades de formación académica, de formación de recursos humanos locales y de otras universidades nacionales, investigación, desarrollo, y otras actividades académico- científicas vinculantes.

5. Bibliografía

- [BCKO08] de Berg, M., Cheong O., van Kreveld, M., Overmars, M., *Computational Geometry: Algorithms and Applications*. 3rd edition, Springer-Verlag, Heidelberg, 2008.
- [DGLH11a] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approaches for MWT and MWPT Problems*. XVII Congreso Argentino de Ciencias de la Computación 2011 (CACIC 2011), 2011
- [DGLH11b] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Using ACO metaheuristic for MWT problem*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [DGLH11c] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Metaheuristic approaches for MWT and MWPT Problems*. XIV Encuentros de Geometría Computacional. Páginas: 79-82. 2011
- [DGLH12] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approximations on Minimum Weight Triangulations and Minimum Weight Pseudo-Triangulations using Ant Colony Optimization Metaheuristic*. *Fundamenta Informaticae*. ISSN: 0169-2968 (Print), 1875-8681 (Online). Volume 119, number 1, pp 1-27.
- [DLMH14] M. G. Dorzán, M. G. Leguizamón, Efrén Mezura-Montes, G. Hernández Peñalver *Approximated algorithms for the Minimum Dilation Triangulation Problem*. *Journal of Heuristics*. DOI 10.1007/s10732-014-9237-2. Print ISSN 1381-1231. Online ISSN 1572-9397. Publisher Springer US. 2014.
- [GDLH11] Gagliardi E. O., Dorzán M. G., Leguizamón M. G. y Hernández Peñalver G.; *Approximations on Minimum Weight Pseudo-Triangulation problem using Ant Colony Optimization*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [GPDGT14] Guasch, M.M; Pierrgallini, M.R ; Dorzán, M.G.; Gagliardi, E.O.; Taranilla, M.T.; “Una herramienta para el análisis y seguimiento de focos epidémicos” en Anales del 17º Concurso de Trabajos Estudiantiles en 43 JAIIO. Páginas: 35-45- Universidad de Palermo, Buenos Aires, Argentina, 2014.
- [GL07] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. *Computational Geometry. Theory and applications*. Elsevier Vol. 38- pages 139-153, 2007.
- [MR06] Mulzer W., Rote G. *Minimum weight triangulation is NP-hard*. In *Proceedings of the 22nd Annual ACM Symposium on Computational Geometry*. 2006.
- [MF04] Michalewicz Z., Fogel D., *How to Solve It: Modern Heuristics*, Springer, 2004.
- [PDG13] Palmero, P.R., Dorzán, M. G., Gagliardi E.O., *Una Herramienta para la Manipulación de Configuraciones Geométricas*, 42º Jornadas Argentinas de Informática e Investigación Operativa - 16º Concurso de trabajos de fin de carrera. Universidad Nacional de Córdoba, Córdoba, Argentina. 2013. ISSN: 1850-2946.
- [TGH11] Taranilla, M. T.; Gagliardi, E. O.; Hernández Peñalver, G. “Approaching Minimum Area Polygonization”. Congreso Argentino de Ciencias de la Computación 2011 Páginas: 31-40, 2011.

Grandes Datos y Algoritmos Eficientes para Aplicaciones de Escala Web

Gabriel H. Tolosa¹, Santiago Bancho¹, Esteban A. Ríssola¹,
Tomás Delvechio¹, Pablo Lavallén¹ y Esteban Feuerstein²
{tolosoft, sbancho, earissola, tdelvechio, plavallen}@unlu.edu.ar; efeurest@dc.uba.ar

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján

²Departamento de Computación, FCEyN, Universidad de Buenos Aires

Resumen

La cantidad y variedad de información disponible online impone constantes desafíos en cuanto a técnicas eficientes para su almacenamiento y acceso. Muchos procesos en múltiples dominios requieren que este acceso se realice bajo restricciones de tiempo (eficiencia) y con parámetros de alta calidad (eficacia). En este escenario existen por un lado, necesidades puntuales de los servicios que recolectan y utilizan información de la más diversa y compleja naturaleza y por el otro, aparecen oportunidades únicas para avances científico/tecnológicos en áreas como algoritmos, estructuras de datos, sistemas distribuidos y procesamiento de datos a gran escala. Ejemplos concretos son las máquinas de búsqueda para la web, las redes sociales y los sistemas que generan grandes cantidades de datos como la telefonía móvil, entre otros.

Esta problemática abre nuevos interrogantes constantemente y, mientras se intentan resolver, aparecen nuevos desafíos. Algunas de estas preguntas tienen que ver con nuevas estructuras de datos y algoritmos altamente eficientes.

En este proyecto se estudian, proponen, diseñan y evalúan estructuras de datos y algoritmos eficientes junto con el análisis de grandes datos que permitan mejorar las prestaciones de los sistemas, tanto en eficiencia y escalabilidad como en eficacia.

Palabras clave: algoritmos eficientes, motores de búsqueda, estructuras de datos, grandes datos.

Contexto

Esta presentación se encuentra enmarcada en los proyectos de investigación “Algoritmos Eficientes y

Minería Web para Recuperación de Información a Gran Escala” del Departamento de Ciencias Básicas (UNLu) y “Modelos y herramientas algorítmicas avanzadas para redes y datos masivos” del Departamento de Computación de la Facultad de Ciencias Exactas y Naturales (UBA).

Introducción

En las últimas décadas diferentes manifestaciones de los conceptos de redes, datos e información están impactando en la sociedad, tomando diferentes formas y con implicancias aún no conocidas en su totalidad. Aplicaciones en Internet (en particular, en la web), teléfonos con capacidades de cómputo considerables, hardware de bajo costo y las redes sociales [12, 11], entre otras conforman un ecosistema en el cual se desarrollan nuevas actividades que ofrecen, además, nuevos desafíos.

En este espacio, dos fenómenos se complementan y retroalimentan. Por un lado, el incremento exponencial de la cantidad de datos accesibles a través de las distintas redes y, por el otro, el número de usuarios y aplicaciones que acceden a éstos. Por esto, existe una necesidad permanente de nuevas ideas algorítmicas y herramientas computacionales que permitan resolver de forma eficiente los problemas que se plantean. Uno de los ejemplos más notables es la web, que ha experimentado en los últimos años un crecimiento en tamaño y complejidad sin precedentes, convirtiéndola en el mayor repositorio de información en el mundo, creando nuevas necesidades de almacenamiento, procesamiento y búsquedas, expandiendo los límites del trabajo en una sola máquina y unos pocos algoritmos al trabajo distribuido, paralelo y altamente eficiente.

En este escenario existen por un lado, necesida-

des puntuales de los servicios que recolectan y utilizan información de la más diversa y compleja naturaleza y por el otro, aparecen oportunidades únicas para avances científico/tecnológicos en áreas como algoritmos, estructuras de datos, sistemas distribuidos y procesamiento de datos a gran escala.

El enfoque más general para acceder a la información en la web es el uso de motores de búsqueda, a partir de consultas basadas en las necesidades de información de los usuarios. De forma simple, los motores de búsqueda intentan satisfacer la consulta de los usuarios realizando procesos de recuperación sobre una porción del espacio web que “conocen”, es decir, que han recorrido, recopilado y procesado [4]. Este proceso cuenta con dos características fundamentales: operan con estrictas restricciones de tiempo, es decir, las consultas deben ser respondidas en pequeñas fracciones de tiempo (milisegundos) y deben ofrecer resultados relevantes a la consulta de los usuarios sobre un escenario altamente heterogéneo. Además, los usuarios no solo *buscan* en la web para satisfacer sus necesidades de información sino que, además, realizan tareas cotidianas (por ejemplo, organizar un viaje, comprar cosas, etc.). Los motores de búsqueda se han convertido en herramientas indispensables y las cuestiones relacionadas con su eficiencia (escalabilidad) y eficacia son temas de muy activa investigación [8].

Esta proliferación de grandes volúmenes de datos y de usuarios en casi todos los ámbitos de la actividad humana ha creado una gran demanda de nuevas y poderosas herramientas para convertir datos en información útil. Surgieron así diferentes aportes desde el área de *machine learning* como patrones de reconocimiento, análisis estadístico de datos, visualización, agrupamientos, redes neuronales, entre otros. Estos conceptos y técnicas, aplicados al ámbito de la web se los conoce como Minería Web (Web Mining) [5], e incluye el estudio de los datos (minería de contenido), el grafo web (minería de la estructura) y el comportamiento de los usuarios (minería de uso). En algunos ámbitos, algunas de estas aplicaciones son llamadas también análisis de Big Data [35] (Datos Masivos o Grandes Datos) ya que en sus procesos ingestan grandes volúmenes de datos de fuentes diversas [24]. En general, ayudan a resolver problemas que demandan soluciones más complejas y que involucran cómputo paralelo, almacenamiento distribuido y necesitan arquitecturas que puedan escalar de manera flexible [29], tanto en

cómputo como almacenamiento. Como las técnicas para descubrimiento de conocimiento son transversales a cualquier disciplina científica, existe un amplio abanico de soluciones de optimización aún no exploradas para el ámbito de los motores de búsqueda a gran escala que pueden ser tratadas siguiendo la metodología y las técnicas propias de la minería de datos. Incluso, algunas soluciones son significativamente más complejas ya que los volúmenes de información son muy grandes, llegan de manera continua y requieren respuestas en tiempo real [26, 27].

Líneas de investigación y desarrollo

En este proyecto se continúan líneas de I+D del grupo que incorporan análisis de grandes datos en aplicaciones de escala web (como un motor de búsqueda web) que permitan aumentar sus prestaciones. Existen oportunidades de investigación en temas poco explorados por la comunidad científica que permiten mejorar y/o rediseñar los algoritmos internos y las estructuras de datos usadas para recuperación de información de gran escala. En especial, las líneas de I+D principales son:

a. Estructuras de Datos

1. Distribuidas: Los sistemas de búsqueda en texto utilizan como estructura de datos básica un índice invertido, formado por un vocabulario (V) con todos los posibles términos y un conjunto de *posting lists* (L) con información acerca de los documentos donde aparece cada término junto con información usada para el ranking. Como los sistemas de búsqueda a gran escala se ejecutan en clusters de computadoras, es necesario distribuir los documentos entre los nodos, ya sea, por documentos [4], por términos [4] o híbridas (2D [14] y 3D [13]). En todos los casos, el particionado y la asignación de particiones a los nodos de búsqueda impacta en la performance. Resultados experimentales muestran que es posible obtener mejoras si se incorpora la arquitectura del cluster (cantidad de nodos, procesadores y núcleos) en la optimización. Además, los nodos de un motor de búsqueda almacenan su porción del índice en memoria (total o parcialmente), lo que modifica los modelos de costos. Esto ofrece oportunidades para aprovechar de mejor manera el espacio a través del uso eficiente de técnicas de compresión, de reordenamiento y de representación de las listas.

2. Escalables: Para poder mantener la eficiencia conforme se incrementa la cantidad de información que generan algunos servicios (por ejemplo, los sitios de microblogging) son necesarios algoritmos y estructuras de datos escalables. Los aspectos principales a tener en cuenta en este escenario son la tasa de ingestión de documentos, la disponibilidad inmediata del contenido y el predominio del factor temporal [7, 3]. Para satisfacer estas demandas, resulta indispensable mantener el índice invertido en memoria principal. Dado que este es un recurso limitado, se trata de mantener solamente aquella información que permita alcanzar prestaciones de efectividad razonables (o aceptables) [9]. Por ejemplo, el control del crecimiento de las estructuras de datos es un enfoque válido para abordar el problema [27]. Siguiendo esta línea, se propone el desarrollo de una familia de algoritmos de invalidación y poda selectiva [22] de las estructuras de datos a partir del monitoreo online de la evolución y dinámica del vocabulario.

3. Algoritmos Eftcientes: Una de la técnicas más utilizadas para mejorar la performance en motores de búsqueda a gran escala es el *caching*, que se basa en la idea fundamental de almacenar en una memoria de rápido acceso los ítems que van a volver a aparecer en un futuro cercano. Existen múltiples niveles de caching en una máquina búsqueda, por ejemplo: resultados [23], posting lists [37], intersecciones [21] y documentos [31]. Nuestro grupo se enfoca en el problema de las intersecciones para la cual se proponen políticas de admisión y reemplazo que consideren el costo de ejecutar una consulta [15]. Por otro lado, integrar diferentes caches permite optimizar el uso de espacio, lo que impacta positivamente en las prestaciones [32].

Una dirección muy interesante es tratar de optimizar la estrategia de caching incorporando información proveniente de redes sociales. En trabajos previos del grupo se ha mostrado que los temas que son tendencia en redes sociales guardan relación con el aumento de la popularidad de una consulta relacionada al mismo [25] y permiten mejorar la performance del cache. Esta línea de trabajo es prometedora ya que el uso de esta clase de información ha mostrado resultados positivos en otros ámbitos (por ejemplo, para mejorar el rendimiento de CDNs).

b. Grandes Datos en Aplicaciones Web

Los motores de búsqueda son probablemente uno de los primeros ejemplo del uso de Grandes Datos. Las demandas de recolección de documentos, almacenamiento, análisis, gestión y búsqueda requieren de sofisticados algoritmos que operan sobre arquitecturas paralelas y distribuidas. Además, la información generada por las búsquedas de los usuarios (consultas, clics, etc.) se convierte en información muy valiosa a partir de la cual es posible encontrar patrones de comportamiento y obtener estadísticas acerca de cómo los usuarios interactúan con los buscadores. Algunos trabajos [16, 18] ya mostraron la potencialidad de estas técnicas.

Esta propuesta global propone optimizar procesos internos de un buscador por lo que se considera que existen oportunidades de optimización que abren nuevos problemas y temas de investigación.

c. Plataforma de Procesamiento Distribuida para Grandes Datos

El los últimos años han aparecido plataformas para procesamiento distribuido en clusters con interfaces de alto nivel que *facilitan* el procesamiento distribuido con el costo de montar capas de software que ofrecen un nivel de abstracción considerable. Los ejemplos clásicos son el sistema de archivos distribuido HDFS [30] y las plataformas Hadoop [34] y Spark [36]. El grupo investiga cómo utilizarlas eficientemente en los problemas antes mencionados.

En el caso de las máquinas de búsqueda, un requerimiento es la indexación distribuida. Los documentos son procesados de forma distribuida y el resultado final debe ser un índice invertido particionado por algún criterio (como se mencionó anteriormente) que pueda ser implementado en un cluster. En los últimos años, además, se han propuesto nuevas estructuras de datos avanzadas que ofrecen un mejor rendimiento en la recuperación (en algunos contextos), como Block-Max [10] y Treaps [19].

Esta línea de investigación se centra en estudiar, diseñar y evaluar algoritmos de construcción de índices sofisticados como los mencionados utilizando estrategias comúnmente utilizadas en el ámbito de Grandes Datos (por ej., sobre Hadoop) y tratar de determinar cómo influyen algunos parámetros como el tamaño de la colección y la arquitectura del cluster a utilizar.

Otro ejemplo, es el procesamiento de flujos de *streams*, por ejemplo, datos provenientes de redes sociales o imágenes de cámaras en tiempo real, entre otros. En este caso, se estudia e intenta determinar las condiciones para ejecutar el procesamiento mencionado en clusters utilizando la plataforma Spark para el procesamiento de imágenes mediante algoritmos de *machine learning*. La idea es determinar la mejor manera de particionar y distribuir el problema para una arquitectura dada cuyos resultados deben cumplir con restricciones temporales.

d. Algoritmos para Redes Sociales

Las redes sociales online se han convertido sin dudas en una de las aplicaciones más populares de Internet, y han modificado la forma en que los usuarios interactúan e intercambian información. Estas redes atraen a millones de usuarios [6, 17, 20] que, de forma implícita, generan estructuras con propiedades emergentes [1] que surgen del comportamiento global. En general, este tipo de redes tienen a nivel estructural una topología libre de escala (muy sesgada y autosimilar), lo que permite estudiar porciones de la red y extraer propiedades generales. Esto posibilita diseñar algoritmos eficientes para compartir y distribuir la información generada. Esto es especialmente interesante si se tiene en cuenta que la red es un ambiente altamente dinámico y de gran escala. En este caso, si se considera que estos servicios son procesos humanos (y no meramente tecnológicos), su mejor comprensión posibilitará aprovechar la inteligencia colectiva para mejorar servicios como las búsquedas web y aplicar a nuevos escenarios.

e. Comunidades

Otra dirección interesante es en el estudio de algoritmos eficientes para la conformación de comunidades o grupos [33]. Esta es una tarea desafiante a gran escala en aspectos que van desde el tamaño y el tipo de interacción hasta las similitudes por contenido. Si bien existen diversos métodos para analizar y modelar este tipo de redes, la necesidad de algoritmos que combinen información estructural con las propiedades de los nodos es un requerimiento para un amplio espectro de potenciales aplicaciones concretas. Algunos de estos problemas tienen aplicación potencial en proyectos de colaboración abierta, en la salud (grupos de personas con patologías similares) o en el caso de catástrofes [28, 2].

Resultados y objetivos

El objetivo principal del proyecto es estudiar, desarrollar, aplicar, validar y transferir modelos, algoritmos y técnicas que permitan construir herramientas y/o arquitecturas para abordar algunas de las problemáticas relacionadas con las búsquedas a gran escala y el procesamiento de grandes datos. Se pretende estudiar los problemas mencionados relacionados con técnicas de optimización para aplicaciones de búsqueda y proponer mejoras arquitecturales que permitan mejorar la eficiencia de un sistema. Se propone profundizar sobre el estado del arte y definir, analizar y evaluar nuevos enfoques incorporando las técnicas de minería de la web a los procesos internos en aplicaciones de escala web. En particular se estudiarán las siguientes líneas principales:

- Estructuras de datos eficientes, en especial aquellas propuestas recientemente a los efectos de evaluar posibles mejoras orientadas a problemas de datos masivos.
- Técnicas de caching, enfocando el problema no solamente en las políticas de reemplazo, sino también en políticas de admisión, tema que no ha tenido suficiente desarrollo aún. Aquí se propone un enfoque mediante el uso de técnicas de Web Mining para establecer y aprovechar propiedades de las consultas.
- Arquitecturas para aplicaciones específicas, diseñando aplicaciones de búsqueda ad-hoc para problemas concretos, donde una solución de propósito general no es la más eficiente. Aquí se deben estudiar cómo las estructuras de datos y los algoritmos de búsqueda se complementan de mejor manera para aumentar la eficiencia del sistema.
- Algoritmos para el tratamiento de datos provenientes de redes sociales de interés, interactuando con los motores de búsqueda.

Específicamente,

- Diseñar y evaluar nuevas técnicas que optimicen procesos internos de un motor de búsqueda, utilizando información proveniente de procesos de minería web, aumentando la eficiencia del sistema. En especial, políticas de reemplazo/admisión para diferentes niveles de caché.

- Diseñar y evaluar estructuras de datos ad-hoc (centralizadas y/o distribuidas) para problemas concretos y siguiendo el mismo criterio que en el caso previo (mejorar la eficiencia).
- Determinar, mediante procesos de minería web, relaciones entre los objetos del sistema (documento y consultas) y los usuarios externos que permitan establecer mecanismos de resolución de las consultas que aporten mejoras de eficacia (mayor precisión) en la obtención de los resultados.
- Estudiar las potencialidades de las plataformas para procesamiento de datos masivos aplicadas a problemas de búsquedas, principalmente para indexación distribuida y optimizar su rendimiento a partir de utilizar diferentes estrategias y configuraciones.
- Estudiar los flujos de información en redes sociales y su interacción con otros sistemas para el armado automático de comunidades de interés, por ejemplo, grupos de personas con intereses médicos (patologías) afines.

Formación de Recursos Humanos

Este proyecto brinda un marco para que algunos docentes auxiliares y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico. En el mismo, hay en finalización una tesis de la maestría en “Exploración de Datos y Descubrimiento de Conocimiento”, DC, FCEyN, Universidad de Buenos Aires.

Actualmente, se están dirigiendo cuatro trabajos finales correspondientes a la Lic. en Sistemas de Información de la Universidad Nacional de Luján en temas relacionados con el proyecto. Además, hay dos pasantes alumnos y un becario CIN (Beca de Estímulo a las Vocaciones Científicas). Se espera dirigir al menos dos estudiantes más por año y presentar dos candidatos a becas de investigación.

Referencias

- [1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74, 2002.
- [2] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi. Online team formation in social networks. In *Proc. of the 21st International Conference on World Wide Web*, WWW '12, New York, NY, USA, 2012. ACM.
- [3] N. Asadi, J. Lin, and M. Busch. Dynamic memory allocation policies for postings in real-time twitter search. *CoRR*, abs/1302.5302, 2013.
- [4] R. A. Baeza-Yates and B. A. Ribeiro-Neto. *Modern Information Retrieval - The concepts and technology behind search*, 2nd ed. Pearson Education Ltd., 2011.
- [5] L. Bing. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Springer, Secaucus, NJ, USA, 2008.
- [6] D. M. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 2007.
- [7] M. Busch, K. Gade, B. Larson, P. Lok, S. Luckenbill, and J. Lin. Earlybird: Real-time search at twitter. In *Proc. of the 28th International Conference on Data Engineering*, ICDE '12. IEEE Computer Society, 2012.
- [8] B. B. Cambazoglu and R. A. Baeza-Yates. Scalability and efficiency challenges in large-scale web search engines. In *Proc. of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM, 2015.
- [9] C. Chen, F. Li, B. C. Ooi, and S. Wu. Ti: An efficient indexing mechanism for real-time search on tweets. In *Proc. of the 2011 ACM SIGMOD International Conference on Management of Data*. ACM, 2011.
- [10] S. Ding and T. Suel. Faster top-k document retrieval using block-max indexes. In *Proc. of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '11. ACM, 2011.
- [11] P. A. Dreyer Jr. and F. S. Roberts. Irreversible k -threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Applied Mathematics*, 2009.
- [12] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.
- [13] E. Feuerstein, V. G. Costa, M. Marín, G. Tolosa, and R. A. Baeza-Yates. 3d inverted index with cache sharing for web search engines. In *18th International Conference, Euro-Par 2012, August 27-31, 2012.*, 2012.
- [14] E. Feuerstein, M. Marín, M. J. Mizrahi, V. G. Costa, and R. A. Baeza-Yates. Two-dimensional distributed inverted files. In *16th International Symposium of String Processing and Information Retrieval, SPIRE'09, August 25-27, 2009.*

- [15] E. Feuerstein and G. Tolosa. Cost-aware inter-section caching and processing strategies for in-memory inverted indexes. In *In Proc. of 11th Workshop on Large-scale and Distributed Systems for Information Retrieval*, LSDS-IR'14, 2014.
- [16] Y. Hu, Y. Qian, H. Li, D. Jiang, J. Pei, and Q. Zheng. Mining query subtopics from search log data. In *Proc. of the 35th international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2012.
- [17] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: Understanding microblogging usage and communities. In *Proc. of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, New York, NY, USA, 2007. ACM.
- [18] P. Kaushik, S. Gaur, and M. Singh. Use of query logs for providing cache support to the search engine. In *International Conference on Computing for Sustainable Global Development (INDIACom)*, 2014.
- [19] R. Konow, G. Navarro, C. L. Clarke, and A. López-Ortiz. Faster and smaller inverted indices with treaps. In *Proc. of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13. ACM, 2013.
- [20] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proc. of the 19th International Conference on World Wide Web*, WWW '10, 2010.
- [21] X. Long and T. Suel. Three-level caching for efficient query processing in large web search engines. In *Proc. of the 14th international conference on World Wide Web*. ACM, 2005.
- [22] A. Ntoulas and J. Cho. Pruning policies for two-tiered inverted index with correctness guarantee. In *Proc. of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007.
- [23] R. Ozcan, I. S. Altingovde, and O. Ulusoy. Cost-aware strategies for query result caching in web search engines. *ACM Trans. Web*, 5(2), May 2011.
- [24] A. Rajaraman and J. D. Ullman. *Mining of Massive Datasets*. Cambridge University Press, New York, NY, USA, 2011.
- [25] S. Ricci and G. Tolosa. Efecto de los trending topics en el volumen de consultas a motores de búsqueda. In *XVII Congreso Argentino de Ciencias de la Computación, CACIC.*, 2013.
- [26] E. Rissola and G. Tolosa. Inverted index entry invalidation strategy for real time search. In *Proc. of the XXI Congreso Argentino en Ciencias de la Computación, CACIC '15*, 2015.
- [27] E. Rissola and G. Tolosa. Improving real time search performance using inverted index entries invalidation strategies. *Journal of Computer Science & Technology*, 16(1), 2016. ISSN: 1666-6038.
- [28] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proc. of the 19th International Conference on World Wide Web*, WWW '10, New York, NY, USA, 2010. ACM.
- [29] E. Schadt et al. Computational solutions to large-scale data management and analysis. *Nature reviews Genetics*, 11(9), 2010.
- [30] K. Shvachko, H. Kuang, S. Radia, and R. Chander. The Hadoop distributed file system. In *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies, MSST2010*, 2010.
- [31] T. Strohman and W. B. Croft. Efficient document retrieval in main memory. In *SIGIR 2007: Proc. of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007.
- [32] G. Tolosa, L. Becchetti, E. Feuerstein, and A. Marchetti-Spaccamela. Performance improvements for search systems using an integrated cache of lists+intersections. In *Proc. of 21st International Symposium of String Processing and Information Retrieval, SPIRE'14*, 2014.
- [33] M. Wang, C. Wang, J. X. Yu, and J. Zhang. Community detection in social networks: An in-depth benchmarking study with a procedure-oriented framework. *Proc. VLDB Endow.*, 8(10), 2015.
- [34] T. White. *Hadoop: The Definitive Guide*. O'Reilly Media, Inc., 1st edition, 2009.
- [35] W. X. Z. Xingquan, W. Gong-Qing, and D. Wei. Data mining with big data. *Knowledge and Data Engineering, IEEE Transactions on*, 26(1), 2014.
- [36] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica. Spark : Cluster Computing with Working Sets. *HotCloud'10 Proc. of the 2nd USENIX conference on Hot topics in cloud computing*, page 10, 2010.
- [37] J. Zhang, X. Long, and T. Suel. Performance of compressed inverted list caching in search engines. In *Proc. of the 17th international conference on World Wide Web*, WWW '08. ACM, 2008.

Indexación y Búsqueda sobre Datos no Estructurados

Norma Herrera, Darío Ruano, Paola Azar, Susana Esquivel

Departamento de Informática

Universidad Nacional de San Luis, Argentina

{nherrera, dmruano, epazar, esquivel}/@unsl.edu.ar

Anabella De Battista, Andrés Pascal

Departamento Ingeniería en Sistemas de Información

FRCU, Universidad Tecnológica Nacional

Entre Ríos, Argentina

{anadebattista, andrespascal22}/@gmail.com

Abstract

Las bases de datos actuales han incluido la capacidad de almacenar datos no estructurados tales como imágenes, sonido, texto, video, etc. La problemática de almacenamiento y búsqueda en estos tipos de base de datos difiere de las bases de datos clásicas, dado que no es posible organizarlos en registros y campos, y aun cuando pudiera hacerse, la búsqueda exacta carece de interés. Es en este contexto donde surgen nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de índices eficientes para estas nuevas bases de datos.

1 Contexto

El presente trabajo se desarrolla en el ámbito de la línea Técnicas de Indexación para Datos no Estructurados del Proyecto Tecnologías Avanzadas de Bases de Datos (22/F414), cuyo objetivo es realizar investigación básica en problemas relacionados al manejo y recu-

peración eficiente de información no tradicional.

2 Introducción

La mayoría de los administradores de bases de datos actuales están basados en el modelo relacional, presentado por Edgard F. Codd en 1970. Bajo el modelo relacional, cada elemento de la base de datos puede ser almacenado como un registro (tupla) y cada registro a su vez dividido en campos (atributos). La mayoría de las consultas que se realizan a una base de datos relacional (conocidas también como bases de datos tradicionales) se corresponden con *búsquedas exactas*, esto significa obtener todos los registros cuyos campos coinciden exactamente con los campos aportados durante la búsqueda. También se pueden realizar búsquedas por rango sobre valores numéricos, y búsquedas de sub-cadenas sobre campos alfabéticos; en estos casos debe existir una relación de orden sobre los campos consultados.

La información disponible en formato digital aumenta día a día su tamaño de manera ex-

ponencial. Gran parte de esta información involucra el uso de datos no estructurados tales como imágenes, sonido, texto, video, etc. Debido a que no es posible organizar estos tipos de datos en registros y campos, las tecnologías tradicionales de bases de datos para almace-

namiento y búsqueda de información no son adecuadas en este ámbito.

Es en este contexto donde surgen nuevos modelos de bases de datos capaces de cubrir las necesidades de almacenamiento y búsqueda de estas aplicaciones. Nuestro interés se basa en el diseño de índices para estas nuevas bases de datos, centrándonos en bases de datos textuales y en espacios métricos.

Bases de Datos Textuales (BDT) Una base de datos de texto es un sistema que mantiene una colección grande de texto, y provee acceso rápido y seguro al mismo. Sin pérdida de generalidad, asumiremos que la base de datos de texto es un único texto T posiblemente almacenado en varios archivos. Las búsquedas en la que el usuario ingresa un *patrón de búsqueda* y el sistema retorna todas las posiciones del texto donde el patrón ocurre, es una de las búsquedas más comunes en este tipo de bases de datos.

Mientras que en bases de datos tradicionales los índices ocupan menos espacio que el conjunto de datos indexado, en las bases de datos de texto el índice ocupa más espacio que el texto, pudiendo necesitar de 4 a 20 veces el tamaño del mismo [8, 14]. Una alternativa para reducir el espacio ocupado por el índice es buscar una representación compacta del mismo, manteniendo las facilidades de navegación sobre la estructura. Pero en grandes colecciones de texto, el índice aún comprimido suele ser demasiado grande como para residir en memoria principal [9, 10]. Por esta razón, el estudio de índices comprimidos y en memoria secundaria para búsquedas en texto es un tema de creciente interés en la comunidad de bases de

datos.

Espacios Métricos El modelo de espacios métricos permite formalizar el concepto de búsqueda por similitud en bases de datos no tradicionales [4].

Un espacio métrico está formado por un conjunto de objetos X y una función de distancia d definida entre ellos que mide cuan diferentes son. La base de datos será un subconjunto finito $L \subseteq X$.

Una de las consultas más comunes en este modelo de bases de datos es la *búsqueda por rango*. En esta búsqueda dado un elemento $q \in X$, al que llamaremos *query* y un radio de tolerancia r , la búsqueda por rango consiste en recuperar los objetos de la base de datos cuya distancia a q no sea mayor que r . Para evitar examinar exhaustivamente la base de datos, se preprocesa la misma por medio de un *algoritmo de indexación* con el objetivo de construir una *índice*, diseñado para ahorrar cálculos en el momento de la búsqueda. En [4] se presenta un desarrollo unificador de las soluciones existentes en la temática.

Bases de datos temporales

En las bases de datos temporales [19, 12] los datos asociados al tiempo forman parte de la relevancia de sus registros. El modelo temporal permite almacenar y recuperar datos que dependen del tiempo. Mientras que las bases de datos tradicionales tratan al tiempo como otro tipo de dato más, este modelo incorpora al tiempo como una dimensión. Se distinguen dos tipos de tiempo, el *tiempo válido*, y el *tiempo transaccional*. El tiempo válido es el período en el cual un hecho existe y el tiempo transaccional es el periodo en el cual el hecho es registrado en la base de datos. Estos tiempos no necesariamente tiene que coincidir, por ejemplo, algunos determinados sucesos del siglo XX pueden haberse ingresado a una base de datos durante el siglo XXI.

Bases de datos métrico-temporales (BDMT)

Este modelo permite almacenar objetos no estructurados con tiempos de vigencia asociados y realizar consultas por similitud y por tiempo en forma simultánea. Formalmente un *Espacio Métrico-Temporal* es un par (U, d) , donde $U = O \times N \times N$, y la función d es de la forma $d: O \times O \rightarrow R^+$. Cada elemento $u \in U$ es una tripla (obj, t_i, t_f) , donde obj es un objeto (por ejemplo, una imagen, sonido, cadena, etc) y $[t_i, t_f]$ es el intervalo de vigencia de obj . La función de distancia d , que mide la similitud entre dos objetos, cumple con las propiedades de una métrica (positividad, simetría y desigualdad triangular).

Un nuevo tipo de consulta son las denominadas métrico-temporales que se definen formalmente en símbolos como:

$$(q, r, t_{iq}, t_{fq})_d = \{o / (o, t_{io}, t_{fo}) \in X \wedge d(q, o) \leq r \wedge (t_{io} \leq t_{fq}) \wedge (t_{iq} \leq t_{fo})\}$$

La consulta implica buscar todos los objetos

o de la parte finita X del universo U que estén a una distancia a lo más r de q , y que su tiempo asociado t coincida (ose solape) con el tiempo de la consulta.

Varios índices métrico-temporales se han propuesto en este ámbito, todos estos índices fueron desarrollados para ser eficientes en memoria principal.

3 Líneas de Investigación

3.1 Índices Comprimidos en Memoria Secundaria para BDT

Como ya mencionamos, el principal problema que surge al indexar una bases de datos de texto es el espacio ocupado por el índice.

Una forma de tratar con este problema es buscar una representación compacta del índice, manteniendo las facilidades de navegación sobre la estructura. Esto significa encontrar una representación que ocupe menos espacio que la representación clásica, pero que

permita navegar sobre el índice sin necesidad de descomprimirlo [6, 7, 9, 10, 13, 15, 16, 18]. Un *trie de sufijos* es un índice que permite resolver eficientemente las operaciones de

búsquedas en texto pero que necesita en espacio 10 veces el tamaño del texto indexado. En [17] se presenta una nueva representación de un trie de sufijos que permite reducir el espacio necesario para almacenar el índice, eliminando la necesidad de mantener los punteros explícitos a los hijos. Esta representación surge como una extensión a árboles r-arios de la técnica presentada en [11] y tiene la ventaja de permitir un posterior proceso de paginado para manejar eficientemente el trie de sufijos en memoria secundaria [20].

Hemos realizado una implementación que mejora en espacio a la anterior en un 40%, sin afectar los tiempos de búsqueda. Esta nueva versión compacta del trie de sufijos consiste en usar códigos *DAC* (*Directly Addressable Variable-Length Code* [3]), para los arreglos que representan la secuencia de saltos y de grados. La navegación sobre esta nueva representación sigue los lineamientos generales propuestos en [17], adaptándolo a los códigos *DAC*.

Estamos trabajando en integrar esta nueva representación con la técnica de paginado propuesta en [17], a fin de lograr un índice comprimido en memoria secundaria. Nos encontramos en la etapa de implementación de esta nueva propuesta.

3.2 Índices en Memoria Secundaria para BDMT

Varios índices métrico-temporales se han propuesto en este ámbito, todos estos índices fueron desarrollados para ser eficientes en memoria principal; dos de ellos son el *H-FHQT* [5] y el *NewH-FHQT* [2].

El *H-FHQT* consiste en una lista de los instantes válidos de tiempo, donde cada celda de la lista contiene un índice *FHQT* [1] con el que indexa todos los objetos vigentes en dicho instante. Esta estructura es eficiente en bases de datos métrico-temporales donde los objetos tienen vigencia en un sólo instante de tiempo. El *NewH-FHQT* está basado también en el uso del *FHQT* como estructura métrica y el enfoque temporal se ha abordado mediante el uso de una línea de tiempo, del mismo modo que en el *H-FHQT*.

Este índice consiste en una lista compuesta por los instantes válidos de tiempo. Para cada instante de la lista que posee objetos vigentes, se construye un FHQT para indexar los objetos. La principal diferencia con el índice antes propuesto se da en la etapa de construcción de los FHQTs. En este caso se van tomando los primeros pivotes disponibles de la lista, que se considera una lista circular, de tal manera que el FHQT del instante i , este construido con pivotes diferentes a los de los instantes $i - 1$ e $i + 1$. La construcción de FHQTs consecutivos con diferentes grupos de pivotes da a la estructura mayor poder de filtrado de elementos desde el punto de vista métrico. Esta idea se plantea debido a que los objetos a indexar tienen un intervalo de vigencia asociado, por lo que pueden estar presentes en varios FHQTs consecutivos. Con este enfoque se logra que los objetos pasen por varios filtros, se eliminen la mayor cantidad de objetos mediante la firma y la desigualdad triangular y se reduzcan así la cantidad necesaria de evaluaciones de la función de distancia al momento de la ejecución de la consulta.

Estos índices se desarrollaron bajo el supuesto de que la memoria principal tiene capacidad suficiente como para mantener tanto el índice como la base de datos. En este contexto nuestro objetivo es adecuarlos para que los mismos también resulten eficientes en memoria secundaria. Cabe señalar que no existe hasta el momento ningún índice en memoria secundaria para este tipo de base de datos.

4 Resultados Esperados

Se espera obtener índices eficientes, tanto en espacio como en tiempo, para el procesamiento de consultas en bases de datos textuales y en espacios métricos. Los mismos serán evaluados tanto analíticamente como empíricamente.

5 Recursos Humanos

El trabajo desarrollado en esta línea forma parte del desarrollo de un Trabajo Final de la Licenciatura, dos Tesis de Maestría y una Tesis de Doctorado, todas ellas en el área temática de Ciencias de la Computación, en la Universidad Nacional de San Luis.

References

- [1] R. Baeza-Yates, W. Cunto, U. Manber, and S. Wu. Proximity matching using fixed-queries trees. In *Proc. 5th Combinatorial Pattern Matching (CPM'94)*, LNCS 807, pages 198–212, 1994.
- [2] A. De Battista, A. Pascal, N. Herrera, and G. Gutierrez. Metric-temporal access methods. *Journal of Computer Science & Technology*, 10(2):54–60, 2010.
- [3] Nieves R. Brisaboa, Susana Ladra, and Gonzalo Navarro. Directly addressable variable-length codes. In *SPIRE*, pages 122–130, 2009.
- [4] E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.
- [5] A. De Battista, A. Pascal, G. Gutierrez, and N. Herrera. Un nuevo índice métrico-temporal: el historical-fhqt. In *Actas del XIII Congreso Argentino de Ciencias de la Computación*, Corrientes, Argentina, 2007.

- [6] P. Ferragina and G. Manzini. Indexing compressed text. *J. ACM*, 52(4):552–581, 2005.
- [7] P. Ferragina, G. Manzini, V. Mäkinen, and G. Navarro. Compressed representations of sequences and full-text indexes. *ACM Trans. Algorithms*, 3(2):20, 2007.
- [8] G. H. Gonnet, R. Baeza-Yates, and T. Snider. *New indices for text: PAT trees and PAT arrays*, pages 66–82. Prentice Hall, New Jersey, 1992.
- [9] R. González and G. Navarro. A compressed text index on secondary memory. In *Proc. 18th International Workshop on Combinatorial Algorithms (IWOCA)*, pages 80–91. College Publications, UK, 2007.
- [10] R. González and G. Navarro. Compressed text indexes with fast locate. In *Proc. 18th Annual Symposium on Combinatorial Pattern Matching (CPM)*, LNCS 4580, pages 216–227, 2007.
- [11] N. Herrera and G. Navarro. Árboles de sufijos comprimidos en memoria secundaria. In *Proc. XXXV Latin American Conference on Informatics (CLEI)*, Pelotas, Brazil, 2009.
- [12] C. S. Jensen. A consensus glossary of temporal database concepts. *ACM SIGMOD Record*, 23(1):52–54, 1994.
- [13] V. Mäkinen and G. Navarro. *Compressed Text Indexing*, pages 176–178. Springer, 2008.
- [14] U. Manber and G. Myers. Suffix arrays: A new method for on-line string searches. *SIAM Journal of Computing*, 22(5):935–948, 1993.
- [15] G. Navarro. Indexing text using the ziv-lempel trie. *Journal of Discrete Algorithms (JDA)*, 2(1):87–114, 2004.
- [16] G. Navarro and K. Sadakane. *Compressed Tree Representations*. Springer, 2nd edition, 2015.
- [17] D. Ruano and N. Herrera. Representación secuencial de un trie de sufijos. In *XX Congreso Argentino de Ciencias de la Computación*, Buenos Aires, Argentina, 2014.
- [18] K. Sadakane. New text indexing functionalities of the compressed suffix arrays. *J. Algorithms*, 48(2):294–313, 2003.
- [19] B. Salzberg and V. J. Tsotras. A comparison of access methods for temporal data. *ACM Computing Surveys*, 31(2), 1999.
- [20] J. Vitter. External memory algorithms and data structures: Dealing with massive data. *ACM Computing Surveys*, 33(2):209–271, 2001.

La Gestión de la Información Científica en Abierto: ¿Estamos Maximizando la Visibilidad Web?

Jose Texier¹, Fernando Emmanuel Frati¹, Fernanda Carmona¹, Alberto Riba¹, Emiliano Peressini¹, Jusmeidy Zambrano¹, Antonio Castro Lechtaler²

¹Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina

²CISTIC/FCE - Universidad de Buenos Aires

{jtexier, fefrati, fbcarmona, ariba}@undec.edu.ar, emiliano.peressini@gmail.com,
jzambrano@undec.edu.ar, acastro@est.iue.edu.ar

RESUMEN

La "Era de la Información" se caracteriza por reconocer a la información como un recurso valioso en todos los ámbitos. En este sentido, la Universidad Nacional de Chilecito (UNdeC) se propone posicionar su producción científica a través de una estrategia de gestión de la información de forma eficaz y eficiente, maximizando su visibilidad y garantizando la generación de productos científicos de calidad. Con esta línea de I/D/I se pretende concienciar a la comunidad universitaria respecto de algunos términos de uso frecuente en las áreas de visibilidad y gestión del conocimiento como lo son: los Repositorios Institucionales y el movimiento del Acceso Abierto.

Palabras clave: producción científica, repositorios institucionales, acceso abierto, recursos, visibilidad web.

CONTEXTO

La línea de I/D/I corresponde al diseño, desarrollo e implementación de proyectos que fortalecerán la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación, de la mayoría, de la producción académica y científica de la UNdeC.

Actualmente se encuentra activo un PDTS (Proyectos de Desarrollo Tecnológico y Social) aprobado el año pasado y titulado "Repositorios digitales con contenidos orientados a las necesidades de escuelas rurales primarias y secundarias". Su ejecución comenzó en abril del 2016 y finaliza en marzo del 2018. También se encuentra activo el proyecto "Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales" correspondiente a la convocatoria "Redes Internacionales 9", promovida por la Secretaría de Políticas Universitarias de la Nación. De igual manera, se tiene un proyecto del 2013-2014 del programa "Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica" de la Secretaría de Ciencia y Tecnología en la UNdeC (FiCyT – UndeC). El título es "Fortalecimiento y visibilidad web de la producción científica de la UNdeC".

En cuanto a la participación del grupo de trabajo en consorcios internacionales, se están desarrollando tres (3) propuestas para vincular la UNdeC con organizaciones internacionales: International Federation of Library Associations and Institutions (IFLA), Latin American

and Caribbean Consortium of Engineering Institution (LACCEI) e Iberoamerican Science and Technology Education Consortium (ISTEC-BIREDIAL).

Por otra parte, se destaca que la UNdeC cuenta con la estructura tecnológica y de RRHH necesarios para ejecutar los proyectos que surjan de la línea de investigación.

1. INTRODUCCIÓN

Gestión/organización del conocimiento es un concepto que ha ganado cierta difusión. Barité y Dahlberg [1], [2] señalan que cuando se habla de *organización del conocimiento* "es importante tener presente que (...) se hace referencia (...) [al] conocimiento socializado, compartido o comunicado, que además ha sido registrado".

Por tanto, la información es la base para acceder al conocimiento. Se entiende al conocimiento como "una entidad abstracta que existe solamente en la mente de un ser humano en tanto sujeto cognoscente (es decir, es lo que yo sé), mientras que la información es el conocimiento comunicado, compartido o socializado" [3]. Esta diferenciación entre ambos conceptos (conocimiento e información) nos permiten comprender por qué es necesario que la información esté organizada para que el acceso y la apropiación [14] por parte de los sujetos les haga partícipes de la sociedad y los empodere como ciudadanos informados. Lo que en palabras de Barité y Dahlberg apunta a la socialización del conocimiento.

Sin embargo, esta situación se hace posible si esa información puede (o debe) registrarse en un soporte físico para convertirse en un documento o recurso bibliográfico con genuina utilidad por

quien lo desee, por lo que los Repositorios Institucionales (RI) bajo la filosofía del Acceso Abierto pueden convertirse en una de las vías más idóneas para lograr la visibilidad web que requiere la Institución.

Se entiende que los **Repositorios Institucionales (RI)** están constituidos por un conjunto de archivos digitales en representación de productos científicos y académicos que pueden ser accedidos por los usuarios [4]. En otras palabras, los RI se entienden como estructuras web interoperables que alojan recursos científicos, académicos y administrativos, tanto físicos como digitales, descritos por medio de un conjunto de datos específicos (metadatos) [5]–[7]. Tienen como propósito recopilar, catalogar, gestionar, acceder, difundir y preservar la información [4], [8]. Vale la pena destacar que los RI son vías de comunicación científica, pero no son canales de publicación. Eso quiere decir que se deben seguir los mismos mecanismos de validación científica existentes hasta ahora a través de las revisiones por pares, pero los autores deben hacer énfasis en mostrar sus publicaciones y datos primarios de sus investigaciones. Se destaca que los RI deben disponer también de Objetos de Aprendizaje (OA), que son entidades generalmente entregadas a través de Internet y diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos [12, 13].

Los RI se configuran dentro de la filosofía del **Acceso Abierto** (en inglés *Open Access - OA*). Esta filosofía tiene como fin asegurar el acceso libre y abierto a la producción científica, es decir, garantizar el acceso a través de Internet sin que los derechos del copyright sean una barrera [9], [10]. Los RI materializan el objetivo del OA porque

la información que se deposita es una producción que tienen como propósito ser accesible, sin restricciones y preservada digitalmente.

Sobre la base de lo expuesto, esta línea de investigación desarrollará proyectos que fortalezcan la difusión y gestión de la comunicación científica y académica de los becarios, docentes e investigadores de cualquier disciplina de la UNdeC. Esto se logrará mediante la socialización y sistematización de un conjunto de conocimientos prácticos que permitan alcanzar una visibilidad web que ayude a mejorar el posicionamiento de la UNdeC en el país y en el mundo. Todo esto bajo el contexto de la filosofía del Acceso Abierto.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- ☐ Repositorios institucionales y bibliotecas digitales.
- ☐ Objetos de aprendizaje.
- ☐ Gestión de la información y el conocimiento.
- ☐ Sistemas de información web y bases de datos.
- ☐ Índices bibliométricos.
- ☐ Interoperabilidad.
- ☐ Preservación digital.
- ☐ Recuperación de la información.
- ☐ Análisis semántico de la información científica.

3. OBJETIVOS

El desarrollo de esta línea en la UNdeC esta permitiendo cumplir con los siguientes objetivos:

- ☐ Difundir el concepto de conocimiento y la distinción de dato e información.
- ☐ Comprender los problemas y las tendencias asociadas a la

producción abierta del conocimiento, el uso y la difusión.

- ☐ Analizar la visibilidad web de la UNdeC.
- ☐ Conocer la mayoría de la producción científica de la UNdeC.
- ☐ Analizar los recursos educativos existentes que cumplan con las necesidades educativas de la UNdeC y la ciudad de Chilecito.
- ☐ Fomentar la formación de habilidades y conocimientos relacionados con los RI.
- ☐ Desarrollar un mapa conceptual de la producción científica de la UNdeC.
- ☐ Entender y valorar el movimiento mundial de Acceso Abierto y señalar las vías de su materialización.
- ☐ Desarrollar las estrategias necesarias para optimizar la visibilidad científica de la UNdeC.
- ☐ Implementar Repositorios de prueba para visualizar los diferentes recursos (educativos, académicos y científicos) que se produzcan en Chilecito.

4. RESULTADOS OBTENIDOS/ESPERADOS

- ☐ Conocimiento de la producción científica, lugar(es) donde está alojada y qué personal la realizó. En julio del 2016 se realizó un relevamiento de los productos científicos con filiación UNdeC, de esta manera, se obtuvo una ponderación del posicionamiento web de la UNdeC.
- ☐ Relevamiento de los diferentes recursos educativos producidos en la UNdeC y en la ciudad de Chilecito, sobre los linemaientos

del proyecto PDTS, anteriormente nombrando.

- ☐ Elaboración de un mapa conceptual de la producción científica.
- ☐ Curso a nivel de postgrado llamado “Gestión de la Información Académica y Científica”.
- ☐ Diseño de un conjunto de estrategias para optimizar la producción científica de la UNdeC.
- ☐ Se desarrolló un proyecto de asignaturas sobre las Bibliotecas Digitales para presentarse en dos universidades [15, 16].
- ☐ Se elaboró un software que permitió la extracción de metadatos de artículos de diferentes fuentes para normalizarla y visualizarla [17, 18].
- ☐ Se implementó un repositorio de prueba que esta gestionando los recursos educativos y objetos de aprendizaje generados dentro del proyecto PDTS.
- ☐ Se están elaborando dos artículos, uno sobre la producción científica de la UNdeC y otro sobre un analizador semántico de los artículos científicos.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por cuatro docentes de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC (acreditadas por CONEAU), dos doctores especializados en repositorios institucionales, bibliotecas digitales, desarrollo de software, cómputo paralelo y tecnología grid. Otra docente que está finalizando su doctorado y dos definiendo su tesis de Maestría en Informática. También participa un alumno

avanzado de grado. En otras palabras, se cuenta con un recurso humano con habilidades y formación académica en las diversas áreas de la propuesta, asegurando la concreción de la línea, por ejemplo, uno es especialista en RI y OA, con investigación en esa área y otra esta en formación (maestría) sobre el área de los Objetos de Aprendizaje. Adicionalmente, se destaca que dos están categorizados en el programa de incentivos.

Los integrantes son docentes de las asignaturas Programación I, Sistemas I, Arquitecturas Paralelas, Teoría de la Computación y Herramientas de Ingeniería de Software. Estas asignaturas contemplan la aprobación mediante la participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

6. BIBLIOGRAFÍA

- [1] M. Barité, *Diccionario de organización y representación del conocimiento*. Montevideo, 2000.
- [2] I. Dahlberg, “Knowledge organization: its scope and possibilities,” *Knowledge organization*, vol. 20, no. 4.
- [3] A. M. Martínez-Tamayo and J. C. Valdez, *Indicación y Clasificación en Bibliotecas*, Primera edición. Argentina: Alfagrama, 2009.
- [4] J. Texier, “Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior,” presented at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, Mexico, 2013, p. 9.
- [5] C. A. Lynch, “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age,” *ARL*., 2003. [Online]. Available: <http://www.arl.org/resources/pubs/br/>

- br226/br226ir.shtml. [Accessed: 28-Oct-2013].
- [6] J. Tramullas and P. Garrido, "Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones," *El Profesional de la Información*, vol. 15, no. 3, pp. 171–181, 2006.
- [7] H. Van de Sompel, S. Payette, J. Erickson, C. Lagoze, and S. Warner, "Rethinking Scholarly Communication," *D-Lib Magazine*, vol. 10, no. 9, Sep. 2004.
- [8] M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martinez, and A. Pinto, "SEDICI – Desafíos y experiencias en la vida de un repositorio digital," *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [9] P. Suber, "Ensuring open access for publicly funded research," *BMJ*, vol. 345, 2012.
- [10] D. Torres-Salinas, N. Robinson-García, and A. Cabezas-Clavijo, "Compartir los datos de investigación en ciencia: introducción al data sharing," *Profesional de la Información*, vol. 21, no. 2, pp. 173–184, 2012.
- [11] P. Suber, "Open access, impact, and demand," *BMJ*, vol. 330, no. 7500, pp. 1097–1098, May 2005.
- [12] D. Wiley, "Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy," presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE, 2001.
- [13] E. Morales, F. Garcia, A. Barron, A. Berlanga, and C. Lopez, "Propuesta de Evaluación de Objetos de Aprendizaje," presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE), Barcelona, 2005.
- [14] Kalman J. (2001). Saber lo que es la Letra. Informe. México: DIE-CINVESTAV-IPN.
- [15] Texier, J., Zambrano, J., & Carmona, F. B. (2016). Las Bibliotecas Digitales en el Currículum de las carreras de Ciencias de la Computación: una propuesta posible para Argentina y Venezuela. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [16] Texier, J. (2016). Los repositorios institucionales y su importancia en las Universidades Venezolanas. *Revista Scitus*. Venezuela.
- [17] Villarreal, G. L., Terruzzi, F. A., De Giusti, M. R., Lira, A. J., & Texier, J. D. (2016). Fostering the institutional repository through policies and interoperability with online services: the case of Universidad Nacional de La Plata. *Scholarly and Research Communication*, 6(1).
- [18] Texier, J., Zambrano, J., & Riba, A. (2016). Normalización de los LACCEI Proceedings a través de un proceso ETL.

Minería de Datos Aplicada a la Educación: Modelo de Deserción Universitaria en la Universidad Tecnológica Nacional Facultad Regional Rosario

Luciano Valía, José Rostagno, Juan Miguel Moine,
Cristian Bigatti, Fabiana María Riva, Eduardo Amar

*Departamento de Ingeniería en Sistemas de Información
Facultad Regional Rosario
Universidad Tecnológica Nacional*

*lucianovalia@hotmail.com, joserostagno@hotmail.com, juanmiguelmoine@gmail.com, cristianbigatti@gmail.com,
fabianamriva@gmail.com, eduardoamar000@gmail.com*

Resumen

La deserción universitaria es uno de los mayores problemas que atraviesan las Universidades de nuestro país. La Minería de Datos es una disciplina que puede colaborar con esta situación, brindando un conjunto de técnicas y métodos para detectar patrones de comportamiento en los alumnos y accionar en consecuencia. Por este motivo, abordamos un estudio cuyo objetivo es el descubrimiento de patrones de deserción universitaria en la carrera Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Rosario.

Palabras clave: *deserción universitaria, minería de datos, estudiantes universitarios, patrones de comportamiento.*

Contexto

Este trabajo fue realizado en el marco del Proyecto de Investigación y Desarrollo: Modelo de Minería de Datos para la identificación de perfiles de deserción universitaria en alumnos de ISI (PID-UTN 3802).

Introducción

En la actualidad, la deserción universitaria representa un problema importante que enfrentan las universidades de nuestro País y del exterior. Se estima que en la Argentina alrededor del 80% de los alumnos que ingresan a la universidad desertan [1]. La Universidad Tecnológica Nacional Facultad

Regional Rosario, al igual que la mayoría de las universidades de nuestro País, no está exenta a esta situación en su carrera Ingeniería en Sistemas de Información.

Los sistemas de información en la universidad almacenan día a día una gran cantidad de datos, tanto del estado académico como de la situación socioeconómica de los alumnos, los cuales pueden ser explorados por disciplinas como la Minería de Datos para obtener nuevo conocimiento. Minería de Datos (también llamado Explotación de Información) es el proceso de extraer conocimiento útil, comprensible y novedoso de grandes volúmenes de datos, siendo su principal objetivo encontrar información oculta o implícita que no es posible obtener mediante métodos estadísticos convencionales. El resultado del proceso es un conjunto de patrones (modelos), los cuales pueden ser convertidos en información valiosa para la toma de decisiones [2].

En general, el problema de deserción universitaria ha sido analizado en función de dos grandes grupos de variables que caracterizan a los alumnos. En primer lugar desde el punto de vista académico, tomando variables como el grado de avance en la carrera, medido como el número de materias aprobadas y el promedio de calificaciones del alumno [3,4] y en segundo lugar trabajando exclusivamente con variables socioeconómicas como el máximo nivel de estudios alcanzados por los padres, situación laboral y edad [5,6,7].

Si bien algunos incorporan el estudio simultáneo de variables académicas y socioeconómicas los mismos fueron realizados con información proveniente del sistema de gestión académica, señalando sus autores la mala calidad de la misma (desactualizada e incompleta), especialmente aquella referida a aspectos sociales y económicos [8,9]. En [10] se demuestra que a partir de un relevamiento de 8 estudios previos, sólo 2 de ellos fueron realizados con datos de buena calidad. Esto se muestra gráficamente en la Figura 1, donde se ve que el 75% de las investigaciones han utilizado datos de mala calidad, incompletos y desactualizados.

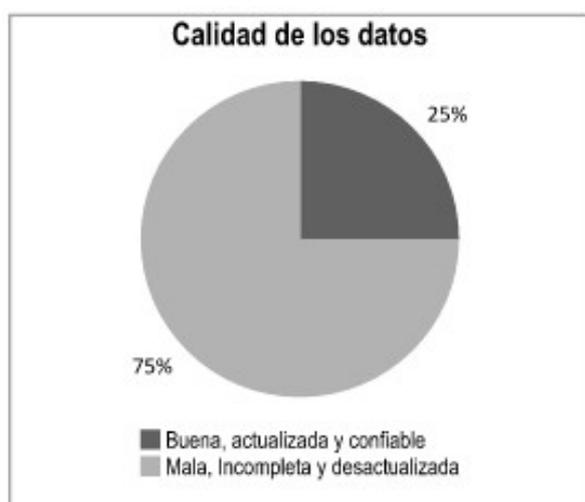


Fig. 1. Calidad de los datos utilizados en estudios previos

En este proyecto se propone el empleo de técnicas y métodos de Minería de Datos para la identificación de patrones de deserción en alumnos de los primeros años de la carrera Ingeniería en Sistemas de Información en la Universidad Tecnológica Nacional Facultad Regional Rosario, utilizando información actualizada proveniente de encuestas presenciales.

Líneas de investigación y desarrollo

En el marco de este proyecto se han definido las siguientes líneas de investigación:

- La evaluación y comparación de diferentes estudios previos realizados en materia de deserción universitaria utilizando técnicas de

Minería de Datos, con el fin de establecer una discusión sobre un criterio de deserción.

- La definición de variables socioeconómicas y académicas involucradas en el proceso de deserción de los alumnos. La recolección de datos y su posterior procesamiento mediante un análisis exploratorio que permita conocer la naturaleza de los mismos.
- La creación y evaluación de modelos de Minería de Datos, tanto predictivos como descriptivos, que permitan mejorar la toma de decisiones relacionadas con esta problemática.

Resultados obtenidos/esperados

En [10] hemos presentado un estudio comparativo de trabajos previos sobre la problemática de deserción universitaria en diferentes universidades de nuestro País y del exterior.

Además, a partir de una recolección de datos realizada en el año 2015, se ha logrado elaborar un estudio de las variables socioeconómicas y académicas que caracterizan a los alumnos de la UTN Facultad Regional Rosario [11]. Entre los resultados más interesantes se pudo detectar que los hábitos de estudio (cantidad de horas de estudio y asistencia a clase de consulta) varían en función del año de cursado y de la localidad de procedencia (Figura 2). También se encontró un patrón interesante en los alumnos que tienen becas, donde sus rendimientos son inferiores a quienes no la tienen o reciben subsidio del Estado (Figura 3). Además, se pudo observar que el desempeño académico (cantidad de materias aprobadas) varía según el nivel de educación de los padres (Figura 4), y no resulta significativamente influenciado por la situación laboral o el tipo de estudio secundario del alumno.

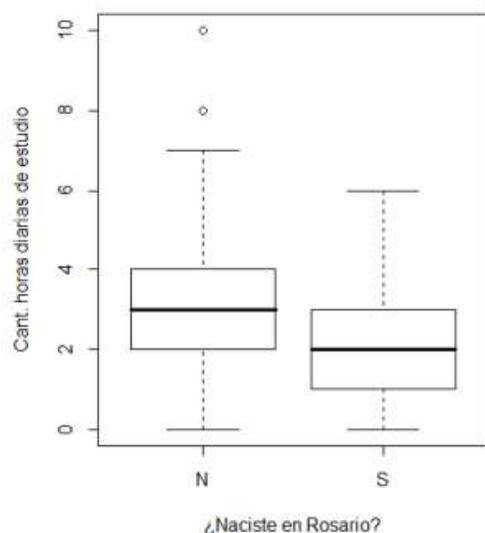


Fig. 2. Nació en Rosario vs Cantidad de horas de estudio.

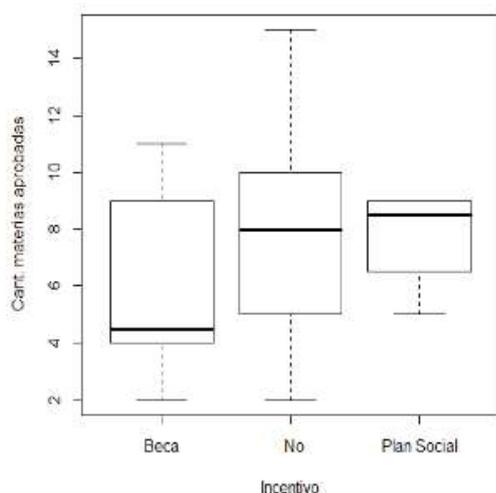


Fig. 3. Recibe incentivo vs Cantidad de materias aprobadas.

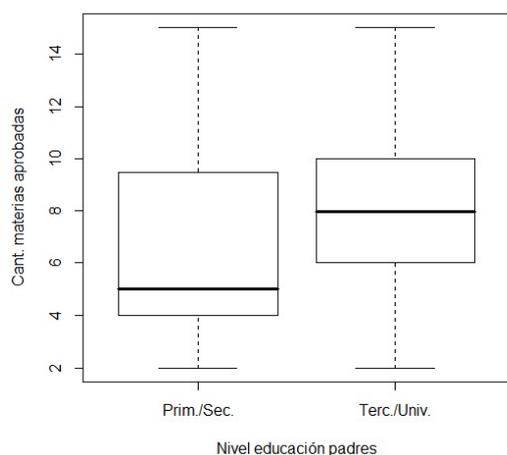


Fig. 4. Nivel de estudio padres vs Cantidad de materias aprobadas

Siguiendo esta línea de trabajo se espera poder recolectar información sobre el estado académico actual de los alumnos relevados, para comenzar con la elaboración de modelos predictivos y descriptivos que permitan estimar la deserción universitaria en nuestra institución. Para esta tarea, se considerará inicialmente que un alumno desertó si luego de dos años de haber sido encuestado (es decir, al inicio del ciclo 2017) no se encuentra inscripto en alguna materia de la carrera.

Formación de los Recursos Humanos

En el marco de este proyecto de investigación se encuentran trabajando alumnos y graduados de la Universidad Tecnológica Nacional Facultad Regional Rosario bajo la coordinación de docentes e investigadores de la institución. Se han realizado investigaciones sobre temáticas actuales de Minería de Datos, para su posterior divulgación tanto en jornadas abiertas a la comunidad como en cátedras de la carrera Ingeniería en Sistemas de Información.

Referencias

1. Lorenzano, C. "La Deserción Universitaria en la Universidad de Tres de Febrero, Universidad Tres de Febrero". Recuperado el 2 de marzo 2015, de <http://www.untref.edu.ar/documentos/AutoevaluacionLadesercion.pdf>
2. Hernández Orallo, Ramirez Quintana y Ferri Ramirez (2004) "Introducción a la Minería de Datos". Editorial Pearson Prentice Hall. España. ISBN 84-205-4091-9.
3. Balestieri, C. & otros (2013) "Modelo neuronal para la estimación del riesgo de deserción de alumnos de grado". EST 2013, 42 JAIIO.
4. Karamouzis, S. T., & Vrettos, A. (2008) "An artificial neural network for predicting student graduation outcomes". World Congress on Engineering and Computer Science (pp. 991-994).

5. Formia, S. (2013) “Evaluación de técnicas de Extracción de Conocimiento en Bases de Datos y su aplicación a la deserción de alumnos universitarios” (Tesis Especialidad). Universidad Nacional de La Plata.
6. Giovagnoli, P. I. (2002) “Determinantes de la deserción y graduación universitaria: una aplicación utilizando modelos de duración”. Documentos de Trabajo. Universidad Nacional de La Plata. Recuperado el 16 de febrero de 2016 de <http://sedici.unlp.edu.ar/handle/10915/3436>.
7. Valero, S., & otros (2010) “Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos”. Universidad Tecnológica de Izúcar de Matamoros. Recuperado el 2 de marzo de 2015, de <http://www.utim.edu.mx/~svalero/docs/e1.pdf>.
8. Pereira, R. T., Romero, A. C., & Toledo, J. J. (2013) “Descubrimiento de perfiles de deserción estudiantil con técnicas de minería de datos”. Vínculos, vol10, pág. 373-383.
9. Fischer, E. (2012) “Modelo para la Automatización del Proceso de Determinación de Riesgo de Deserción en Estudiantes Universitarios”. (Tesis doctoral). Universidad de Chile.
10. Rostagno, José; Moine, Juan Miguel; Bigatti, Cristian Germán. (2016) “Abordaje del fenómeno de deserción universitaria con técnicas de minería de datos”. Revista Rumbos Tecnológicos. Volumen 8. Universidad Tecnológica Nacional Facultad Regional Avellaneda.
11. Valía, Luciano; Moine, Juan Miguel; Rostagno, José y Otros (2016) “Hacia un modelo de deserción universitaria: Análisis exploratorio de variables socioeconómicas y académicas de alumnos de primer y segundo año de la carrera Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional Facultad Regional Rosario”. IV Congreso Nacional de Ingeniería Informática y Sistemas de Información. Universidad Católica de Salta.

Minería de Datos en Respecto de la Incidencia del Clima sobre el Desperfecto en el Alumbrado Público

Sergio Quiroga, Mag. Alejandra Malberti, Mag. Raúl Klenzi

Instituto de Informática / Departamento de Informática /Facultad de Cs. Exactas Físicas y Naturales / Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", Rivadavia, San Juan, Teléfonos: 4260353 4260355 - 4260394 - 4264721 - 4234129, Fax 0264-4234980, <http://www.exactas.unsj.edu.ar> {sergiooquiroga,amalberti,rauloscarklenzi}@gmail.com

Resumen

El presente estudio corresponde a una extensión de un trabajo previo “Minería de datos en la detección de desperfectos en el alumbrado público” (WICC 2015), propone extraer información que permita descubrir la incidencia de factores climáticos en la cantidad y tipología de fallas habituales que se producen en el alumbrado público. Los datos provienen de reclamos de usuarios, y otros inherentes a condiciones climáticas particulares registradas en las fechas tratadas. Previo al análisis de los datos por medio de Minería de Datos se consideran aspectos tales como ruido, datos ausentes, y volatilidad, entre otros. Las direcciones correspondientes a los reclamos se normalizan para poder realizar tareas de geolocalización de Google, con la finalidad de detectar zonas de conflicto. En esta propuesta se aplica la metodología CRISP-DM y se utiliza la herramienta Knime analytics, en especial la API de Google Maps.

Palabras clave:

Minería de datos -Geolocalización- Knime – Google Maps

Contexto

La línea de investigación que permite elevar la presente propuesta está contenida en el proyecto bianual “La Ciencia de Datos en grandes colecciones de datos”, aprobado por CICITCA-UNSJ y sujeto a evaluación externa. Los datos procesados provienen de “reclamos” generados en el call-center de una empresa local, que se encarga del mantenimiento del alumbrado público en la ciudad de San Juan. Estos reclamos son generados por los vecinos que detectan fallas en el alumbrado público y solicitan, vía telefónica, su reparación. Cada reclamo es transcripto en una planilla Excel por el operador receptor del llamado, quien registra la dirección postal o catastral correspondiente a la ubicación de la luminaria, o conjunto de luminarias, que presentan desperfectos. Los datos recepcionados son tratados por medio de la aplicación de técnicas de minería de datos.

Introducción

El presente trabajo involucra a una empresa que se dedica al mantenimiento del alumbrado público en diversos departamentos de la provincia de San Juan. Esta empresa se ha propuesto mejorar el servicio que brinda a sus usuarios, especialmente en lo que refiere

a la detección temprana de problemas en el alumbrado consciente que la mayor parte de ellos son causados por inclemencias climáticas. Para ello se tienen en cuenta los reclamos realizados por los usuarios: alumbrado público apagado (APA), alumbrado público prendido de día (APP), lámparas apagadas (LA), lámpara prendida de día (LPP), lámpara que prende y apaga (LPA), cables cortados (CC), lámparas tapadas por las ramas (LTxR) y globos apagados (GA); así como los datos recopilados sobre condiciones climáticas ocurridas en fechas determinadas – velocidad del viento, temperatura, presión atmosférica entre otros. Esta organización busca optimizar sus procesos de tal forma de lograr rapidez y eficiencia en la prestación del servicio a fin de alcanzar satisfacción de sus usuarios. Es así que, para facilitar la identificación de las zonas afectadas, se utilizará la dirección del reclamo para realizar su geolocalización en el mapa. En las Figuras 1 se muestra en Google Maps las ubicaciones de diversos reclamos, en este caso la dirección corresponde a la calle número y orientación. En la figura 2 podemos ver otra forma de visualizar los reclamos en este caso además la muestra corresponde solo a las direcciones que hacen esquina, es decir aquellas en las que no se contaba ni con el número ni orientación.



Figura 1: Ubicación de reclamos de usuarios en Google Maps



Figura 2: Otra forma de visualizar los reclamos

En la Figura 3 se observa los colores diferentes corresponden a los distintos motivos.

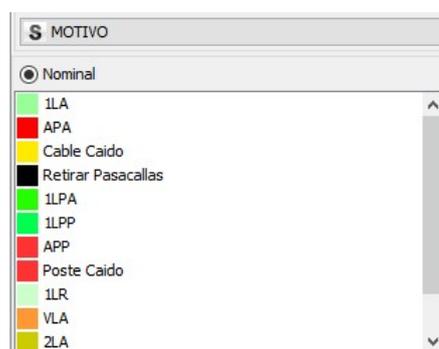


Figura 3: Listado de la asignación de color a los distintos motivos de reclamos

“La geolocalización es la capacidad para obtener la ubicación geográfica real de un objeto, como un radar, un teléfono móvil o un ordenador conectado a Internet...El término geolocalización está estrechamente relacionado con el uso de sistemas de posicionamiento, pero puede distinguirse de estos por un mayor énfasis en la determinación de una posición significativa (por ejemplo, una dirección de una calle) y no sólo por un conjunto de coordenadas geográficas.” [Wikipedia]. La API de Google Maps que ofrece Google permite la geolocalización en tiempo real y así trabajar con datos geográficos implementados a nivel mundial. Además, posee un sistema inteligente para la búsqueda de direcciones, lo que permite lograr buenas aproximaciones aún con direcciones mal escritas. Esto también puede resultar en alguna debilidad en cuanto a direcciones similares, ya que puede ocurrir que algunas de ellas aparezcan en lugares

erróneos, por tanto, es necesario entender el modo de trabajo de Google Maps. Si bien es limitado en cuanto a lo que se pueda desarrollar internamente, hay que comprender que para realizar posicionamiento se maneja con coordenadas a través de su latitud y longitud.

Google Maps utiliza para funcionar los lenguajes HTML CSS y Javascript. Los mapas son solo imágenes que están colocadas en una página HTML. El objetivo de la API es enviar información de las coordenadas al mapa a través de AJAX¹ y devuelve como resultado la imagen con la posición.

El API consiste de archivos JavaScript que contienen las clases, métodos y propiedades que se usan para el comportamiento de los mapas.

Las coordenadas están expresadas usando números decimales separados por coma. La latitud siempre precede la longitud. Google maps no se limita a cierta cantidad de decimales, sin embargo, según unas pruebas realizadas en la web <http://ojalesa.mbnet.fi>² concluye que:

5 a 6 decimales: es el máximo que debemos usar para ser específicos

4 decimales: para algún detalle en el mapa

3 decimales: es bueno para centrar ciudades

2 decimales: es apropiado para centrar países o estados, tal vez 3 por Mónaco

La abundante cantidad de datos recogidos son tratados por medio de técnicas adecuadas, entre las que se encuentra la Minería de Datos (MD). Para Witten (Witten, 2005) la Minería de Datos es el proceso de extraer

conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. Este proceso conlleva técnicas de análisis de datos destinadas a extraer patrones, describir tendencias y regularidades, predecir comportamientos y, en general, aprovechar un gran volumen de información digitalizada. Para tal fin se utiliza el software Knime Analytics, una plataforma open source (de código abierto), *...diseñado para descubrir el potencial oculto de los datos, usando la minería para obtener nuevas ideas o predecir nuevos futuros...*³.

En la Figura 4 se muestra el proceso para visualizar ubicaciones de reclamos en Google Maps, realizado en Knime.

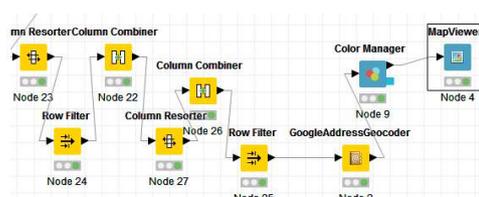


Figura 4 : proceso Knime para visualizar las ubicaciones de los reclamos en Google Maps

Líneas de Investigación, Desarrollo e Innovación

En el marco del laboratorio se llevan adelante diferentes trabajos de investigación aplicada, caracterizada por el tipo de datos observados y analizados. Según la variedad y tipología de datos se está trabajando en análisis de series temporales, aplicaciones en Data Streammining, y reconocimiento de perfiles de usuarios, intentando llevarlas a plataformas paralelas en cluster de computadoras.

pequeños paquetes de datos. Fuente: aprenderaprogramar.com Copyright 2006-2017

² Página de ejemplos y testeos de la API de Google Maps: <http://ojalesa.mbnet.fi/exam/decimal.html>

³ Cita de <https://www.knime.org/>

¹ Para agilizar los tiempos de espera al servidor web surgió Ajax (inicialmente Asynchronous JavaScript And XML, aunque hoy día ya no es una tecnología ligada a XML con lo cual no pueden asociarse las siglas a estos términos). Tecnología que busca evitar las demoras propias de las peticiones y respuestas del servidor mediante la transmisión de datos en segundo plano usando un protocolo específicamente diseñado para la transmisión rápida de

Se utiliza la metodología CRISP-DM4, este modelo del proceso de Minería de Datos, para realizar tratamiento de grandes cantidades de datos, con implementaciones de algoritmos de minería de datos,

Se utiliza la metodología CRISP-DM5, este modelo del proceso de Minería de Datos, se divide en seis pasos Compresión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue

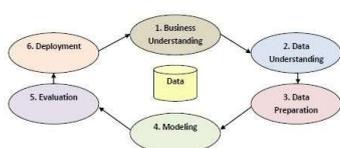


Figura 5: Modelo conceptual de CRISP-DM

En nuestro caso, se aplican tareas de: 1) ingreso de datos, ajustando los parámetros de los diferentes módulos de Knime al formato de los datos de entrada 2) preprocesamiento, filtrados, ajustes y adaptación de formatos de datos, como de diferentes estrategias, siendo estas: segmentación, clasificación entre otras.

Objetivos y Resultados

El objetivo que se persigue es proporcionar zonas en las que se detectan fallas en el alumbrado público debido a las inclemencias, a partir de la aplicación de estrategias de minería de datos, análisis de factores climáticos e incorporación y procesamiento de información relativa a la geolocalización.

Al momento de la elaboración de este documento, se han obtenido los siguientes resultados:

- Profundización en el conocimiento de las tareas de minería de datos, y de plataformas open source.

- Tratamiento de bases de datos provenientes de distintas fuentes.

- Detección y solución de inconsistencias en los datos.

- Estudio de técnicas y herramientas de geolocalización.

- Geolocalización de direcciones de reclamos en Google Maps.

Así mismo, se ha trabajado con los módulos contenidos en KNIME GoogleAddressGeocoder y MapzenGeocoder permitiendo éste último una mayor carga de registros por unidad de tiempo, aunque con una precisión de ubicación menor asociadas a las variables latitud y longitud de una referencia cartográfica.

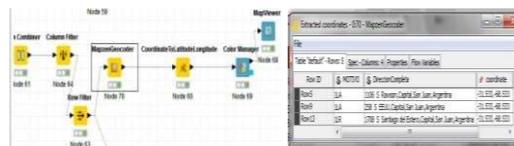


Figura 6

Esto hace, como se puede apreciar en la figura 6, que puntos con diferente dirección postal sean ubicados en un mismo punto de coordenadas.

Por el contrario, el módulo Google AddressGeocoder genera la salida que se puede apreciar en la Figura 7 y donde cada uno de las direcciones postales posee su correspondiente coordenada.

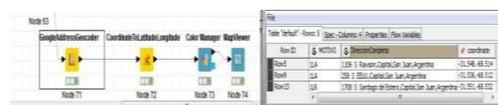


Figura 7

Por otro lado, la herramienta permite desde un módulo de combinación de columnas de datos, alcanzar el formato necesario con que deben ser ingresadas las direcciones postales en cada uno de

⁴Cross Industry Standard Process for Data Mining

⁵Cross Industry Standard Process for Data Mining

los módulos o APIs que transforman este dato a un sistema de coordenadas de latitud y longitud como muestra la figura 8.



Figura 8

Formación de Recursos Humanos

Las tareas desarrolladas en el ámbito del proyecto han permitido la conclusión y desarrollo actual de muchos trabajos finales de grado pertenecientes a alumnos de las carreras contenidas en el ámbito del Departamento Informática de la Facultad de Ciencias Exactas Físicas y Naturales.

Referencias

- Mark Salvador, Ron Resmini (auth.), Guido Cervone, Jessica Lin, Nigel Waters (eds.) **-Data Mining for Geoinformatics. _ Methods and Applications-** Springer New York (2014)
- Harvey J. Miller, Jiawei Han, **Geographic Data Mining and Knowledge Discovery**, Second Edition (Chapman & Hall CRC Data Mining and Knowledge Discovery Series)-CRC Press (2009)
- Hernández Orallo, José; Ramírez Quintana, Ma José; Ferri Ramírez, César. (2004) **Introducción a la Minería de Datos**. Edit. Pearson educación
- Microsoft (2013) Prueba y validación (minería de datos) <http://msdn.microsoft.com/esAR/library/ms174493.aspx>

- Molina L. (2002) **Data mining: torturando a los datos hasta que confiesen**.

<http://www.uoc.edu/web/esp/art/uoc/molina1102/molina1102.html>.

- Pérez López, César; Santín González, Daniel. (2006) **Minería de Datos Técnicas y Herramientas**. Edit. Alfaomega Grupo Editor

- Pyle, Dorian; Kaufmann, Morgan. (1999) **Data Preparation for Data Mining**. Edit. Morgan Kufmann Publishers Inc.

- Witten, Ian H.; Eibe, Frank. (2005) **Data Mining Practical Machine Learning Tools and Techniques**. Edit. Elsevier Inc.

- **Geolocalización**. (s.f.). En Wikipedia. Recuperado el 1 de febrero de 2017 de <https://es.wikipedia.org/wiki/Geolocalizaci%C3%B3n>

North, M. (2012). *Data mining for the masses* (pp. 91-100). Global Text Project.

Minería de Datos para Análisis del Microbioma Humano

Cristóbal R. Santa María*, Victoria Santa María**, Laura Avila*, Luis López* Juan Otaegui*, Marcelo Soria***

*DIIT-UNLaM, **Instituto Lanari-FMed-UBA, ***FAUBA Florencio

Varela 1903 San Justo Pcia. de Buenos Aires 54-011-44808952

csanta_maria@ing.unlam.edu.ar

vcrtstmr@hotmail.com

laura_avila75@yahoo.com.ar

llopez@ing.unlam.edu.ar

soria@agro.uba.ar

juancarlosotaegui@yahoo.com.ar

Resumen

Se expone la línea de investigación que lleva adelante el Grupo de Investigación y Desarrollo en Data Mining del Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM. Se detallan los resultados del proyecto de investigación “Aplicaciones de Data Mining al Estudio del Microbioma Humano”, C169 del Programa de Incentivos. La línea de trabajo intenta aportar procedimientos computacionales adecuados para analizar la relación clínica entre el microbioma intestinal y la presencia de patologías tales como el cáncer de colon y la enfermedad de Crohn. El trabajo hasta aquí realizado comprende la obtención de una muestra de microbiomas de pacientes desde el repositorio de NCBI, la identificación bacteriana a partir del gen marcador y la determinación de la distribución de frecuencias por especies en cada paciente. Se continuó luego con el agrupamiento de pacientes por enterotipos y la evaluación clínica de las categorías obtenidas.

Palabras Clave: Microbioma Secuencias Clasificación Diagnóstico

Contexto

Desde los comienzos de su vida el cuerpo humano es colonizado por bacterias,

arqueas, hongos y virus. Esta comunidad de microorganismos se denomina microbioma y contiene diez veces más células que las del propio cuerpo humano. La cantidad de genes presentes en total es varios órdenes de magnitud mayor que la del genoma humano. Si bien esto se conoce desde hace largo tiempo, la imposibilidad de cultivo en laboratorio de la mayoría de esos microorganismos ha dificultado hasta ahora análisis profundos. La nueva generación de tecnologías de secuenciación de ADN ha permitido comenzar a estudiar las características del microbioma humano según la edad de la persona, la localización geográfica, los hábitos alimentarios y la presencia de enfermedades. El objetivo principal de estos estudios metagenómicos es analizar la estructura y la dinámica de comunidades microbianas, para establecer cómo se relacionan sus miembros entre sí, cuáles son las sustancias que producen y que consumen, y especialmente cuáles son sus interacciones con las células humanas próximas y cómo se modifica la comunidad en presencia de enfermedades. El estudio por medio de la identificación de un gen marcador como el 16s rRNA pretende evaluar características ecológicas como la riqueza y la diversidad, mientras que el análisis del metagenoma como un todo, identificando las secuencias obtenidas por comparación con una base

de datos genética previamente armada, permite agrupar los genes identificados por funciones metabólicas asociadas con la presencia de enfermedades. Los estudios que emplean tales técnicas se multiplican velozmente y existen a nivel mundial proyectos de investigación como el Metagenomics of the Human Intestinal Tract (MetaHIT) o el Human Microbiome Project (HMP). En nuestro país en el marco del Plan Nacional de Ciencia, Tecnología e Innovación (Argentina Innovadora 2020), dentro del área de Salud, se ha comenzado a desarrollar una Plataforma Tecnológica de Genómica y Bioinformática que facilitará estudios similares. El objetivo general es entender el funcionamiento del microbioma humano a partir del procesamiento y análisis de muestras de secuencias de ADN, y desarrollar nuevas herramientas para analizar y caracterizar el curso de patologías poniendo énfasis en el cáncer de colon y en la enfermedad de Crohn.

Introducción

El trabajo computacional consiste en obtener los datos secuenciados de una muestra integrada por varios microbiomas. Cada microbioma debe cotejarse con una base de datos correspondiente a un gen marcador para encontrar la distribución de frecuencias de los microorganismos identificados por tal gen. Alternativamente el conjunto de secuencias del microbioma puede compararse con otra base de datos de funciones genéticas para agrupar los genes integrantes por función y así obtener la distribución de frecuencias según las funciones metabólicas que las secuencias integrantes revelan [2]. En cualquier caso una vez formadas las matrices que representan por fila las distribuciones de cada microbioma individual estos pueden agruparse en clusters. Cada fila del conjunto representa a un paciente y en la

base de datos esa instancia es un vector donde cada componente corresponde a una especie o género distinto, y el valor que adopta es la cantidad presente del respectivo microorganismo o grupo (taxón). Armados con las distancias y los encadenamientos adecuados, los diferentes clusters constituyen los enterotipos. Las características de cada agrupamiento logrado deben cotejarse con las apreciaciones clínicas de los pacientes que lo integran, ya obtenidas por otras vías diagnósticas, para apreciar el punto hasta el cual resultan útiles en la evaluación médica de la patología investigada.

Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo pretende estudiar en detalle la aplicación de métodos computacionales supervisados y no supervisados sobre los microbiomas para clasificar y predecir patologías. Comprende tanto el enfoque a través del gen marcador, el caso de los resultados que aquí se presentan, como el enfoque a partir de la información de funcionalidad metabólica aportada por el metagenoma. Se intentan alcanzar varios objetivos:

- Dominar la tecnología de almacenamiento, comparación y distribución funcional según las secuencias obtenidas del microbioma intestinal de pacientes por vía preferente de videocolonoscopia o alternativamente por materia fecal.

- Determinar los métodos computacionales más convenientes para los agrupamientos de microbiomas de pacientes de forma que revelen óptimamente sus características clínicas.

- Realizar lo propio respecto de algoritmos de predicción entrenados y testeados para la evaluación clínica.

-Dejar allanado el camino para la aplicación experimental de todos estos métodos a muestras de pacientes locales obtenidas por investigadores del grupo.

-En tal sentido, obtener muestras de pacientes en el medio local, analizar las características del protocolo médico a seguir, enviarlas a secuenciar y aplicar los procedimientos ya probados sobre muestras no propias.

Resultados y Objetivos

Un primer aspecto a resolver fue el del hardware y sistema operativo necesario para el trabajo. Con vistas al desarrollo completo de la línea de investigación dentro del Departamento de Ingeniería e Investigaciones Tecnológicas de la UNLAM se decidió adaptar un servidor con ocho procesadores Intel-I7 y 16 GB de memoria RAM para el trabajo del grupo, ya que el volumen de memoria y la capacidad de proceso requeridas por los conjuntos de datos utilizados en la bibliografía y los que potencialmente pudieran formarse luego a partir de muestras propias, lo hacían necesario. También se decidió dotar al servidor con el sistema BioLinux, de uso habitual para este tipo de trabajos pues posee una fácil interacción con paquetes de software libre, eficientes y probados, en la investigación en biología computacional.

Se trabajó con una muestra de 11 pacientes, 7 sanos y 4 afectados por la enfermedad de Crohn [3], y para realizar las asignaciones taxonómicas y funcionales se utilizó el software de reciente desarrollo, SUPERFOCUS, que efectúa ambas tareas. Para la asignación funcional usa la base de datos SEED [4]. Las salidas de SUPERFOCUS son varios archivos de texto, uno con las asignaciones funcionales, tres con la información de subsistemas y otro más con la información

taxonómica. Para consolidar la información obtenida de todos los bloques se programaron dos scripts en R para procesar la información taxonómica y la funcional. Ambos scripts leen los archivos de una muestra, agregan la información y producen como salida dos tipos de matrices. Para la información taxonómica se producen una serie de matrices con la asignación taxonómica en una dimensión y la muestra en la otra. Existen tablas para cada una de las categorías taxonómicas usadas en biología: reino, phylum, clase, orden, familia, género y especie. En el caso de la información funcional, la salida está constituida por tablas para la función específica y tres más para cada uno de los subsistemas que define el proyecto SEED. Con el archivo de asignaciones taxonómicas se inició el proceso utilizando el software R para armar la matriz de distancias. Para medir la distancia entre dos microbiomas cuyas distribuciones de frecuencias estadísticas de especies son conocidas hay que medir las distancias entre ambas distribuciones. Para ello se utiliza la distancia de Jensen-Shanon [5] Siendo P y Q dos distribuciones conocidas puede definirse $R = \frac{1}{2}(P + Q)$ y su entropía $\kappa = H(R)$. Luego de algunas consideraciones [5] resulta $H(R) \approx \frac{1}{2}H(P) + \frac{1}{2}H(Q)$ y se define la expresión

$$D_F^2 = 2H(R) - H(P) - H(Q)$$

Entonces

$$\sqrt{D_F^2} = (2H(R) - H(P) - H(Q))^{\frac{1}{2}}$$

cumple con las condiciones matemáticas para ser una distancia entre P y Q [10].

De acuerdo a ello queda

$D_F = \left(\sum_{i=1}^N p_i \log p_i + q_i \log q_i \right)^{\frac{1}{2}}$ como expresión de la distancia entre dos distribuciones de frecuencias.

Esta distancia es la que usa el algoritmo PAM (Partitioning around medoids) utilizado en este caso directamente de la

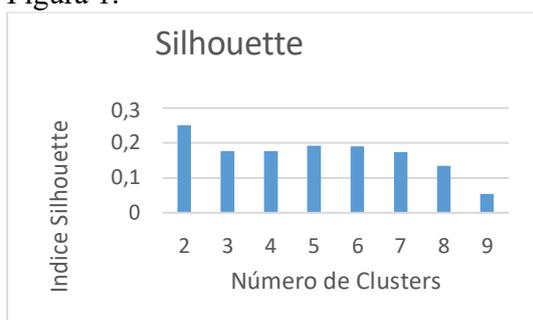
biblioteca cluster de R. El medoide es el elemento para el cual la disimilitud promedio con todos los objetos en el conglomerado es mínima.

En [2] se utiliza para definir el número óptimo de clusters, el índice de Calinski-Harabasz [6] que se construye de la siguiente forma:

$$C = \frac{\frac{B_k}{k-1}}{\frac{W_k}{n-k}}$$

Aquí B_k es la suma de las distancias al cuadrado de todos los elementos i y j que no pertenecen al mismo cluster, W_k es la suma de los cuadrados de las distancias de todos los elementos i y j que pertenecen al mismo cluster, n es el número de elementos a clasificar y k la cantidad seleccionada de clusters. Sin embargo dada la cantidad exigua de microbiomas disponibles en la muestra, en el presente estudio se evaluó solo el agrupamiento obtenido con el índice Silhouette para lo cual se realizaron los agrupamientos considerando distintos números de clusters. Los resultados pueden apreciarse en la Figura 1.

Figura 1.



Se decidió entonces considerar el agrupamiento en 2 enterotipos y realizar sobre él una reducción de las 277 dimensiones de la muestra a través de un análisis de componentes principales. Se vio entonces que las primeras dos componentes explican el 53% y el 21% respectivamente de la información.

El análisis de correlación lineal de estas dos componentes CP1 y CP2 con las variables originales arrojó los siguientes datos significativos volcados respectivamente en las Tablas 1 y 2

Tabla 1

CP1(53%)	Alta Correlación
[Ruminococcus]_obeum	-0,92
Bacteroides_helcogenes	0,89
Bacteroides_salanitronis	0,83
Bacteroides_thetaiotaomicr..	0,92
Bacteroides_vulgatus	0,97
Bacteroides_xylanisolvens	0,96
Kitasatospora_setae	0,81
Porphyromonas_gingivalis	0,96
Prevotella_denticola	0,85

Tabla 2

CP2(21%)	Alta Correlación
butyrate-producing_bacteri..	0,75
Clostridium_saccharolyticu..	0,75
Roseburia_hominis	-0,75
Roseburia_intestinalis	-0,79

Las Tablas 1 y 2 permiten ver que el 74% de la información aportada por el agrupamiento realizado se explica predominantemente por la presencia de las bacterias citadas que tienen alta correlación con las componentes principales 1 y 2. Si bien desde el punto de vista clínico este conocimiento puede representar un aporte importante no bastó, en este caso para que los dos enterotipos obtenidos se correspondieran con la clasificación de sano o enfermo previamente determinada por otros procedimientos diagnósticos. La Figura 3 muestra los enterotipos obtenidos. El Enterotipo 1 es coloreado en azul mientras que el Enterotipo 2 es amarillo. En la Figura 4 se ven los mismos agrupamientos pero ahora la diferencia de color entre los casos señala la presencia o ausencia de la enfermedad.

Figura 3

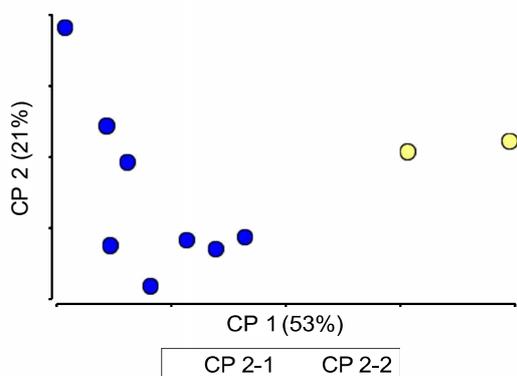
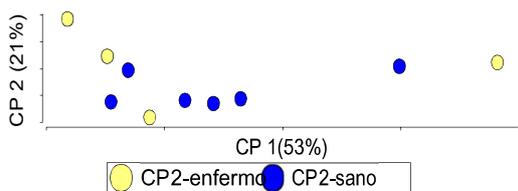
Clusters según Componentes Principales 1 y 2

Figura 4

Clusters por Componentes Principales y Estado de Salud

Los resultados obtenidos confirman que si bien la utilización del gen marcador no bastó para lograr una clasificación adecuada, en 3 de los 4 casos parece confirmarse la manifestación de la Enfermedad de Crohn asociada con la baja presencia de las especies *Bacteroides* y *Clostridium* que aparecen en alta correlación directa con los bajos valores de las CP1 y CP2. [7] Se concluye entonces en la necesidad de ampliar el trabajo por vía del estudio metagenómico de biomarcadores funcionales. Para realizarlo en vez de hallar los enterotipos por clasificación taxonómica habrá que establecerlos por grupos ortólogos (OG) que codifican para distintas enzimas o proteínas que pertenecen a distintas vías metabólicas. [6]

Formación de Recursos Humanos

En el equipo de trabajo participan un magister y un especialista en data mining, un doctor en biología, dos médicos, 2 ingenieros en sistemas, una matemática y un estudiante de ingeniería informática. Está en curso una tesis de maestría.

Referencias

- [1] Ngom-Bru, Catherine and Barretto, Caroline. Gut microbiota: methodological aspects to describe taxonomy and functionality. *Briefings in Informatics*. Vol3 NO 6. 747-750. 2012
- [2] Arumugam, M et al. Enterotypes of the human gut microbiome. *Nature* 2011 may 12; 473(7346): 174-180. doi:10.1038/nature09944
- [3] Morgan, XC. et al. (2012) Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* 2012, 13:R79
- [4] <http://theseed.org>
- [5] Endres, D y Schindeling, J. A New Metric for Probability Distributions. *IEEE Transactions on Information Theory*. Vol. 49 NO.7. 2003.
- [6] Calinski, T., and J. Harabasz. "A dendrite method for cluster analysis." *Communications in Statistics*. Vol. 3, No. 1, 1974, pp. 1-27.
- [7] Ray K. IBD. Understanding gut microbiota in new-onset Crohn's disease. *Nat Rev Gastroenterol Hepatol* [Internet]. Nature Publishing Group; 2014; 11(5):268.

Minería de Datos y Big Data. Aplicaciones en Señales y Textos

L. Lanzarini¹, W. Hasperué¹, C. Estrebou¹, F. Ronchetti^{1,2}, A. Villa Monte^{1,2}, G. Aquino^{1,3}, F. Quiroga^{1,2}, M. J. Basgall^{1,3}, L. Rojas⁴, J. Corvi¹, C. Luna¹, P. Jimbo⁵, A. Fernandez⁶, C. Puente⁷, J. A. Olivas⁸, A. Rosete⁹

¹ Instituto de Investigación en Informática LIDI, Facultad de Informática, UNLP, La Plata, Argentina
Facultad de Informática, UNLP, La Plata, Argentina

² Becario postgrado UNLP ³ UNLP, CONICET, III-LIDI, La Plata, Argentina

⁴ Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco, Ushuaia, Argentina

⁵ Dpto. Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

⁶ Dpto de Economía, Universitat Rovira i Virgili, Reus, España

⁷ Escuela Técnica Superior de Ingeniería ICAI, Universidad Pontificia Comillas, Madrid, España

⁸ Dpto de Tecnología y Sistemas de la Información, Universidad de Castilla-La Mancha, Ciudad Real, España

⁹ Universidad Tecnológica de La Habana “José Antonio Echeverría” (CUJAE), La Habana, Cuba

{laural, whasperue, cesarest, fronchetti, avillamonte, gaquino, fquiroga, mjbasgall}@lidi.info.unlp.edu.ar
{luisf.09, julieta.corvi, carla.lunagennari}@gmail.com, pjimbo@pcpsolutions.com, aurelio.fernandez@urv.net,
cristina.puente@icai.comillas.edu, joseangel.olivas@uclm.es, rosete@ceis.cujae.edu.cu

CONTEXTO

Esta presentación corresponde al Subproyecto “Sistemas Inteligentes” perteneciente al proyecto “Cómputo paralelo de altas prestaciones. Fundamentos y evaluación de rendimiento en HPC. Aplicaciones a sistemas inteligentes, simulación y tratamiento de imágenes” (Periodo 2014–2017) del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes para la resolución de problemas de Minería de Datos y Big Data utilizando técnicas de Aprendizaje Automático. Los sistemas desarrollados se aplican particularmente al procesamiento de señales y textos.

Con respecto al procesamiento de Señales el énfasis está puesto en el análisis de videos con el objetivo de identificar acciones humanas que faciliten la interfaz hombre/máquina y en la detección de patrones de movimiento de los objetos presentes.

En el área de la Minería de Datos se está trabajando, por un lado, en la generación de un modelo de fácil interpretación a partir de la extracción de reglas de clasificación que permita justificar la toma de decisiones y, por

otro lado, en el desarrollo de nuevas estrategias para tratar grandes volúmenes de datos.

Con respecto a Minería de Textos se han desarrollado métodos capaces de extraer las palabras clave de documentos independientemente del lenguaje. Además, se han desarrollado estrategias para resumir documentos a través de la extracción de párrafos.

Palabras clave: Estrategias adaptativas, Reconocimiento de Patrones, Minería de Datos, Minería de Textos, Big Data.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos tipos de estrategias adaptativas. Los resultados obtenidos han sido medidos en la solución de problemas pertenecientes a distintas áreas. A continuación se detallan los resultados obtenidos durante el último año.

1.1. PROCESAMIENTO DE SEÑALES

En el III LIDI, desde hace varios años se viene trabajando en el procesamiento de señales de audio y video. Como resultado de estas investigaciones se han diseñado e

implementado técnicas originales aplicables al reconocimiento tanto de gestos dinámicos como de medidas biométricas. En relación a esta línea, actualmente se están desarrollando los siguientes temas:

Reconocimiento de gestos

El reconocimiento de lengua de señas es un campo de investigación relativamente nuevo cuyo objetivo final es traducir de la lengua de señas a una lengua escrita. Esto implica poder tomar un video en donde una persona habla en lengua de señas, y reconocer la posición de la persona, de su cara y sus manos, la expresión de su rostro, la forma de sus manos y también la de sus labios si la seña requiere pronunciar la palabra para desambiguar. Con esa información, se debe reconocer la seña realizada, para luego con la información de una secuencia de señas generar una traducción a una lengua escrita.

En esta área, se publicó un método para clasificar señas en videos pre-segmentados que abarca todas las etapas del reconocimiento [1, 2]. Éste método no utiliza la información secuencial de la seña, es decir, no utiliza la información temporal. No obstante, los resultados de los experimentos muestran que aún con esa dificultad se pueden clasificar correctamente el 96% de las señas del conjunto de prueba. De esta forma se pudo determinar que la información temporal no es tan importante para el reconocimiento de señas, al menos en bases de datos de estas características.

Para los experimentos se utilizaron dos conjuntos de datos recolectados por nuestro grupo, LSA16 y LSA64. El primero, LSA16, contiene 800 imágenes con 16 clases de formas de mano y el segundo, LSA64, está formado por 3200 videos de 64 clases de señas dinámicas. Los detalles de la base de señas dinámicas LSA64 también han sido publicados [3].

Actualmente se está trabajando en mejorar las etapas de detección y segmentación de la mano y el reconocimiento de su forma aplicando redes convolucionales profundas y otras técnicas relacionadas.

Detección de patrones de movimiento en video

El análisis automático de video con el objetivo de detectar patrones de movimiento es de suma utilidad en distintas áreas. Este tema combina el procesamiento de imágenes digitales con la minería de datos ya que se deben analizar automáticamente la estructura de la escena, las actividades que en ella se están desarrollando y los patrones de movimiento de los objetos involucrados con el propósito de detectar situaciones anómalas.

Se espera poder contribuir al diseño y desarrollo de nuevas estrategias adaptativas aplicables al análisis de videos. Los resultados de esta investigación pueden aplicarse en distintas áreas tales como seguridad, a través de la detección automática de situaciones de riesgo o amenazas en escenas captadas a través de sistemas de video-vigilancia o salud a través de la identificación de comportamientos en personas que padecen enfermedades que alteran su movimiento corporal.

1.2. MINERÍA DE DATOS

Obtención de Reglas de Clasificación

Esta línea de investigación está centrada en la obtención de un conjunto de reglas de clasificación con tres características principales: precisión adecuada, baja cardinalidad y facilidad de interpretación [4,5]. Esto último está dado por el uso de un número reducido de atributos en la conformación del antecedente que, sumada a la baja cardinalidad del conjunto de reglas, permite distinguir patrones sumamente útiles a la hora de comprender las relaciones entre los datos y tomar decisiones.

Como resultado se ha desarrollado un nuevo método de extracción de reglas de clasificación que hace uso de una variante original de la técnica de optimización basada en cúmulos de partículas PSO inicializada a través de una red neuronal competitiva LVQ. Los resultados de su aplicación a un conjunto de 13 bases de datos de repositorio han sido satisfactorios.

Como área de transferencia tecnológica se ha analizado la situación de dos compañías financieras al momento de determinar el riesgo en una operación de otorgamiento de crédito para consumo. Se trata de operaciones con montos muy inferiores a los préstamos hipotecarios que requieren tomar decisiones rápidas ya que generalmente son acordados con los clientes a través de un servicio en línea. En ambos casos se han obtenido conjuntos de reglas con una precisión aceptable y una cardinalidad sensiblemente menor a los métodos convencionales. Volviendo a la necesidad de tomar una decisión rápida, este modelo ofrece una gran ventaja con respecto al mecanismo habitual. Para más detalles consultar [6,7].

Aplicaciones en Big Data

En esta línea se trabaja sobre el procesamiento en *streaming* y en *batch* de grandes volúmenes de datos en formato texto. Para esto se están desarrollando estrategias que aplican técnicas de machine learning que presentan la característica de ser iterativas, operando sobre el conjunto completo de los datos ó sobre los datos de un flujo, brindando resultados en tiempos de respuestas cortos los cuales se adaptan de manera dinámica a la llegada de nuevos datos.

Estas técnicas dinámicas se están empleando bajo el paradigma MapReduce, adecuado para procesamiento paralelo y distribuido. Para el tratamiento de un flujo de datos, se utiliza el enfoque de ventana deslizante temporal manejando el tamaño de la misma de manera dinámica en función de la frecuencia de llegada de los datos y el tiempo de respuesta de la tarea iterativa a realizar sobre ellos, permitiendo que cada dato sea utilizado por el proceso iterativo la mayor cantidad de veces posibles [8,9].

Los temas que se abordan en esta línea abarcan el procesamiento del lenguaje natural, la detección de tópicos, el análisis de sentimiento y procesamiento de datos relacionados al comercio realizado con criptomonedas.

1.3. MINERIA DE TEXTOS

Extracción de palabras clave en documentos de texto

Esta línea de investigación tiene su eje central en el estudio y aplicación de distintos métodos de representación de documentos así como de distintas técnicas adaptativas aplicables en la resolución de problemas de extracción de palabras clave, tarea de sumo interés ya que permite caracterizar un documento facilitando su búsqueda y clasificación [10].

En esta línea se está trabajando en un método de identificación de palabras clave a partir de documentos de texto en español utilizando redes neuronales como estimadores de probabilidad [11].

Este método define una representación vectorial para los términos, para luego aplicar un proceso de filtrado gramatical con el fin de remover términos inválidos. Es importante destacar que ésta es la única parte del método que es dependiente del idioma en cuestión [12].

Una vez obtenida la representación se utiliza un *ensemble* de redes neuronales para construir un modelo de clasificación. Dado un documento a ser analizado, el modelo determina para cada término la probabilidad de ser palabra clave. Estas probabilidades son utilizadas para construir un ranking de términos, lo que proporciona la flexibilidad de seleccionar los mejores N términos.

Actualmente se continúa el desarrollo del método para mejorar su precisión y ampliar su dominio de aplicación.

Síntesis automática de documentos

En esta línea de investigación se desarrollan técnicas capaces de representar y modelizar uno o varios documentos de texto con el objetivo de construir una versión más corta de ellos en forma automática preservando su información. Esto resulta de sumo interés ya que permite obtener el contenido principal de un documento en menos tiempo del que

llevaría hacerlo en forma manual a partir del texto completo.

En [13] se diseñó una estrategia para extraer el criterio utilizado por una persona al resumir un texto. Luego, se lo aplicó a otros documentos logrando obtener un resumen similar al que se hubiera conseguido en forma manual. Para ello, se utilizaron los capítulos de una tesis escrita en LaTeX a los cuales se les calculó un conjunto de métricas conocidas. Luego a través de una técnica de optimización basada en cúmulos de partículas se identificó el aporte de cada métrica en la construcción del resumen esperado.

Actualmente se están llevando a cabo dos trabajos utilizando un conjunto de artículos científicos de acceso libre. Por un lado se está trabajando en el agrupamiento de los documentos para descubrir relaciones entre las métricas que los representan. Por otro se está realizando la comparación de los resúmenes formados por las sentencias causales de un documento y los resúmenes obtenidos de aplicar cada una de las métricas por separado.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio de técnicas de optimización y redes neuronales artificiales para la obtención de reglas de tipo IF-THEN.
- Estudio de técnicas de segmentación de objetos en movimiento presentes en un video.
- Estudio de técnicas de agrupamiento aplicables a la detección de patrones de movimiento.
- Representación y clasificación de configuraciones de manos para el lenguaje de señas.
- Clasificación de señas dinámicas.
- Problemas de clasificación con severo desbalance de clases y métodos. Algoritmos aplicables a los mismos.
- Estudio de distintos métodos de

caracterización de textos haciendo énfasis en su estructura, longitud, idioma y formalidad en la redacción.

- Métodos estructurados y no estructurados aplicables a la representación de documentos. Representación de documentos de texto utilizando métricas.
- Estudio de técnicas para resumen automático de documentos.
- Implementación de técnicas en el paradigma de MapReduce
- Estudios de performance de los algoritmos desarrollados

3. RESULTADOS OBTENIDOS

- Desarrollo de un método de extracción de reglas de clasificación con énfasis en la reducción de la complejidad del modelo aplicable a riesgo crediticio.
- Desarrollo de una representación de términos y un modelo de clasificación con el fin de identificar palabras clave en un documento.
- Desarrollo de un modelo de clasificación de señas segmentadas y comparación de su desempeño con otros modelos del estado del arte.
- Desarrollo de una estrategia dinámica empleando el paradigma MapReduce para aplicar algoritmos iterativos sobre los datos, los cuales arrojan resultados muy similares a los que se obtienen con las mismas tareas ejecutadas de manera secuencial pero utilizando el conjunto de datos completo.
- Determinación de coeficiente de Hurst en transacciones de Bitcoins.
- Identificación de las partes relevantes de un documento.
- Análisis y comparación de resúmenes extractivos de documentos.
- Caracterización de documentos por medio de su agrupamiento.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de los temas involucrados en esta línea de investigación, en los últimos 5 años se han finalizado 4 tesis de doctorado, 2 tesis de maestría, 3 tesis de especialista y 9 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 5 tesis de doctorado, 1 tesis de especialista y 3 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Ronchetti F., Quiroga F., Estrebou C., Lanzarini L., Rosete A. *Sign language recognition without frame-sequencing constraints: A proof of concept on the argentinian sign language*. Publicado en Ibero-American Conference on Artificial Intelligence IBERAMIA 2016 (pp. 338-349)
- [2] Ronchetti, F., Quiroga, F., Estrebou, C.A., Lanzarini. *Handshape recognition for Argentinian Sign Language using ProbSom*. Journal of Computer Science & Technology, vol. 16, N° 1, págs. 1-5, ISSN 1666-6038, 2016.
- [3] Ronchetti, F., Quiroga, F., Estrebou, C.A., Lanzarini, L.C., Rosete, A. . *LSA64: An Argentinian Sign Language Dataset*, publicado en el XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) (pp. 794-803).
- [4] Lanzarini, L., Villa Monte, A., Ronchetti, F.: *SOM+PSO. A Novel Method to Obtain Classification Rules*. Journal of Computer Science & Technology (JCS&T), Vol. 15, No 1, pp. 15-22. ISSN 1666-6038. Abril 2015.
- [5] Lanzarini L., Villa Monte A., Aquino G., De Giusti A. *Obtaining classification rules using lvqPSO*. Advances in Swarm and Computational Intelligence. Lecture Notes in Computer Science. Vol 6433, pp. 183-193. ISSN 0302-9743. Springer-Verlag Berlin Heidelberg. Junio 2015.
- [6] Lanzarini L., Villa Monte A., Fernandez Bariviera A., Jimbo Santana P. *Obtaining Classification Rules Using LVQ+PSO: an application to Credit Risk*. Scientific Methods for the Treatment of Uncertainty in Social Sciences. Advances in Intelligent Systems and Computing. Springer-Verlag Berlin Heidelberg. vol. 377. pp 383-391. ISSN 2194-5357. 2015.
- [7] Lanzarini L., Villa Monte A., Fernandez Bariviera A., Jimbo Santana P. *Simplifying Credit Scoring Rules using LVQ+PSO*. : The International Journal of Systems& Cybernetics. Emerald Group Publishing Limited. vol. 46. Pp. 8-16. ISSN 0368-492X. 2017.
- [8] Basgall, M. J., Hasperué, W., Estrebou C., Naiouf M. *Clustering de un flujo de datos usando MapReduce*. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). Pp. 682-691. ISBN 978-987-733-072-4. Octubre 2016.
- [9] Basgall, M. J., Hasperué, W., Estrebou C., Naiouf M. *Data stream treatment using sliding windows with MapReduce*. Journal of Computer Science & Technology. Vol. 16. ISSN 1666-6038. pp. 76-83. 2016.
- [10] Aquino G., Lanzarini L. *Keyword identification in spanish documents using neural networks..* Journal of Computer Science and Technology, JCS&T. ISSN: 1666-6046. Volumen 15. Número 2. pp. 55-60. Noviembre 2015.
- [11] Aquino G., Hasperué W., Lanzarini L. *Keyword Extraction using Auto-associative Neural Networks*. XX Congreso Argentino de Ciencias de la Computación (CACIC 2014) - ISBN 978-987-3806-05-6. pp.562-570. 2014.
- [12] Aquino, G, Hasperué, W, Estrebou, C, Lanzarini, L. *A Novel Language-Independent Keyword Extraction Method*. Publicado en el Libro Computer Science & Technology Series – XIX Argentine Congress of Computer Science - Selected Papers., 2014. pp.221-232
- [13] Villa Monte A., Lanzarini L., Rojas L., Olivás Varela J.A. *Document summarization using a scoring-based representation*. 2016 XLII Latin American Computing Conference (CLEI). 2016, pp. 1-7.

Modelo de Análisis de Información Desestructurada Utilizando Técnicas de Recopilación y Minería Web

Karina B. Eckert^a, Fabián E. Favret^b, Roberto Suénaga^c, Tokuji Kairiyama^d

Universidad Gastón Dachary

Avda. López y Planes 6519, Posadas, Misiones-Argentina. Tel: +54 (0376) – 447699

{karinaeck^a, fabianfavret^b}@gmail.com, {rsuenaga^c, kairiyama^d}@ugd.edu.ar

Resumen

Debido al gran avance en las Tecnologías de la Información y las Comunicaciones (TICs) se ha facilitado y simplificado de manera significativa el acceso, el procesamiento y el almacenamiento de los datos las organizaciones en general. Estas tecnologías han cambiado el paradigma del análisis de la información, ya que hoy en día el problema no es la escasez, sino la excesiva cantidad de datos disponibles que, claramente, no todos son de utilidad para el proceso de Toma de Decisiones (TD). Otro inconveniente, además del gran volumen de información que se necesita analizar, es que los datos disponibles están desestructurados y distribuidos, dificultando aún más la tarea de detectar aquellos que pueden ser útiles. Este proyecto tiene como objetivo desarrollar una herramienta para búsqueda de información para el proceso de TD.

Palabras clave: Toma de decisiones, minería web, análisis información desestructurada

Contexto

La presente investigación se encuentra en ejecución, acreditado en la Secretaría

de Investigación y Desarrollo de la UGD por resolución N° 07/A/17.

Introducción

Sistemas de Apoyo para el Proceso de TD

Un sistema eficiente para dar apoyo al proceso de TD debe contener las funcionalidades necesarias para el análisis y recopilación de requerimientos de usuario, recolección de la información en base al análisis eficiente de la misma y la presentación adecuada de los resultados.

Un gran número de proyectos de desarrollo de aplicaciones fracasan por no realizar una adecuada definición, especificación, y administración de los requerimientos.

Problemas como la falta de definiciones del usuario, requerimientos incompletos y el mal manejo de los cambios en los requerimientos son factores determinantes [1]. Existen varias técnicas propuestas para recopilar requerimientos como entrevistas y cuestionarios, prototipos o simulaciones del posible sistema o incluso casos de uso [2]. Claramente todas estas técnicas requieren la interacción con el usuario para determinar de manera correcta las necesidades de información. Una vez que se han establecidos los requerimientos del usuario, el sistema debe comenzar a

buscar la información que pueda ser de utilidad para el proceso de TD.

Claramente, para este tipo de actividad, la fuente más importante de información es la Web. La gran cantidad de información desestructurada y distribuida en Internet es una complicación al momento de decidir qué datos son relevantes y cuáles no. Para ello existen múltiples técnicas y métodos que derivan del área del aprendizaje automático y que se utilizan para hacer Minería Web (MW) [3][4].

La MW tiene como objetivo descubrir información útil e intentar extraer conocimiento de la estructura de hipervínculos, de los contenidos de las páginas y de los datos de uso. La MW utiliza muchas técnicas de Minería de Datos (MD) [5], no es puramente una aplicación de MD tradicional debido a la heterogeneidad y la naturaleza semi-estructurada o no estructurada de los datos de la Web. Sobre la base de los tipos principales de datos utilizados en el proceso de minería, las tareas de MW se pueden clasificar en tres tipos: la minería de la estructura de la Web, la minería de contenidos Web y la minería del uso de la Web [3].

El proceso de MW es similar al proceso de MD, la diferencia radica en el proceso de obtención de los datos. En la MD, los datos se colectan y luego se guardan en un almacén de datos. En la MW, la recopilación de datos puede ser una tarea demasiado compleja, sobre todo para determinar la estructura y luego analizar el contenido, tarea que implica el rastreo de un gran número de fuentes de información.

Existen múltiples algoritmos y técnicas que se utilizan para determinar la relevancia de la información y ellos se basan en extracción de conocimiento de los datos. En este sentido, la obtención de reglas que describen tanto las preferencias

como el comportamiento de los usuarios se realiza mediante el análisis de uso de los recursos web, por ejemplo mediante reglas de asociación [6][7] y minería de patrones secuenciales [8][9][10].

Este tipo de técnicas intentan encontrar regularidades en los datos sin tener una referencia explícita externa que dirija la búsqueda. Es decir, se trabaja directamente con los datos resultantes de la interacción del usuario con los recursos web.

También existen algoritmos como los árboles de decisión [11], naive bayes [12][13] y las máquinas de vectores de soporte [14] que son técnicas basadas en robustos principios matemáticos y que han sido muy utilizados para la clasificación de datos. Estos algoritmos, así como algunas redes neuronales [15], requieren la utilización de etiquetas asociadas a los datos. Otro grupo de algoritmos muy utilizados en el análisis de información son los denominados algoritmos de clustering o agrupamiento [16][17]. Este tipo de técnicas tienen como objeto agrupar datos por similitud utilizando alguna función para medir la distancia entre ellos. La idea subyacente es que los datos similares deben estar cerca ya que comparten características similares y es por ello que se intenta medir la mencionada distancia que los separa.

Muchos de estos algoritmos son utilizados para recuperar información de la Web, como en la utilización de Crawlers (programas que recorren la estructura de los hiperlinks), analizadores de hiperlinks y contenido web y analizadores de estructura de la Web. La implementación de estas técnicas otorga valor significativo al proceso de búsqueda, análisis y selección de la información útil para el proceso de TD.

Al obtener información de la Web, es necesaria que sea presentada de manera

adecuada al usuario. Claramente, cuando más precisa y clara sea la forma de mostrar la información, se podrá sacar mayor provecho de la información. Uno de los inconvenientes al presentar la información de relevancia es la integración que presenta la misma [3]. Debido a diversas fuentes de información en la Web, los diferentes sitios suelen utilizar diferentes palabras o términos para expresar la misma información o similar. Con el fin de hacer uso de los datos extraídos de varios sitios, se necesita integrar mediante técnicas semánticas la información de estos sitios. De esta manera se intenta que coincidan los datos que son semánticamente lo mismo pero expresan de manera diferente en distintos sitios [18].

Otro aspecto a considerar en la presentación de la información al usuario es la visualización de datos, que tiene por objeto comunicar la información de forma clara y eficaz a través de la representación gráfica [5]. Se pueden utilizar las técnicas de visualización para exponer conocimiento que de otra manera no sería fácilmente observable mediante el examen de los datos en bruto.

Si bien existen muchas técnicas de visualización de los datos [19] como la proyección geométrica, la visualización jerárquica, la basada en íconos o en relaciones de datos, entre otras, todas tienen como objetivo destacar los aspectos relevantes de la información. De esta manera, las herramientas de visualización son parte esencial en los sistemas para dar apoyo al proceso de TD.

Sector productivo Té

La falta de información útil para el proceso de TD afecta a todas las organizaciones, pero las más vulnerables son las pequeñas que están inmersas en economías regionales, caracterizadas por ser productoras de bienes primarios y no

tener recursos destinados para adaptarse y/o reconvertirse para hacer frente a los cambios del entorno.

El Clúster de Té de Misiones o Conglomerado Productivo Tealero de Misiones se constituye en el año 2006 en el marco del programa FONTAR (Fondo Tecnológico Argentino) donde mediante el Plan de Mejora de la Competitividad (PMC) se analizó las necesidades del sector, se detectaron los factores críticos de éxitos y se propuso diferentes líneas de acción con sus propuestas superadoras.

En el año 2011 se realizó una actualización del PMC donde se analizó el enfoque estratégico contrastando la mirada de los actores locales del té de Misiones con la percepción de los compradores nacionales e internacionales, que en base al diagnóstico surgieron nuevas estrategias para cada eslabón.

Entre las necesidades detectadas dentro del sector del té, se destaca tener acceso a información sobre: nuevos mercados, alternativas sobre agregación de valor al té y conocimiento del mercado internacional. La información permitirá tomar decisiones acerca del análisis para el reposicionamiento del té en el mercado local e internacional, la estrategia comercial de toda la cadena productiva y definir una caracterización del producto en la región que permita reposicionar respecto al mercado regional e internacional.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto abarca cinco etapas:

1. Obtención de información respecto a las técnicas de recopilación de necesidades, búsqueda automática, exploración y MW, y herramientas de TD.. Esta etapa se divide en dos actividades principales: (i) relevamiento

de información específica sobre las técnicas de MW, búsqueda automática y sistemas de soporte para las decisiones y (ii) análisis sistematizado de la información relevada a fin de determinar el estado del arte de los algoritmos y técnicas sobre nuevas formas de búsqueda de información.

2. Recopilación de información detallada del sector productivo del té, incluyendo un análisis de la situación actual y del contexto (programas de mejora, alternativas de financiamiento, proyectos vigentes, etc.). Se estudian técnicas y algoritmos de identificación y análisis de información contextualizada a requerimientos específicos (demandas sectoriales).

3. Estudio, análisis y desarrollo de las herramientas que serán integradas en el modelo de recolección de información. La etapa se enfoca en la relación, influencia y resultados del proceso de integración que contendrá la herramienta a desarrollar. Específicamente esta etapa tiene dos actividades prioritarias: (i) un estudio integral de las herramientas a desarrollar e implementar y (ii) el desarrollo e implementación de las técnicas relacionadas con la captura de requerimientos, la búsqueda automática y la presentación de la información al usuario.

4. Formulación de los modelos de negocio a partir de la utilización de la herramienta en el ámbito productivo del té de Misiones. En esta etapa se pretende que el empleo de la herramienta desarrollada aporte características de valor en los nuevos modelos de negocio que puedan ser utilizados en el sector productivo del té.

5. Evaluación de los resultados obtenidos a partir de la implementación de la herramienta para la obtención de información útil en el ámbito del sector productivo del té. En esta etapa se

trabjará con el sector productivo del té para determinar el impacto potencial que generen los modelos de negocio desarrollados.

Objetivos

Objetivo general

Desarrollar y utilizar un modelo para el proceso de toma decisiones basado en la integración de técnicas de recopilación, exploración y análisis de información.

Objetivos Específicos

- Estudiar y analizar técnicas de manejo de avanzado de datos (clustering, reconocimiento de patrones, modelos descriptivos y predictivos).
- Definir y analizar los modelos de negocio actuales del sector productivo del té de Misiones.
- Analizar métodos de búsqueda y exploración de información en la Web.
- Integrar métodos de recopilación, exploración y análisis de información en una herramienta informática de entorno web.
- Utilizar la herramienta para obtener información para la toma de decisiones del sector productivo del té, que permitan generar escenarios potencialmente convenientes para mejorar la competitividad.

Formación de Recursos Humanos

El equipo del trabajo de investigación está compuesto por un Dr. en Ingeniería de Sistemas y Computación, un magister en administración de empresas especialización en marketing, dos maestrando, uno en redes de datos y el otro en tecnologías de la información.

Director del Proyecto:

Dr. Marcelo Karanik

Co-Director:

Ing. Roberto Suénaga

Docentes-Investigadores:

Ing. Karina Eckert

Ing. Fabián Favret

Mg. Tokuji Kairiyama

Referencias

- [1] I. Sommerville, *Software Engineering*. United Kingdom: Addison-Wesley, 2005.
- [2] M. A. Chaves, “La ingeniería de requerimientos y su importancia en el desarrollo de proyectos de software,” *InterSedes*, vol. 6, no. 10, 2005.
- [3] Bing Liu, *Web Data Mining, Exploring Hyperlinks, Contents, and Usage Data*, First Edition. Chicago: University of Illinois, 2007.
- [4] P. K. P. Ranout and A. S. P. Sharma, *Web Mining-Concept, Classification and Major Research Issues: A Review*, vol. 4, 2 vols. 2016.
- [5] H. J. J. Pei and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
- [6] A. R. T Imieliński and A. Swami, “Mining association rules between sets of items in large databases,” *Acm sigmod record*, vol. 22, 1993.
- [7] Z. Z. R. Kohavi and L. Mason, “Real world performance of association rule algorithms,” *ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. ACM*, 2001.
- [8] R. Srikant and A. Rakesh, “Mining sequential patterns,” *Data Eng. Proc. Elev. Int. Conf. IEEE*, 1995.
- [9] J. Ayres, “Sequential pattern mining using a bitmap representation,” *Proc. Eighth ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. ACM*, 2002.
- [10] J. R. A. McCallum, “Using reinforcement learning to spider the web efficiently,” *ICML*, vol. 99, 1999.
- [11] J. Ross Quinlan, *C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning)*. 1992.
- [12] P. D. M. Pazzani, “On the optimality of the simple Bayesian classifier under zeroone loss,” *Mach. Learn.*, vol. 29, p. 3, 1997.
- [13] K. R. B. Becker and D. Sommerfield, “Improving simple bayes,” *Proc Eur. Conf. Mach. Learn. ECML '97*, 1997.
- [14] B. E. B. I. M. Guyon and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” *Proc. Fifth Annu. Workshop Comput. Learn. Theory ACM*, 1992.
- [15] S. S. Haykin, *Neural Networks and Learning Machines: A Comprehensive Foundation*, 3rd ed. USA: Pearson, 2009.
- [16] A. K. J. R. C. Dubes, *Algorithms for clustering data*. Prentice-Hall, 1988.
- [17] X. R. D. Wunsch, “Survey of clustering algorithms,” *IEEE Trans. Neural Netw.*, vol. 16, no. 3, 2005.
- [18] P. R. P. Heiko, “Semantic Web in data mining and knowledge discovery: A comprehensive survey,” *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 36, pp. 1–22, 2016.
- [19] D. King, “Introduction to the Mining, Analysis and Visualization of Web Content and Usage Minitrack,” *49th Hawaii Int. Conf. Syst. Sci. HICSS IEEE*, 2016.

Modelo para Predecir la Cantidad de Graduados de Ingeniería de UTN Aplicando Técnicas de Minería de Datos

Carrizo Claudio, Saldarini Javier, Ribotta Gabriela, Cardona Fernando, Marotti Juan Ignacio
 Grupo de Investigación Base de Datos - Facultad Regional San Francisco
 Universidad Tecnológica Nacional
 Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
 {cjarrizo77, saldarinijavier, garibotta, ferdcardona, nachomarotti} @gmail.com

RESUMEN

Desde la década del 60, la población mundial ha crecido en forma exponencial, este crecimiento ha generado problemas en sistemas energéticos, sanitarios, telecomunicaciones, infraestructura, etc.

La ingeniería cumple un factor fundamental para el desarrollo económico y el bienestar social de un país. En el ámbito de la República Argentina, será necesario contar, entre otros, con una mayor cantidad de graduados en carreras de ingeniería para incrementar el desarrollo industrial, la innovación productiva y la expansión económica.

En nuestro país, la Universidad Tecnológica Nacional (UTN) aporta el 42,75% de los ingenieros, por lo que esta cifra la posiciona como una de las principales Universidades formadora de ingenieros de Argentina.

El propósito del proyecto consiste en la elaboración de un modelo que permita, por un lado, predecir la cantidad de graduados de carreras de ingeniería en el ámbito de la UTN, y por otro lado, que este mismo modelo permita identificar

cuáles son los patrones que determinarán la graduación de estudiantes de ingeniería en un plazo promedio de 8 años.

Los resultados de este proyecto representarán un aporte para la gestión académica en lo respecta a la planificación, seguimiento y control de las cohortes de los estudiantes de carreras de ingeniería.

Palabras clave: Modelo Predictivo - Minería de Datos - Ingeniería - UTN

CONTEXTO

La Unidad Científico Tecnológica donde se enmarca el presente proyecto es el Grupo de Investigación "Base de Datos". Esta Unidad desarrolla sus líneas de investigación en concordancia con las áreas prioritarias correspondientes a la carrera de Ingeniería en Sistemas de Información de la UTN Facultad Regional San Francisco, las cuales son:

- Gestión de procesos de negocios
- Ingeniería de software

- Gestión y tecnologías de las organizaciones
- Calidad de Software
- Seguridad de la información
- Bases de datos

Estas áreas se encuentran formalizadas a través de la Resolución de Consejo Directivo N 353/2016 de la Facultad Regional San Francisco.

El presente proyecto de investigación y desarrollo se encuentra homologado y financiado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional bajo el código UTN3936 y según la Disposición SCTyP N° 380/15, el periodo de ejecución es desde Enero de 2016 hasta Diciembre de 2018, el mismo está incluido en el Programa I&D + i de Tecnología de las Organizaciones de la Universidad Tecnológica Nacional.

1. INTRODUCCIÓN

Desde el año 1960, la población mundial ha crecido más del doble [1], esto ocasionó un incremento en la demanda de servicios como agua potable, energía, comunicaciones, infraestructura, etc. Este crecimiento además trajo serios inconvenientes en lo que respecta al cuidado del medio ambiente, desafíos importantes en lo que hace a telecomunicaciones y el colapso de sistemas energéticos y sanitarios.

La ingeniería es un factor clave y prioritario para el desarrollo económico y el bienestar social [2]. Sin dudas, los ingenieros son una parte fundamental a la

solución de estos problemas, a través del desarrollo tecnológico y la innovación. Según se puede observar [3], los países que apuesten a la formación de ingenieros, serán aquellos con mayor índice de industrialización y desarrollo. También hay aportes de datos estadísticos que indican que en China hay 1 ingeniero cada 2000 personas, en Alemania 1 cada 2300, en Brasil 1 cada 6000 y finalmente en Argentina 1 de cada 6600.

En nuestro país, la evolución de la industria, mercados internacionales aun no explotados, y una creciente vinculación entre empresas y universidades, hacen que las ingeniería sean esenciales para consolidar el desarrollo industrial, relacionar conocimiento con innovación productiva, y disminuir los niveles de dependencia tecnológica [4]. En este sentido, es necesario incrementar la cantidad de graduados de ingeniería por cantidad de habitantes, según está establecido en el "Plan Estratégico de Formación de Ingenieros (PEFI)", cuyo objetivo principal es lograr 1 ingeniero cada 4000 habitantes [4].

Según un informe realizado por la Secretaría de Políticas Universitarias (SPU), la Universidad Tecnológica Nacional (UTN) forma alrededor del 42.75% de los ingenieros que se gradúan en nuestro país, esta cifra marca la importancia relativa que tiene la UTN respecto de las demás universidades en carreras de Ingeniería y la posiciona a como la Universidad de Ingeniería más grande del país [5].

En vista de la necesidad de contar con un mayor número de graduados de Ingeniería en nuestro país, el objetivo principal del proyecto consiste en elaborar un modelo que permita predecir

estimativamente la cantidad de graduados de ingeniería que puede proveer la UTN para el año 2021. Además este modelo permitirá identificar cuáles son los patrones que determinarán la graduación de estudiantes de ingeniería en un plazo promedio de 8 años.

Para llevar adelante este proyecto se ha seleccionado la temática "Minería de Datos", específicamente lo que respecta a técnicas de minería de datos de tipo predictivas [6]. La minería de datos se origina a principios de los años 80 y es un campo de las ciencias de la computación referido al proceso del descubrimiento de patrones en grandes volúmenes de conjuntos de datos. Está compuesta por técnicas, las cuales provienen de la rama de la inteligencia artificial y de la estadística, y no son más que algoritmos que se aplican sobre un conjunto de datos para obtener resultados [7]. Estos algoritmos se clasifican en Supervisados o Predictivos (entre los más representativos se encuentran los modelos de regresión, árboles de decisión, redes neuronales) y No Supervisados o del descubrimiento del conocimiento (entre los más representativos se encuentran los la segmentación y el análisis clúster). . Dentro de los algoritmos supervisados o predictivos, según lo expuesto por investigaciones de diferentes autores [8] [9] [10], la técnica de minería de datos que más se adapta para la construcción de un modelo predictivo es la de árboles de decisión con sus algoritmos C4.5 y ID3.

Existen en el mercado herramientas de minería de datos que permiten el análisis de datos y construcción de modelos predictivos, algunas de ellas son de tipo comerciales, como por ejemplo, SPSS Clementine [11], SAS Enterprise Miner [12], mientras que otras son de tipo Open

Source, como ejemplo, Weka [13] y Rapid Miner [14], entre otras.

Si bien existen muchos trabajos de investigación referidos a la predicción del rendimiento y deserción universitaria aplicando técnicas de minería de datos [15] [16], hay muy pocos trabajos enfocados en la obtención de modelos que permitan, por un lado, predecir la cantidad de graduados en un plazo de tiempo determinado para el contexto universitario y más específicamente para carreras de ingeniería, y por otro, que también estos modelos permitan identificar cuáles son los patrones que determinarán que los estudiantes de ingeniería puedan graduarse en un plazo promedio de 8 años.

El modelo resultante se transformará en un instrumento para la gestión, seguimiento, planificación y control en el área académica, posibilitando no sólo la predicción de la cantidad de graduados en carreras de ingeniería, sino que también, constituirá una herramienta de utilidad para ejecutar acciones concretas tomando como referencia los patrones identificados en el modelo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Una de las áreas prioritarias establecidas por la especialidad Sistemas de Información en la política de desarrollo de Ciencia y Tecnología de la Facultad Regional San Francisco es la de Base de Datos. Dentro de este contexto, el presente proyecto aborda la temática "Minería de Datos".

Los ejes de trabajo para esta línea de investigación se detallan a continuación:

- Caracterización y selección de las técnicas de minería de datos más adecuadas para llevar a cabo la predicción.
- Caracterización y selección de las herramientas que permitan analizar y procesar los datos para llevar a cabo la predicción.
- Construcción del modelo predictivo a través de la herramienta y técnica de minería de datos seleccionada.

3. RESULTADOS OBTENIDOS/ESPERADOS

El resultado esperado es un modelo que permita determinar estimativamente la cantidad de graduados en carreras de Ingeniería que puede proveer la UTN para el año 2021 e identificar los patrones que permiten determinar la graduación de estudiantes de ingeniería en un plazo promedio de 8 años. Para poder lograr este resultado, se han propuesto los siguientes objetivos:

- Construir el perfil del estudiante de ingeniería de UTN.
- Caracterizar y seleccionar las técnicas de minería de datos más adecuadas para llevar a cabo la predicción.
- Caracterizar y seleccionar las herramientas que permitan analizar y procesar los datos para llevar a cabo la predicción.
- Determinar la cantidad de graduados de ingeniería de UTN para el año 2021
- Identificar los patrones que permiten que los estudiantes puedan graduarse en

carreras de ingeniería en un plazo de 8 años.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está constituido por un director de proyecto, un co- director, dos docentes investigadores y un alumno de la especialidad Sistemas de Información que está cursando actualmente el quinto nivel de la carrera Ingeniería en Sistemas de Información y antes de fin de año comenzará con su tesina de grado.

Uno de los docentes investigadores tiene a su cargo la cátedra Gestión de Datos en la especialidad Sistemas de Información, en donde en una de las unidades del contenido curricular se trata la temática Minería de Datos. Los demás docentes que participan del proyecto también están dictando cátedras que tienen relación con la temática Base de Datos. Por su parte, el alumno mencionado anteriormente, viene participando y desarrollando tareas de investigación desde el año 2015. Dos de los docentes están desarrollando actualmente su tesis de maestría en calidad de software.

5. BIBLIOGRAFÍA

- [1] Día Mundial de la Población. Recuperado el 02 de febrero de 2016. https://www.unicef.org/lac/flash/DW/dia_mundial_de_la_poblacion.htm
- [2] La Ingeniería y su papel fundamental en la sociedad. Recuperado el 15 de febrero de año 2016. <http://www.ingenieriayempresa.org/ingenieria-sociedad/>

- [3] ¿Por qué faltan?. Recuperado el 21 de Marzo de 2016. <http://www.revistanueva.com.ar/portal/verNota/248>
- [4] Plan Estratégico de Formación de Ingenieros (PEFI). Recuperado el 4 de abril de 2016. <http://pefi.siu.edu.ar/>
- [5] La UTN forma más del 40% de los ingenieros que se gradúan en el país. Recuperado el 20 de abril de 2016. <https://www.frba.utn.edu.ar/dia-de-la-ingenieria-la-utn-forma-mas-del-40-de-los-ingenieros-que-se-graduan-en-el-pais/>
- [6] María Pérez Márquez. Minería de datos a través de ejemplos. México, 2015. Edit. Alfaomega.
- [7] Weiss y Indurkha. Predictive Data Mining: A Practical Guide, San Francisco, 1998.
- [8] D. Heredia y W. Nieto. Generación de un modelo predictivo para determinar el desempeño académico en la asignatura fundamentos de programación II del programa de Ingeniería de Sistemas. Colombia, 2011
- [9] S. Valero Oreal, A. Salvador Vargas y M. García Alonso. Minería de datos: predicción de la deserción es colar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos, 2010
- [10] E. Rodallegas R., G. Torres y B. Gaona C. Modelo Predictivo para la determinación de causas de reprobación mediante minería de datos. México, 2010.
- [11] Software SPSS Clementine de IBM Recuperado el 10 de marzo de 2017. <http://www-03.ibm.com/software/products/es/spss-modeler>
- [12] Software SAS Enterprise Miner Recuperado el 20 de mayo de 2016. https://www.sas.com/en_us/software/enterprise-miner.html
- [13] Weka 3: Data Mining Software in Java. Recuperado el 24 de mayo de 2016. <http://www.cs.waikato.ac.nz/ml/weka/>
- [14] Data Science Behind Every Decision. Recuperado el 6 de junio de 2016. <https://rapidminer.com/>
- [15] Yegny Karina Amaya Torradoa, Edwin Barrientos Avendañoa, Diana Judith Heredia Vizcaíno. Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos. Colombia.
- [16] Sergio Celis, Luis Moreno, Patricio Poblete, Javier Villanueva, Richard Weber. Un modelo analítico para la predicción del rendimiento académico de estudiantes de ingeniería. Chile, 2015.

Personalización de la Educación a través de Sistemas Recomendadores Dinámicos

María Emilia Charnelli^{1,3}, Laura Lanzarini², Javier Díaz¹

¹ Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

² Instituto de Investigación en Informática LIDI (III-LIDI).
Facultad de Informática. Universidad Nacional de La Plata.

³ Becario postgrado CONICET.

mcharnelli@linti.unlp.edu.ar, laural@lidi.info.unlp.edu.ar, javierd@info.unlp.edu.ar

Resumen

La línea de investigación actual continúa con el estudio, diseño y desarrollo de nuevas técnicas adaptativas, pertenecientes al área de la Analítica del Aprendizaje, que contribuyan en la toma de decisiones. El énfasis está puesto en la construcción de un Sistema Recomendador inteligente con capacidad para asistir en ámbitos educativos.

Como trabajo previo se analizaron y se aplicaron técnicas no supervisadas y técnicas de visualización de datos masivos con el objetivo de obtener un modelo que representara a cada uno los alumnos de la Facultad de Informática de la UNLP. Esto incluye su interacción con recursos educativos de acceso libre y su interacción con las redes sociales. Dichos perfiles serán utilizados para caracterizar su comportamiento actual y como insumo del algoritmo del sistema recomendador que se desarrollará para poder asistir a los alumnos en forma automática sobre materiales de lecturas adicionales a fin de que puedan alcanzar el comportamiento esperado.

Palabras clave: Analítica del aprendizaje, Sistemas Recomendadores, Personalización de la educación, Minería de Datos Educativa.

Contexto

Este trabajo de investigación se encuadra en el Proyecto de Incentivos acreditado “Internet del futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro” desarrollado en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI.

Introducción

La Analítica del Aprendizaje (Learning Analytics) es un campo de investigación emergente que analiza información referida a personas que están aprendiendo y sus contextos, con el propósito de comprender y optimizar el aprendizaje y los entornos en los que dicho aprendizaje sucede. En los últimos años, las instituciones educativas se han embarcado en su propia exploración de grandes conjuntos de datos para mejorar los índices de retención y proporcionar una experiencia personalizada y de mayor calidad para los estudiantes. La aplicación de técnicas de Minería de Datos en el ámbito educativo ha permitido caracterizar a los distintos actores que intervienen en los procesos de enseñanza y aprendizaje [1]. A medida que el aprendizaje se traslada a la Web, crece la

cantidad de datos que pueden utilizarse para mejorar y personalizar este proceso.

A continuación se describen las tareas realizadas recientemente.

Minería de Datos Masivos

A partir de las líneas de investigación presentadas en WICC 2016 [2], se continuó trabajando sobre los datos académicos y personales de los alumnos de la Facultad de Informática de la UNLP.

Además de la información personal y académica obtenida del sistema de gestión y los entornos de enseñanza y aprendizaje, resulta de interés obtener información adicional sobre la interacción de los alumnos con las redes sociales [3]. Para ello, se analizaron las publicaciones realizadas por los alumnos en grupos de Facebook. Estos grupos fueron creados por los mismos alumnos para discutir temas sobre la Facultad. Los datos fueron recolectados a través de la API Graph de Facebook e involucran más de 3000 publicaciones y más de 1500 alumnos que participan creando, comentando, compartiendo o poniendo “me gusta” a una publicación. Los contenidos extraídos son de texto libre y no están estructurados por lo que se debieron utilizar técnicas de minería de texto. El objetivo central consistía en descubrir cuáles eran los temas que más se trataban en los grupos de Facebook. Se analizó y utilizó BTM (Biterm Topic Model) [4] que es una técnica de aprendizaje no supervisado que descubre los tópicos que caracterizan a un conjunto de documentos breves. En este contexto, cada publicación y comentario es considerado como un documento. Estos resultados ayudan a encontrar cuáles son los temas de más interés de los alumnos desde otro contexto. Y de esta forma lograr modelar a los usuarios de acuerdo a su comportamiento en las redes

sociales. Los resultados de este trabajo fueron publicados en [5].

Modelado de Perfiles de Alumnos y Sistemas Recomendadores

Uno de los puntos centrales de esta línea de investigación consiste en modelar perfiles dinámicos a través de información provista por los entornos de enseñanza y aprendizaje que usan las diferentes cátedras de la facultad, información de los préstamos de la biblioteca, actividad en las redes sociales, entre otros [6] [7] [8] [9]. Estos perfiles permitirán proporcionar información para un sistema de recomendación [10] [11] [12]. Un algoritmo recomendador funciona de forma precisa si dispone de perfiles de usuarios bien construidos. Para esto los sistemas necesitan contar con la mayor cantidad de información posible de los usuarios para proveerles una recomendación razonable. En el ámbito educativo un perfil de usuario es una colección de información personal asociada a un usuario específico. Esta información incluye: habilidades cognitivas, estilos de aprendizaje, intereses, preferencias e interacciones con el sistema. Las redes sociales conforman un ámbito externo a la institución educativa pero disponen de información valiosa acerca de los intereses de los estudiantes. Es un tema de investigación actual los métodos que permiten extraer la información implícita acerca de los comportamientos de los alumnos en las plataformas sociales para obtener modelos dinámicos capaces de adaptarse fácilmente a los cambios de la información y que contribuyan a la toma de decisiones en ámbitos educativos.

Actualmente se está desarrollando un sistema recomendador que se aplicará como una extensión de los entornos de aprendizaje, que utilizando la información

de los perfiles obtenidos podrá recomendar materiales educativos. Se estudiaron y se evaluarán diferentes técnicas de Sistemas Recomendadores como los basados en contenido, basados en filtrado colaborativo, y técnicas híbridas.

Líneas de Investigación, Desarrollo e Innovación

- Estudio de distintas técnicas de aprendizaje no supervisado aplicables a grandes volúmenes de datos.
- Revisión y análisis de técnicas específicas de Learning Analytics.
- Estudio y desarrollo sobre modelado de perfiles dinámicos.
- Estudio y análisis de algoritmos de recomendación.
- Desarrollo de un sistema recomendador de materiales educativos utilizando la información de los perfiles obtenidos.

Resultados y Objetivos

- Detección y generación de características representativas de los alumnos desde diferentes contextos: académicos, personales y su interacción con redes sociales.
- Construcción de un modelo de alumnos a partir de sus temas de interés en las redes sociales.
- Desarrollo de un sistema recomendador inicial que utilice la información modelada de los alumnos para sugerirles recursos educativos abiertos.

Formación de Recursos Humanos

Dentro de los temas involucrados en esta línea de investigación, actualmente hay 2 investigadores realizando su doctorado, y uno de ellos a su vez, finalizando una maestría.

Referencias

- [1] Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- [2] Díaz, Francisco Javier; Lanzarini, Laura Cristina; Charnelli, María Emilia. Personalización de la educación a través de la creación de perfiles dinámicos de los alumnos. *XVIII Workshop de Investigadores en Ciencias de la Computación*. 2016, Entre Ríos.
- [3] KAYA, Tugberk; BICEN, Huseyin. The effects of social media on students' behaviors; Facebook as a case study. *Computers in Human Behavior*, 2016, vol. 59, p. 374-379.
- [4] Cheng, X., Yan, X., Lan, Y., Guo, J.: Btm: Topic modeling over short texts. *IEEE Transactions on Knowledge and Data Engineering* 26 (2014) 2928–2941
- [5] Charnelli, M E. Lanzarini, L. Diaz, F. Modeling Students through Analysis of Social Networks. *XXII Congreso Argentino de Ciencias de la Computación*. San Luis, 2016.
- [6] W. Paireekreng and T. Prexawanprasut, "An integrated model for learning style classification in university

students using data mining techniques", Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2015 12th International Conference on, Hua Hin, 2015, pp. 1-5.

[7] Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Kloos, C. D. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, 47, 139-148.

[8] Chrysafiadi, K., & Virvou, M. (2015). Student Modeling for Personalized Education: A Review of the Literature. In *Advances in Personalized Web-Based Education* (pp. 1-24). Springer International Publishing.

[9] Mota, P., Melo, F., & Coheur, L. (2015, May). Modeling Students Self-Studies Behaviors. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1521-1528). International Foundation for Autonomous Agents and Multiagent Systems.

[10] Drachsler, H., Verbert, K., Santos, O. C., & Manouselis. Panorama of recommender systems to support learning. In *Recommender systems handbook* (pp. 421-451). Springer US, 2015.

[11] Tejeda-Lorente, Á., Bernabé-Moreno, J., Porcel, C., Galindo-Moreno, P., & Herrera-Viedma, E. (2015). A Dynamic Recommender System as

Reinforcement for Personalized Education by a Fuzzly Linguistic Web System. *Procedia Computer Science*, 55, 1143-1150.

[12] LI, Yanyan, et al. Designing a Learning Recommender System by Incorporating Resource Association Analysis and Social Interaction Computing. *En State-of-the-Art and Future Directions of Smart Learning*. Springer Singapore, 2016. p. 137-143.

Procesamiento y Recuperación en Bases de Datos Masivas

Luis Britos, Veronica Gil-Costa, Fernando Kasián, Verónica Ludueña, Romina Molina, Marcela Printista, Nora Reyes, Patricia Roggero, Guillermo Trabes
LIDIC, Dpto. de Informática, Fac. de Cs. Físico Matemáticas y Naturales, Universidad Nacional de San Luis
lbritos, gvcosta, fkasian, vlud, mprinti, nreyes, proggero}@unsl.edu.ar
romy00@gmail.com, guillermotrabes@hotmail.com

Edgar Chávez

Centro de Investigación Científica y de Educación Superior de Ensenada, México

elchavez@cicese.mx

Claudia Deco

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario

deco@fceia.unr.edu.ar

Resumen

En la actualidad es cada vez más evidente la necesidad de procesar conjuntos de datos, de manera tal de poder obtener información útil a partir de ellos. Los sistemas de información demandan no sólo poder realizar búsquedas eficientes sobre distintos tipos de datos, tales como texto libre, audio, video, secuencias de ADN, etc., sino también poder manejar grandes volúmenes de estos datos. Dada una consulta, el objetivo de un sistema de recuperación de información es obtener lo que podría ser útil o relevante para el usuario, usando una estructura de almacenamiento especialmente diseñada para responderla eficientemente.

En esta línea de investigación el principal objetivo es desarrollar herramientas eficientes para sistemas de información sobre bases de datos masivas, conteniendo datos multimedia. Por lo tanto, se analizan nuevas técnicas que permitan una buena interacción con el usuario, nuevas estructuras de datos (índices) capaces de manipular eficientemente datos multimedia y que puedan utilizarse para administrar bases de datos multimedia masivas. Además, se busca desarrollar herramientas que permitan soportar tanto la recolección como el procesamiento de grandes volúmenes de datos multimedia.

Palabras Claves: *recuperación de información, computación de alto desempeño, bases de datos masivas.*

1. Contexto

Esta línea de investigación se encuentra enmarcada dentro del Proyecto Consolidado 3-30114 de la Universidad Nacional de San Luis (UNSL) y en el Programa de Incentivos (Código 22/F434): “Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos”, dentro de la línea “Recuperación de Datos e Información”, desarrollada en el Laboratorio de Investigación y Desarrollo en Inteligencia

Computacional (LIDIC) de la UNSL.

En este contexto, se pretende contribuir a la incorporación de información no estructurada en los procesos de toma de decisiones y resolución de problemas, no considerados en los enfoques clásicos. Por lo tanto, el objetivo principal es diseñar e implementar índices eficientes para conjuntos masivos de datos multimedia, que puedan servir de apoyo a sistemas de recuperación de información sobre distintos tipos de datos no estructurados. Para ello, se debe considerar que son necesarios índices más eficientes para memorias jerárquicas, dinámicos, con E/S eficiente, escalables (capaces de manejar grandes volúmenes de datos), considerando además técnicas de computación de alto desempeño (HPC).

2. Introducción y Motivación

Con el uso masivo de internet, estamos en presencia de un fenómeno donde la aceleración tanto del crecimiento del volumen de datos capturados y almacenados, como la creciente variación en los tipos de datos requeridos, hace que las técnicas tradicionales para el procesamiento, análisis y obtención de información útil deban ser redefinidas para formular nuevas metodologías de abordaje.

Los sistemas de computación tradicionales hacen uso intensivo de información estructurada; es decir, datos generados con un formato específico. En estos casos, la estructura o formato de esta información puede ser fácilmente interpretada y directamente utilizada por un programa de computadora. Pero el hecho de restringirse al uso de este tipo de información conduce, muchas veces, a representar una vi-

sión parcial del problema y dejar fuera información que podría ser relevante para la resolución efectiva del mismo. En este contexto gran parte de la información que se requiere para la toma de decisiones y la resolución de problemas de índole más general proviene de información no estructurada.

Habitualmente, se utilizan diferentes métodos de acceso o índices [2] para responder eficientemente a consultas para recuperación de información sobre bases de datos multimedia, principalmente por la gran cantidad de datos con los que se trabaja. Los índices pueden tener distintas características que los hacen indicados para aplicaciones reales: eficientes, dinámicos, escalables, resistentes a la *maldición de la dimensión*, entre otras. Un enfoque prometedor para sistemas de recuperación usando búsqueda por similitud es “la búsqueda basada en contenidos”, la cual usa el dato multimedia mismo. Para calcular la similitud entre dos objetos multimedia, se debe definir una función de distancia. Dicha función mide la disimilitud entre dos objetos.

El concepto de espacios métricos da un marco formal, independiente del dominio de la aplicación, para definir el concepto de búsqueda por similitud. Un espacio métrico está compuesto por un universo U de objetos y una función de distancia $d: U \times U \rightarrow \mathbb{R}^+$, que satisface las propiedades que la hacen una métrica. Las consultas por similitud sobre una base de datos S son básicamente de dos tipos: *búsqueda por rango* y *búsqueda de los k vecinos más cercanos*. La función de similitud (distancia) mide el mínimo esfuerzo (costo) necesario para transformar un objeto en otro. Dependiendo de los tipos de datos multimedia reales el cálculo de dicha función puede ser muy costoso. En particular, para ahorrar cálculos de distancia es importante que dicha distancia satisfaga la desigualdad triangular.

Si la base de datos S posee n objetos, las consultas se pueden responder llevando a cabo n evaluaciones de distancia. Sin embargo, en la mayoría de las aplicaciones, las distancias son costosas de computar (por ej.: comparación de huellas digitales). En conjuntos masivos de datos la búsqueda secuencial es impráctica y, en general, los repositorios de datos multimedia son grandes volúmenes de datos. Para responder a las consultas con la menor cantidad de cálculos de distancia se debe preprocesar la base de datos para construir un índice. En algunos casos, es probable que la base de datos, el índice, o ambos, no puedan almacenarse en memoria principal. Por lo tanto, para lograr eficiencia, se debe minimi-

zar el número de operaciones de E/S, considerar la jerarquía de memorias, en algunos casos admitir respuestas no exactas y utilizar técnicas paralelas.

Esta propuesta se enfoca en obtener herramientas de recuperación de información, desarrollando nuevas técnicas y aplicaciones que soporten la interacción con el usuario, diseñando estructuras de datos (índices), capaces de manipular eficientemente grandes volúmenes de datos no estructurados y facilitando la realización de diferentes consultas, de modo de acercar las bases de datos multimedia al nivel de desarrollo de las bases de datos tradicionales.

3. Líneas de Investigación

Se pretende investigar sobre distintos aspectos de los sistemas de recuperación de información multimedia sobre grandes volúmenes de datos: el diseño de nuevos índices, representaciones que reflejen características de interés de los objetos, distintas consultas sobre estos tipos de bases de datos y eficiencia al considerar grandes volúmenes de datos.

Procesamiento y Modelado de Grandes Volúmenes de Datos

Con el continuo crecimiento de la Web y las redes sociales, es cada vez mayor la cantidad de datos que se generan. Además de estos datos de aspecto social, también están los datos científicos generados por diferentes disciplinas como la biología, la física, la astronomía y la medicina. Estos datos provenientes de diferentes fuentes generalmente deben ser procesadas y analizadas en el menor tiempo posible.

Para abordar esta problemática, desde la UNSL, se han realizado investigaciones que analizan la posibilidad de utilizar arquitecturas heterogéneas basadas en dispositivos programables (FPGAs - Field Programmable Gate Array) y en sistemas sobre chip (SoCs - System-on-a-chip) para representar datos multimedia y procesar búsquedas de elementos complejos [5]. Estas plataformas tienen la característica de ser re-configurables y se pueden adaptar a medida para resolver un problema específico. De esta manera, logran reducir los tiempos de procesamiento. Adicionalmente, reportan un menor consumo de energía que las unidades de procesamiento gráfico (GPUs).

Por otro lado, el procesamiento masivo de datos debe ser acompañado de un modelo de computación que permita organizar las comunicaciones y los cómputos. En la literatura existen varios mode-

los de computación que intentan generar un puente entre la arquitectura y las aplicaciones. Sin embargo, muchos de estos modelos fallan debido a que su complejidad crece con el número de parámetros que utilizan para caracterizar las arquitecturas. En este sentido, el modelo BSP (Bulk Synchronous Parallel) [12] ha sido ampliamente aceptado y utilizado por gran parte de la comunidad científica, debido a que es un modelo simple que permite abstraer los costos de las arquitecturas. Este modelo BSP, originalmente fue propuesto para clusters de computadoras conectadas mediante una red de comunicación. En el año 2011, se presentó el modelo Multi-BSP [13] para arquitecturas multi-core.

Recientemente hemos utilizado estos modelos (BSP y Multi-BSP) para modelar algoritmos MapReduce. En particular, el objetivo es identificar los componentes que más costo tienen y el efecto de las sincronizaciones impuestas por el modelo [11]. Actualmente, se está estudiando el efecto que tienen las barreras de sincronización en arquitecturas multi-core, para evaluar la jerarquía más adecuada en que las sincronizaciones de componentes atómicos reducen los costos adicionales impuestos por el modelo de computación.

Esta línea de investigación involucra a dos tesis de doctorado, una beca de CONICET y dos tesis de maestría (una finalizada).

Índices para Bases de Datos Multimedia

Los índices que resultan apropiados, para luego realizar búsquedas sobre bases de datos multimedia, son los índices métricos [2]. Todos ellos aprovechan que la distancia satisface la propiedad de desigualdad triangular para ahorrar algunos cálculos de distancia. Esta propiedad permite estimar la distancia entre cualquier objeto de consulta q y los objetos de la base de datos, si se mantienen algunas distancias de los elementos de la base de datos a objetos distinguidos. Los distintos enfoques se diferencian en si esos objetos distinguidos son *pivotes* o *centros*. Si son pivotes se almacenan las distancias de todos los objetos de la base de datos a ellos. Si son centros se particiona el espacio en zonas denominadas *particiones compactas*, por cercanía a los centros y se almacena un radio de cobertura para determinar la zona de cada centro.

En nuestro caso, nos interesa poder diseñar buenos índices que consideren:

Dinamismo: Los índices pueden construirse de manera estática, si los objetos de la base de datos

se conocen de antemano. En estos índices denominados *estáticos* las búsquedas se realizan luego de construido el índice. Por el contrario, si no se pueden tener los objetos de antemano y la única manera de construir el índice es a medida que se incorporan los elementos a la base de datos; es decir, de manera incremental, se considera que las búsquedas pueden realizarse en cualquier momento. Esta clase de índices se denominan *dinámicos*. Los índices estáticos, por conocer a toda la base de datos, pueden seleccionar los mejores objetos distinguidos para una estructura de datos determinada. En cambio, en los índices dinámicos esto no es posible.

Jerarquía de Memorias: Otro aspecto importante para buscar una solución es saber si se puede trabajar en memoria principal o, por el contrario, si por ser conjuntos de datos masivos se deberá trabajar en otros niveles de la jerarquía de memorias. En caso que el índice deba alojarse en memoria secundaria, se deben minimizar la cantidad de cálculos de distancia y también el número de operaciones de E/S.

Computación de Alto Desempeño: En algunos casos, si no se logra la eficiencia deseada mediante la optimización del índice en sí mismo, se pueden aplicar técnicas de computación de alto desempeño con el fin de acelerar los tiempos de respuesta a las consultas.

Exactitud de la Respuesta: Otra manera de acelerar la respuesta a una consulta por similitud es admitir una respuesta aproximada, permitiendo que la misma sea de menor calidad o menos exacta, pero muy rápida.

Dimensionalidad Intrínseca: los índices para búsquedas por similitud, al trabajar sobre el modelo de espacios métricos, pueden también sufrir de la llamada *maldición de la dimensión*; es decir, los índices se degradan a medida que la dimensión de los espacios aumenta. Existen índices que se comportan mejor en espacios difíciles (dimensión intrínseca mediana a alta) y otros que son adecuados para espacios fáciles (dimensión intrínseca baja).

Considerando que nos interesa trabajar sobre conjuntos de datos masivos, los volúmenes de información con los que se debe trabajar (millones de imágenes en la Web) hace necesario que los índices sean almacenados en memoria secundaria. En este caso, para lograr eficiencia, no sólo se debe considerar que en las búsquedas se realice el menor número de cálculos de distancia sino también, dado el costo de

las operaciones sobre disco, se efectúe la menor cantidad posible de operaciones de E/S. Por ello, en esta línea nos hemos dedicado a diseñar índices especialmente adaptados para trabajar en memoria secundaria, que tengan buen desempeño principalmente en las búsquedas. Así, se ha diseñado e implementado una versión paralela del *Conjunto Dinámico de Clusters* (DSC) [8]. Este índice, basado en la *Lista de Clusters* (LC), está especialmente diseñado para

memoria secundaria y es completamente dinámico, admite inserciones y eliminaciones y tiene un buen desempeño en las búsquedas, principalmente en la cantidad de operaciones de E/S. DSC ha demostrado ser muy competitivo frente a otras de las buenas estructuras del estado del arte. Por lo tanto, se buscará aplicar y comparar distintas estrategias de paralelización con el fin de determinar la más adecuada.

Por otro lado, tomando como base al índice para búsquedas aproximadas *Lista de Permutaciones Agrupadas* (LPA), que combina un algoritmo basado en *Permutaciones* con una LC [4], se ha propuesto una nueva versión dinámica de la LPA que permite realizar búsquedas por similitud aproximadas sobre conjuntos de datos masivos [3]. Esta nueva versión de la LPA es consciente que trabaja en memoria secundaria y no sólo considera minimizar los costos en cantidad de distancias calculadas, sino también en cantidad y tipo de operaciones de E/S.

Existen en la actualidad pocas medidas que permitan reflejar adecuadamente la dimensionalidad intrínseca de los espacios métricos [2]. Sin embargo, si se pudiera calcular la dimensionalidad intrínseca de un espacio métrico con cierta confiabilidad, se podría elegir el índice que tuviera mejor desempeño en las búsquedas para esa dimensión en particular. Por lo tanto, se han propuesto nuevas medidas de evaluación de la dimensionalidad intrínseca y se las ha evaluado experimentalmente junto a otras medidas ya conocidas, para ver cuál de ellas puede reflejar de manera más confiable la dimensionalidad intrínseca de un espacio métrico [7].

En esta línea de investigación se están desarrollando dos tesis de maestría.

Sistema Administrador para Bases de Datos Multimedia

A pesar de que las operaciones más comunes sobre bases de datos multimedia son las búsquedas por rango o de k -vecinos más cercanos, existen otras operaciones de interés tales como las distintas variantes del *join* por similitud. La operación de *join* por similitud se considera una de las operaciones

que debería brindar típicamente un sistema administrador para bases de datos multimedia.

Existen diferentes variantes para el *join* por similitud, dependiendo del criterio de similitud Φ utilizado, pero ellas tienen en común que se aplican entre dos bases de datos A y B , ambos subconjuntos del mismo universo del espacio métrico U que modela a la base de datos multimedia. El resultado de cualquiera de las variantes del *join* por similitud entre A y B obtendrá el conjunto de pares formados por un objeto de A y otro de B , tales que entre ellos se satisface el criterio de similitud Φ considerado. Las variantes más conocidas son: el *join* por rango, el *join* de k -vecinos más cercanos y el *join* de k pares de vecinos más cercanos; entre otras.

Formalmente, dadas $A, B \subseteq U$, se define el *join por similitud* entre A y B ($A \star_{\Phi} B$) como el conjunto de todos los pares (x, y) , donde $x \in A$ e $y \in B$; es decir, $(x, y) \in A \times B$, tal que $\Phi(x, y)$ es verdadero (se satisface el criterio de similitud Φ entre x e y). Al resolver el *join* por similitud es posible que ambas, una o ninguna de las bases de datos posean un índice; o que ambas bases de datos se indexen conjuntamente con un índice diseñado para el *join*. Calcular cualquiera de las variantes del *join* por similitud de manera exacta es muy costoso [9], así como la pena analizar posibilidades de obtener una respuesta aproximada al *join*, más rápidamente, aunque siempre buscando buena calidad en la respuesta.

PostgreSQL es el primer sistema de base de datos que permite realizar consultas por similitud sobre algunos atributos, particularmente indexa para búsquedas de k -vecinos más cercanos (índices *KNN-GiST*). Estos índices pueden ser usados sobre texto, comparación de ubicación geoespacial, etc. Sin embargo, los índices *K-NN GiST* proveen plantillas sólo para índices con estructura de *árbol balanceado* (*B-tree*, *R-tree*), pero el "balance" no siempre es bueno para los índices que se utilizan en búsquedas por similitud [1]. Además, no se dispone de este tipo de consultas para todo tipo de datos métricos. Así, es importante proveer un DBMS para bases de datos métricas que maneje todos los posibles datos métricos y las operaciones de interés sobre ellos [6].

Más aún, dado que las respuestas a consultas de *join* suelen ser conjuntos muy grandes de pares de objetos y muchos de esos pares son muy similares entre sí, se planea introducir sobre las operaciones de *join* la posibilidad de diversificar las respuestas [10]; es decir, un operador de *join* por similitud que asegure un conjunto más pequeño, más diversifica-

do de respuestas útiles y, de ser posible, más rápido de obtener. Estos desarrollos, entre otros, permitirán tener un DBMS con mayores posibilidades de aplicación en sistemas de información reales.

Esta línea corresponde a una tesis de maestría.

4. Resultados

Se ha propuesto, diseñado y evaluado un sistema CBIR sobre una plataforma SoC basada en FPGAs [5]. Se ha implementado la versión paralela del índice DSC, que trabaja con grandes volúmenes de datos, diseñado especialmente para memoria secundaria, que admite inserciones y eliminaciones de elementos y que permitirá responder eficientemente a lotes de consultas por similitud. Se ha obtenido una versión para memoria secundaria y dinámica de la LPA [3]. Se continúa trabajando en la extensión de PostgreSQL para que brinde facilidades de soporte a más tipos de consultas por similitud, sobre distintos tipos de datos y que considere opciones de respuesta aproximada, como así también la posibilidad de obtener una respuesta diversificada en el caso de los joins por similitud.

5. Formación de Recursos

En esta línea se están realizando las siguientes tesis de posgrado en Ciencias de la Computación:

Tesis de Doctorado: (1) “Simulación Semi-asíncrona DEVs de Motores de Búsqueda Web” y (2) “Predicción de Tiempo de Ejecución de Consultas para Motores de Búsqueda Web”.

Tesis de Maestría: (1) “Estructuras Eficientes sobre Datos Masivos para Búsquedas en Espacios Métricos”, (2) “Cómputo Aproximado del Grafo de Todos los k -Vecinos”, (3) “Simulación Paralela Aproximada sobre S4 para Motores de Búsqueda en la Web”, (4) “Recuperación de Imágenes sobre Plataformas de Sistemas de Cómputo de Alta Productividad” y (5) “Sistema Administrador para Bases de Datos Métricas”.

Referencias

- [1] E. Chávez, V. Ludueña, and N. Reyes. Revisiting the VP-forest: Unbalance to improve the performance. In *Proc. de las JCC08*, page 26, 2008.
- [2] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM*, 33(3):273–321, Sept. 2001.
- [3] K. Figueroa, C. Martínez, R. Paredes, N. Reyes, and P. Roggero. Dynamic list of clustered permutations on disk. *Computer Science and Technology*, 201–211, 2016.
- [4] K. Figueroa and R. Paredes. List of clustered permutations for proximity searching. In *Similarity Search and Applications*, volume 8199 of LNCS, 50–58. Springer Berlin Heidelberg, 2013.
- [5] V. Gil-Costa, R. Molina, R. Petrino, C. Soza Paez, A. Printista, and J. Dondo Gazzano. *Hardware Acceleration of CBIR System with FPGA-Based Platform*, 138–170. Advances in Computer and Electrical Engineering. IGI Global, 2016.
- [6] F. Kasián and N. Reyes. Búsquedas por similitud en PostgreSQL. In *Actas del XVIII CACIC*, 1098–1107, Bahía Blanca, Argentina, Oct. 2012.
- [7] G. Navarro, R. Paredes, N. Reyes, and C. Bustos. An empirical evaluation of intrinsic dimension estimators. *Information Systems*, 64:206–218, 2017.
- [8] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48–78, 2016.
- [9] R. Paredes and N. Reyes. Solving similarity joins and range queries in metric spaces with the list of twin clusters. *JDA*, 7:18–35, March 2009. doi:10.1016/j.jda.2008.09.012.
- [10] L. F. D. Santos, L. Olmes Carvalho, W. D. Oliveira, A. J.M. Traina, and C. Jr. Traina. Diversity in similarity joins. In *Similarity Search and Applications*, volume 9371 of LNCS, 42–53. Springer International Pub., 2015.
- [11] H. Senger, V. Gil-Costa, L. Arantes, C. A. C. Marcondes, M. Marín, L. M. Sato, and F. A.B. da Silva. Bsp cost and scalability analysis for mapreduce operations. *Concurrency and Computation: Practice and Experience*, 28(8):2503–2527, 2016. cpe.3628.
- [12] Leslie G. Valiant. A bridging model for parallel computation. *Commun. ACM*, 33(8):103–111, Aug. 1990.
- [13] Leslie G. Valiant. A bridging model for multi-core computing. *J. Comput. Syst. Sci.*, 77(1):154–166, January 2011.

Propuesta de Procesos Complementarios para un Sistema de Recuperación de Información

Rey, M., Kuna, H., Rambo, A., Canteros, A., Cantero, A., Martini, E., Corrales, N., Rauber, F.

1. Depto. de Informática, Facultad de Ciencias Exactas Quím. y Naturales, Universidad Nacional de Misiones.

hdkuna@gmail.com

RESUMEN

Con la finalidad de mejorar el proceso de búsqueda de información para investigadores por medio de la utilización de un Sistema de Recuperación de Información (SRI) específico de las ciencias de la computación, es necesario definir métodos que trabajen sobre el análisis de los datos existentes para generar operaciones que incrementen la relevancia de los resultados a presentar a los usuarios. Entre las alternativas de técnicas aplicables se destacan: el tratamiento de tópicos clave, técnicas de clustering y de análisis de probabilidad basadas en Bayes, entre otras. En el presente trabajo se exponen propuestas de procesos a través de las que se busca demostrar que es posible utilizar distintas técnicas para generar procesos que, a partir de datos disponibles en el sistema, proporcionen información para mejorar la calidad de los resultados y operaciones internas de la herramienta, siendo este el objetivo principal de la presente línea de investigación.

Palabras clave: *información científica, meta-buscador, clustering, topic modelling, análisis de probabilidad.*

CONTEXTO

Esta línea de investigación articula el “Programa de Investigación en Computación” de la Facultad de Ciencias Exactas Químicas y Naturales (FCEQyN) de la Universidad Nacional de Misiones (UNaM) con el Grupo de Investigación Soft Management of Internet and Learning (SMILe) de la Universidad de Castilla-La Mancha, España.

1 INTRODUCCION

1.1 Antecedentes

En trabajos anteriores se ha presentado un meta-buscador orientado a la recuperación de información científica correspondiente al área de ciencias de la computación [1]. A lo largo del proceso de desarrollo de esta herramienta se han integrado componentes y módulos cuya función consiste en maximizar la relevancia de los resultados a presentar al usuario [2].

Sin embargo, se considera que el desarrollo e integración de procesos cuya operatoria explote los datos almacenados por el meta-buscador es el siguiente paso en el desarrollo de la solución. Considerando que el objetivo central de cualquier SRI es maximizar la relevancia de los resultados a presentar a su usuario, la evolución del meta-buscador desarrollado pasa por la integración de funciones anexas a las básicas de búsqueda y evaluación [3]. Procesos como los planteados en el presente trabajo se consideran de gran utilidad para la recuperación de información de mayor utilidad.

1.2 Gestión de datos del SRI

En trabajos anteriores se han detallado algunos de los inconvenientes detectados en el meta-buscador. En la mayoría de los casos se trató de operaciones en las que se identificaba un problema común, la falta de datos almacenados internamente por el SRI tanto como resultado de su operatoria como de la interacción con el usuario. Se determinó que sin esta capacidad el desarrollo de procesos complementarios a los básicos del SRI sería una tarea de gran complejidad [2].

En este sentido, se implementó un esquema para la representación de los meta-datos de

los elementos con los que habitualmente opera el meta-buscador. Además de desarrollar la funcionalidad de almacenamiento de los datos involucrados en cada operación de búsqueda realizada.

Como resultado, el SRI progresivamente almacena datos que posteriormente y en una instancia de procesamiento off-line se complementan a través de procesos de ETL (Extracción, Transformación y Carga por su sigla en inglés). Tales operaciones, parten de los registros básicos de las entidades almacenadas y obtienen los datos necesarios para completar los perfiles definidos para cada una de ellas en la base de datos (BD) del SRI [3].

1.3 Procesos complementarios para el SRI y tecnologías relacionadas

Con un mayor volumen de datos en la BD del meta-buscador, se propuso comenzar con el desarrollo de procesos complementarios al mismo que permitieran incrementar la relevancia de los resultados a presentar al usuario y optimizar su funcionamiento en general.

Esta es una alternativa que se reconoce en otros SRI, operando sobre diferentes contextos y conjuntos de datos. En ellos, se reconocen diversas técnicas que permiten mejorar la experiencia del usuario con el sistema. Un ejemplo es el caso de recomendación de productos en una tienda virtual en base a búsquedas y compras previas del usuario, utilizando técnicas de clustering y filtrado basado en perfiles y datos de opiniones [4]. En las plataformas de contenido como pueden ser Netflix o Spotify se reconocen herramientas para recomendación u armado de listas de reproducción basado en intereses demostrados por el usuario o sus contactos [5, 6]. La estimación de utilidad de resultados con base en una consulta ingresada por el usuario, contextualizando la misma sobre selecciones previas que hubieran resultado satisfactorias para otros usuarios [7] y su revisión a partir de la consideración de los perfiles de usuario [8, 9] son otro tipo de soluciones aplicables en

las que se encuentran implementaciones en la actualidad. De igual manera, la detección de patrones de uso o navegación sobre sitios web pueden ser utilizados para ofrecer contenido específico a un usuario [10, 11]. En esta última alternativa, la identificación de outliers e inliers [12] es un aspecto que hace a la calidad de la solución que percibe el usuario, y dada la naturaleza de los datos se deberá evaluar la vinculación a métodos como los prototipos difusos propuestos por Zadeh [13].

En el ámbito de SRI que operan con información científica se reconocen iniciativas similares. En algunas se han planteado métodos para aumentar la precisión en la asignación de algunos metadatos a las publicaciones [14] permitiendo mejorar su proceso de clasificación, además de la asignación automática de categorías en las que una BD pueda ordenar su catálogo a través del análisis de referencias bibliográficas [15]. Con respecto al tratamiento de términos clave (keywords), su reconocimiento en documentos científicos es de utilidad para la estimación del contenido del mismo y facilitar su clasificación a través de diferentes métodos, como, por ejemplo: la relación con los términos presentes en los títulos y nombres de fuentes de publicación [16]. En otras soluciones se reconoce el uso del coeficiente TF-IDF (Term Frequency - Inverse Document Frequency) [17, 18] junto a técnicas de clustering para generar métodos de clasificación automática para contenido web [19]. Finalmente, se reconocen procesos de recomendación de resultados basados en técnicas de topic modelling en conjunto con otras de análisis de probabilidad [20, 21].

A partir de la enumeración anterior y otros ejemplos similares existentes en la bibliografía, se puede considerar que el desarrollo de procesos complementarios para un SRI forma parte de la evolución del mismo. De esta manera, se plantea al objetivo principal del presente trabajo, el desarrollo de procesos complementarios que generen un impacto directo en la

experiencia del usuario con un metabuscador que opera sobre documentos científicos del área de ciencias de la computación.

2 LÍNEAS DE INVESTIGACION, DESARROLLO E INNOVACIÓN

Para un investigador científico la búsqueda de información en internet implica la utilización de herramientas especializadas que permitan obtener un mayor volumen de resultados relevantes y de estrecha relación con las consultas que ejecute sobre las mismas. En este sentido, la integración de procesos complementarios en un metabuscador de propósito específico como el producto de la presente línea de investigación, constituye un área de trabajo de sumo interés ya que guarda estrecha relación con la efectividad de sus operaciones.

En este contexto, la utilización de técnicas y métodos de probada efectividad en otras áreas como inteligencia artificial, explotación de información y análisis de grandes volúmenes de datos, son los pilares del desarrollo de los procesos mencionados. Su integración permitirá incrementar la relevancia y calidad integral de los resultados a presentar al usuario.

3 RESULTADOS Y OBJETIVOS

3.1 Procesos planteados

Inicialmente se han planteado procesos centrados en la presentación de resultados al usuario. La primera propuesta consiste en métodos para agrupar los resultados considerando el área temática sobre la cual el usuario con base en análisis de tópicos y métodos de probabilidad bayesiana. Mientras que la segunda propuesta abarca la presentación de resultados recomendados a partir de los elementos de la BD del SRI, implementando técnicas específicas para este tipo de operaciones.

3.1.1 Proceso 1

La propuesta en este caso comienza por reconocer las palabras clave que se encuentren en los resultados de la ejecución

de una búsqueda y determinar cuál es su probabilidad de ocurrencia en las diferentes áreas temáticas con las que opera la herramienta. Concretamente se pretende aumentar la precisión en la clasificación de los resultados a una determinada área temática a fin de poder organizar mejor el listado final en base a las preferencias del usuario.

En este sentido, cobran importancia las keywords de los resultados, ya que pueden tomarse como indicadores temáticos representativos del contenido del documento resultante de la búsqueda. La utilización de éstas como recurso para un proceso de clasificación de los resultados que presenta el meta-buscador permite la utilización de un amplio conjunto de técnicas como ser: clustering, determinación de frecuencia ocurrencia de términos sobre una colección de documentos y topic modelling. El proceso que se presenta en esta sección se encuentra en una etapa de diseño, se ha determinado que el mismo opere sobre los resultados que se presentan al usuario una vez ejecutada una búsqueda por el SRI. Sobre ese listado se aplicarían técnicas de topic modelling para la extracción de términos clave que serán utilizados en un proceso de clasificación para determinar el área temática a la que pertenecen. Como resultado se podrán aplicar diferentes técnicas para ponderar los resultados que resulten más cercanos al área de interés del usuario.

Con respecto a su implementación, se ha avanzado en la generación de un prototipo de módulo integrado al SRI que captura los resultados obtenidos del proceso de búsqueda y extrae los tópicos más representativos de cada uno de ellos. Actualmente se está generando con base en la taxonomía definida por la ACM una colección de documentos base para cada área y subárea en la que se divide la disciplina. El paso siguiente consiste en la extracción de tópicos de cada conjunto de documentos a fin de establecer un conjunto de keywords relacionado con cada ítem de

la taxonomía. Una vez que se cuente con estos elementos, se procederá con la selección de las técnicas a través de las cuales se realizará la clasificación de los resultados, incluyendo en el análisis a los grupos antes mencionados: clustering, modelos basados en probabilidad y técnicas de análisis de frecuencia de términos.

3.1.2 Proceso 2

Este proceso tiene por objetivo realizar una recomendación de datos de autores relacionados con la temática de la consulta ingresada por el usuario. Para ello utiliza la BD interna del SRI, obteniendo un conjunto de perfiles de autores que se pueden considerar influyentes en el área de la consulta. De esta manera, el meta-buscador podría presentar resultados propios mientras se ejecutan las consultas sobre las fuentes externas, incrementando la interacción con el usuario y disminuyendo los tiempos de espera.

El proceso en cuestión se encuentra en una fase de diseño, evaluando las técnicas aplicables según el estado del arte referido a sistemas de recomendación. En este sentido, entre las opciones disponibles se presentan inicialmente aquellas basadas en contenido, en las que una opción es utilizar perfiles de usuario generados a partir del contenido que reconocen de utilidad en el uso de la herramienta y así recomendar nuevo contenido en base a tales preferencias. Por otra parte, existen métodos denominados colaborativos, en estos la recomendación se genera a partir de una comparación entre perfiles de usuario que sean considerados similares, por lo tanto, un elemento que haya sido de utilidad para un usuario sería de potencial interés para otros usuarios con un perfil análogo, generando así la recomendación. Finalmente, se reconocen métodos que toman características de los mencionados previamente buscando unificar capacidades y disminuir el impacto de problemas que pueden presentar en su operatoria. De esta manera, los métodos híbridos, se presentan como una alternativa flexible en aquellos casos en los que las recomendaciones no

pueden producirse únicamente a partir de contenido o en forma colaborativa.

En la actualidad se están evaluando en detalle los escenarios en los que cada método obtiene los mejores resultados, además de analizar las técnicas necesarias para su implementación, esperando obtener una propuesta de método a corto plazo para iniciar su desarrollo.

3.2 Trabajos en curso y a futuro

En el marco de la presente investigación la prioridad actual del desarrollo se centra en la finalización del diseño e implementación de los procesos complementarios. En relación a esta actividad, el desarrollo de un módulo de gestión de usuarios, que permita la definición de perfiles de los usuarios del SRI se presenta como una necesidad en un futuro cercano. Por otro lado, la incorporación de técnicas de detección de outliers e inliers y vinculando estos conceptos a los prototipos difusos propuestos por Zadeh, realizando esta detección en los perfiles tanto de los usuarios como de las otras entidades almacenadas en la BD del meta-buscador, con el objetivo de detectar elementos tales como: intrusos, mal uso del SRI, datos erróneos en los datos de las entidades, así como nuevas tendencias de búsqueda.

4 FORMACION DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Investigación en Computación” de la FCEQyN de la UNaM, con diez integrantes relacionados con las carreras de Ciencias de la Computación de la UNaM. De los cuales tres están realizando su tesis de grado, tres se encuentran realizando una maestría. La línea y el equipo de investigación se vinculan con el Grupo de Investigación SMILe de la Universidad de Castilla-La Mancha, España.

5 BIBLIOGRAFIA

1. Kuna, H., et al: Avances en la Construcción de un Sistema de

- Recuperación de Información para Información Científica en Ciencias de la Computación. XVIII Workshop de Investigadores en Ciencias de la Computación (2016).
2. Rey, M., Kuna, H., et al: Propuesta de Esquemas de Perfiles para la Recuperación de Datos Científicos para un Sistema de Recuperación de Información del Área de Ciencias de la Computación. XXII Congreso Argentino de Ciencias de la Computación, San Luis, Argentina (2016).
 3. Rey, M., Kuna, H., Martini, E., Canteros, A., Cantero, A., Rambo, A., Biale, C., Corrales, N.: Un Metabuscador como Plataforma para el Desarrollo de Procesos de Explotación de Datos Científicos. IV Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación. Corrientes, Argentina (2016).
 4. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*. 7, 76–80 (2003).
 5. Gomez-Uribe, C.A., Hunt, N.: The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Trans. Manage. Inf. Syst.* 6, 13:1–13:19 (2015).
 6. Sander Dieleman: Recommending music on Spotify with deep learning, <http://benanne.github.io/2014/08/05/spotify-cnns.html> (2014).
 7. Baeza-Yates, R., Hurtado, C., Mendoza, M.: Query Recommendation Using Query Logs in Search Engines. In: *Current Trends in Database Technology - EDBT 2004 Workshops*. pp. 588–596. Springer, Berlin, Heidelberg (2004).
 8. Sugiyama, K., Hatano, K., Yoshikawa, M.: Adaptive Web Search Based on User Profile Constructed Without Any Effort from Users. In: *Proceedings of the 13th International Conference on World Wide Web*. pp. 675–684. ACM, New York, NY, USA (2004).
 9. Fawcett, T., Provost, F.J.: Combining Data Mining and Machine Learning for Effective User Profiling. In: *KDD*. pp. 8–13 (1996).
 10. Somlo, G.L., Howe, A.E.: Adaptive Lightweight Text Filtering. In: *Advances in Intelligent Data Analysis*. pp. 319–329. Springer, Berlin, Heidelberg (2001).
 11. Pazzani, M., Billsus, D.: Learning and Revising User Profiles: The Identification of Interesting Web Sites. *Machine Learning*. 27, 313–331 (1997).
 12. Hawkins, D.M.: Identification of outliers. Taylor & Francis (1980).
 13. Zadeh, L.A.: A note on prototype theory and fuzzy sets. *Cognition*. 12, 291–297 (1982).
 14. Gómez Núñez, A.J.: Una aproximación multimetodológica para la clasificación de las revistas de Scimago Journal & Country Rank (SJR), (2016).
 15. Gómez-Núñez, A.J., Vargas-Quesada, B., de Moya-Anegón, F., Glänzel, W.: Improving SCImago Journal & Country Rank (SJR) subject classification through reference analysis. *Scientometrics*. 89, 741 (2011).
 16. Garland, K.: An experiment in automatic hierarchical document classification. *Information Processing & Management*. 19, 113–120 (1983).
 17. Salton, G.: Developments in automatic text retrieval. *science*. 253, 974 (1991).
 18. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. *Information processing & management*. 24, 513–523 (1988).
 19. Muñoz, G., del Cisne, M.: Módulo para Clasificación Automática y Temática de Páginas Web., (2012).
 20. Hernández, A., Tomás, D., Navarro Colorado, B.: Una Aproximación a la Recomendación de artículos científicos según su grado de especificidad. *Procesamiento del Lenguaje Natural*. 91–98 (2015).
 21. Blei, D.M.: Probabilistic topic models. *Communications of the ACM*. 55, 77–84 (2012).

Selección de Biomateriales Utilizados en Implantes Dentales Aplicando Técnicas de Minería de Datos

N. Ganz¹, H. Kuna², A. Ares¹

1. Laboratorio de Ciencia de los Materiales, Instituto de Materiales de Misiones, Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones.

2. Depto. de Informática - Facultad de Ciencias Exactas Químicas y Naturales, Universidad Nacional de Misiones.

{nancy.bea.ganz}@gmail.com

RESUMEN

El gran volumen de datos existente en el sector de la salud dificulta la toma de decisiones por parte de los especialistas, debido a que no se aplican técnicas que aprovechen al máximo la información disponible, ocasionando la dificultad de reconocer patrones de comportamiento y extraer conocimiento oculto de los datos almacenados. Además, la no predicción del comportamiento, basado en el conocimiento previo, puede acarrear un alto porcentaje de fracaso, más aún cuando se trata de un campo tan primordial como el de la salud. De aquí, surge la necesidad de aplicar técnicas de minería de datos, debido a que son capaces de extraer patrones, de predecir comportamientos, regularidades y, de sacar provecho a la información automatizada. El objetivo principal de esta línea de investigación es la identificación de factores que contribuyen al éxito o al fracaso de los implantes dentales, a través de la aplicación de técnicas de minería de datos. Se busca determinar cuáles son las condiciones óptimas que debe tener el paciente, el implante, la técnica quirúrgica utilizada por el profesional implantólogo y el seguimiento postoperatorio. Logrando una taxonomía según el origen de fabricación y según el tipo de biomaterial utilizado en la industria del implante.

Palabras Clave: Biomateriales, Minería de Datos, Implantes Dentales, Oseointegración.

CONTEXTO

Esta línea de investigación se lleva a cabo en el Laboratorio de Ciencia de los Materiales del Instituto de Materiales de Misiones (IMAM), de la Facultad en Ciencias Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM), en el marco de un plan de tesis doctoral. Está financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) a través de una “Beca Interna Doctoral” otorgada por Resolución D N° 4869.

1. INTRODUCCIÓN

En la Argentina, como en otros países del mundo se fabrican biomateriales para diferentes aplicaciones con el objetivo de restaurar las funciones del cuerpo humano, como es el caso de los implantes dentales. Sin embargo, es necesario contar con información suficientemente calificada y accesible sobre: tipo de implantes que se fabrican en el país y cuáles se importan, empresas fabricantes a nivel nacional, proveedores internacionales, propiedades de los biomateriales, rendimiento de los mismos en servicio (vida útil), causas del

fallo, características y condiciones de salud de los pacientes que requieren de estos biomateriales, regulaciones provinciales y nacionales para el uso de las mismas, accesibilidad, etc. Para esto, es necesario relevar información de diferentes fuentes, como revistas científicas, bases de datos, Internet, empresas nacionales e internacionales, hospitales, sanatorios, médicos, investigadores en biomateriales, pacientes, normas, leyes vigentes, ministerios de salud y de educación nacional y provinciales, etc.

La carencia de un registro digital provincial o nacional de implantes dentales, que contenga datos sobre enfermedades sistémicas, condiciones del paciente a la hora de la intervención, características del implante utilizado, datos del procedimiento de la fase quirúrgica y datos del seguimiento postoperatorio, hace dificultosa la tarea de investigación para extraer conocimiento desconocido sobre patrones que podrían llegar a influir en el proceso de oseointegración o éxito del implante.

1.1 Biomateriales

Los biomateriales se pueden definir como materiales biológicos comunes tales como piel, madera, o cualquier elemento que reemplace la función de los tejidos o de los órganos vivos.[1] En otros términos, un biomaterial es una sustancia farmacológicamente inerte diseñada para ser implantada o incorporada dentro del sistema vivo.[2]

Los biomateriales se implantan con el objeto de reemplazar y/o restaurar tejidos vivientes y sus funciones, lo que implica que están expuestos de modo temporal o permanente a fluidos del cuerpo, aunque en realidad pueden estar localizados fuera del propio cuerpo,

incluyéndose en esta categoría a la mayor parte de los materiales dentales que tradicionalmente han sido tratados por separado.[3]–[5]

Debido a que los biomateriales restauran funciones de tejidos vivos y órganos en el cuerpo, es esencial entender las relaciones existentes entre las propiedades, funciones y estructuras de los materiales biológicos, por lo que son estudiados bajo tres aspectos fundamentales: materiales biológicos, materiales de implante y la interacción existente entre ellos dentro del cuerpo. Dispositivos como miembros artificiales, amplificadores de sonido para oído y prótesis faciales externas, no son considerados como implantes.[5]–[7]

El éxito de un biomaterial o de un implante depende de tres factores principales: propiedades y biocompatibilidad del implante, condiciones de salud del receptor, y habilidad del cirujano que realiza el implante.[8]–[10]

Los requisitos que debe cumplir un biomaterial son: a) Ser biocompatible; b) No ser tóxico, ni carcinógeno; c) Ser químicamente estable e inerte; d) Tener una resistencia mecánica adecuada; e) Tener un tiempo de fatiga adecuado; f) Tener densidad y peso adecuados; g) Tener un diseño de ingeniería perfecto; h) Ser relativamente barato y reproducible.

Los usos quirúrgicos de los biomateriales son múltiples, por ejemplo, para implantes permanentes: a) en el sistema esquelético muscular, para uniones en las extremidades superiores e inferiores (hombros, dedos, rodillas, caderas, etc.) o como miembros artificiales permanentes; b) en el sistema cardiovascular, corazón (válvula, pared, marcapasos, corazón entero), arterias y venas;

c) en el sistema respiratorio, en laringe, tráquea y bronquios, diafragma, pulmones y caja torácica; d) en sistema digestivo: esófago, conductos biliares e hígado; e) en sistema

genitourinario, en riñones, uréter, uretra, vejiga; f) en sistema nervioso, en marcapasos; g) en los sentidos: lentes y prótesis de córneas, oídos y marcapasos caróticos; h) otras aplicaciones se encuentran por ejemplo en hernias, tendones y adhesión visceral; i) implantes cosméticos maxilofaciales (nariz, oreja, maxilar, mandíbula, dientes), pechos, testículos, penes, etc.[1]–[11]

1.2 Antecedentes

En los últimos años, el campo de la minería de datos ha tenido muchos avances respecto a la aplicación y desarrollo de técnicas en el sector de la salud, para la predicción de enfermedades y para la toma de decisiones en base al análisis de grandes cantidades de datos. Por ejemplo: [12] Extrae datos de Historias Clínicas Electrónicas y utiliza la combinación de dos algoritmos, el J48 (Técnica supervisada, Árboles de Decisión) y el Simple K-Means (Técnica no supervisada, Agrupamiento), para contribuir al diagnóstico de la hipertensión arterial, obtuvieron diversos patrones de comportamiento con relación a los factores de riesgo a sufrir hipertensión. [13] Compara Árbol de Decisión, Clasificación Bayesiana, métodos predictivos como K-Nearest Neighbors (KNN) y Redes Neuronales para la predicción de enfermedad cardíaca. [14] Utiliza Redes Neuronales y reglas de asociación, para la detección y clasificación de tumores en mamografía digital. [15] Utiliza Redes Neuronales para discriminar grupos entre pacientes que se encontraban en terapia intensiva, y un Perceptrón Lineal entrenado con un procedimiento de selección de características (señales respiratorias y cardíacas). [16] Utiliza Redes Bayesianas y Árboles de Decisión, además evalúa la utilidad de la metodología

bayesiana en la predicción y el diagnóstico médico de enfermedades complejas (cardiovasculares).

Existen muchos trabajos de minería de datos aplicados a detección temprana y tratamiento de enfermedades oncológicas, así como en la identificación y reducción de riesgos y prevención de estas enfermedades. Sin embargo, no se han encontrado trabajos que apliquen técnicas de minería de datos específicamente al campo de los biomateriales, como en el caso de los injertos o implantes dentales, en [17] se muestra un análisis estadístico de regresión logística múltiple para determinar los factores que influyen en el éxito de los implantes dentales.

La detección de datos anómalos usando técnicas de Minería de Datos permite detectar grupos de datos que pueden ser de especial interés en temas relacionados por ejemplo con el análisis de los biomateriales.[18]

Ante la necesidad de brindar una aproximación sistemática para la implementación de proyectos de Minería de Datos, diversas empresas han especificado un proceso de modelado, diseñado para guiar al usuario a través de una sucesión formal de pasos. Estas son: SEMMA, CRISP-DM (metodología a utilizar en este trabajo de investigación para plasmar la resolución del problema)[19] y P3TQ.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación tiene por objetivo el estudio de algoritmos de aprendizajes automático, con el fin de simular un sistema con inteligencia principalmente debido a que el entorno de estudio está afectado por variables biológicas, las cuales cambian constantemente,

logrando así el diseño de un sistema (a través de una metodología híbrida) que incorpore el aprendizaje y permita la auto-adaptación a nuevas circunstancias e incluso permitir la incorporación de nuevas variables de estudios. Además, estudiar las propiedades mecánicas, químicas y físicas de los biomateriales utilizados en la implantología dental, para una mejor comprensión de la funcionalidad y resistencia de los implantes.

3. RESULTADOS ESPERADOS

Se pretende obtener un sistema basado en los patrones resultantes, con el objetivo de que al ingresar nuevos casos o condiciones sistémicas de pacientes, se pueda obtener algún resultado en el que indique el posible nivel de oseointegración que tendría ese paciente.

Este análisis e implementación es viable debido a los avances en el campo de la computación de alta performance, ya que permiten nuevas oportunidades en la simulación de sistemas biológicos y aplicaciones bioinformáticas, biología computacional y química computacional. El uso de base de datos de gran tamaño permitirá generar candidatos de pacientes potenciales y gran capacidad de procesamiento para mayor exactitud de cálculo, la idea es demostrar empíricamente que el modelo o metodología híbrida de algoritmos planteados, se adapte adecuadamente al estudio de caso.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte del plan de Tesis Doctoral de la Lic. Nancy B. Ganz, becaria CONICET, desde 2016. Cursa el Doctorado en Ciencias Aplicadas de la Facultad en Ciencias

Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM). Tesis denominada “Aplicación de la Minería de Datos para la selección de Biomateriales”, bajo la dirección del Dr. Kuna, Horacio Daniel y codirección de la Dra. Ares, Alicia Esther.

Además, en esta investigación se trabaja en conjunto con expertos en el área de:

- Minería de Datos: *Dr. Kuna, Horacio Daniel.*
- Biomateriales: *Dra. Ares, Alicia Esther.*
- Especialistas en Implantología Oral:
 - Odontólogo, Especialista en Implantología y Rehabilitación Compleja: *Aimone, Gabriel Alejandro.*
 - Odontólogo, Especialista en Patología Bucal e Implantología Oral: *Padula Diego Hernán.*

5. BIBLIOGRAFÍA

- [1] G. S. Duffó, Biomateriales: una mejor calidad de vida, *Colección Ciencia Joven, EUDEBA*, Editorial Universitaria Buenos Aires, 2007.
- [2] B. D. Ratner, A. S. Hoffman, F. J. Schoen, and J. E. Lemons, *Biomaterials Science. An introduction to Materials in Medicine, Elsevier Academic Press*, San Diego, California - USA, 1996.
- [3] R. L. Macchi, *Materiales dentales, Editorial Médica Panamericana*, Buenos Aires, 2007.
- [4] M. C. Piña Barba, *La física en la medicina, 2da. Edici. México, Secretaría de Educación Pública*, 1998.
- [5] M. V. Regi, *Biomateriales: Repuestos para el Cuerpo Humano*, Madrid, España, Discurso de Ingreso Real Academia de Ingeniería, 2004.

- [6] Univ. de Clemson, The history of the Annual International Biomaterials Symposium & Annual Meeting of the Society of Biomaterials. [Online]. Available: <http://www.clemson.edu/centers-institutes/cwhall/index.html>. [Accessed: 10-Mar-2017].
- [7] R. H. Alvarez, Válvulas cardíacas protésicas: Revisión actualizada, *Revista de Posgrado de la Via Cátedra de Medicina*, vol. 137, pp. 19–32, 2004.
- [8] G. P. Kothiyal and A. Srinivasan, Trends in Biomaterials, *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2016.
- [9] B. Valdez, M. Schorr, E. Valdez, and M. Carrillo, Biomateriales para la rehabilitación del cuerpo humano, *Ciencia y Desarrollo en Internet*, 2005.
- [10] M. Cristina and P. Barba, Los biomateriales y sus aplicaciones, Inst. de Inv. en Materiales, Universidad Nacional Autónoma de México. pp. 55–58.
- [11] T. R. Cuadrado, Biomateriales y Dispositivos Biomedicos hacia la Sofisticación y el Reuso, Instituto de Investigación en Ciencia y Tecnología de Materiales, Facultad de Ingenierías, Universidad Nacional del Mar de Plata, CONICET. [Online]. Available: <http://www.hpc.org.ar/images/revista/169-v3p86.pdf>. [Accessed: 10-Mar-2017].
- [12] F. Dávila Hernández and Y. Sánchez Corales, Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas, *Rev. Cuba. Informática Médica*, vol. 4, no. 2, pp. 174–183, 2012.
- [13] J. Soni, U. Ansari, D. Sharma, and S. Soni, Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction, *Int. J. Comput. Appl.*, vol. 17, no. 8, pp. 43–48, 2011.
- [14] M. Antonie, A. Coman, and O. R. Zaiane, Application of Data Mining Techniques for Medical Image Classification, *Proc. Second Int. Work. Multimida Data Min.*, pp. 94–101, 2001.
- [15] C. Arizmendi, E. Romero, R. Alquezar, P. Caminal, I. Díaz, S. Benito, and B. Giraldo, Data mining of patients on weaning trials from mechanical ventilation using cluster analysis and neural networks, *2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 4343–4346, 2009.
- [16] G. Solarte and Y. Castro, Modelo híbrido para el diagnóstico de enfermedades cardiovasculares basado en inteligencia artificial, *Rev. Tecnura*, vol. 16, no. 33, pp. 35–52, 2012.
- [17] J. E. B. Tamez, F. N. Zilli, L. A. Fandiño, and J. M. Guizar, Factores relacionados con el éxito o el fracaso de los implantes dentales colocados en la especialidad de Prosthodontia e Implantología en la Universidad de La Salle Bajío, *Revi. Esp. Cirugía Oral y Maxilofac.*, pp. 1–9, 2016.
- [18] H. Kuna, R. G. Martinez, and F. Villatoro, Automatic Outliers Fields Detection in Databases, *J. Model. Simul. Syst. HyperSciences Publ.*, vol. 3, no. 1, p. 14–20., 2012.
- [19] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer and R. Wirth, CRISP-DM 1.0: Step-by-step data mining guide, 2000.

Sistema de Apoyo a las Decisiones (DDS) para la Productividad de las Universidades: Implementación de Tableros de Control

Guadalupe Pascal¹; Esteban Grillo¹; Diego Servetto¹ y Andrés Redchuk¹

¹Instituto de Investigaciones de Tecnología y Educación (IIT&E) – Centro Asociado CIC
Facultad de Ingeniería - Universidad Nacional de Lomas de Zamora

Contacto: guadapascal@gmail.com; eagg1992@gmail.com;
diegoservetto@gmail.com; andres.redchuk@gmail.com

Resumen

La complejidad creciente implica la necesidad de adquirir métodos más eficientes para el tratamiento de los datos de una organización, en consecuencia, las organizaciones deben optimizar sus sistemas de gestión. Sin embargo, no fue antes de 1990 cuando el concepto de Inteligencia de Negocios se arraigó el mundo empresarial. Ahora es tiempo del mundo académico: las Universidades no pueden escapar a este escenario.

La línea de investigación y desarrollo presentada propone relevar los requerimientos claves de las unidades inter-institucionales encargadas de la toma de decisiones con el propósito de crear tableros de control dinámicos, interactivos y estratégicos.

En el trabajo, se exponen los resultados obtenidos de las tres primeras etapas de implementación del Data Warehouse y los objetivos específicos de la próxima etapa de explotación.

Palabras clave: TIC, Business Intelligence, Universidad, Tableros de Control.

Contexto

El Instituto de Investigaciones en Tecnología y Educación (IIT&E) de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (FI-UNLZ) trabaja desde el año 2006 en actividades de Investigación y Desarrollo orientadas a generar conocimiento que contribuya en la Educación Superior mediante la incorporación sistematizada de las TIC.

En este sentido, el crecimiento ha sido acelerado, creciente e ininterrumpido. En el año 2013, los objetivos propuestos han sido consolidar una base de datos confiable y oportuna de la FI-UNLZ bajo el proyecto “Las Tecnologías de la Información y la Comunicación: Aplicación en la Gestión de Instituciones Universitarias”. El proyecto; aprobado, acreditado, incentivado y ampliado; ha superado los objetivos y en el 2016 se define la línea de Investigación y Desarrollo específica. Finalmente, en el 2017, se pone en marcha el Proyecto “Sistemas de Apoyo a las Decisiones (DSS) para la productividad: aplicación en la Gestión Académica en Universidades” del Programa “Enseñanza de la ingeniería: Desarrollo y Evaluación de modelos, estrategias y tecnologías para mejorar los indicadores académicos y la eficiencia organizacional”.

Actualmente el IIT&E desarrolla sus actividades como Centro Asociado a la Comisión de Investigaciones Científicas, dependiente del Ministerio de Ciencia Tecnología e Innovación de la Provincia de Buenos Aires, lo cual imparte positivamente en la investigación, el desarrollo y la transferencia de los resultados al medio local.

Introducción

En estos tiempos, la gestión del conocimiento es de suma importancia en el planeamiento de cualquier organización (Bertoglio et al, 2014). En este sentido, la articulación con las TICs cumple un rol fundamental ya que, con el objetivo de mejorar el manejo operativo, favorecen la toma de decisiones estratégicas y

mejoran el desempeño de todas las funciones (Minnaard et al, 2015).

Sin embargo, la complejidad creciente, implica la necesidad de adquirir métodos más eficientes para el tratamiento de los datos de una organización a lo largo de la cadena de valor y en consecuencia, las organizaciones deben optimizar sus sistemas de gestión (Pascal et al, 2017).

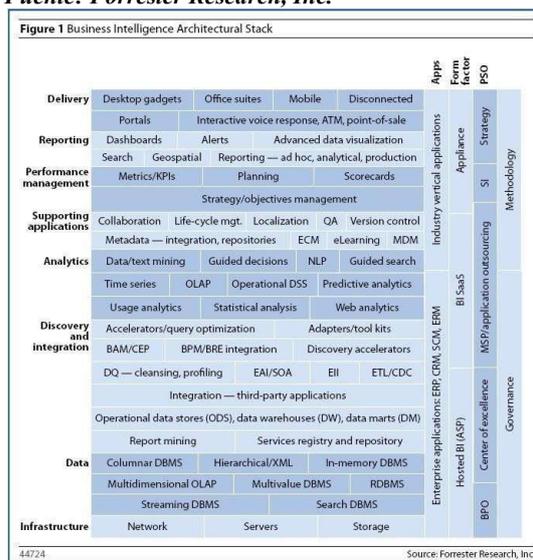
Las universidades no pueden escapar de este escenario, debido a la complejidad en el intercambio de información en organizaciones como las instituciones universitarias, es necesario contar con un sistema en donde los datos se encuentren disponibles de manera tal que proporcionen información acerca de indicadores claves, como por ejemplo, el rendimiento académico (Minnaard et al, 2016; Moguerza et al 2017).

El conjunto de tecnologías capaces de satisfacer dicha necesidad se agrupan mediante la Inteligencia de Negocios. Se reconoce como Business Intelligence (BI) o inteligencia de negocios, en cualquier industria o sector económico, al conjunto de metodologías que permiten recopilar, depurar y transformar los datos para dar soporte a los procesos de toma de decisiones (Kumar y Chadha, 2012). El término tiene sus inicios con Luhn en 1958, cuando utilizó Business Intelligence System para referirse a un sistema automático que acepta información original y disemina los datos adecuada y rápidamente a los lugares correctos. Desde entonces, y con la velocidad característica de la gestión del conocimiento y las tecnologías, no ha parado de aumentar su alcance de referencia. A mediados de los años 60, la información era representada mediante conjunto de archivos planos (Luque Ruiz, 2002). Sin embargo, para la década siguiente, los Sistemas de Apoyo a la Toma de Decisiones ya ofrecían soluciones interactivas y específicas para procesos directivos. Es entonces, hacia la década del 90 que se inicia un paradigma en torno al Data Warehouse (Armendáriz et al, 2016). El gran aporte del Data Warehouse radica en el ambiente computacional donde los usuarios son puestos directamente con los datos que ellos necesitan

para tomar mejores decisiones (Ponniah, 2010).

Actualmente, y en constante avance, la inteligencia de negocios dispone de múltiples técnicas y tecnologías para mejorar los resultados futuros; en la Figura 1 se observa un diagrama de dichas tecnologías.

Figura 1: Business Intelligence Architectural Stack – Fuente: Forrester Research, Inc.



Estas tecnologías, dan lugar a diversas funciones como la generación de informes, el procesamiento analítico en línea, la analítica, la minería de datos, la gestión del rendimiento empresarial, la evaluación comparativa, la minería de texto y el análisis predictivo (Elena, 2011).

Como se ha visto, son muchas las herramientas que han nutrido el concepto de Inteligencia de Negocios, sin embargo, fue recién en 1990 cuando el concepto resulta aceptado por las empresas y las organizaciones han entendido la importancia de tomar mejores decisiones con información (Simpson y Siguaw, 2000; Schreidber et al, 2006). Este hecho, ha posicionado a Howard Dresner, socio de la firma Gartner Group, como el padre de la Inteligencia de Negocios (Armendáriz, 2016). Trasvasar algunas de las técnicas exitosas en el mundo empresarial al mundo académico es una tarea aún pendiente. (González de Lena, 2015). Estas técnicas, que requieren de profundo conocimiento de TIC, gestión de bases de datos, estadística multivariante, entre otras, nos permiten descubrir las relaciones

entre variables, nuevos tipos de asociaciones y anomalías de la información educacional recabada por los sistemas de recogida de datos existentes en las universidades públicas argentinas y aumentar el conocimiento que tenemos de los alumnos (Greenacre 2008). El rendimiento académico de un alumno depende de múltiples factores, tales como variables sociales, ambientales, psicológicas, económicas. La identificación de patrones en los grandes volúmenes de datos, de forma de determinar aquellas variables que puedan predecir el fracaso temprano y además puedan perfilar a aquellos grupos de riesgo, permitirá mejorar el desempeño de todos los alumnos y optimizar el uso de los recursos públicos (Campos, 2016).

Líneas de Investigación, Desarrollo e Innovación

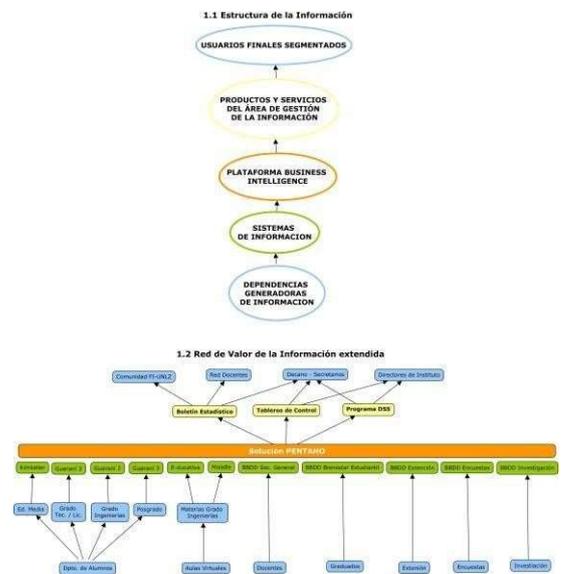
El eje del trabajo, está orientado a la toma de decisiones basada en información medible y verificable. En este sentido, se utilizarán técnicas cuantitativas de la inteligencia de negocios aplicadas a la mejora de la gestión académica de la FI UNLZ y se explotarán la Plataforma de BI PENTAHO, de la Unidad Académica.

Así mismo, partirá del aprovechamiento del know how adquirido por la institución en actividades de investigación y de vinculación con el medio.

La línea de I+D se encargará relevar los requerimientos claves de las unidades inter-institucionales encargadas de la toma de decisiones con el propósito de crear tableros de control dinámicos, interactivos y estratégicos. La naturaleza de los datos es estrictamente cuantitativa; se desarrollará mediciones sistemáticas y se garantizará la calidad de la información mediante la incorporación fundamentada de análisis estadísticos.

La metodología de la investigación, inicialmente, será de carácter descriptivo correlacional para poder estudiar la relación entre las variables clave para la gestión académica (desgranamiento, rendimiento académico de las carreras, rendimiento

1. RED DE VALOR DE LA INFORMACIÓN DE LA FACULTAD DE INGENIERÍA DE LA UNLZ



académico por ciclos de la carrera, o por materias articuladas, etc.), con la finalidad de determinar con KPI (indicador clave de rendimiento)

Resultados y Objetivos

Hasta el momento, el equipo de trabajo ha desarrollado y avanzado en línea con la arquitectura de los datos. En la primera etapa, se ha consolidado el conjunto de datos del Data Warehouse referido al comportamiento de las cátedras, por ser considerado el principal lugar de desempeño de los alumnos. En una segunda etapa, se incorporaron los datos referidos al entorno virtual, por tratarse del segundo ámbito de desempeño del alumno. La tercera etapa, vigente desde el año 2016, consiste en la incorporación de los datos del resto de las dependencias de la unidad académica al Data Warehouse institucional. En la Figura 2 se puede observar la red de valor de información de la FI-UNLZ.

Figura 2: Red de Valor de la Información en la FI-UNLZ, Fuente: Pascal, Servetto, Lobo Mirasson, Luna, 2017

En esta instancia, el objetivo general es continuar con la optimización de la actual plataforma de Bussines Intelligence de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora, mediante la construcción de tableros de control; a partir del relevamiento y análisis de los

requerimientos de las autoridades de la institución, con el objeto de proveer las herramientas necesarias para la toma de decisiones.

Los objetivos específicos para la siguiente etapa son:

Relevar los requerimientos claves de determinadas unidades inter-institucionales encargadas de la toma de decisiones; participar en la potencial ampliación de la arquitectura del Data Warehouse y el diseño del proceso de Data Scrubbing (limpieza de datos); Elaborar los KPIs necesarios y construir un tablero de control piloto.

Formación de Recursos Humanos

La línea de Investigación y Desarrollo presentada, con su carácter protagónico en el Proyecto “Sistemas de Apoyo a las Decisiones (DSS) para la productividad: aplicación en la Gestión Académica en Universidades” se encuentra bajo la Dirección del Esp. Marcelo Estayno y la co-dirección de la Dr. Claudia Minnaard. El programa al que pertenece cuenta con docentes investigadores, personal técnico, personal de apoyo y becarios (Becas CIN y Becas LomasCYT); entre ellos se relacionan constantemente mediante actividades de docencia y transferencia copertivamente. Así mismo, se puede garantizar que, en particular, este proyecto favorecerá la función docente y administrativa de la Unidad Académica, debido a que los resultados buscan contribuir en la calidad educativa.

Finalmente, caben destacar, nueve tesis de posgrado dirigidas por el equipo de investigación y defendidas con éxito recientemente (dos de ellas tesis doctorales, siete de ellas tesis de maestría, en Ecuador, Chile, España y Argentina).

Referencias

- Bertoglio, R., Corizzo, M. M., Pascal, G., Servetto, D., & Steinman, B. (2014). Las Tecnologías de la Información y Comunicación: aplicación en la gestión de Instituciones Universitarias. In *XVI Workshop de Investigadores en Ciencias de la Computación*.
- Pascal, G., Servetto, D., Mirasson, U. L., & Luna, Y. (2017). Aplicación de Business Intelligence para la toma de decisiones en Instituciones Universitarias. Implementación de Boletines Estadísticos en la Universidad Nacional de Lomas de Zamora (UNLZ). *Revista Electrónica sobre Tecnología, Educación y Sociedad*, 4(7).
- Minnaard, C., Servetto, D., Lobo Mirassón, U. y Pascal, G. (2015). La información y la tecnología para la toma de decisiones: aplicación Data WareHouse en instituciones Universitarias. In *XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015)*.
- Minnaard, C., Servetto, D., Pascal, G., & Mirasson, U. L. (2016). Nuevas dimensiones y métricas en la información para la toma de decisiones: Aplicación Data WareHouse en Instituciones Universitarias. *Revista Iberoamericana de Producción Académica y Gestión Educativa*.
- Moguerza, J., Fernández-Muñoz, J.J., Redchuk, A., Cardone-Riportella, C. y Navarro-Pardo E. (2017). Factor structure and stability of a quality questionnaire within a postgraduate program. *Análes de Psicología*, Vol 33, N° 2.
- Elena, C. (2011). Business intelligence. *Journal of Knowledge Management, Economics and Information Technology*, 1(2), 1-12.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A. & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research*, 99(6), 323-337.
- Simpson, P. y Siguaw J. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22(3), 199-213.
- Armendáriz, R. N., Urdiales, M. G. V., Corral, J. J. V., Salcido, M. H. T., Favela, J. A. A., & Ávila, R. M. L. (2016). Evolución de la inteligencia de negocios. *CULCyT*, (57).
- Kumar, V., & Chadha, A. (2012). Mining association rules in student’s assessment data.

International Journal of Computer Science Issues, 9(5), 211-216.

- Luque Ruiz, I., & Luque Ruiz, I. (2002). Bases de datos: desde Chen hasta Codd con Oracle. *Alfaomega*.
- Ponniah, P., & Ponniah, P. (2010). *Data warehousing fundamentals for IT professionals* (2nd ed.). Hoboken, N.J.: John Wiley & Sons.
- Gonzalez de Lena, María Teresa. Diciembre 2015. La metodología Seis Sigma como herramienta de mejora de la calidad en la gestión de programas académicos. Tesis Doctoral. Universidad Rey Juan Carlos de Madrid. Directores: Javier Martínez Moguerza y Andrés Redchuk.
- Greenacre, Michael. 2008. *La Práctica del Análisis de Correspondencia*. Traducido por Jordi Comas Angelet. Fundación BBVA.
- Campos de Laire, Luis. Mayo 2016. Modelo Predictor de la Deserción en Pregrado de Alumnos de Primer Año en la Universidad Autónoma de Chile. Tesis de Maestría. Universidad Rey Juan Carlos de Madrid. Director: Andrés Redchuk.

Técnicas de Análisis de Información en Neurociencias Aplicadas

Juan I. Larregui^(1,6), Juan A. Biondi^(1,2,4), Gerardo Fernández^(2,4), Marcela Schumacher^(2,4),
David Orozco^(4,5), Liliana R. Castro^(2,3,4), Silvia M. Castro^(1,4,6), Osvaldo Agamennoni^(2,4)

juan.larregui@cs.uns.edu.ar, juan.biondi@uns.edu.ar, gerardo.fernandez@uns.edu.ar,
schumachermarcela@gmail.com, davidorozco@live.com, lcastro@uns.edu.ar, smc@cs.uns.edu.ar,
oagamen@uns.edu.ar

- ⁽¹⁾ VyGLab, Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur
⁽²⁾ Instituto de Investigación en Ingeniería Eléctrica (IIIE)
⁽³⁾ Dpto. de Matemática, Universidad Nacional del Sur
⁽⁴⁾ Grupo de Investigación y Desarrollo en Procesos Cognitivos,
 Universidad Nacional del Sur
⁽⁵⁾ Clínica Privada Bahiense, Bahía Blanca
⁽⁶⁾ Instituto de Ciencias e Ingeniería de la Computación (ICIC)

RESUMEN

Determinar cuál es la información más relevante de un conjunto de datos es una tarea fundamental en su proceso de análisis. En grandes conjuntos de datos multidimensionales es crucial poder determinar cuáles son las dimensiones más relevantes que caracterizan el conjunto de datos con una mínima pérdida de información en dicha transformación; esto es relevante tanto desde el punto de vista del análisis computacional como desde el visual. Para ello, se aplican técnicas de *reducción dimensional*. Otro aspecto importante es la determinación de características distintivas de los subconjuntos para su posterior diferenciación. Este proceso es conocido como *extracción de características*.

El seguimiento de los movimientos oculares (*eye tracking*) permite evaluar la información adquirida por una persona durante la lectura u observación de una imagen. Durante ambos procesos, el ojo realiza movimientos en las cuales la información es adquirida para luego ser interpretada por el cerebro. El *eye-tracker* registra y graba tanto información relativa a los movimientos como el tiempo que necesita la persona para procesar la información. Esta

técnica es utilizada en disciplinas muy diversas vinculadas a la evaluación de los procesos cognitivos que se desarrollan en el ser humano. Los *eye-tracker* son en general dispositivos de un costo muy elevado y que requieren una calibración para su uso. Dependiendo de la aplicación que se le quiera dar a la información extraída o de la patología del sujeto de estudio, a veces no es posible realizar este proceso.

Nuestro grupo de investigación actualmente trabaja en técnicas de *reducción dimensional*, *visualización* y *extracción de características* además de en la utilización de estas técnicas para diseñar y desarrollar *eye-trackers* de bajo costo, independientes de la posición y libres de calibración.

Palabras clave: *eye-tracking*, *reducción dimensional*, *extracción de características*, *análisis visual*.

CONTEXTO

Parte de este trabajo se desarrolla en el contexto de la línea de investigación Identificación y Desarrollo en Procesos Cognitivos. El grupo está integrado por becarios e investigadores del Dpto. de Matemática, del Dpto. de Ingeniería Eléctrica y de Computadoras pertenecientes al Instituto

de Investigaciones en Ingeniería Eléctrica (IIIE) y del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Dpto. de Ciencias e Ingeniería de la Computación, todos ellos pertenecientes a la Universidad Nacional del Sur.

La línea de Investigación presentada está inserta en el Proyecto PICT 2013 0403 *Evaluación del desempeño cognitivo a través del comportamiento ocular*, Categoría: Plan Argentina Innovadora 2020. Tipo: A, dirigido por el Dr. Osvaldo Agamennoni y financiado por la ANPCyT. También se encuentra inserta en el proyecto acreditado *Análisis Visual de Grandes Conjuntos de Datos (24/N037)*, dirigido por la Dra. Silvia Castro y financiado por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur.

1. INTRODUCCIÓN

El *eye-tracker* es un dispositivo no invasivo que permite, mediante la grabación de los movimientos oculares, el diagnóstico temprano de enfermedades neurodegenerativas [1], [6], [9]. La detección temprana de diversas enfermedades neurodegenerativas permite, en muchos casos, prolongar una calidad de vida razonable para el paciente y el entorno familiar, como así también reducir el costo asistencial.

Lograr avances en el conocimiento sobre el modelado de los movimientos oculares en el proceso de lectura y de observación de una escena o imagen nos permitirá ahondar en el conocimiento de las estrategias que utiliza el cerebro cuando procesa información. Definir nuevos y mejores modelos que permitan tener en cuenta aspectos fisiológicos del sistema óculo-motor permitirá efectuar mejores predicciones sobre tales comportamientos. Dado que esta información es inherentemente multi-dimensional, resulta necesario contar con técnicas que permitan conocer cuáles de ellas son las que mejor la representan; además, es de suma importancia poder encontrar patrones distintivos que nos

permitan diferenciar distintas patologías que podrían estar afectando al cerebro.

Entre las diversas patologías, hay algunas en las que los pacientes no soportan el uso de medidas de fijación (por ejemplo en autismo). En estos casos, es de suma importancia que los dispositivos de adquisición no requieran su uso, haciendo necesario contar con dispositivos que no requieran calibración.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la línea de investigación planteada en este trabajo se pretende avanzar en la búsqueda de metodologías que permitan una evaluación objetiva (no dependiente del profesional que la efectúa), sin someter a estrés a la persona evaluada, y no invasiva (se registran los movimientos oculares con una cámara de video de alta velocidad) que posibiliten detectar Deterioro Cognitivo Leve (DCL) o incipiente [1]. De esta manera se pretende facilitar la determinación de deterioro cognitivo normal por edad del producido por otras patologías neurodegenerativas y, posteriormente, el seguimiento temporal periódico del paciente. Para esto, se prevé el uso de técnicas de análisis multirresolución y de reducción dimensional de grandes conjuntos de datos y la extracción de características distintivas de la misma para su posterior análisis y visualización. Para esto se plantea utilizar técnicas de *machine learning*, en particular *sparse autoencoders* [12] y análisis mediante la transformada *curvelet* [11], de manera de encontrar las dimensiones representativas de un conjunto de datos. Además, se plantea la visualización de esta información como parte esencial del proceso.

Constituye otra línea de investigación la utilización de técnicas de *extracción de características* en imágenes para el diseño y el desarrollo de dispositivos *eye-tracker* de bajo costo, independientes de la posición y libres de calibración.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el grupo de trabajo se desarrolló un método orientado al diagnóstico y medición del deterioro cognitivo temprano a partir del modelado del movimiento ocular durante el proceso de lectura ([2], [5], [6], [7], [8], [9]). Se desarrollaron diversos modelos del comportamiento ocular que permiten evaluar la manera en que personas sanas y pacientes en una fase muy temprana de la enfermedad procesan información durante la lectura. Para ello se diseñó el material de estímulo apropiado que permitió medir el desempeño de la memoria ejecutiva, de la memoria de trabajo, de la memoria semántica y de la memoria de recuperación (*retrieval memory*).

A partir de este trabajo de investigación se analizaron distintos aspectos de la problemática del DCL o incipiente. Se estudiaron los efectos de la predictibilidad de las palabras actualmente fijada y de las entrantes (que no son necesariamente las inmediatas siguientes) sobre la duración de la fijación. Cuando una oración es leída por lectores sanos, se generan expectativas y se realizan predicciones sobre las palabras entrantes reduciendo luego el tiempo de fijación ocular sobre tales palabras. Tales fijaciones permiten evaluar el funcionamiento de la memoria de trabajo, de la memoria semántica y de la memoria de recuperación, entre otras facultades. Esta capacidad predictiva se distorsiona, por ejemplo, durante los primeros estadios de la enfermedad de Alzheimer [7].

Se analizó el efecto de la previsibilidad contextual de una oración en el comportamiento del movimiento ocular de los pacientes con enfermedad de Alzheimer leve, en comparación con los controles de la misma edad, utilizando la técnica de *eye-tracking* y modelos lineales de efecto mixto [8]. Los resultados de este trabajo indican que es posible hacer el diagnóstico temprano del deterioro cognitivo típico de una enfermedad tipo Alzheimer [1] y, por otro lado, que existe

un amplio margen de variación en dicho comportamiento sobre el cual es importante concentrar el estudio, con el fin de evaluar la posibilidad de hacer una detección antes de llegar a la etapa Prodrómica.

Nuestro objetivo es utilizar esta capacidad de abordaje sobre temáticas de modelado, representación y visualización de la información al caso particular del sistema oculomotor a los efectos de vincular los aspectos sistémicos funcionales con los fisiológicos y cognitivos.

Una de las primeras actividades previstas es definir una serie de marcadores que muestren claramente el desempeño cognitivo vinculado a los aspectos atencionales, de funcionamiento de la memoria de trabajo, la memoria ejecutiva y la memoria semántica. Se estudiarán diversas formulaciones de los mismos con el fin de contar con un conjunto que permita describir, de la mejor manera posible, los distintos aspectos de la capacidad cognitiva de una persona. Para ello se utilizarán también técnicas que permitan cuantificar la incertidumbre de los modelos [15] a los efectos de mejorar progresivamente las estimaciones a medida que se adquieran nuevos datos y poder estimar con la mayor precisión posible un DCL ocasionado por la EA.

Con el objetivo de facilitar el análisis de la información disponible y su desarrollo temporal por parte de los profesionales de la salud, se estudiarán y evaluarán distintas alternativas de visualización de la misma ([3], [4]).

De acuerdo al estado del arte, la información es obtenida por medio de dispositivos costosos que requieren de cooperación del sujeto bajo estudio para ser calibrados. En contextos prácticos, esta fase de calibración no siempre es posible debido a la naturaleza de la enfermedad del paciente. La habilidad de medir automáticamente variables relacionadas al comportamiento utilizando sensado basado en *visión computacional* puede ser valiosa a la hora de permitir obtener información relacionada con

el comportamiento del sujeto bajo estudio sin requerir intervención humana sustancial. Tecnologías como *behaviour imaging* pueden ser utilizadas en varios casos en un instrumento de medición dado que aportan herramientas de bajo costo computacional para manejar grandes cantidades de datos de video e información de otras fuentes obtenida durante las sesiones. En particular, puede permitir la síntesis, visualización y comparación de la información observada entre poblaciones y a través del tiempo, de una manera que no resulta posible mediante métodos manuales [13]. Para explorar esta línea de investigación sin renunciar a tomar en cuenta todos los casos particulares donde un *eye-tracker* no está disponible y/o el usuario no puede cooperar en la fase de calibración, puede ser utilizado un dispositivo menos preciso pero de bajo costo, no invasivo, seguro y libre de calibración.

4. FORMACIÓN DE RECURSOS HUMANOS

El Dr. Agamennoni y la Dra. Liliana Castro son especialistas en el área de desarrollo de modelos determinísticos y de los límites de incertidumbre de los mismos. La Dra. Silvia Castro trabaja en el área de Visualización. El Dr. Gerardo Fernández trabaja en el modelado de los movimientos oculares y su utilización en el desarrollo de técnicas de diagnóstico de deterioro cognitivo. La Farm. Marcela Schumacher realiza tareas vinculadas con la evaluación de la técnica de diagnóstico temprano de deterioro cognitivo y el Dr. David Orozco realiza actividades vinculadas con los aspectos médicos. El Ing. Juan Larregui trabaja en análisis multiresolución y procesamiento de imágenes. El Ing. Juan Biondi trabaja en aspectos computacionales del desarrollo de los tests y en el diseño y desarrollo de las visualizaciones adecuadas y en el desarrollo de técnicas de reducción dimensional.

Tesis en Desarrollo

Tesis de Doctorado en Ciencias de la Computación

- Juan I. Larregui. Tema: *Reconstrucción 3D en Tiempo Real*. Dirección: Dra. Silvia Castro.
- Juan A. Biondi. *Desarrollo de Modelos del Comportamiento Ocular*. Dirección: Dra. Silvia Castro, Dr. Osvaldo Agamennoni.

Becarios

- Juan I. Larregui. Becario CONICET, Beca Interna Doctoral.
- Juan A. Biondi. Becario PICT.

Cursos de Pre/Post grado relacionados con el tema de la línea de Investigación dictados por integrantes del Grupo de Trabajo

Cursos de Pregrado

- *Wavelets de primera generación: Una introducción*. Materia optativa para los estudiantes de la Licenciatura en Matemática. Dto. de Matemática. UNS.
- *Procesamiento de Imágenes*. Materia optativa para los estudiantes de la Ingeniería en Sistemas de Computación. UNS.
- *Teoría de Sistemas Lineales*. UNS.

Cursos de Posgrado

- *Introducción a la teoría de wavelets y sus aplicaciones*. Liliana Castro. Secretaría de Posgrado y Educación Continua. UNS.
- *Aplicaciones de wavelets* Seminario. Liliana Castro. Secretaría de Posgrado y Educación Continua. UNS.
- *Modelación de Sistemas*. Osvaldo Agamennoni. Secretaría de Posgrado y Educación Continua. UNS.
- *Procesamiento y Análisis de Imágenes*. Silvia Castro, Secretaría de Posgrado y Educación Continua. UNS.
- *Visualización de Grandes Conjuntos de Datos*. Silvia Castro, Luján Ganuza.

Secretaría de Posgrado y Educación Continua. UNS.

- *Introducción al Análisis Funcional y a la Teoría de Operadores*. Liliana Castro. Dto. de Matemática. UNS.

5. BIBLIOGRAFÍA

- [1] Alberta, M.S., DeKosky, S.T., Dickson, D., Duboise, B., Feldman, H., Fox, M., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M. C., Thies, B., Phelps, C.H. *The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease*. *Alzheimer's & Dementia*. Vol 7, pp: 270–279. 2011.
- [2] Alvarez, M., Castro, L., Agamennoni, O. *Set Membership Estimation Theory for Wiener Modelling using HLCPWL Functions*. *International Journal of Modelling, Identification and Control*. Vol 14, pp 13-26. 2011.
- [3] Card, S., Mackinlay, J., Shneiderman, B., *Readings in Information Visualization – Using Vision to Think*, M. K., 1999.
- [4] Escarza, S., Larrea, M., Urribarri, D., Martig, S., Castro, S., *Integrating Semantics in the Visualization Process*. En *Scientific Visualization: Interactions, Features, Metaphors*, Dagstuhl FOLLOW-UPS Series Hans Hagen Ed., Schloss Dagstuhl-Leibniz-Zentrum für Informatik. Dagstuhl, Alemania, 2, 92-102, Vol2, 2011.
- [5] Fernández, G., Biondi, J., Castro, S., Agamennoni, O. *Pupil size behavior during on line processing of sentences*. *Journal of Integrative Neuroscience*, 2017.
- [6] Fernández, G., Castro, L., Schumacher, M., Agamennoni, O. *Diagnosis of mild Alzheimer Disease through the Analysis of Eye Movements during Reading*. *Journal of Integrative Neuroscience*. 2015.
- [7] Fernández G., Laubrock J., Mandolesi P., Colombo O., Agamennoni O. *Registering eye movements during reading in Alzheimer disease: difficulties in predicting upcoming words*. *Journal of Clinical and Experimental Neuropsychology*. Vol. 36, Iss. 3, pp: 302-316. 2014.
- [8] Fernandez G., Manes F., Rotstein N., Colombo O., Mandolesi P., Politi L., Agamennoni O. *Lack of contextual-word predictability during reading in patients with mild Alzheimer disease* *Neuropsychologia*, Vol 62, pp 143-151. 2014.
- [9] Fernández G., Schumacher M., Mandolesi P., Colombo O., Castro L., Agamennoni O. *Eye movement behavior during reading in patients with probable Alzheimer disease*. Presentado en la 17th European Conf. of Eye Movements (ECEM), Lund, Suecia, 2013.
- [10] Kliegl, R., Nuthmann, A., Engbert, R. *Tracking the mind during reading: The influence of past, present, and future words on fixation durations*. *Journal of Experimental Psychology: General*, 135, 12-35. 2006.
- [11] Larregui, J.I., Castro, L.R., Castro, S.M., *Curvelet Transform for Bovine Iris Segmentation*, V Congreso de Matemática Aplicada, Computacional e Industrial, 2015.
- [12] NG, Andrew. *Sparse autoencoder*. CS294A Lecture notes. 2011.
- [13] Rehag, J. M. (2011, June). *Behavior Imaging: Using Computer Vision to Study Autism*. In *MVA* (pp. 14-21).
- [14] Sigut, J., Piñeiro, J., González, E., Torres, J. *An expert system for supervised classifier design: Application to Alzheimer diagnosis*. *Expert Systems with Applications* 32 927–938. 2007.
- [15] Uusitalo L., Lehtikoinen A., Helle I., Myrberg K. *An overview of methods to evaluate uncertainty of deterministic models in decision support*. *Environmental Modelling & Software* Vol: 63, pp: 24-31. 2015.

Tecnologías de Procesamiento de Datos Masivos

Ramiro Rivera, Luciano Bracco, Valentín Costa, Facundo Coto, Patricia Cristaldo, Lautaro Ramos, Natalia Rapesta, Juan Pablo Núñez, Soledad Retamar, Anabella De Battista

Grupo de Investigación en Bases de Datos, Departamento Ingeniería en Sistemas de Información,
Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional

Entre Ríos, Argentina

{riverar, braccol, costav, cotof, cristaldop, ramosl, rapestan, nunezjp, retamars,
debattistaa}@frcu.utn.edu.ar

Norma Edith Herrera

Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina

nherrera@unsl.edu.ar

Resumen

Las grandes cantidades de datos que se producen en la actualidad, sumadas a su heterogeneidad y la velocidad con que se generan, hacen que las herramientas tradicionales de análisis de datos no resulten adecuadas para su recopilación, almacenamiento, gestión y análisis. En este contexto surge el término Big Data, en referencia a características como gran volumen, velocidad y variedad de producción de los datos, y a las herramientas que se utilizan para encontrar valor en los mismos. La posibilidad de hallar patrones y tendencias en estas grandes cantidades de datos impacta directamente en la toma de decisiones en áreas tan diversas como salud, genética, agro, predicciones climáticas, redes sociales, marketing, finanzas, educación, entre otras.

En este artículo se presentan los tópicos de interés del proyecto *Minería de Datos: su aplicación a repositorios de datos masivos*.

Palabras clave: Big Data, minería de datos, clustering, agrupamiento, streaming, gestión de proyectos.

Contexto

El presente trabajo se desarrolla en el ámbito del proyecto *Minería de Datos: su aplicación a repositorios de datos masivos (UTI3781TC)* del Grupo de Investigación en Bases de Datos, perteneciente al Departamento Ingeniería en Sistemas de Información de la Universidad Tecnológica

Nacional, Facultad Regional Concepción del Uruguay.

1. Introducción

El término Big Data surge en referencia a conjuntos de datos cuyo tamaño supera la capacidad de procesamiento de las herramientas tradicionales de bases de datos. En general se habla de Big Data o Análisis de Big Data como sinónimos, ya que no sólo se desea hacer referencia a la gran cantidad y complejidad de los datos, sino también a las herramientas utilizadas para procesarlos y extraer conocimiento útil.

Algunas definiciones indican que Big Data puede definirse a partir de las siguientes características [1]:

- *Volumen:* órdenes superiores a Terabytes de datos.
- *Variiedad:* distintos tipos de datos provenientes de diversas fuentes que pueden organizarse tanto en forma estructurada como no estructurada.
- *Velocidad:* referido a la velocidad de generación de los datos o a la rapidez con la que se generan y procesan los datos.
- *Variabilidad:* referido a la inconsistencia que puede presentar los datos en ocasiones, dificultando las tareas de análisis.
- *Valor:* gracias a la posibilidad de tomar decisiones al responder preguntas que antes no era posible, ofrece a la organización una ventaja estratégica.

Estos grandes repositorios de datos se generan desde fuentes tan diversas como redes sociales, instrumentos científicos, dispositivos móviles o redes de sensores, entre otros, y representan nuevos desafíos para su almacenamiento, tratamiento, distribución y análisis, ya que además de ser datos de gran volumen su complejidad es creciente. En este contexto cobra gran relevancia la posibilidad de hallar patrones en estos datos, pero más aún la posibilidad de explicar dichos patrones, ya que impacta directamente en la toma de decisiones aplicable en áreas tan diversas como salud, genética, agro, predicciones climáticas, redes sociales, marketing, finanzas, educación, entre otras. En este conexto, una actividad interesante es la detección de agrupamientos en repositorios de datos masivos y complejos.

Como se mencionó anteriormente, en la actualidad no es un obstáculo la capacidad de recopilar datos, en cambio sí lo es la capacidad de gestionar, analizar, sintetizar, visualizar y descubrir conocimiento en los datos recopilados de manera oportuna y en una forma escalable. Debido a que no es posible procesar estos datos tan masivos y complejos con herramientas tradicionales, han surgido nuevos algoritmos especialmente diseñados que aprovechan las características del procesamiento paralelo.

La Minería de Datos involucra e integra técnicas de diferentes disciplinas tales como tecnologías de bases de datos y data warehouse, estadística, aprendizaje de máquina, computación de alta performance, computación evolutiva, reconocimiento de patrones, redes neuronales, visualización de datos, recuperación de información, procesamiento de imágenes y señales, y análisis de datos espaciales o temporales. En este proyecto se estudian procesos de Minería de Datos desde una perspectiva de bases de datos, con enfoque en técnicas eficientes y escalables a conjuntos de datos masivos.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo principal de nuestro

proyecto de investigación es el estudio de técnicas de Minería de Datos aplicables a repositorios de datos masivos, atendiendo principalmente a su eficiencia y escalabilidad [2]. Los tópicos en los que se trabaja actualmente son: el tratamiento de datos en *streaming* y el análisis y comparación del funcionamiento de algoritmos de clustering y clasificación aplicables a datos masivos para posteriormente proponer mejoras a los algoritmos existentes o bien, nuevos algoritmos.

En la gestión de las actividades del proyecto se emplea la Metodología Fundacional para Ciencias de Datos, que consta de diez pasos y es algo similar a otras metodologías reconocidas para Minería de Datos, pero que enfatiza varias de las nuevas prácticas en ciencias de datos como el uso de grandes volúmenes de datos, la incorporación de análisis de texto dentro del modelo predictivo y la automatización de algunos procesos [3]. A continuación se detallan algunas de las actividades realizadas.

2.1. Análisis Bibliométrico

Se trabaja en análisis bibliométrico tradicional y alternativo, midiendo el impacto de publicaciones científicas en sus distintas modalidades de difusión. Actualmente se está elaborando un análisis cuantitativo de publicaciones de autores de instituciones argentinas en la bases de datos SCOPUS de Elsevier [4], accedida desde la Biblioteca Electrónica de Ciencia y Tecnología del Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación. En esta ocasión para las búsquedas se utilizó la palabra clave "pesticida". En algunos casos se aplicó como filtro que las publicaciones correspondiesen a Argentina para identificar y reunir los trabajos en los que al menos uno de los autores incluyera la mención de una institución argentina en los datos de afiliación institucional, a fin de poder comparar con la cantidad de publicaciones del resto del mundo. Los accesos a los datos se realizaron mediante de la API de Scopus [5] y mediante scripts desarrollados en R [6].

2.2. Algoritmos de clustering

El estudio de algoritmos de agrupamiento o *clustering* aplicables a datos masivos incluyó, en una primera etapa, la búsqueda bibliográfica sobre el funcionamiento y propósito general de esta clase de algoritmos y los diferentes tipos y características de los mismos. Posteriormente se analizaron las problemáticas que conlleva la masividad de los datos sobre los algoritmos de clustering, como por ejemplo la pérdida de eficiencia debido a la “Maldición de la Dimensionalidad”. Se realizó una búsqueda sobre algoritmos desarrollados para sortear los problemas vinculados a los grandes conjuntos de datos, considerando tanto algoritmos específicos para la problemática como implementaciones de algoritmos tradicionales en plataformas de procesamiento de grandes cantidades de datos.

Luego del relevamiento se decidió trabajar con Halite [7], un algoritmo de agrupamiento novedoso diseñado para abordar las problemáticas de la alta dimensionalidad y el gran volumen de los conjuntos de datos. Se estudió en profundidad el funcionamiento del algoritmo, y se realizó una implementación del mismo a fin de poder comprobar su funcionamiento con conjuntos de datos de prueba del repositorio de la Universidad de California en Irvine (UCI, <https://goo.gl/c8s2kg>).

Actualmente se trabaja en una comparativa de Halite con algoritmos de clustering tradicionales [8] implementados en la herramienta Spark [9].

2.3. Algoritmos de clasificación

Debido a la gran cantidad de problemas de aplicación, la clasificación es uno de los temas más estudiados dentro del campo de la minería de datos. La tarea de clasificación consiste en la construcción de un modelo basado en instancias de datos clasificadas, que sea capaz de predecir o clasificar nuevos ejemplos de datos cuya etiqueta o clase se desconoce. La tarea de construir este modelo

a partir de los denominados “datos de entrenamiento” no es una tarea sencilla debido a que éstos suelen ser ruidosos o a que algunas de las características más importantes son desconocidas, sumado a los costos que puede acarrear el cálculo de la similitud entre los datos de entrenamiento y el aprendizaje del modelo. Hemos abordado el estudio de los principales métodos de clasificación entre los que se pueden mencionar [9]:

- El *algoritmo de Naive Bayes* ha demostrado ser fácil de implementar, poseer eficiencia computacional y en la tasa de clasificación y obtener resultados precisos cuando el número de registros es grande.
- *Árboles de Decisión*: en esta técnica se divide el conjunto de datos de entrenamiento en dos o más partes utilizando alguno de sus atributos. Este proceso se repite hasta que solo haya datos de la misma clase en cada rama del árbol. La elección del atributo se puede llevar a cabo mediante el cálculo de la entropía o del índice Gini. Las principales ventajas de esta técnica se encuentran en que los modelos son fáciles de interpretar, tiene una implementación simple, soporta valores continuos y discretos y es tolerante al ruido presente en los datos. Una de las desventajas señaladas es que no tiene buen desempeño cuando el conjunto de datos de entrenamiento es pequeño o ruidoso.
- *Redes Neuronales*: entre los algoritmos más conocidos en esta área se encuentran los llamados Mapas Auto-organizados de Kohonen. Entre las principales ventajas se menciona la facilidad de uso, no necesita ser reprogramada y es aplicable a una amplia gama de problemas; mientras que las principales desventajas son la complejidad en determinar la cantidad de neuronas y capas necesarias, el tiempo de procesamiento elevado y que pueden tener un aprendizaje lento.
- *Algoritmos Genéticos*: estas técnicas inspiradas en la teoría de la evolución natural han sido aplicadas para la clasificación y debido a que realizan una

búsqueda global e independiente del dominio, los convierte en una herramienta robusta, escalable y aplicable en distintas etapas del proceso de extracción de conocimiento.

Adicionalmente, y debido a la evolución de las tecnologías de Big Data, un argumento que justifica su escalabilidad es que su funcionamiento facilita las implementaciones paralelas. Sin embargo una de las principales dificultades es el proceso de evaluación de individuos, aspecto en el que se continúa investigando.

2.4. Sistema de procesamiento de streaming de datos

Otra parte de los esfuerzos del grupo se hallan abocados al estudio del procesamiento de datos en streaming, tema que cobra cada vez más protagonismo, tanto a nivel académico como por su capacidad de aportar a la Inteligencia de Negocios de las organizaciones. Ya no es suficiente con ser capaces de procesar grandes cantidades de datos extraídos de repositorios o generados por las organizaciones, sino que deben ser procesados de manera rápida, o en “real time”, además de generar información precisa. Los datos en streaming pueden provenir de diversas fuentes, como archivos de registros generados por los clientes que utilizan sus aplicaciones móviles o web, compras electrónicas, información de redes sociales, operaciones bursátiles o servicios geoespaciales. Algunos casos en los que resulta útil el análisis de streaming de datos son la detección de fraudes, monitoreo de sistemas, intercambios comerciales y demás. El procesamiento de streams en tiempo real está diseñado para analizar y actuar en función de información a medida que la misma se genera, mediante el uso de consultas continuas (consultas del tipo SQL que operan sobre ventanas temporales e informacionales) [11]. Esto requiere un cambio de paradigma en cuanto al almacenamiento, obtención y procesamiento

de la información. Las “bases de datos tradicionales” no fueron concebidas para este propósito por lo que debe hacerse uso de otras herramientas que otorguen la potencia y versatilidad que se requieren. Con este fin se busca la definición de una arquitectura capaz de procesar estos streams. Los componentes de la arquitectura deben ser capaces de interconectarse entre sí, proveer una alta tolerancia a fallas y permitir una escalabilidad elevada. A su vez, resultan atractivas aquellas herramientas basadas en el Software Libre, que se hallan respaldadas en el conocimiento colectivo de su comunidad. La propuesta en la que se trabaja es el desarrollo de la arquitectura de un sistema de procesamiento de datos en streaming, capaz de responder con una latencia máxima de 30 segundos a partir de un volumen de 100 eventos/seg.

En este desarrollo se utilizan las siguientes herramientas:

- *Kafka*: broker de mensajería, utilizado para centralizar la recepción de información sobre los eventos que se produzcan.
- *Zookeeper*: mecanismo de sincronización distribuido, que mantiene el estado y configuración de las demás piezas de software del sistema.
- *Docker*: tecnología de contenerización, distribución de aplicaciones y virtualización, su objetivo es garantizar sencillez de despliegue y posibilidad de escalado de la arquitectura.
- *Storm*: sistema distribuido de procesamiento de eventos en streaming, capaz de definir los “camino” y “transformaciones” que sufren los eventos para poder extraer datos de interés para la organización.
- *Redis*: Base de datos NoSQL, del tipo clave-valor, utilizada para permitir la reconfiguración del sistema sin necesidad de Down-times.

3. Resultados obtenidos y esperados

Con este proyecto se espera proponer modificaciones o mejoras a los algoritmos de clustering o de clasificación existentes para datos masivos. Además se espera obtener una herramienta eficiente en el análisis de datos en streaming. Mediante la aplicación de la Metodología Fundacional para Ciencias de Datos se espera obtener resultados favorables en cada uno de los proyectos que se desarrollen.

4. Formación de Recursos Humanos

Este proyecto dio inicio a una nueva línea de investigación dentro del Grupo de investigación en Bases de Datos de la Fac. Reg. Concepción del Uruguay de la U.T.N.. Tres de los investigadores del proyecto están desarrollando tesis de maestría. En el proyecto colaboran dos becarios graduados con beca de iniciación a la investigación, que tienen previsto la realización de posgrados en el área temática del proyecto. Además participan en el proyecto cuatro becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación. Tres de ellos están realizando su práctica supervisada en la temática de análisis de streaming.

5. Referencias

- [1] Fan Wei and Bifet Albert. Mining big data: Current status, and forecast to the future. SIGKDD Explor. Newsl.,14(2):1–5, apr 2013.
- [2] Larose Daniel T. Discovering Knowledge in Data: An Introduction to Data Mining. Wiley-Interscience, 2004.
- [3] Foundational Methodology for Data Science. IBM. Disponible en <http://goo.gl/w8tT9J> . Accedido el 14 de Marzo de 2017.
- [4] SCOPUS. <http://www.scopus.com>. Accedido 03/2017.
- [5] Scopus API. <http://http://goo.gl/umz8m6>. Accedido 03/2017.
- [6] R Project. <https://www.r-project.org/> Accedido 03/2017.
- [7] Robson L. F. Cordeiro, Christos Faloutsos, and Caetano Traina Junior. 2013. Data Mining in Large Sets of Complex Data. Springer Publishing Company, Incorporated.
- [8] J.H. Orallo, M.J.R. Quintana, and C.F. Ramírez. Introducción a la minería de datos. Editorial Pearson, 2004. ISBN: 84 205 4091 9. 2004.
- [9] Spark Streaming. <http://spark-project.org> Accedido 03/2017.
- [10] Lyubchik, L.,Grinberg, G. Real time recursive preference learning to rank from data stream. 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP) Data Stream Mining & Processing (DSMP), IEEE First International Conference on. :280-285 Aug, 2016
- [11] Tyler Akidau. The world beyond batch: Streaming 101. <https://goo.gl/xhPVZQ>. Accedido 03/2017.

Transformando Datos de Biodiversidad en Linked Data

Marcos Zarate^{1,2}, German Braun³, Samuel Almonacid²

¹ Laboratorio de Investigación en Informática (LINVI), Facultad de Ingeniería
Universidad Nacional de la Patagonia San Juan Bosco.

² Centro para el Estudio de Sistemas Marinos, Centro Nacional Patagónico Consejo
Nacional de Investigaciones Científicas y Técnicas (CESIMAR-CENPAT-CONICET)

³ Grupo de Investigación en Lenguajes e Inteligencia Artificial Departamento de Teoría
de la Computación

Facultad de Informática, Universidad Nacional del Comahue
{zarate,almonacid}@cenpat-conicet.gob.ar, german.braun@fi.uncoma.edu.ar

Resumen

La biodiversidad es esencial para la vida en la tierra y motiva muchos esfuerzos en la recopilación de datos sobre especies, que son utilizados por investigadores para el estudio de los seres vivos. Sin embargo, dados que estos datos se extraen desde diferentes lugares geográficos y se almacenan en distintos formatos, su recuperación, combinación e integración es aún un problema abierto.

El objetivo general de este trabajo de investigación es desarrollar una arquitectura para convertir y publicar datos de biodiversidad utilizando tecnologías de la Web Semántica, en particular los principios establecidos por la iniciativa Linked Open Data (LOD) para compartir y relacionar información. Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional de la Patagonia San Juan Bosco, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

Palabras clave: Linked Data, Web Semántica, Biodiversidad, RDF

Contexto

El LINVI es el Laboratorio de Investigación en Informática dependiente de la Facultad de Ingeniería de la

Universidad Nacional de la Patagonia San Juan Bosco. La especialidad del laboratorio es la investigación, desarrollo, servicios de vinculación y formación de recursos humanos en Informática, en particular esta línea de investigación se incluye dentro del proyecto *Clasificación de Información en BigData mediante la utilización de Técnicas de Inteligencia Artificial y Análisis de Redes Sociales*.

Este trabajo también está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes y Web Semántica*, cuya duración es de cuatro años, y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de dos becas internas doctorales de 5 años de duración, finalizando en abril de 2019 y 2020.

Por otra parte, el CESIMAR unidad ejecutora dependiente del CONICET desarrolla líneas de investigación y forma recursos humanos orientados a comprender el funcionamiento de los ecosistemas marinos, y brindar soporte científico y tecnológico para el manejo y la conservación de los recursos del mar. Estas entidades cuentan con Convenios de Colaboración mutua y han realizado

trabajos conjuntos en temáticas relacionadas con los sistemas de tiempo real, inteligencia artificial, redes y procesamiento de imágenes.

1. Introducción

Las colecciones biológicas contienen información heredada irremplazable sobre nuestra biosfera, que es esencial para comprender como la biodiversidad está cambiando en una era de impactos humanos sin precedentes. Dichos análisis solo son prácticos si se digitalizan, integran y hacen disponibles en línea los datos de las colecciones biológicas de todo el mundo. Estas tareas son un foco importante del campo de la informática y, aunque presentan muchos desafíos, también prometen ofrecer beneficios significativos para la ciencia que estudia la biodiversidad y sus disciplinas asociadas. En los últimos años, la comunidad informática de la biodiversidad ha hecho grandes progresos hacia el logro de este objetivo mediante la creación de vocabularios comunes compartidos, como Darwin Core (DwC) [1] y mecanismos de publicación como el Integrated Publishing Toolkit (IPT) [2]. Gracias a estas y otras iniciativas nacionales e internacionales, ahora existen cientos de millones de registros de biodiversidad de todo el mundo.

Aunque se han conseguido logros sustanciales en llegar a un consenso sobre los estándares de datos, formatos de metadatos y mecanismos de intercambio por parte de las organizaciones de normalización como el Open Geospatial Consortium, Inc. (OGC¹), la Organización Internacional de Normalización (ISO²) y

el World Wide Web Consortium (W3C³), todavía hay una serie de vacíos que impiden la plena interoperabilidad entre los sistemas de información de biodiversidad. Los científicos son desafiados por el volumen y la heterogeneidad de los tipos de datos y formatos [3], y la dificultad de descubrir, acceder e integrar conjuntos de datos desde múltiples fuentes [4,5].

La Web Semántica [6] y sus tecnologías asociadas proporcionan una solución natural a estos problemas al permitir una red de datos y conocimientos vinculados donde todos los objetos de datos son identificados de manera única y las relaciones entre ellos están explícitamente definidas [6,7]. Por consiguiente, cada vez son más reconocidas las ventajas de las tecnologías asociadas a la Web Semántica y a LOD, no solo en la investigación sobre la diversidad biológica y sus disciplinas conexas (por ejemplo, [8,9,10,11]), sino también en las ciencias de la vida [12,13,14].

En la actualidad uno de los portales de biodiversidad más utilizados por la comunidad científica internacional pertenece a la organización intergubernamental Global Biodiversity Information Facility (GBIF⁴), cuyo objetivo es estructurar nodos nacionales en una red accesible vía Web de manera libre y gratuita. GBIF nace en 2001 y comprende en la actualidad 53 países y 43 organizaciones internacionales.

El nodo Argentino de GBIF es el Sistema Nacional de Datos Biológicos (SNDB⁵), una iniciativa del Ministerio de

¹ <http://www.opengeospatial.org>

² <http://www.iso.org/iso/home.html>

³ <https://www.w3.org/>

⁴ <http://www.gbif.org/>

⁵ <http://datos.sndb.mincyt.gob.ar/>

Ciencia, Tecnología e Innovación Productiva, conjuntamente con el Consejo Interinstitucional de Ciencia y Tecnología (CICYT) y enmarcada dentro del Programa de Grandes Instrumentos y Bases de Datos. Sin embargo, aunque GBIF soporta interoperabilidad a través de su estándar para datos de biodiversidad Darwin Core, en la actualidad, estos datos no están publicados siguiendo los criterios establecidos por la iniciativa LOD, lo que representa una desventaja a la hora de compartir e integrar información. El objetivo de esta línea de investigación es estudiar e implementar modelos y arquitecturas que permitan la publicación de datos de biodiversidad en la Web de Datos. En el ámbito de este trabajo proponemos analizar diferentes tecnologías y formalismos que se requieren para poder ofrecer una fuente de datos enlazada y abierta, con información de biodiversidad, como así también enfoques para su integración y herramientas para consumir estos datos que se beneficien de su forma estandarizada de representación y de la posibilidad de enlazar nuevas fuentes de datos en tiempo de ejecución. En particular, nos centraremos en la preparación y conversión de los datos a RDF [15], la definición de las URIs y el interenlazado de nuestras fuentes actuales. Como trabajo futuro, planeamos desarrollar un prototipo de aplicación Web específica para visualizar estos datos y así, aumentar la visibilidad fomentando la colaboración entre grupos interdisciplinarios.

El presente trabajo se estructura de la siguiente forma. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación, el problema que se estudia y los objetivos. En la sección 3

indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática

2. Líneas de Investigación, Desarrollo e Innovación

El proyecto de investigación *Clasificación de Información en BigData mediante la utilización de Técnicas de Inteligencia Artificial y Análisis de Redes Sociales* tiene como objetivo evaluar técnicas existentes e implementar desarrollos experimentales que permitan clasificar, ordenar, jerarquizar y analizar información sobre grandes volúmenes utilizando tecnologías semánticas.

Por otro lado, el proyecto de investigación *Agentes Inteligentes y Web Semántica* tiene varios objetivos generales. Uno de ellos es el de desarrollar conocimiento especializado en el área de Interoperabilidad Semántica de la Información. En este sentido, se estudian, entre otras, el desarrollo de agentes de información cuyo ambiente de trabajo es la Web, además de técnicas de representación de conocimiento, razonamiento automático y modelado ontológico.

Estas líneas de investigación confluyen en el estudio de formalismos y tecnologías para cubrir las necesidades emergentes de compartir, actualizar e integrar el conocimiento de sistemas computacionales pre-existentes. Particularmente, hemos escogido experimentar en la publicación de datos de biodiversidad para generar datos abiertos y enlazados con otras fuentes disponibles en la Web. Para ello, se utilizarán conjuntos de datos científicos existentes de investigadores colaboradores del Centro Nacional

Patagónico (CENPAT-CONICET), con la finalidad de contar con expertos que nos permitan realizar y estructurar los mecanismos de clasificación y evaluar los resultados de los desarrollos experimentales en estos conjuntos de datos.

3. Resultados Obtenidos y Trabajo Futuro

Inicialmente, se realizó un relevamiento de las tecnologías disponibles para nuestra arquitectura, y se puntualizó sobre las plataformas D2RQ [16], Jena [17], OpenRefine [18] y GraphDB [19], entre otras. De este análisis, se determinó que las plataformas más convenientes para un primer prototipo son OpenRefine y GraphDB. El primero soporta extensiones para la creación de triplas RDF⁶ a partir de una gran variedad de formatos de entrada, tales como CSV, hojas de cálculo, JSON y el mismo formato RDF. Además, permite explorar y depurar los conjuntos de datos, aplicar transformaciones y definir vocabularios asociados a los diferentes campos, de una manera amigable. Por otro lado, GraphDB es un repositorio semántico que permite almacenar las tripletas generadas por OpenRefine y también trabajar con motores de inferencia y de consultas SPARQL [20] sobre estos datos estructurados.

La Fig. 1 muestra la arquitectura inicial para la conversión de los datos. Dicha arquitectura acepta datos de biodiversidad en formato tabular estándar, incluyendo DwC-A. El proceso de depuración y exploración de datos involucra la definición de URIs y los diferentes

vocabularios basados, principalmente, en el estándar Darwin Core⁷ y aquellos para la descripción de datos personales, filiatorios y de recursos, tales como FOAF⁸ y Dublin Core⁹, respectivamente. Este proceso, inicialmente manual, genera una salida en formato JSON, para procesar los siguientes repositorios de manera automática.

Actualmente, se está trabajando en la vinculación de nuestros repositorios, como un proceso de interenlazado interno, de depuración y exploración, y con otros repositorios RDF externos. En todos los casos, se está experimentando con LIMES¹⁰ y SILK¹¹, las cuales son herramientas automáticas para la detección de enlaces en la Web de Datos. Finalmente, está previsto diseñar un prototipo para la visualización y consulta de los datos generados y enlazados, además de su ontología relacionada. Dicho prototipo estará basado en la herramienta para modelado ontológico crowd [21].

4. Formación de Recursos Humanos

Dos de los autores de este trabajo están inscriptos en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur, mientras que uno de ellos se encuentra inscripto en el Doctorado en Ingeniería Mención en Procesamiento de Señales e Imágenes en

⁶ <http://refine.deri.ie/>

⁷ <http://rs.tdwg.org/dwc/rdf/dwcterms.rdf>

⁸ <http://xmlns.com/foaf/spec/>

⁹ <http://dublincore.org/documents/demi-terms/>

¹⁰ <http://aksw.org/Projects/LIMES.html>

¹¹ <http://silkframework.org/>

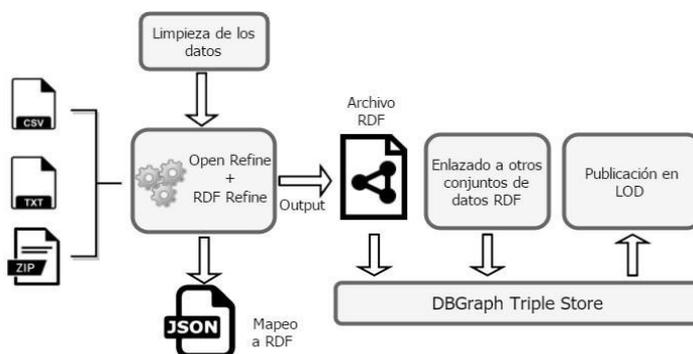


Fig. 1. Arquitectura inicial para la transformación de datos de biodiversidad a RDF.

la Universidad Tecnológica Nacional Facultad Regional Buenos Aires. Los mismos cuentan con beca interna doctoral del CONICET. El presente trabajo se enmarca dentro de la investigación realizada para el desarrollo de las tesis doctorales, en donde se investigan las diversas formas de integrar conjuntos de datos de biodiversidad de diversos repositorios digitales usando tecnologías de la Web Semántica. El tema que se presenta viene profundizándose mediante el estudio continuo y con la presentación de diferentes trabajos en reuniones científicas en donde se muestran los avances realizados y los posibles resultados que se esperan de la investigación.

Referencias

- [1] John Wieczorek, David Bloom, Robert Guralnick, Stan Blum, Markus Doring, Renato Giovanni, Tim Robertson, and David Vieglais. Darwin core: an evolving community-developed biodiversity data standard. *PLoS one*, 7(1):e29715, 2012. Author. (year, month). Title. Presented at Conference title. [Type of Medium]. Available: site/path/file
 - [2] Tim Robertson, Markus Doring, Robert Guralnick, David Bloom, John Wieczorek, Kyle Braak, Javier Otegui, Laura Russell, and Peter Desmet. The gbif integrated publishing toolkit: facilitating the efficient publishing of biodiversity data on the internet. *PLoS One*, 9(8):e102623, 2014.
 - [3] Anne E Thessen and David J Patterson. Data issues in the life sciences. *ZooKeys*, 150(150):15–51, 2011.
 - [4] CL Chandler, RC Groman, A Shepherd, MD Allison, D Kinkade, S Rauch, PH Wiebe, and DM Glover. Using controlled vocabularies and semantics to improve ocean data discovery. In *AGU Fall Meeting Abstracts*, 2013.
 - [5] Tanu Malik and Ian Foster. Addressing data access needs of the long-tail distribution of geoscientists. In *Geoscience and Remote Sensing Symposium (IGARSS)*, 2012 IEEE International, pages 5348–5351. IEEE, 2012.
 - [6] Tim Berners-Lee, James Hendler, Ora Lassila, et al. The semantic web. *Scientific american*, 284(5):28–37, 2001.
 - [7] Tom Heath and Christian Bizer. *Linked data: Evolving the web into a global data space*. Synthesis lectures on the semantic web: theory and technology, 1(1):1–136, 2011.
 - [8] Roderic DM Page. Biodiversity informatics: the challenge of linking data and the role of shared identifiers. *Briefings in bioinformatics*, 9(5):345–354, 2008.
 - [9] Cynthia S Parr, Robert Guralnick, Nico Cellinese, and Roderic DM Page. Evolutionary informatics: unifying knowledge about the diversity of life. *Trends in ecology & evolution*, 27(2):94–103, 2012.
 - [10] Joshua S Madin, Shawn Bowers, Mark P Schildhauer, and Matthew B Jones. Advancing ecological research with ontologies. *Trends in ecology & evolution*, 23(3):159–168, 2008.
 - [11] Andrew R Deans, Matthew J Yoder, and James P Balhoff. Time to change how we describe biodiversity. *Trends in ecology & evolution*, 27(2):78–84, 2012.
 - [12] Robert Stevens, Carole A Goble, and Sean Bechhofer. Ontology-based knowledge representation for bioinformatics. *Briefings in bioinformatics*, 1(4):398–414, 2000.
 - [13] Judith A Blake and Carol J Bult. Beyond the data deluge: data integration and bio-ontologies. *Journal of biomedical informatics*, 39(3):314–320, 2006.
 - [14] Huajun Chen, Tong Yu, and Jake Y Chen. Semantic web meets integrative biology: a survey. *Briefings in bioinformatics*, 14(1):109–125, 2013.
 - [15] Ora Lassila and Ralph R Swick. Resource description framework (rdf) model and syntax specification. 1999.
 - [16] Christian Bizer and Andy Seaborne. *Treating non-rdf databases as virtual rdf graphs*.
 - [17] Brian McBride. *Jena: A semantic web toolkit*. IEEE Internet computing, 6(6):55–59, 2002.
 - [18] Ruben Verborgh and Max De Wilde. *Using OpenRefine*. Packt Publishing Ltd, 2013.
 - [19] Ralf Hartmut Guting. *Graphdb: Modeling and querying graphs in databases*. In *VLDB*, volume 94, pages 12–15, 1994.
 - [20] Eric Prud, Andy Seaborne, et al. *Sparql query language for rdf*. 2006.
- Christian Gimenez, German Braun, Laura Cecchi, and Laura Fillotrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *Proc. of the 2nd Simposio Argentino de Ontologías y sus Aplicaciones (SAOA)* co- located at *Jornadas Argentinas de Informatica (JAIIO)* - to appear , 2016.

Un Método para la Construcción y Mantenimiento de una CMDB para Áreas de Informática Internas de Empresas de Servicios: Gestión de Fuentes de Información

Cecilia Massano, Carlos Salgado, Alberto Sanchez, Mario Peralta
 Departamento de Informática Facultad de Ciencias Físico-Matemáticas y
 Naturales Universidad Nacional de San Luis
 Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
 e-mail: ceciliamassano@gmail.com, {csalgado, alfanego, mperalta}@unsl.edu.ar

Resumen

Existen diferentes factores que conllevan a la necesidad de dirigirnos hacia el desarrollo de servicios de calidad en la industria del software. Diversos modelos de calidad plantean la necesidad de gestionar los cambios en los servicios. Es condición *sine equa non* la gestión de los elementos de configuración que soportan la prestación de un servicio, en pos de contar con toda la información necesaria para la gestión de cambios y diseño de servicios[1]. Esto es un desafío para los diseñadores de servicios, que se encuentran ante la necesidad de contar con información actualizada respecto a la gestión de dichos elementos de configuración. El acceso a la información que signifique un cambio en el diseño de los servicios en tiempo y forma, la elección de una granularidad adecuada para la gestión de estos elementos de configuración, la definición de una forma sistemática para establecer las relaciones entre los mismos, son algunos de los problemas a los que se enfrentan los diseñadores de servicios.

Este trabajo pretende brindar un método para la construcción y mantenimiento de una CMDB

(Configuration manager data base – Base de datos de gestión de configuración), centrando nuestra atención en la gestión de las fuentes de información.

Palabras Clave: CMDB, Servicio TI, Elementos de Configuración, Fuente de Información.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Esta investigación surge en el marco del Proyecto de Tesis de la Maestría en Calidad de Software (Ord. 017/09-CD) de la Ing. María Cecilia Massano, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Dicho Proyecto de Tesis se titula: *Un método para la*

construcción y mantenimiento de una CMDB para áreas de informática internas de empresas de servicios.

Introducción

En la industria del software se presenta la necesidad de demostrar que la construcción de productos de calidad es una realidad constante, un factor de competitividad fundamental. De la misma manera, los sectores de la industria dedicados a la prestación de servicios están trabajando en pos de prestar servicios de calidad. Reason et al[2] enumeran tres razones fundamentales por la cual es necesario diseñar los servicios en las organizaciones: una razón económica, una razón social y una razón social.

Las empresas que se dedican a la prestación de servicios, y organizaciones que cuentan con áreas de sistemas dedicadas a brindar servicios a otras áreas, dirigen sus esfuerzos a brindar servicios de calidad a partir de la aplicación de algunos modelos de calidad conocidos en la industria. ISO 20000[3]e ITIL®[4], CMMI Services®[5]. Son ejemplos de modelos que tienen por objetivo establecer un marco de trabajo que asegure la calidad de los servicios prestados por una entidad.

Particularmente, ITIL® en su versión 3 plantea la necesidad de generar y mantener una base de datos de administración de configuración, conocida como CMDB. Una buena gestión del cambio es un factor clave para tener éxito en esta tarea. Este trabajo presenta una propuesta respecto a cómo construir y mantener una CMDB de forma eficiente y asegurando que la información esté actualizada, centrándose en la gestión de las fuentes

de información.

Aunque existe diversa bibliografía referida a formas de construcción de una CMDB, estas pierden el énfasis en la necesidad del mantenimiento en un entorno de constante cambio. No resuelve la problemática, únicamente la propuesta de un conjunto de pasos a ejecutar en un proyecto de implementación de CMDB. Es necesario tener un plan de acción con guías que estén preparadas para todos estos cambios de contexto. Vamos a centrar nuestra atención en las fuentes de información, como parte fundamental de este método de mantenimiento de una CMDB.

Líneas de Investigación, Desarrollo e Innovación

En base a lo expresado anteriormente, se presenta la situación de la industria que pretende justificar la propuesta, y un acercamiento a los modelos de calidad que enmarcan este trabajo.

A. Situación de la industria: problemáticas y beneficios de la gestión de una CMDB

Dando un contexto temporal y espacial al problema planteado, se puede aseverar que el mercado ha ido mutando, siendo en la actualidad, la tecnología, un servicio básico para garantizar el crecimiento económico de las empresas. Según una encuesta del Ceissi Argentina, un 84% de las empresas en el país consideran que las tecnologías modifican el modelo de su negocio y plantean la necesidad de realizar inversiones en los años venideros, en pos de mejorar su software e infraestructura tecnológica, para adecuarse a estas nuevas condiciones del mercado.[6]

Este trabajo se centra en las organizaciones enfocadas en la prestación de servicios de TI, ya sea para clientes internos o externos. La problemática plantea que para brindar servicios de TI de calidad, es necesario generar y mantener una CMDB donde se visualice su estructura. Para esto, es fundamental la elección y correcta gestión de las fuentes de información. Así se podrán identificar los elementos de configuración que forman un servicio TI.

Esto se basa en los recurrentes fracasos de las organizaciones en la implementación y mantenimiento de las CMDB. Existen numerosos artículos que avalan esta aseveración de distintas empresas que presentan sus propias experiencias. Para mencionar alguno, la corporación BDNA enuncia en su informe Data Quality Issues in de the CMDB publicado en 2013[7], cuáles son los 5 problemas identificados en la calidad de los datos de las CMDB, los cuales pueden ser: inconsistentes, duplicados, irrelevantes, incompletos, desactualizados o tener conflictos

En el informe, la corporación explica que a partir del estudio de datos de diferentes clientes ha determinado que este conjunto de problemas no está directamente ligado a los diferentes involucrados (clientes, proveedores, usuarios), y la forman en que llevan la implementación de una CMDB, sino simplemente a la complejidad intrínseca y el dinamismo del mundo de las tecnologías de la información en el que nos vemos inmersos actualmente. Esta complejidad está dada, entre otros puntos, por la amplia diversidad de fuentes de datos para armar una CMDB.

Contar con una CMDB tiene múltiples

beneficios. Se citan algunos de diversas fuentes:

- Es necesaria para controlar, mantener y mejorar el ambiente de un conjunto de servicios
- Contiene todos los activos y configuraciones de una organización y los servicios que la componen
- Contiene información y documentación para dar soporte a los procesos de gestión de servicios
- Asegura la integración entre estos procesos
- Otorga información respecto de la infraestructura de los servicios útil para la toma de decisiones [8]
- Es una base para tener una buena gestión de cambios [9]
- Colabora con la gestión de incidentes y problemas, y la gestión de la entrega del servicio[10]

Luego de analizar y estudiar cada uno de los puntos mencionados como beneficio de contar una CMDB, y teniendo en cuenta la experiencia en distintos grupos de trabajo en la gestión de servicios de TI, es posible agregar que es un beneficio fundamental. Trabajar con una CMDB permite realizar análisis de impacto ante un cambio en un servicio de TI, para determinar si otros servicios y activos serán afectados ante dichos cambios.

B. Normas y modelos de calidad de servicios: ISO 20000®, CMMI Services® e ITIL® como modelos de calidad de servicios TI

La calidad en la prestación de servicios puede decirse que tiene sus orígenes en la ISO/IEC 20000-1®. Esta certificación es una norma internacional que establece los requisitos para certificar la prestación de servicios de TI.

Existe también actualmente el modelo propuesto ITIL, en su versión 2011. Este modelo no es certificable por una organización, pero existe una evaluación en diferentes niveles que puede rendir cada persona de una organización.

En tercer lugar, se incluye como marco de referencia el modelo de CMMI en su constelación de servicios, llamada CMMI Service® versión 1.3.

Resultados y Objetivos

En base a lo expresado anteriormente, se presenta un resumen de la propuesta.

A. Gestión de fuentes de información diversas de una CMDB

1) Macro-proceso: Gestión de activos y configuración

Dentro de la etapa de Transición del servicio, ITIL 2011® incluye el proceso de Gestión de activos y configuración [12]. Tomando la propuesta de Becker, Olsina y Papa en [13], se genera un modelo para trabajar la generación y mantenimiento de una CMDB. Comienza asumiendo que los procesos de ITIL® son macro-procesos. Cada macro-proceso está compuesto por un conjunto de procesos que contienen actividades a desarrollar para lograr los objetivos plantados. El macro-proceso de Gestión de activos y configuración tiene por objetivo mantener los elementos de configuración y sus relaciones para dar soporte a un servicio de TI[12].

El método para la construcción y mantenimiento de una CMDB para áreas de informática internas de empresas de servicios, incluye los siguientes procesos dentro del macro- proceso de Gestión de activos y configuración: Planificación de la gestión de una CMDB, definición de

roles involucrados, gestión de fuentes de información, gestión del diseño de un servicio TI, gestión de activos y gestión de configuración. Este trabajo se centra en la gestión de fuentes de información.

2) Gestión de fuentes de información de CMDB

Se presenta una propuesta para trabajar con las fuentes de información que alimentan y hacer crecer una CMDB que consta de las siguientes actividades:

Analizar fuentes de información disponibles: Establecer cuáles son las fuentes de información para la detección de cambios en la CMDB.

Definir prioridad de las fuentes de información: Para lograr un equilibrio entre la capacidad de trabajo del equipo de Responsables de CMDB y la urgencia de los cambios.

Definir periodicidad de detección de elementos de configuración de fuente de información: Determinar tiempos periódicos o hitos que generan modificaciones en la CMDB.

Determinar roles involucrados: Para establecer con quién relevar los cambios en la CMDB, y a quién consultar ante una posible modificación en la infraestructura de un servicio.

Establecer el o los eventos vinculados a una fuente de información que pueden implicar un cambio en la CMDB: Para identificar los hitos debemos reconocer los eventos que los determinan, disparadores de cambios en la CMDB.

Establecer la capacidad de trabajo del equipo responsable por la CMDB: Para trabajar con estas fuentes en función del tiempo disponible, de manera organizada.

Definir una forma de registro de

cambios pendientes en la CMDB: Definir el mecanismo para registrar los cambios pendientes, el responsable, el momento de actualización, los servicios afectados y la priorización de pendientes.

B. Fuentes de información

Se listan un conjunto de posibles fuentes de información, como un resultado de la investigación planteada:

- Proyectos de desarrollo de software
- Proyectos de innovación de negocio
- Proyectos de modificación de infraestructura tecnológica
- Resultados de pruebas de contingencia
- Cambios mayores a servicios que pueden afectar otros servicios
- Nuevas definiciones respecto a cómo modelar en la CMDB
- Gestión de cambios

C. Avances de la investigación propuesta

La gestión de las fuentes de información para encontrar elementos de configuración antes propuesta está trabajándose y siendo probada en un área interna de sistemas de una empresa de servicios. Es objetivo del trabajo de tesis que enmarca esta investigación, abarcar algunos de estos aspectos, así como brindar nuevas líneas de investigación vinculadas a esta temática.

Formación de Recursos Humanos

Bajo esta línea de investigación se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Ing. María Cecilia Massano, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

Referencias

- [1] A. Parthiban, “CMDB Implementation - A tale of two extremes - White paper”, 2016.
- [2] B. Reason, L. Lovlie y M. B. Lu, *Service Design for Bussines. A practical Guide to Optimizing the Customer Experience*, USA: Wiley, 2016.
- [3] ISO/IEC, *ISO/IEC 20000-2:2012 Parte 2: Código de buenas prácticas.*, 2012.
- [4] Arraj y Valerie, “ITIL The Basics”, 7 2013.
- [5] CMMI Product Team, *CMMI (r) for Services, Version 1.3*, Hanscom AFB: Carnegie Mellon University, 2010.
- [6] cessi ArgentiIna, “Cámara de Empresas de Software & Servicios Informáticos de la República Argentina”, 2016.
- [7] BDNA Corporation, “Data Quality Issues in de the CMDB”, 2013.
- [8] K. Farrell y B. Ethridge, “What do you need from a configuration management database,» 15 5 2007.
- [9] A. Parthiban, “CMDB Implementation. A tale of two extremes”, 2016.
- [10] Axelos, “Axelos Global Best Practice”, 2013.
- [11] B. Orand, *Foundations of IT Service Managment*, 2da ed., J. Villarreal, Ed., North Charleston SC: Createspace Independent Publishing Platform, 2011, p. 348.

**Computación Gráfica,
Imágenes y Visualización**

Análisis de Capturas de Movimientos para la Animación de Humanos Virtuales

Dana Urribarri^{1,2}, Martín Larrea^{1,2}, Silvia Castro^{1,2}

¹Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab)
Departamento de Ciencias e Ingeniería en Ciencias de la Computación (DCIC)
Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

²Instituto de Ciencias e Ingeniería de la Computación (ICIC),
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

{dku, mll, smc}@cs.uns.edu.ar

Resumen

El análisis del movimiento humano (HMA) se refiere al análisis e interpretación de los movimientos humanos en el tiempo. Durante décadas, fue un campo de investigación que atravesaba varias áreas: biología, psicología, multimedia, etc. En el campo de la visión por computadora, el HMA emergió gracias al video y a la aparición de sofisticados algoritmos de dominio público. Las tecnologías de captura de movimiento (Mo-Cap) han agregado al HMA la posibilidad de analizar el movimiento a partir de una representación en 3D del esqueleto. Por otro lado, hoy en día, los ambientes sintéticos habitados por humanos virtuales (HHVV) son habituales en un sinnúmero de aplicaciones. Sin embargo, crear un humano digital es una tarea sumamente compleja. Dado que estamos acostumbrados a cómo luce hasta el último detalle de un humano, cualquier imperfección en el HV es altamente perceptible y produce el rechazo de quien lo observa. La teoría del valle inquietante sostiene que cuanto más cerca se está de lograr algo artificialmente humano, mayor es el nivel de rechazo que hay en los observadores humanos. Un mejor entendimiento de los factores que hacen al movimiento humano reconocible y aceptable es de gran valor en las aplicaciones que requieren realismo en los movimientos de los personajes virtuales.

Palabras clave: *Animación, Mo-Cap, Computación gráfica*

Contexto

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores.

1. Introducción

El análisis del movimiento humano (*HMA*) se refiere al análisis e interpretación de los movimientos humanos en el tiempo [1,3]. Durante décadas, fue un campo de investigación que atravesaba varias áreas: biología, psicología, multimedia, etc. En el campo de la visión por computadora, el *HMA* emergió gracias al video y a la aparición de sofisticados algoritmos de dominio público. Las tecnologías de captura de movimiento (*Mo-Cap*) han agregado al *HMA* la posibilidad de analizar el movimiento a partir de una representación en 3D del esqueleto [14].

Por otro lado, hoy en día, los ambientes sintéticos habitados por humanos virtuales (HHVV) son habituales en un sinnúmero de aplicaciones [6,15,16,24]. Sin embargo, crear un humano digital es una tarea sumamente compleja. Dado que estamos acostumbrados a

cómo luce hasta el último detalle de un humano, cualquier imperfección en el HV es altamente perceptible y produce el rechazo de quien lo observa [2,7,9]. La teoría del valle inquietante sostiene que cuanto más cerca se está de lograr algo artificialmente humano, mayor es el nivel de rechazo que hay en los observadores humanos [22].

Un mejor entendimiento de los factores que hacen al movimiento humano reconocible y aceptable es de gran valor en las aplicaciones que requieren realismo en los movimientos de los personajes virtuales [4,7].

Actualmente existen diversas técnicas para realizar animaciones interactivas en tiempo real [21]; éstas técnicas difieren en el *trade-off* que ofrecen entre la cantidad de control sobre el movimiento, la exactitud y naturalidad del movimiento resultante y el tiempo de cálculo requerido. Elegir la técnica adecuada depende de las necesidades de la aplicación.

La animación realista de un HV es un problema desafiante. Los procesos biomecánicos y fisiológicos que ocasionan el movimiento son difíciles de entender y replicar.

En aplicaciones *offline* los escenarios de movimiento están predefinidos y los animadores planean por adelantado cada detalle del movimiento de los personajes. Sin embargo, en las aplicaciones interactivas esto es imposible, ya que las acciones ocurren dinámicamente y dependen de múltiples factores como la interacción del usuario o el estado actual del mundo. La animación de tales HHVV debe realizarse en tiempo real para permitir la interacción con el ambiente y con otros HHVV.

Es por esto que en muchas aplicaciones se utilizan *Mo-Caps* almacenados en bases de datos que posteriormente se trasladan a los modelos de HHVV para animarlos. Teniendo en cuenta que se debe almacenar una gran cantidad de *Mo-Caps* para obtener diversidad de movimientos y que estos pueden aplicarse solo en escenarios previamente planeados, surge la necesidad de contar con métodos

alternativos para sintetizar humanos que se comporten naturalmente. Una estrategia tradicionalmente empleada es la utilización de las capturas de movimiento conjuntamente con métodos algorítmicos; sin embargo estos últimos aproximan burdamente las restricciones físicas del cuerpo y del entorno y por lo tanto generan artefactos visuales e intersecciones entre los objetos.

El movimiento del cuerpo humano se puede describir desde varios puntos de vista, por ejemplo el mecanismo del movimiento en el espacio y el tiempo, la expresividad cualitativa del movimiento, la trayectoria del movimiento en el espacio, el ritmo y la coordinación del movimiento, entre otras características.

Lograr que los HHVV se muevan de manera aceptable es un desafío que requiere identificar las principales características de los movimientos reales y modelar estos movimientos de manera que permitan ser reproducidos en la animación de HHVV.

2. Líneas de Investigación, Desarrollo e Innovación

En el contexto descripto, hay varias líneas de trabajo que es necesario atacar para enfrentar este desafío:

- En primera instancia es necesario que identificar las propiedades que, además de la trayectoria, hacen al movimiento humano. ¿Qué hace que dos rutinas que ejecutan la misma secuencia de movimientos se perciban de forma diferente?
- Por otro lado, es necesario contar con herramientas que permitan analizar comparativamente diferentes repeticiones de una secuencia de movimientos. Este análisis puede llevar a identificar secuencias correctamente ejecutadas, medir la experiencia de una persona realizando un movimiento, identificar cuáles son las falencias en la realización de una rutina, etc.

- En cuanto a la animación de HHVV, contar con herramientas de comparación permite identificar cuáles son los puntos débiles de los movimientos sintéticos y tomar medidas para corregirlos.

3. Resultados Esperados

El objetivo de esta línea de investigación es analizar el movimiento del cuerpo humano para identificar, entender y modelar las características que deben tener los movimientos sintéticos para minimizar el rechazo y resultar aceptables al público.

Este proyecto se centrará en el análisis comparativo de capturas de movimientos en el dominio específico del karate. Se espera distinguir automáticamente las secuencias realizadas por atletas expertos de atletas con niveles de experiencia menores. En particular se buscará analizar las propiedades cualitativas de los movimientos capturados, establecer distintos parámetros que caractericen las secuencias de movimientos y, finalmente, encontrar técnicas que permitan comparar distintas secuencias de movimientos. De esta forma se espera lograr un análisis comparativo de movimientos realizados por atletas expertos, intermedios y novatos.

4. Bibliografía

- [1] Jasbir Arora and Karim Abdel-Malek. *Human Motion Simulation: Predictive Dynamics*. Academic Press, 1st edition, 2013.
- [2] James E. Cutting and Lynn T. Kozlowski. *Recognizing friends by their walk: Gait perception without familiarity cues*. *Bulletin of the Psychonomic Society*, 9:353–356, 1977.
- [3] Yu Ding, Ken Prepin, Jing Huang, Catherine Pelachaud, and Thierry Artières. *Laughter animation synthesis*. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '14*, pages 773–780, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.
- [4] Rukun Fan, Songhua Xu, and Weidong Geng. *Example-based automatic music-driven conventional dance motion synthesis*. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):501–515, 2012.
- [5] K. Grammer, V. Keki, B. Stribel, M. Atzmüller, and B. Fink. *Bodies in motion: A window to the soul*. In *Evolutionary Aesthetics*, pages 295–323, 2003.
- [6] S. Hagler, D. Austin, T.L. Hayes, J. Kaye, and M. Pavel. *Unobtrusive and ubiquitous inhome monitoring: A methodology for continuous assessment of gait velocity in elders*. *IEEE Transactions on Biomedical Engineering*, 57(4):813–820, 2010.
- [7] Ludovic Hoyet, Kenneth Ryall, Katja Zibrek, Hwangpil Park, Jehee Lee, Jessica Hodgins, and Carol O’Sullivan. *Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies*. *ACM Transactions on Graphics*, 32(6):204:1–204:11, November 2013.
- [8] Perttu Hämäläinen, Sebastian Eriksson, Esa Tanskanen, Ville Kyrki, and Jaakko Lehtinen. *Online motion synthesis using sequential monte carlo*. In *Proc. SIGGRAPH 2014, SIGGRAPH '14*, 2014.
- [9] K. L. Johnson and L. G. Tassinary. *Perceiving sex directly and indirectly: Meaning in motion and morphology*. *Psychological Science*, 16(11):890–897, 2005.
- [10] K. L. Johnson and L. G. Tassinary. *Compatibility of basic social perceptions determines perceived attractiveness*. In *Proc. of the National*

- Academy of Sciences, volume 104, pages 5246–5251, 2007.
- [11] G. Johansson. *Visual perception of biological motion and a model for its analysis*. Perception & Psychophysics, 14:201–211, 1973.
- [12] Dohyung Kim, Minsu Jang, Youngwoo Yoon, and Jaehong Kim. *Classification of Dance Motions with Depth Cameras Using Subsequence Dynamic Time Warping*. In Proceedings of the 2015 8th International Conference on Signal Processing, Image Processing and Pattern Recognition (SIP) (SIP '15). IEEE Computer Society, Washington, DC, USA, 5-8. 2015.
- [13] Chern Hong Lim, Ekta Vats, and Chee Seng Chan. *Fuzzy human motion analysis*. Pattern Recognition. 48, 5 (May 2015), 1773-1796. 2015.
- [14] Liliana Lo Presti and Marco La Cascia. *3D skeleton-based human action classification*. Pattern Recognition. 53, C (May 2016), 130-147. 2016.
- [15] Nadia Magnenat-Thalmann and Zerrin Kasap. *Virtual humans in serious games*. In Proceedings of the 2009 International Conference on CyberWorlds, CW '09, pages 71–79, Washington, DC, USA. IEEE Computer Society. 2009.
- [16] J. Music, M. Cecic, and M. Bonkovic. *Testing inertial sensor performance as hands-free human-computer interface*. WSEAS Transactions on Computers, 8:715–724, April 2009.
- [17] M. Peskin and F. N. Newell. *Familiarity breeds attraction: Effects of exposure on the attractiveness of typical and distinctive faces*. Perception, 33(2):147–158, 2004.
- [18] F. Pollick, J. Kay, K. Heim, and R. Stringer. *Gender recognition from point-light walkers*. Journal of Experimental Psychology: Human Perception and Performance, 31(6):1247–1265, 2005.
- [19] Michalis Raptis, Darko Kirovski, and Hugues Hoppe. *Real-time classification of dance gestures from skeleton animation*. In Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '11), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, 147-156. 2011.
- [20] Gillian Rhodes. *The evolutionary psychology of facial beauty*. Annual Review of Psychology, 57(1):199–226, 2006.
- [21] H. van Welbergen, B.J.H. van Basten, A. Egges, Z.M. Ruttkay, and M.H. Overmars. *Real time character animation: A trade-off between naturalness and control*. In M. Pauly and G. Greiner, editors, Eurographics - State-of-the-Art-Report, pages 45–72, Munich, 2009. Eurographics Association. ISSN: 1017-4656.
- [22] Katsu Yamane, Yuka Ariki, and Jessica K. Hodgins. *Animating non-humanoid characters with human motion data*. In Zoran Popovic and Miguel A. Otaduy, editors, Symposium on Computer Animation, pages 169–178. Eurographics Association, 2010.
- [23] Katja Zibrek, Ludovic Hoyet, Kerstin Ruhland, and Rachel McDonnell. *Evaluating the effect of emotion on gender recognition in virtual humans*. In Proceedings of the ACM Symposium on Applied Perception, SAP '13, pages 45–49, New York, NY, USA, 2013.
- [24] V. B. Zordan, A. Majkowska, B. Chiu, and M. Fast. *Dynamic response for motion capture animation*. ACM Transactions on Graphics, 24:697–701, 2005.

El Uso de la Realidad Virtual Inmersiva en Terapias Motrices

N. Jofré, G. Rodríguez, Y. Alvarado, J. Fernández, R. Guerrero

Laboratorio de Computación Gráfica / Dpto. Informática / FCFMyN /
Universidad Nacional de San Luis

Ejército de los Andes 950, Tel: 0266 4420823, San Luis, Argentina
{npasinetti, gbrodriiguez, ymalvarado, jmfer, rag}@unsl.edu.ar

Resumen

La Realidad Virtual (RV) se ha convertido en una herramienta poderosa con una infinidad de áreas de aplicación en la vida cotidiana, permitiendo abordar tópicos cada vez más complejos entre los cuales se encuentran los vinculados al área de la salud.

Debido principalmente a su característica inmersiva, la realidad virtual ha logrado tener un gran impacto psicológico y físico en las personas gracias a que puede engañar al cerebro y a su vez al sistema motor humano para que experimente y reaccione a sensaciones que no son reales, pero que lo parecen. Esto habilita a recrear situaciones destinadas a tratar y/o estudiar problemas relacionados a la salud física de las personas.

A diferencia de los tratamientos tradicionales, la RV permite incluso alcanzar un grado de exposición superior al que sería posible en vivo, dado que la manipulación de un entorno virtual facilita enfatizar aquellas situaciones a recrear; es por ello que la sensación de inmersión es un aspecto clave.

Esta propuesta de trabajo pretende aplicar RV inmersiva acompañada de dispositivos que permitan capturar la motricidad del cuerpo humano con el fin de utilizarla como una herramienta alternativa para la terapia física en pacientes.

Palabras clave: Realidad Virtual, Computación Gráfica, Interfaces Humano-Computadoras, Interfaz Natural de Usuario, Terapia Física.

Contexto

La propuesta de trabajo se lleva a cabo dentro de la línea de Investigación “*Procesamiento de Información Multimedia*” del proyecto “*Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos*”. Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

Introducción

Durante los últimos años, la tecnología de RV se ha desarrollado significativamente. En cuanto a las aplicaciones de RV, estas comenzaron como simples aplicaciones gráficas hechas para el entretenimiento y el aprendizaje hasta llegar a utilizarse en la actualidad ampliamente en muchas ramas profesionales [1, 2, 3].

Esta revolución tecnológica ha influenciado diversos aspectos en la vida del ser humano, facilitando y mejorando las tareas cotidianas. En pleno auge de innovación, las herramientas tecnológicas han venido a intentar ayudar aún (en ciertos aspectos) a uno de los mayores inconvenientes del ser humano: la salud. Actualmente, las tecnologías avanzadas relacionadas a los gráficos por computadora se encuentran interviniendo en diversas áreas de la salud desde estudios rutinarios, educación médica, operaciones quirúrgicas, diseño e impresión de prótesis 3D, hasta el tratamiento de enfermedades [4, 5].

Es conocido que el uso de técnicas de RV afecta a la percepción e interacción del ser humano, ya que éste se desenvuelve eficientemente en un entorno virtual en tiempo real haciendo uso de sus sentidos y habilidades naturales. Adicionalmente, es posible lograr que a largo plazo el empleo de la RV altere el cuerpo físico mediante el uso de dispositivos de seguimiento corporal. Ambas capacidades pueden ser utilizadas para asistir en las terapias de pacientes, tanto físicas como psicológicas [6, 7].

Uno de los usos de la RV más frecuentemente mencionados en la bibliografía es el tratamiento de condiciones psicológicas tales como fobias y ansiedades por medio de la exposición que se conoce como **Terapia de Exposición a la Realidad Virtual** (VRET - *Virtual Reality Exposition Therapy*). Existen muchos casos de uso exitoso de tales herramientas terapéuticas, como por ejemplo en el tratamiento de la *acrofobia* y el *trastorno de estrés postraumático* [8, 9].

Por otro lado, la terapia física basada en Realidad Virtual es uno de los desarrollos recientes más innovadores y prometedores en tecnología de rehabilitación, en la cual el usuario interactúa con imágenes generadas por ordenador a la vez que puede moverse y manipular objetos virtuales, y realizar otras acciones que incrementan la sensación de inmersión en el ambiente simulado logrando generar la sensación de presencia en el mundo virtual [10].

Los pacientes sometidos a terapia física tradicional pasan por una serie de sesiones realizando ejercicios para ayudar a mejorar el rango de movimiento en las regiones afectadas del cuerpo debido a una enfermedad o lesión. Sin embargo, los pacientes encuentran estas tareas repetitivas y aburridas y terminan por no completar el programa de terapia prescrita. Se ha demostrado que los ejercicios de terapia basada en juegos han llevado a un aumento de las tasas de cumplimiento [11].

En el caso de estudio [12] se describen mejoras en la percepción visual, el equilibrio y

la movilidad luego de once sesiones de rehabilitación con juegos de deportes que usan sensores de movimientos. Recientemente, los videojuegos han sido introducidos en la práctica clínica logrando ser una herramienta beneficiosa para promover las habilidades motoras en pacientes. Si bien, algunos autores sostienen que el contenido y los mecanismos de las intervenciones de rehabilitación rara vez se han definido objetivamente, las oportunidades ofrecidas por los videojuegos y los sensores de movimiento han logrado captar el interés de los terapeutas [13].

En [14] los autores sostienen que los niños con Síndrome de Down tienen puntuaciones más bajas en tareas de equilibrio y agilidad que los niños con otros trastornos. Para estos casos y de acuerdo a investigaciones realizadas es posible mejorar el equilibrio con videojuegos de *Wii-fit* como una terapia eficiente basada en Realidad Virtual.

Es evidente que la aplicación de sistemas de realidad virtual como intervenciones terapéuticas para mejorar las habilidades de movimiento en pacientes con impedimentos motores es un área prometedora de investigación en rehabilitación. Para lograr un sistema de terapia física exitosa se debe considerar que la rehabilitación motriz consiste en el análisis y enseñanza correcta (en calidad y cantidad) de ciertos movimientos motrices; para ello es necesario disponer de los dispositivos de seguimiento corporal adecuados para cada caso.

Con el fin de lograr representaciones creíbles para las distintas terapias mencionadas, es necesario el desarrollo de aplicaciones realistas que incluyan técnicas de visualización digital avanzadas, simulación de escenarios virtuales con entidades estáticas y dinámicas (entre ellas la generación de humanos virtuales) y simulaciones de propiedades físicas; todos estos aspectos son investigados por los autores del presente artículo [15, 16, 17, 18]. Adicionalmente, se requiere el uso de sistemas inmersivos [19] y dispositivos de interacción natural tales como los guantes de datos [20], los sensores de seguimiento corporales, sensores de

cabeza, entre otros.

Líneas de Investigación, Desarrollo e Innovación

En función de lo anteriormente expresado, resulta que la realidad virtual se convierte en una herramienta poderosa que posibilita asistir y/o mejorar las terapias físicas actuales. Considerando que la motricidad es el dominio que el ser humano es capaz de ejercer sobre su propio cuerpo, la idea de esta propuesta es instruir en la enseñanza de ese dominio. Para ello si bien la motricidad es una característica integral ya que intervienen todos los sistemas del cuerpo, resulta necesario establecer una clasificación más específica. Por lo tanto si se tienen en cuenta los movimientos motrices y los dispositivos de RV que permiten detectarlos, se propone clasificar las aplicaciones para terapias del sistema motor humano en dos líneas de investigación.

- *Motricidad Gruesa.*

Los sistemas que detectan motricidad gruesa deben considerar los movimientos drásticos y pronunciados; aquellos que involucran grupos musculares grandes con una significativa aplicación de fuerza, distancia y movimiento. La motricidad gruesa comprende todo lo relacionado con el desarrollo del cuerpo, coordinando desplazamientos y equilibrio, por lo tanto en general su detección se refiere a considerar aquellas acciones realizadas con la totalidad del cuerpo debiendo procesar los grandes grupos musculares y posturales del cuerpo humano.

En consecuencia, debido a su característica de general, la terapia física con RV involucra el uso de dispositivos que permiten detectar la totalidad del cuerpo como por ej. *Kinect*, *Wii*, entre otros [21, 22].

- *Motricidad Fina.*

Los sistemas de terapia para motricidad fina trabajan con movimientos voluntarios mucho más precisos que implican pequeños grupos de músculos. El control de la motricidad fina es la coordinación de músculos, huesos y nervios para producir movimientos pequeños y precisos. Un sistema de motricidad fina debería por ejemplo incluir los mecanismos para que un paciente pueda recoger un pequeño elemento con el dedo índice y el pulgar. Usualmente los sistemas de motricidad fina consideran a todas aquellas acciones que se realizan básicamente con las manos, a través de coordinaciones óculo-manuales. Sin embargo, la motricidad fina también incluye la acción de los pequeños grupos musculares de la cara y los pies. Las aplicaciones de terapia física con RV para motricidad fina requieren de dispositivos de detección de movimiento más precisos, dentro de los cuales se encuentra el *guante de datos* (dispositivo capaz de adquirir información sobre la flexión de los dedos, orientación de la mano, etc.), *guante háptico* (dispositivo capaz de retornar fuerza y/o tacto), y otros dispositivos de control gestual (*Leap Motion*, entre otros) [23, 24, 25].

Resultados y Objetivos

El grupo de trabajo, además de pertenecer a un proyecto de investigación de la UNSL, ha desarrollado tareas dentro del marco de un Proyecto de la Comunidad Europea. Asimismo, se ha obtenido la aprobación de 4 proyectos por parte de la Secretaría de Políticas Universitarias para la concreción de las actividades ya iniciadas.

Como consecuencia, se han desarrollado 2 trabajos finales de carreras asociados al uso de guantes de datos y sensores corporales dentro de un sistema de RV inmersivo, así como también la elaboración de aplicaciones de innovación y desarrollo asociadas a la temática planteada.

Actualmente las acciones se encuentran focalizadas en la incorporación de nuevas estrategias que permitan alcanzar una mejor percepción e interacción al mismo tiempo que se logre alterar el cuerpo físico.

Formación de Recursos Humanos

Los trabajos realizados han permitido la concreción de trabajos de fin de carrera de la Licenciatura en Ciencias de la Computación, así como también la definición de tres trabajos de tesis de Maestría en Ciencias de la Computación, actualmente en ejecución.

Adicionalmente se ha obtenido una beca de iniciación a la investigación y una beca de perfeccionamiento, otorgadas por la Secretaría de Ciencia y Técnica de la UNSL.

Referencias

- [1].J. Vince. *Introduction to Virtual Reality*. Springer London, 2011.
- [2].J. Linowes. *Unity Virtual Reality Projects*. Packt Publishing, 2015.
- [3].F. Steinicke. *Being Really Virtual: Immersive Natives and the Future of Virtual Reality*. Springer International Publishing, 2016.
- [4]. J.D. Westwood. *Medicine Meets Virtual Reality: The Convergence of Physical & Informational Technologies: Options for a New Era in Healthcare*. Studies in health technology and informatics. IOS Press, 1999.
- [5].R.J. Lamson. VR immersion therapy for treating psychological, psychiatric, medical, educational and self-help problems, July 30 2002. US Patent 6,425,764.
- [6].Levin M. F., Weiss P.L., Keshner E.A. *Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles*. Phys Ther. 2015, 95(3), 415-25.
- [7].D.J. Reinkensmeyer and V. Dietz. *Neurorehabilitation Technology*. Springer International Publishing, 2016
- [8].Russell A. McCann at al. Virtual reality exposure therapy for the treatment of anxiety disorders: An evaluation of research quality. *Journal of Anxiety Disorders*, 28(6):625 – 631, 2014
- [9].T. D. Parsons and A. Rizzo. Virtual Human Patients for Training of Clinical Interview and Communication Skills. In *International Conf on Disability, VR and Associated Technology*, Maia, Portugal, September 2008.
- [10]. Davide Corbetta, Federico Imeri, and Roberto Gatti. Rehabilitation that incorporates vr is more effective than standard rehabilitation for improving walking speed, balance and mobility after stroke: a systematic review. *Journal of Physiotherapy*, 61(3):117-124, 2015.
- [11]. Shawn N. Gieser, Peter Sassaman, Eric Becker, and Fillia Makedon. Pot hunter: A virtual reality game for analyzing range of motion. In *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, PETRA '15, pages 3:1 {3:4, New York, NY, USA, 2015. ACM.
- [12]. Judith E Deutsch, Megan Borbely, Jenny Filler, Karen Huhn, and Phyllis Guarrera-Bowlby. Use of a low-cost, commercially available gaming console (wii) for rehabilitation of an adolescent with cerebral palsy. *Physical Therapy*, 88(10):1196, 2008.
- [13]. John Whyte and Tessa Hart. Its more than a black box; its a russian doll: Defining rehabilitation treatments. *American Journal*

- of Physical Medicine Rehabilitation*, 82(8):639-652, 2003.
- [14]. Yee-Pay Wuang, Ching-Sui Chiang, Chwen-Yng Su, and Chih-Chung Wang. Effectiveness of virtual reality using wii gaming technology in children with down syndrome. *Research in Developmental Disabilities*, 32(1):312 - 321, 2011.
- [15]. Yoselie Alvarado et al. Visualización digital avanzada: desafío científico y tecnológico. In *XIII Workshop de Investigadores en Ciencias de la Computación*, pag 344-348, 2011.
- [16]. Yoselie Alvarado et al.. Simulación de entidades en realidad virtual. In *XIV Workshop de Investigadores en Ciencias de la Computación*, 2012.
- [17]. Yoselie Alvarado et al.. Simulación de personajes conversacionales virtuales dentro de un entorno de realidad virtual. In *XV Workshop de Investigadores en Ciencias de la Computación*, 2013.
- [18]. Yoselie Alvarado et al.. Vida artificial y personajes virtuales. In *XVI Workshop de Investigadores en Ciencias de la Computación*, 2014.
- [19]. Yoselie Alvarado et al.. A virtual reality computing platform for real time 3d visualization. In *XVIII Congreso Argentino de Ciencias de la Computación*, 2012.
- [20]. L. N. Jofré et al.. Non-verbal communication interface using a data glove. In *2016 IEEE Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI)*, pag. 1–6, Nov 2016.
- [21]. <http://www.xbox.com/es-AR/xbox-one/accessories/kinect>.
- [22]. <https://www.nintendo.es/Wii/Wii-94559.html>.
- [23]. <http://www.5dt.com/data-gloves/>.
- [24]. <http://www.cyberglovesystems.com/cybertouch/>.
- [25]. <https://www.leapmotion.com/>.

Evaluación de Algoritmos de Refocalización para Radiancias de Arquitectura Plenóptica 1.0

Gustavo Bizai, Facundo Peiretti, Adrián Salvatelli, Bartolomé Drozdowicz

Laboratorio de Sistemas de Información, Facultad de Ingeniería,
Universidad Nacional de Entre Ríos
Ruta 11 Km.10 Oro Verde, Entre Ríos, Argentina
E-mail: {gbizai, asalvatelli, bdrozdo}@bioingenieria.edu.ar

RESUMEN

La fotografía plenóptica ha adquirido especial relevancia en los últimos años. Sus cualidades intrínsecas que permiten elegir a posteriori focos y puntos de vista conforman una potente herramienta, tema de estudio del PIDUNER N°6146. Este proyecto, constituido de diferentes líneas de trabajo, contempla actividades de desarrollo de algorítmica para la generación de imágenes multifoco a partir de imágenes plenópticas. En este trabajo se presentan resultados del ensayo de distintos algoritmos implementados.

Palabras clave: Campos de Luz, Fotografía Plenóptica, multifoco, multiperspectiva.

CONTEXTO

La línea de I+D+i presentada es parte del Proyecto de Investigación y Desarrollo PIDUNER N°6146 “Sistema de Información Plenóptica como medio diagnóstico para Lámparas de Hendidura”, financiado por la Universidad Nacional de Entre Ríos y ejecutado en el ámbito de la Facultad de Ingeniería (FI-UNER).

1. INTRODUCCIÓN

La fotografía denominada “plenóptica” ha adquirido especial interés en los últimos años [1]. Si bien sus orígenes se remontan a principios del siglo pasado, no fue hasta la aparición de alguna capacidad de cálculo y

tecnología para la fabricación de pequeñas lentes que cobró relevancia [2-5]. Las posibilidades que brinda, atractivas en primera instancia desde el punto de vista recreacional, abren un abanico de aplicaciones donde interesa la visualización en tres dimensiones. Idealmente, la función plenóptica [4] actúa como la relación entre los objetos físicos y sus correspondientes imágenes retinales y por eso se dice que es intermediaria entre el mundo y el ojo. La función plenóptica representa los parámetros de la energía radiante del entorno. Es una función 7 dimensional (posición en el espacio (3D), dirección (2D), frecuencia (1D) y tiempo (1D)) que contiene la información visual disponible para un observador en cualquier punto en el espacio. Cada elemento de este espacio 7D corresponde a un único rayo de luz. En 1996 Levoy y Hanrahan [6] propusieron una parametrización que es un subconjunto 4D de la función plenóptica y tomaron de Gershun [7] el término “*light field*” (campo de luz, LF) para definir la radiancia en el espacio como función de la posición y dirección. Dicha reducción de parámetros sucede cuando se trabaja con imágenes estáticas con canales de color separados y con parametrizaciones especiales [8]. Con el LF es posible, a posteriori de la captura, obtener imágenes reenocadas en una región de interés, visualizar la escena desde diferentes ángulos y generar mapas de profundidades de la escena. El PIDUNER propone aplicar esta metodología al diagnóstico de patologías del fondo ocular. Tal es el caso del edema macular, donde la posibilidad de contar con imágenes de la

topografía del fondo ocular es una herramienta de soporte a la toma de decisiones.

El proyecto tiene tres líneas de trabajo, una de las cuales es el desarrollo de la algorítmica para la obtención de imágenes multifoco, multiperspectiva y mapa de profundidades. El presente trabajo presenta los resultados del ensayo de tres algoritmos implementados. Por el momento no se cuenta con imágenes plenópticas del fondo ocular, pero sí hay avances en la configuración de algunas ópticas que han permitido capturar imágenes de escenas pequeñas con buena calidad, preservando las propiedades de radiancia.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La investigación pertenece al Área Computación Gráfica, Imágenes y Visualización. La Fotografía Plenóptica es una forma de Fotografía Computacional que se presenta como una alternativa superadora de la fotografía convencional. La posibilidad de discriminar la dirección desde donde provienen los rayos permite, luego de procesos computacionales, visualizar la escena en diferentes profundidades y perspectivas.

3. RESULTADOS OBTENIDOS

Se realizaron tomas con una cámara plenóptica Lytro F01, la primera generación de esta marca. La F01 posee una matriz de 378 x 379 microlentes ordenadas en geometría hexagonal, de 11 x 11 píxeles cada una y un patrón Bayer para codificar los canales de color RGB. La calibración y posterior decodificación de las capturas fue realizada mediante el uso de una caja de herramientas para Matlab de dominio público denominada "LFToolbox (v0.4)", creada por Donald Dansereau [9][10]. Como resultado de la decodificación de las capturas se obtuvieron LFs en formato de matriz 5 dimensional

$LF(x,y,u,v,ch)$, siendo (x,y) la posición de las microlentes, (u,v) la posición de los píxeles dentro de cada microlente y ch el canal de color RGB. Además, se emplearon algunas imágenes crudas capturadas con otros dispositivos, de diferentes dimensiones (x,y) y (u,v) , también de dominio público [11][12]. Para construir su LF 5D correspondiente, se implementó una función que codifica el LF a partir de la imagen cruda. Una vez obtenidos los LFs, se procedió a ensayar tres algoritmos: Desplazamiento-Suma, Planar Frecuencial y Planar Radiancia [13]. Se diseñaron experimentos para evaluar el rendimiento de los algoritmos desde el punto de vista de sus tiempos de cómputo y capacidad de enfoque en diferentes profundidades, preservando la calidad de la captura original.

Los algoritmos

Desplazamiento-Suma (*'desplaz_suma'*): se basa en el Operador Fotográfico [14] del que se deduce que la refocalización a una determinada distancia es función de un factor de desplazamiento alfa. Las $u \times v$ imágenes de perspectiva son desplazadas según dicho factor y posteriormente promediadas.

Planar Radiancia (*'RefocusRadiancia'*): la refocalización en otro plano diferente del foco óptico original se entiende como una transformación de traslación de la radiancia [15]. Esto significa una operación de sesgado de la radiancia en una parametrización posición-pendiente, conocida como "corte" de la radiancia. En términos de la implementación, se arma una grilla 4D a partir del LF, se aplica la operación de sesgado según un parámetro beta, se interpolan los valores del LF a sus nuevas posiciones y se renderiza la imagen reenfocada.

Planar Frecuencial (*'RefocusEspectro'*): se aplica la Transformada de Fourier a la radiancia y se trabaja en el dominio frecuencial. De manera equivalente al filtro Planar Radiancia, el reenfoque es equivalente

a un corte del hiperespacio transformado [14] [16]. Se obtiene un plano transformado 2D a partir de tomar de la FFT del LF 4D, un pixel en la nueva posición (uu,vv) para cada lenticula (XX,YY) según una relación que involucra un parámetro theta. Una vez interpolado dicho plano, se le aplica la Transformada Inversa de Fourier 2D para obtener la imagen refocalizada.

Tiempos de Cómputo

Se midieron los tiempos de cómputo en una laptop con las siguientes capacidades: Processor: Intel® Core™ i7-6700HQ CPU @ 2.60GHz 2.59GHz. RAM Memory: 16.0 GB (15.8 GB usable). System type: 64-bit Operating System, x64-based processor. Matlab R2014b. La imagen de prueba, tomada con la Lytro F01, resulta de una serie de imágenes a diferentes distancias de separación, según el esquema mostrado en Figura 1.

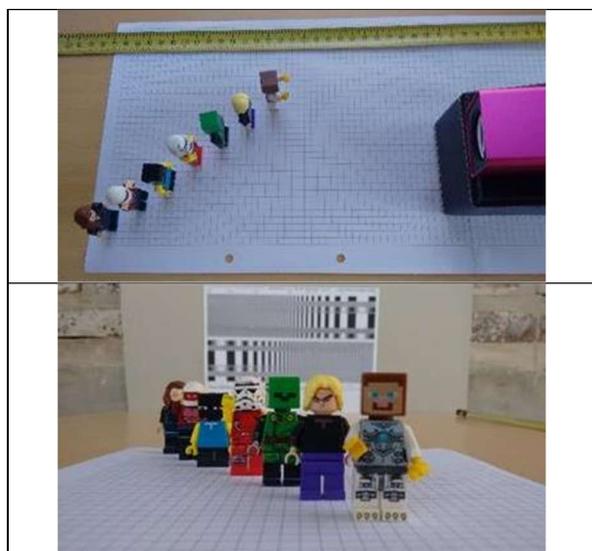


Figura 1.

La primera prueba consistió en correr cada algoritmo sobre la imagen de la Figura 1, con diferentes parámetros alfa, beta y theta respectivamente, para evaluar si los tiempos se ven influenciados. El Gráfico 1 muestra que en ningún caso los tiempos se ven influenciados de manera apreciable por los valores de sus parámetros, excepto los casos de alfa=1 y beta=0, cuando la imagen reenfocada coincide con la captura original en foco óptico. Por otro

lado, puede destacarse la gran diferencia en los tiempos promedio que existe entre el algoritmo *RefocusRadiancia* respecto de los otros dos.

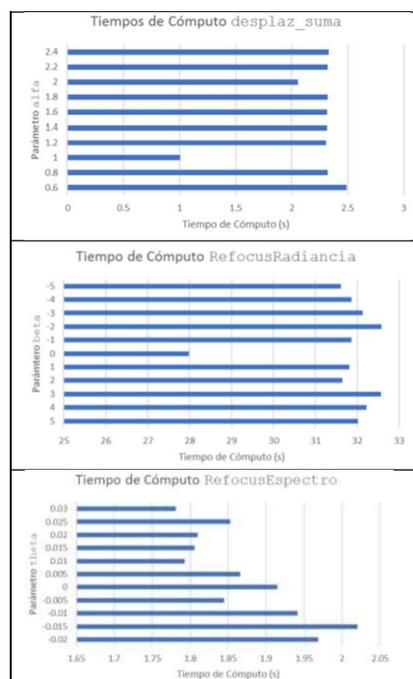


Gráfico 1.

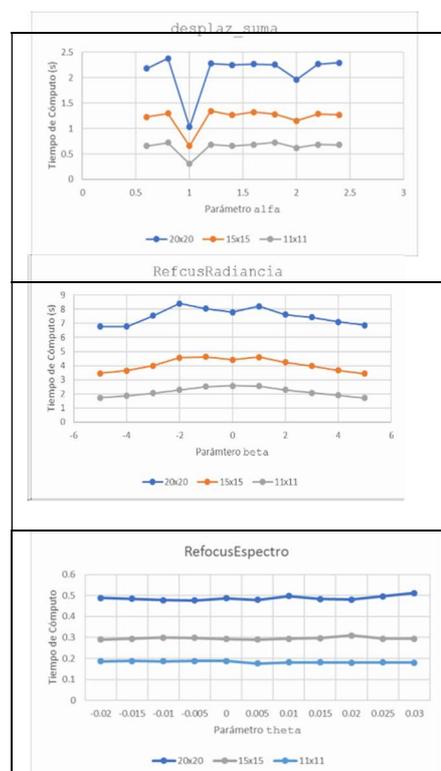


Gráfico 2.

A continuación se ensayaron los algoritmos con imágenes de igual cantidad de microlentes (114 x 114) pero diferente cantidad de pixeles

por lenticula: 20x20, 15x15 y 11x11. Se observan los resultados en Gráfico 2.

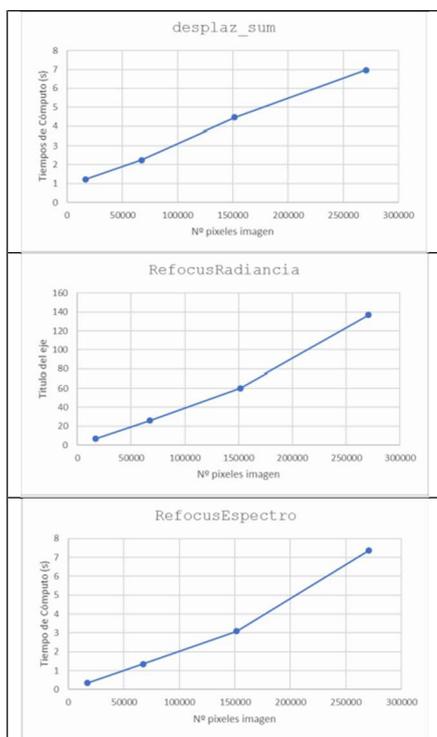


Gráfico 3.

Se puede intuir una relación de linealidad entre estas variables, a mayor cantidad de píxeles por lenticula, mayores son los tiempos de cómputo. A su vez, estos últimos no presentan variación notable con los cambios en los parámetros alfa, beta, theta, al igual que en el caso anterior. Finalmente se estudió la influencia del tamaño de la matriz de microlentes: diferentes cantidades de microlentes con igual número de píxeles por lenticula. Para ello se fue recortando una imagen de mayor tamaño a la de Lytro F01, capturada con Lytro Illum, de libre disposición en [11]. El Gráfico 3 muestra los resultados. Si bien a primera vista pareciera que se conserva la linealidad entre tiempos de cómputo y tamaño de imagen, obsérvese que para las imágenes de mayor tamaño (433x625) los tiempos se disparan para los casos de Planar Radiancia (*RefocusRadiancia*) y Planar Frecuencial (*RefocusEspectro*).

Capacidad de Enfoque (Refocalización)

A los efectos de visualizar de manera más objetiva las regiones en foco, se sustrae a la imagen reenfocada su correspondiente imagen de foco extendido, la cual se obtiene al disminuir la apertura de las microlentes individuales. Si bien esto introduce ruido debido a la disminución de rayos que se promedian para contribuir a la imagen enfocada, se observa que la región donde la sustracción es cercana a cero, es la región en foco. En la Figura 2 se muestra un ejemplo.

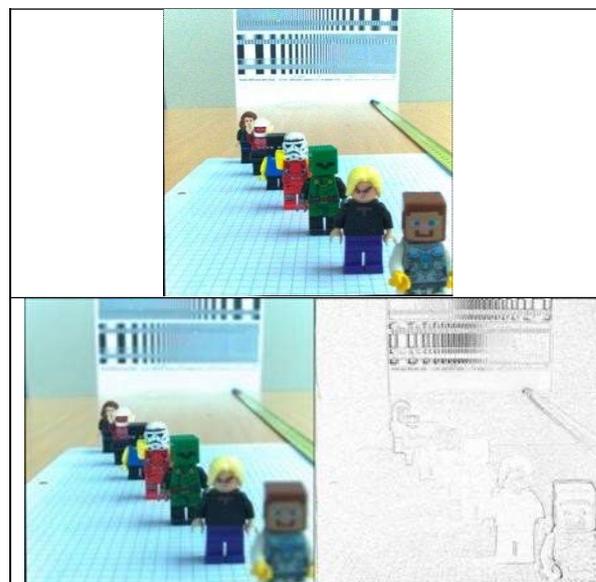


Figura 2. Sup.: Imagen en foco extendido por disminución de 4 píxeles de lado para cada lenticula de 11x11 (se promedian los 3x3 píxeles centrales). Inf.: para reenfoque con Desplazamiento-Suma para $\alpha=1.4$ y complemento de la imagen diferencia entre este reenfoque y la imagen de foco extendido. La región de reenfoque está entre el muñeco verde y el rojo.

Hay resultados interesantes de capturas con arreglos de lentes que permiten buen enfoque de pequeñas estructuras. Esto es alentador en aplicaciones tales como oftalmología y dermatología. En la Figura 3 se observan algunos ejemplos de tomas con LytroF01. Se está trabajando en un análisis objetivo de los resultados, utilizando la sustracción anteriormente explicada o la aplicación de filtros resaltabordes, pues las regiones en foco presentan buena definición mientras que las desenfocadas se ven borrosas. Otra herramienta de análisis es el mapa de

profundidad, para comparar la resolución en profundidad del reenfoque de los algoritmos.

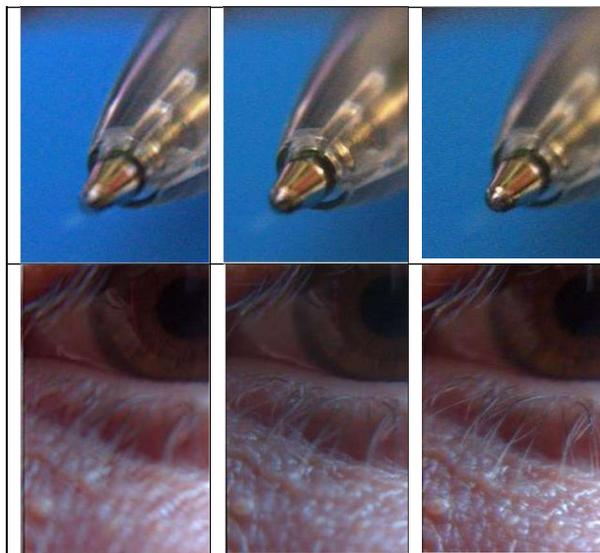


Figura 3. Imágenes de LytroF01 con óptica agregada para foco macro. Columna izquierda: foco lejos; centro: foco intermedio; derecha: foco cerca. Reenfocadas con Planar Frecuencia. Ambas series fueron generadas con el algoritmo *RefocusEspectro* con rangos valores de Theta de -0.02, 0, 0.02.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está conformado por especialistas del área de sistemas de información, inteligencia artificial y bioingeniería. Relacionadas con el Proyecto de Investigación se han generado tres Tesinas de Grado (una aprobada y las otras en curso) y una de Maestría (en etapa de elaboración de tesis).

5. BIBLIOGRAFÍA

- [1] Navarro Fructuoso H 2014 Captura y reproducción de imágenes 3D mediante sistemas de Imagen Integral *Tesis Doctoral Universidad de Valencia*.
- [2] F. Ives, "Parallax stereogram and the process of making same" U.S. Patent No. 725,567.1903.
- [3] G. Lippmann, "Epreuves réversibles donnant la sensation du relief," *J. Phys. Théor. Appl.* 7, 821–825 (1908).
- [4] Adelson E and Bergen J 1991 The plenoptic function and the elements of early

vision Computational models of visual processing 91 n.1 pp 3–20

[5] Adelson E. H. y Wang J. Y. A., "Single lens stereo with a Plenoptic Camera", *IEEE transactions on pattern analysis and machine intelligence*, vol 14, N°2, Feb. 1992.

[6] Levoy M and Hanrahan P 1996 Light field rendering *SIGGRAPH ACM* pp 31–42

[7] Gershun A 1939 The light field. Moscow, 1936. *Journal of Mathematics and Physics* XVIII, 51–151 Translated by P. Moon and G. Timoshenko

[8] Georgiev, T.: New results on the Plenoptic 2.0 camera. *Signals, Systems and Computers, Conference Record of the Forty-Third Asilomar Conference*. Pp 1243–1247 (2009)

[9] <http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4>

[10] Donald G. Dansereau, Oscar Pizarro, Stefan B. Williams; Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 1027-1034

[11] http://www-personal.acfr.usyd.edu.au/ddan1654/LFTtoolbox0.3_Samples1.zip

[12] <http://graphics.stanford.edu/projects/lfmicroscope/>

[13] Gustavo Bizai, Facundo Peiretti, Adrián Salvatelli, Alejandro Hadad, Diego Evin, Bartolomé Drozdowicz. "Algoritmos para el tratamiento de imágenes de campos de luz". Artículo número 8501. XVIII Workshop de Investigadores en Ciencias de la Computación WICC 2016, 14 y 15 de abril de 2016, Facultad de Ciencias de la Administración, UNER. Concordia, Entre Ríos

[14] Ng R 2006 Digital light field photography Ph.D. dissertation, Stanford University

[15] Andrew Lumsdaine ; Todor G. Georgiev ; Georgi Chunev. Spatial analysis of discrete plenoptic sampling. *Proc. SPIE 8299, Digital Photography VIII*, 829909 (January 24, 2012)

[16] Donald G. Dansereau, Oscar Pizarro, Stefan B. Williams. Linear Volumetric Focus for Light Field Cameras. *ACM Transactions on Graphics (TOG)*. Volume 34 Issue 2 Article No. 15, February 2015

Evaluación del Error en la Detección de Puntos de Borde en Imágenes SAR Polarimétricas

Daniel Monferrán¹, Juliana Gambini² y Alejandro C. Frery³

¹Instituto de Industria, Universidad Nacional de General Sarmiento, Pcia. de Buenos Aires, Argentina. - Dpto. de Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional San Martín Pcia. de Buenos Aires, Argentina dmonferr@ungs.edu.ar

²Dpto. de Ingeniería Informática, Instituto Tecnológico de Buenos Aires, Buenos Aires, Argentina - Dpto. de Ingeniería en Computación, Universidad Nacional de Tres de Febrero Pcia. de Buenos Aires, Argentina. mgambini@itba.edu.ar

³LaCCAN, Universidad Federal de Alagoas, Maceió, AL, Brazil, acfrery@gmail.com

RESUMEN

El Radar de Apertura Sintética polarimétrico (PolSAR - Polarimetric Synthetic Aperture Radar) es ampliamente utilizado en teledetección porque permite capturar imágenes terrestres de alta resolución. La interpretación automática de imágenes PolSAR es una tarea muy difícil porque éstas contienen un gran volumen de información y además se encuentran contaminadas con ruido speckle. Las características de este ruido hacen necesario utilizar métodos estadísticos para el procesamiento digital de este tipo de imágenes. En esta línea de investigación se pretende evaluar el error que se comete al calcular las posiciones de los puntos de borde dentro de la imagen, utilizando la distribución Wishart compleja y experimentos de Montecarlo en imágenes PolSAR simuladas.

Palabras claves: Imágenes PolSAR, Segmentación automática de Imágenes, Distribución Wishart.

CONTEXTO

Esta línea de investigación se inserta en el proyecto titulado Interpretación automática de Imágenes SAR Polarimétricas por medio de Modelos Estadísticos y Contornos Activos, correspondiente al Doctorado en Ciencia y Tecnología de la Universidad Nacional de General Sarmiento. Las otras dos instituciones

involucradas son el Depto. de Ingeniería Informática del Instituto Tecnológico de Buenos Aires y la Universidad Federal de Alagoas, Brasil.

1. INTRODUCCIÓN

El Radar de Apertura Sintética polarimétrico ha demostrado ser una herramienta eficaz para la teledetección geofísica. Se trata de una tecnología que aporta información complementaria con la de otras modalidades de teledetección. Los sistemas PolSAR transmiten y reciben trenes de pulsos electromagnéticos con diferentes combinaciones de polarización. El PolSAR no requiere de una fuente de iluminación externa, hace uso de su propia iluminación de tipo coherente y por lo tanto pueden tomarse imágenes de día o noche, pero tienen el problema de que están afectadas por ruido speckle característico de este tipo de iluminación. Este ruido, que no es aditivo ni gaussiano, confiere un aspecto granular a las imágenes y hace necesario el modelado con distribuciones estadísticas apropiadas para su procesamiento y análisis.

Una de las etapas importantes en el análisis de imágenes es encontrar los bordes que delimitan las regiones correspondientes a áreas de la imagen que pueden ser de pastura, de forestación o urbanas.

Los trabajos de investigación destinados a determinar las posiciones de los puntos de borde

en imágenes PolSAR, utilizan diferentes métodos, en (J. Gambini, 2008) se emplean cinco métodos para la detección de bordes, entre ellos, máxima verosimilitud con la distribución GA0, y contornos activos, para la delineación de los bordes usan curvas B-spline. En el artículo (R. Fjortoft, 1998) se propone un detector de borde para las imágenes SAR que produce un mapa de resistencia de borde sobre píxeles de la imagen, en (J.-M. Beaulieu, 2004) se desarrolla un proceso jerárquico de optimización paso a paso para la segmentación empleando el método de máxima verosimilitud, utilizan el modelo Gaussiano multivariado, la distribución Wishart y la distribución K. En (C. D'Elia, 2014) se propone la combinación de un algoritmo de segmentación basado en campos aleatorios de Markov estructurados en árbol y la clasificación orientada a objetos.

Los algoritmos diseñados para determinar las posiciones de los puntos de borde en este tipo de imágenes deben mantener un equilibrio entre la precisión de los resultados obtenidos y el costo computacional.

Para identificar puntos de borde en este tipo de imágenes, es necesario modelar los datos con una distribución estadística apropiada y luego estimar los parámetros correspondientes. En el artículo (A. D. C. Nascimento M. M., 2014), los autores utilizan muestras con un número reducido de datos, con el propósito de reducir el costo computacional. Esto posee la desventaja de que muestras pequeñas pueden no entregar suficiente información para identificar la presencia de borde. Por esta razón, resulta relevante evaluar la magnitud del error teniendo en cuenta el tamaño de las mismas.

El objetivo de esta línea de investigación es estudiar métodos para determinar la magnitud del error que se introduce cuando se calculan puntos de borde sobre una imagen de PolSAR, en aquellos casos en que se utiliza la distribución Wishart compleja (A. D. C. Nascimento A. C., 2014), (H. Skriver, 2001) y la función de verosimilitud para la estimación de las posiciones de los puntos de borde.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para la estimación de la posición del punto de borde se utilizan imágenes SAR polarimétricas

simuladas, formadas por dos regiones que representan una zona de pastizal y una zona urbana, generadas con la distribución Wishart compleja con diferentes parámetros para cada región. Luego, se toma una muestra con datos en un entorno de la posible ubicación del punto de borde y se maximiza la función de verosimilitud variando las particiones de la muestra, la posición donde se da el máximo es el punto de transición y por lo tanto la posición del punto de borde.

La distribución Wishart compleja está indexada por dos parámetros: la matriz de covarianzas Σ , y el número equivalente de looks L con que la imagen fue formada, el cual es el mismo para toda la imagen.

La función de densidad de la distribución Wishart compleja está dada por:

$$f(Z, \Sigma, L) = \frac{L^{mL} |Z|^{L-m}}{|\Sigma|^L \Gamma_m(L)} \exp(-L \operatorname{tr}(\Sigma^{-1}Z))$$

donde, Σ es la matriz de covarianzas, L es el número equivalente de looks, m es el número de canales de polarización, $\Gamma_m(L)$ es la función Gamma polarimétrica, $|\cdot|$ es el operador determinante y $\operatorname{tr}(\cdot)$ es el operador traza.

Dada una partición A, B de los datos sobre un segmento, la función logaritmo de la verosimilitud está dada por.

$$l(j) = \sum_{k=1}^j \ln f(Z_k, \Sigma_A, L) + \sum_{k=j+1}^N \ln f(Z_k, \Sigma_B, L)$$

donde j es la posición del pixel dentro del segmento, Σ_A y Σ_B son valores estimados de las matrices de covarianza de los conjuntos A y B que satisfacen la siguiente condición:

$$\begin{aligned} Z_k &\in A \text{ para } k = [1, \dots, j], \\ Z_k &\in B \text{ para } k = [j + 1, \dots, N] \end{aligned}$$

Finalmente, el estimador de la posición del punto de borde:

$$\hat{j}_{ML} = \operatorname{argmáx}_j(l(j))$$

En este trabajo utilizamos un experimento de Monte Carlo, para evaluar la precisión en el cálculo de los puntos de borde. Con un número total de 400 ensayos para cada muestra.

El experimento consiste en:

1. Generar la imagen simulada de 600 x 600 píxeles conformada por dos regiones contiguas de 300 x 600 píxeles caracterizadas por las siguientes matrices de covarianza:

$$\Sigma_p = \begin{bmatrix} 360932 & 11050 + j3759 & 63896 + j1581 \\ - & 98960 & 6593 + j6868 \\ - & - & 208843 \end{bmatrix}$$

$$\Sigma_u = \begin{bmatrix} 962892 & 19171 - j3759 & -154638 + j191388 \\ - & 56707 & -5798 + j16812 \\ - & - & 472251 \end{bmatrix}$$

Número de looks $L = 6$.

Las matrices de covarianza utilizadas Σ_p y Σ_u se obtuvieron a partir de muestras de zonas tomadas de una imagen capturada por el sensor de E-SAR sobre Weßling, Bayern, Alemania y corresponden a una zona de pastizal y una zona urbana, respectivamente.

Los elementos omitidos en las dos matrices se pueden obtener como complejos conjugados de sus respectivos simétricos. En la Figura 1 puede observarse una imagen PolSAR simulada de 203 x 60 píxeles y un segmento de $N = 101$ píxeles (representado en rojo) de donde se toma la muestra.



Figura 1 (80,21)

2-Se trazan segmentos de ancho un píxel y de un mismo largo N sobre la imagen, los datos de los píxeles bajo cada segmento constituyen cada una de las muestras. La Figura 2 permite apreciar los segmentos en verde y las marcas rojas de las posiciones estimadas de puntos de borde en el proceso de detección.

3-La cantidad de píxeles que aporta cada región se denota como (a,b) con, N : número total de píxeles del segmento, a : número de píxeles pertenecientes a la región A, b : número de píxeles pertenecientes a la región B.

Por ejemplo, la Figura 1 representa una muestra de 101 píxeles, el par $(80,21)$ indica que 80 píxeles pertenecen a la región A, 21 píxeles pertenecen a la región B y para esta relación no se encontró el verdadero punto de borde en ningún experimento.

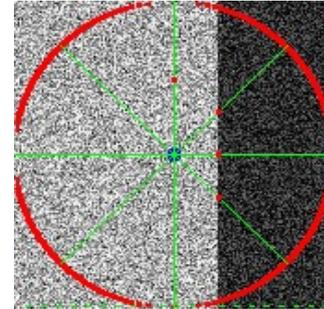


Figura 2.

4- Se determina el punto de borde P en cada segmento maximizando la función de verosimilitud en función de la posición del punto de cambio j , variando j entre $0.1N$ y $0.9N$ como se muestra en la Figura 3.

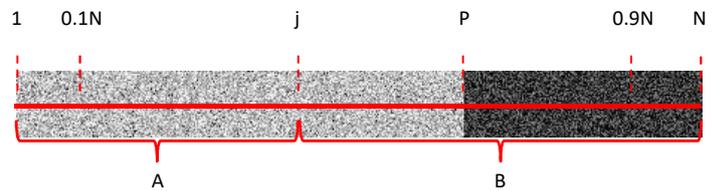


Figura 3. $0.1N \leq j \leq 0.9N$

3. RESULTADOS OBTENIDOS

El error porcentual calculado para cada par (a,b) se obtiene como:

$$\frac{\text{nro. de ensayos con error}_{(a,b)}}{\text{nro. total de ensayos}_{(a,b)}} * 100$$

Donde $\text{nro. total de ensayos}_{(a,b)} = 400$

Las Figuras 4 y 5 muestran los errores porcentuales que resultan de los ensayos para distintas relaciones (a,b) , para muestras de 101 píxeles y 201 píxeles, respectivamente.

Para los valores resaltados en círculo, en las Figuras 4 y 5, se tiene que a/b si $a < b$ ó b/a si $b < a$, en todos los casos el resultado nos da aproximadamente 0.3. Se interpreta que la variación del error no depende del tamaño de estas muestras y si de la relación en el aporte de píxeles de cada región.

Si $\frac{a}{b} = 0,3$ y $b = N - a$, implica que
 $a = \frac{3}{10}(N - a)$, operando queda $\frac{13}{10}a = \frac{3}{10}N$,

finalmente $a = \frac{3}{13}N \cong 0.23N$, es decir que el aporte de cada región a la muestra tiene que ser al menos un 23% del total de datos de la muestra.

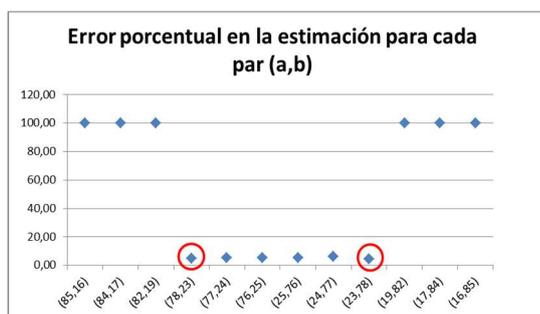


Figura 4 Muestras de 101 píxeles

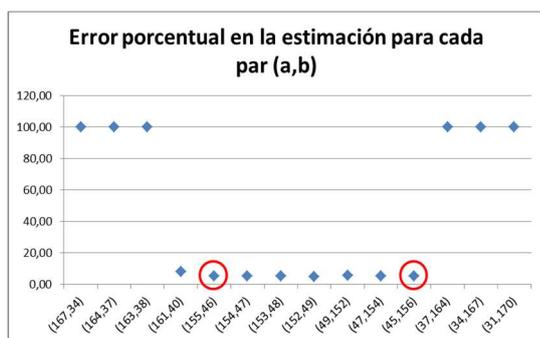


Figura 5 Muestras de 201 píxeles

En un estudio posterior se puede evaluar cómo influye la relación entre matrices de covarianzas de las dos regiones sobre la distribución del error mostrada en la Figura 4. También se puede analizar cómo se distribuye el valor del error para cada relación (a,b).

4. FORMACION DE RECURSOS HUMANOS

Este trabajo forma parte de la tesis de doctorado del Ing. Daniel Monferran.

5. BIBLIOGRAFIA

- A. D. C. Nascimento, A. C. (Agosto de 2014). Bias correction and modified profile likelihood under the Wishart complex distribution. IEEE Transactions on Geoscience and Remote Sensing, 52(8), 4932-4941.
- A. D. C. Nascimento, M. M. (Febrero de 2014). Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(2), 648-663.
- C. D'Elia, S. R. (Abril de 2014). SAR Image Classification Through Information-Theoretic Textural Features, MRF Segmentation, and Object-Oriented Learning Vector Quantization. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(4), 1116 - 1126.
- H. Skriver, J. S. (2001). Polarimetric edge detector based on the complex Wishart distribution. IEEE 2001 International Geoscience and Remote Sensing Symposium, 7.
- J. Gambini, M. M.-B. (2008). Accuracy of edge detection methods with local information in speckled imagery. Statistics and Computing, 8(1), 15-26.
- J.-M. Beaulieu, R. T. (Octubre de 2004). Segmentation of textured polarimetric SAR scenes by likelihood approximation. IEEE Transactions on Geoscience and Remote Sensing, 42(10), 2063-2072.
- R. Fjortoft, A. L. (Mayo de 1998). An Optimal Multiedge Detector for SAR Image Segmentation. IEEE Transactions on Geoscience and Remote Sensing, 36(3), 793-802.
- A. D. C. Nascimento, A. C. (Agosto de 2014). Bias correction and modified profile likelihood under the Wishart complex distribution. IEEE Transactions on

Framework para la Generación de Templates en Sistemas de Catálogos de Realidad Aumentada

N. Mangiarua, C. Montalvo, F. Petrolo, D. Sanz, N. Verdicchio, E. Lobatto,
A. Rosenthal, M. Becerra, S. Igarza, J. Ierache

Universidad Nacional de La Matanza, Departamento de Ingeniería e
Investigaciones Tecnológicas.

Grupo de Realidad Aumentada Aplicada
Florencio Varela 1903, La Matanza, Buenos Aires, Argentina
jierache@unlam.edu.ar

Resumen

En este trabajo presentamos el desarrollo de un sistema de aumentación de metacontenidos sobre las bases del sistema de catálogo virtuales aumentados [1],[2] a fin de mejorar la usabilidad del sistema original para usuarios no expertos. Se propone pues la incorporación del concepto de template de aumentación de la realidad, para los catálogos de realidad aumentada. Los mismos permitirán predefinir la cantidad y tipos de contenidos, junto con sus transformaciones geométricas y el orden en el que aparecen en un determinado template. Así, al aplicarlo a un catálogo aumentado ayudará a mantener un formato uniforme entre todos sus marcadores, así como también simplificará la carga del material a aumentar por parte de usuarios no expertos en el área.

Palabras clave: Realidad Aumentada, Catálogo Virtual Aumentado, Template, herramienta de generación de contenidos aumentados.

Contexto

La investigación es desarrollada por el grupo de investigación de Realidad Aumentada Aplicada del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, en el marco del proyecto PROINCE C-202. El grupo se encuentra financiado a través de las becas otorgadas por la institución.

Introducción

Se plantea como hipótesis de trabajo la simplificación de la labor del usuario al momento de subir y publicar su propio contenido aumentado pudiendo generar plantillas de creación de catálogos aumentados. Estos catálogos están compuestos por el conjunto de imágenes seleccionadas por el usuario, a las que le incorporará contenidos de realidad aumentada. Los usuarios podrán compartir sus catálogos con un grupo de usuarios, sin la necesidad de tener que recurrir a herramientas específicas de aplicación en RA.

Este equipo de investigación ha desarrollado diversas aplicaciones, como juegos de tablero [3],[4] y otras

herramientas para la generación de materiales didácticos [5],[6] para el área educativa, o sistemas de tarjeta aumentada basados en conocimiento para la asistencia médica en emergencias [7],[8].

La línea de investigación que desarrolla el grupo de realidad aumentada pretende hacer accesible la tecnología de Realidad Aumentada (a partir de aquí denominada RA) a personas sin conocimientos técnicos y contribuir a enriquecer sus proyectos con contenido aumentado.

El proyecto de investigación aplicada se encuadra en el contexto de la aplicación de tecnologías de Realidad Aumentada en la vida cotidiana de las personas, generando así un impacto en la sociedad que contribuye a la participación de los mismos en el ámbito tecnológico.

La RA permite la fusión de datos virtuales y objetos sobre el mundo físico, enriqueciendo la percepción de la realidad, aumentándola [9]. Se suele comparar esta tecnología con Realidad Virtual, pero no representan el mismo concepto. Realidad Virtual implica la inmersión en un ambiente completamente virtual mientras que RA se mantiene en el mundo real mientras se despliegan y muestran contenidos virtuales superpuestos sobre él [10].

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación aplicada se orienta al desarrollo de un módulo adicional para el Sistema de Catálogos Virtuales Aumentados [1], [2] extendiendo así el alcance de las

funcionalidades del mismo. Estudios de usabilidad sobre el sistema original han dejado en evidencia la necesidad de facilitar el flujo de trabajo para usuarios no expertos. Se propone entonces la incorporación del concepto de una plantilla (template), para los catálogos de realidad aumentada. Los mismos permitirán predefinir la cantidad y tipos de contenidos, junto con sus transformaciones geométricas (posición, rotación, escala) y su orden (de aparición en el editor) De esta manera, al ser aplicado a un catálogo aumentado en el momento de su creación, ayudará a mantener un formato uniforme entre todos sus marcadores, así como también simplificará la carga del material a aumentar por parte de usuarios no expertos del sistema, facilitando la personalización del catálogo y reduciendo los tiempos para su publicación.

Tanto el uso del catálogo, como la nueva funcionalidad de creación de templates demuestran la necesidad de incorporar funcionalidades para permitir al usuario compartir su catálogo o template con otros. Por este motivo, se plantea agregar funciones para la gestión de grupos de usuarios, a los cuales cada usuario podrá solicitar acceso, y así compartir catálogos y templates con los restantes miembros.

Resultados y Objetivos

La primera fase del proyecto contempla como objetivo el desarrollo del módulo para la creación de templates para la plataforma web del Sistema de Catálogos Virtuales Aumentados.

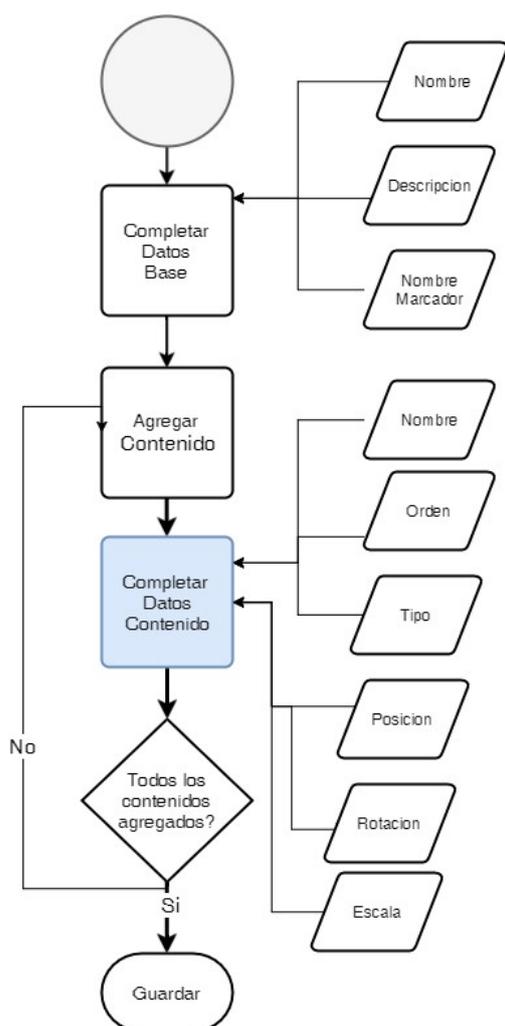


Fig. 1 Diagrama de Actividades

Este módulo le permitirá al usuario crear un nuevo template definiendo un nombre y una descripción al mismo, agregando distintos tipos de contenidos con su orden y transformación geométrica.

Adicionalmente se investigarán y presentarán resultados de acuerdo a los modelos de datos iniciales, agregando un nombre opcional para el marcador de RA (la imagen que es reconocida y utilizada para ubicar el contenido aumentado)

como muestra la figura [1].

Dicho nombre ayuda a mantener el contexto de aplicación del catálogo aumentado. Por ejemplo, si se creara un template para un catálogo aumentado que define un juego de mesa, los marcadores podrían llamarse “casilleros”. De igual manera se agrega un nombre opcional a cada contenido.

En figura [2] se presenta un prototipo de interfaz para satisfacer esta función de creación de template.

Fig. 2 Prototipo de Pestaña para la generación de templates.

Además, en este módulo se contará con funciones adicionales como las de eliminar un template y la gestión de grupos de usuarios que podrán tener visibilidad de las plantillas ya creadas.

En la segunda fase del proyecto se desarrolla la funcionalidad de la plataforma móvil del Sistema de Catálogos Virtuales Aumentados para explotar los diversos contenidos agregados a los correspondientes templates. Si bien parte de la información

predefinida en un template es transparente a la aplicación móvil, algunas mejoras arquitecturales menores serán necesarias para asegurar la escalabilidad y flexibilidad incorporadas.

La tercera fase del proyecto se destina a las pruebas en distintos dominios, definiendo y llevando a cabo casos de prueba, buscando campos de aplicación específicos sobre los cuales explotar un catálogo de realidad aumentada, enfocado principalmente en el contexto social, como ser las áreas del arte, gobierno digital .

Luego, a partir de los resultados obtenidos, evaluar y corregir errores que hayan surgido, realizando los ajustes necesarios para refinar el desempeño del módulo generador de templates desarrollado para el Sistema de Catálogos Virtuales Aumentados en una nueva iteración de desarrollo.

Dado el fuerte componente de usabilidad del subsistema de templates, se prevé la realización de numerosos e importantes cambios durante esta etapa, que puedan complementar y satisfacer las necesidades identificadas durante las pruebas del sistema de catálogos original.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por un investigador formado, dos investigadores en formación, dos graduados, y cinco alumnos del departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM, trabajando en el área de Realidad Aumentada. La formación de

los integrantes tanto en el área técnica como en el mecanismo formal de investigación y publicación conforma una importante porción de nuestro trabajo.

Adicionalmente el grupo participa en los distintos eventos científicos tecnológicos abiertos a la comunidad que organiza la Universidad tales como Expo Proyecto y la Bienal de Arte, como así también en congresos internacionales y nacionales, como CACIDI[7], IWBBIO [8], HCIST, MCCSIS, CACIC, TE&ET [4],[5].

Referencias

1. Ierache J, Mangiarua,N., Bevacqua, S.,Verdicchio, N., Becerra,M., Sanz, D.,Sena, E., Duarte,N., Igarza, S., (2014), “Sistema de Catálogo para la Asistencia a la Creación, Publicación, Gestión y Explotación de Contenidos Multimedia y Aplicaciones de Realidad Aumentada”. XX Congreso Argentino de Ciencias de la Computación Red de Universidades con Carreras de Informática
2. Ierache, J., Mangiarua, N. Bevacqua, S., Verdicchio, N., Becerra, M, Sanz, D. , Sena, M. , Ortiz, F. , Duarte, N. , Igarza, S. (2015). “Development of a Catalogs System for Augmented Reality Applications”. World Academy of Science, Engineering and Technology, International Science Index 97, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(1), 1 - 7.

3. Ierache, J., Mangiarua, N., Verdicchio, N., Sanz, D., Montalvo, D., Petrolo, Igarza, S. (2015). "Sistema de Catálogo Virtual Aumentado. Integración de Framework Especializado orientado a juegos didácticos" XI Congreso de Tecnología en Educación y Educación en Tecnología
4. Verdicchio, N Sanz, D Igarza S, Mangiarua N, Montalvo, C Petrolo, F Barth, J Ierache, J "Sistema de Catálogo Virtual Aumentado Integración de Framework Especializado Orientado a Juegos Didácticos". TE&ET 2016. XI Congreso de Tecnología en Educación en Tecnología, 9 y 10 de Junio Morón FICCTE-UM, pp 597-604, ISBN 978-987-3977-30-5.
<http://sedici.unlp.edu.ar/handle/10915/53517>
5. Mangiarua N., Ierache J., Bevacqua S., Becerra M., Verdicchio N., De la Llave E., Duarte N., Sanz D., Ortiz F., Igarza S., "Herramienta de Realidad Aumentada para la explotación de material didáctico tradicional". TE&ET 2014. IX Congreso de Tecnología en Educación en Tecnología. 1a ed. - Chilecito: UNdeC, 2014. E-Book. ISBN 978-987-24611-1-978-987-24611-1-9. CDD 370.1 .Fecha de catalogación: 12/06/2014, 250-254
6. Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Ortiz F, Igarza S. 2014. "Herramienta de Realidad Aumentada para facilitar la enseñanza en contextos educativos mediante el uso de las TICs". Revista Latinoamericana de Ingeniería de Software, 1(1):1-3, ISSN 2314-2642
7. Ierache, J., N. Mangiarua, N. Verdicchio, D. Sanz, C. Montalvo, F. Petrolo and S. Igarza, "Augmented. Card System Based on Knowledge for Medical Emergency Assistance". I IEEE CACIDI Congreso Argentino de Ciencias de la Informática y Desarrollos de la Investigación Dic 2016 IEEE Xplore Digital Library (en prensa) ISBN 978-1-5090-2938-9 2016
8. Ierache, J., Verdicchio, N., Duarte, N., Montalvo, C., Petrolo, F., Sanz, D., Barth, J Mangiarua, N., Igarza, S., "Augmented Reality Card System for Emergency Medical Services", IWBBIO 2016 (International Work-Conference on Bioinformatics and Biomedical Engineering) Proceedings Extended abstracts 20 - 22 abril 2016 Granada (SPAIN), pp 487-494, ISBN 978-84-16478-75-0
9. Manresa Yee, M. Abásolo, R Más Sansó and M Vénere (2011). "Realidad virtual y realidad aumentada. Interfaces avanzadas." ISBN 978-950-34-0765-3
10. Cristina Manresa Yee, María José Abásolo, Ramón Más Sansó and Marcelo Vénere (2011). "Realidad virtual y realidad aumentada. Interfaces avanzadas." ISBN 978-950-34-0765-3

Identificación No Supervisada de Fauna Marina Mediante Procesamiento de Imágenes Fotográficas

Débora Pollicelli^{1,2}, Mariano Coscarella^{1,3}, Claudio Delrieux⁴

¹*CESIMAR-CONICET, Centro para el Estudio de Sistemas Marinos, Consejo Nacional de Investigaciones Científicas y Técnicas, CCT CENPAT, Bv. Almirante Brown 2915, 9120, Puerto Madryn, Chubut, Argentina*

²*LINVI, Departamento de Informática, Facultad de Ingeniería, UNPSJB, Bv. Almirante Brown 3051, 9120, Puerto Madryn, Chubut, Argentina*

³*Departamento de Biología General, Facultad de Ciencias Naturales, UNPSJB, 9120, Puerto Madryn, Chubut, Argentina*

{deborapollicelli, mcoscarella}@gmail.com

⁴*Laboratorio de Ciencias de las Imágenes, Departamento de Ingeniería Eléctrica y Computadoras, Universidad Nacional del Sur y CONICET, 8000 Bahía Blanca, Argentina – cad@uns.edu.ar*

RESUMEN

Las técnicas de procesamiento inteligente de imágenes han experimentado en la última década un desarrollo muy importante, debido a la conjunción de varios factores tecnológicos.

Por un lado, se han propuesto técnicas y metodologías de estimación robusta de los parámetros geométricos relevantes en fotografías oblicuas (y por lo tanto una reconstrucción 3D plausible de las mismas), por otro lado, se han desarrollado nuevos modelos de reconocimiento de patrones e identificación no supervisada, aplicados con éxito en una gran variedad de contextos y áreas de intervención.

Estos avances permiten construir la hipótesis fundamental de esta línea de trabajo, según la cual es posible investigar y desarrollar técnicas y metodologías de procesamiento inteligente de imágenes para la identificación de individuos de fauna marina.

En el caso particular de los cetáceos, en los que las muescas en aletas dorsales son utilizadas para realizar la identificación individual, es necesario el procesamiento de imágenes fotográficas oblicuas. Una solución de este tipo poseería un sinnúmero de ventajas. Por un lado, la intervención no es directa y por lo tanto no es traumática para los individuos bajo estudio. Por otro lado es de aplicación masiva, es decir puede utilizarse para resolver problemas similares en distintas especies.

Palabras clave: Análisis de imágenes, foto-

identificación, cetáceos, delfín *Cephalorhynchus commersonii*, aprendizaje automatizado, reconocimiento de patrones.

CONTEXTO

El presente trabajo se enmarca dentro de las tareas de investigación del Laboratorio de Mamíferos Marinos del CENPAT-CONICET, cuyas actividades están centradas en el estudio de la dinámica poblacional, ecología de comunidades, dieta, reproducción comportamiento y genética de especies predadores tope del Mar Argentino, con énfasis en mamíferos marinos.

Estas actividades se llevan adelante desde hace 30 años, contando con una extensa base de datos fotográfica de tonina overa (*Cephalorhynchus commersonii*).

Dicho acervo fotográfico requiere una intensiva tarea manual para ser utilizada, por lo que surge como tarea innovadora la investigación y desarrollo de algoritmos de análisis de imágenes que permitan la identificación no supervisada de individuos de la fauna marina con base en dichas fotografías.

Dentro de dicha línea, se están llevando adelante varios trabajos de investigación, en particular el desarrollo de tesis doctoral *Algoritmos para la identificación no supervisada de fauna marina mediante imágenes fotográficas* en conjunto con el Laboratorio de Ciencias de las Imágenes (LCI)

del Departamento de Ingeniería Eléctrica y Computadoras, Universidad Nacional del Sur y CONICET.

INTRODUCCIÓN

El manejo de las poblaciones silvestres requiere estimaciones precisas y exactas de los parámetros que determinan la dinámica poblacional. Esto es particularmente crítico para las especies marinas superiores, en las que los efectos antrópicos, la fragmentación de hábitat, el cambio climático, y otros factores pueden ser inferidos de los cambios observados en éstos parámetros.

Una de las técnicas que provee una gran cantidad de información consiste en el marcado de individuos (ej. colocando precintos u otros dispositivos), con la expectativa de poder ubicar e identificar el mismo individuo en un tiempo posterior.

Uno de los principales supuestos de este análisis es que el comportamiento de individuos marcados no se ve modificado por la marca. Este supuesto es crítico y puede no mantenerse en caso de que la marca aplicada sea muy conspicua, proporcione una ventaja adicional a la supervivencia del animal, o provea algún tipo de beneficio reproductivo.

Una alternativa a la marcación artificial son las marcas naturales que presentan los animales.

Estas marcas permiten reconocer a los individuos mediante señas particulares que varían con la especie estudiada. El desarrollo de trabajos de investigación a largo plazo (especialmente en vertebrados terrestres, exceptuando las aves) mostró que muchos animales salvajes pueden ser individualizados mediante características o marcas naturales. La obtención de fotografías de cada animal, y su posterior análisis en gabinete proporciona información fehaciente sobre cuáles fueron los animales presentes en una sesión de muestreo particular.

La foto-identificación ha permitido a investigadores de todo el mundo reconocer individuos mediante marcas naturales desde la década de 1970. El reconocimiento individual

es una herramienta que permite obtener una gran cantidad y variedad de información biológica en la que se incluyen estimaciones de parámetros poblacionales tales como la mortalidad, la supervivencia y el ciclo reproductivo [1, 2]. Cuando se realizan sesiones fotográficas de más de una localidad pueden detectarse patrones de movimiento o migraciones [2].

La información proveniente de las capturas (fotográficas) de los individuos se analiza mediante modelos de captura-recaptura para obtener estas estimaciones [3, 4].



Figura 1: Recaptura fotográfica de ejemplar *Cephalorhynchus commersonii* identificado en la base de datos del LAMAMA-CESIMAR-CONICET. Presenta marcas principales sobre el borde de la aleta dorsal y marcas auxiliares sobre el lomo. Las marcas principales son identificables desde ambos flancos del animal, mientras que las auxiliares son visibles solo por un lado.

Las marcas utilizadas en estos animales son principalmente muescas o variaciones en el patrón de coloración de las aletas dorsales y caudales [1]. Adicionalmente, las cicatrices y las anomalías en el patrón de coloración se utilizan como características auxiliares en la identificación.

La singularidad de algunas formas de daños sufridos por el borde posterior de la aleta a lo largo de la vida de un individuo permite que pueda ser utilizado para identificar al individuo. Este método ha sido aplicado exitosamente en

orcas (*Orcinus orca*), cachalotes (*Physeter catodon*), delfines del indo-pacífico (*Sousa chinensis*), delfines nariz de botella (*Tursiops truncatus*), delfines rotadores de Hawai (*Stenella longirostris*), y ballenas yubartas (*Megaptera novaeangliae*) [1, 5].

La foto-identificación tradicional implica un laborioso proceso manual de búsqueda en cada fotografía en que se observan las características utilizadas para la identificación, en un catálogo de los individuos conocidos [6].

Por dicha razón existe un interés creciente en automatizar este proceso utilizando procesamiento de imágenes. Los programas más utilizados en el reconocimiento individual en delfines son el FinsScan y DARWIN [7]. Aunque estos programas pueden realizar reconstrucciones en 3D con el objeto de compensar diferencias en la orientación del cuerpo del animal, ambos necesitan que el operador registre un punto fijo (landmark) en la aleta y en el caso del FinScan también es necesario registrar la posición de al menos alguna muesca en la aleta.

Sin embargo, en el caso de los delfines pertenecientes al género *Cephalorhynchus*, ninguna de las cuatro especies que habitan los mares del hemisferio sur posee una aleta dorsal que presente una punta prominente, ya que sus aletas son redondeadas en vez de falcadas [6, 8]. En general, ninguno de los sistemas actuales efectúa un procesamiento correcto para las diferentes orientaciones de aleta (pose 3D) antes de realizar la correspondencia de una manera robusta [9]. Según Stweman et al., 2006, estos métodos no son robustos, y para su uso se requiere información adicional para optimizar la registración.

Este último punto es crítico, dado que la técnica de matching es basada en pixels (distancia Euclídea entre imágenes), por lo que aún una leve diferencia de registración lleva a no reconocer adecuadamente animales presentes en la base de datos.

Del mismo modo, cualquier elemento que genere diferencias métricas en dicha evaluación (deterioros, por ejemplo) hacen prácticamente inutilizable el sistema de reconocimiento.

En otras palabras, al no existir una correlación entre la similitud de la forma de las aletas y la métrica utilizada, el reconocimiento es muy propenso a una alta tasa de falsos positivos y negativos.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La hipótesis de trabajo consiste en que es posible aplicar procesamiento inteligente de imágenes para la identificación de individuos de diferentes especies marinas a partir de imágenes fotográficas oblicuas.

Esto permitiría superar las limitaciones mencionadas en los sistemas de identificación actuales.

Los procedimientos a llevar a cabo para cumplir con los objetivos se basan en el uso conjunto de dos grupos de técnicas.

El primer grupo propone obtener una registración 3D basada en puntos característicos (features) que no sean necesariamente landmarks anatómicos, mientras que el segundo grupo busca un mecanismo de reconocimiento basado en mecanismos de análisis de formas con métricas más robustas que la distancia Euclídea. Respecto de la registración 3D, se plantea primero realizar una estimación de los parámetros extrínsecos de la cámara a través del uso de la geometría epipolar [10], buscando evaluar la matriz fundamental de la cámara a través de puntos característicos en la imagen, los cuales se encuentran a través de algoritmos bien establecidos como SURF [11] o SIFT [12]. Con una estimación de los parámetros extrínsecos, es posible determinar en forma robusta la pose 3D encontrando una transformación de mayor cantidad de grados de libertad que las utilizadas en los sistemas mencionados más arriba [13].

Con una registración 3D más robusta, es posible el análisis de formas a través de puntos característicos utilizando análisis por componentes principales [14], ICP (Iterative Closest Point) [15], RANSAC (RANdom SAmple Consensus) [16]. Estos métodos o una

combinación de ellos permiten elaborar modelos geométricos sobre los cuales es posible un análisis de la forma de los objetos para los que existan métricas de similitud adecuadas en este contexto, como el análisis de Procrustes u otros descriptores de forma [17].

También se prevé el uso de técnicas de reconocimiento de patrones como por ejemplo Deep Learning o ADA Boost, para alcanzar tasas de sensibilidad y especificidad adecuadas.

RESULTADOS Y OBJETIVOS

1.1 OBJETIVO GENERAL

La presente línea de investigación propone desarrollar nuevos métodos de segmentación e identificación en imágenes que puedan ser utilizados para la identificación no supervisada de individuos pertenecientes a especies de cetáceos que no pueden ser analizadas con los métodos tradicionales.

Luego se aplicarán estos métodos a repositorios fotográficos de diferentes especies, lo que servirá como insumo para ulteriores investigaciones relacionadas con estimaciones de abundancia y otros estudios ecológicos y ambientales, articulándose con otras investigaciones científicas estratégicas en el mar argentino que incluyen actividades de exploración y conservación.

Además, los datos obtenidos pueden resguardarse en formato digital en repositorios abiertos, lo cual permitirá generar a largo plazo una base de información fotográfica de gran valor científico.

1.2 OBJETIVOS ESPECÍFICOS

Los objetivos particulares son los siguientes:

1. Estudio exhaustivo de la bibliografía reciente en la temática de fotoidentificación, reconocimiento de patrones, reconstrucción 3D, análisis y descriptores de forma, detección de puntos característicos.
2. Análisis del estado del arte en sistemas y metodologías de propósitos similares y sistemas disponibles tanto comerciales como

libres.

3. Investigación y desarrollo de metodologías de reconocimiento e identificación no supervisada, y de estimación de parámetros geométricos a partir de imágenes oblicuas utilizando landmarking, morfometría y técnicas asociadas, así como su integración dentro de un workflow de procesamiento.
4. Estudio de la información relevante en modelos de aprendizaje no supervisado, deeplearning, minería de datos y metodologías de clasificación no supervisada.
5. Estudio e implementación de modelos de identificación y reconocimiento a partir de features fotográficos y landmarks.
6. Análisis de requerimientos en casos de uso y aplicación.
7. Desarrollo de un prototipo inicial que integre los resultados teórico-prácticos obtenidos y que considere casos de uso definidos de acuerdo al contexto de aplicación en particular, utilizando los repositorios fotográficos existentes.
8. Testeo y validación del prototipo. Análisis estadístico de los resultados. Realización de pruebas en campo y con repositorios fotográficos.

Actualmente estamos trabajando en el análisis de imágenes utilizando primero descriptores de Haar Cascade para la detección automática de ROI (principalmente la aleta) y luego los descriptores morfométricos para obtener un vector de características.

FORMACIÓN DE RECURSOS HUMANOS

Debido que esta línea de investigación está ligada al desarrollo de una tesis doctoral, se cuenta con la participación la doctoranda Lic. Pollicelli María Débora junto con el equipo de dirección conformado por dos investigadores especializados en las diferentes disciplinas que abarca el proyecto.

Por un lado, el Dr. Delrieux, especialista reconocido en la temática de procesamiento de imágenes, específicamente imágenes aéreas y satelitales.

Por otra parte, el Dr. Coscarella, especialista en mamíferos marinos, particularmente en el campo de la ecología del comportamiento y genética de poblaciones de cetáceos y en el reconocimiento individual de los mismos mediante métodos tradicionales desde hace más de 15 años.

Ambos complementan los aspectos de formación en el tema a desarrollar, ya que la implementación de los algoritmos al problema planteado requiere una aproximación metodológica y de desarrollo y su implementación no puede realizarse sin considerar los aspectos biológicos del sistema bajo estudio.

REFERENCIAS

- [1] P. S. Hammond, S. A. Mizroch, and G. P. Donovan, *Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters*, vol. 12 of *International Whaling Commission Special Issue Series*. Cambridge: International Whaling Commission, 1990.
- [2] S. Bräger, S. Dawson, E. Slooten, S. Smith, G. Stone, and A. Yoshinaga, “Site fidelity and along-shore range in Hector’s dolphin, an endangered marine dolphin from New Zealand,” *Biological Conservation*, vol. 108, no. 3, pp. 281–287, 2002.
- [3] G. C. White, *Capture-recapture and removal methods for sampling closed populations*. Los Alamos National Laboratory, 1982.
- [4] C. Q. Da Silva, J. Zeh, D. MaDIGaN, J. Laake, D. Rugh, L. BaRaFF, W. KOSKI, and G. Miller, “Capture-recapture estimation of bowhead whale population size using photoidentification data,” *Journal of Cetacean Research and Management*, vol. 2, no. 1, pp. 45–62, 2000.
- [5] B. Würsig and M. Würsig, “The photographic determination of group size, composition and stability of coastal porpoises (*Tursiops truncatus*),” *Science*, vol. 198, pp. 755–756, 1977.
- [6] M. A. Coscarella, *Ecología, comportamiento y evaluación del impacto de embarcaciones sobre manadas de tonina overa Cephalorhynchus commersonii en Bahía Engaño, Chubut*. Ph.d., 2005.
- [7] A. Gilman, T. Dong, K. Hupman, K. Stockin, and M. Pawley, “Dolphin fin pose correction using icp in application to photoidentification,” in *Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference of*, pp. 388–393, IEEE, 2013.
- [8] R. N. P. Goodall, A. R. Galeazzi, S. Leatherwood, K. W. Miller, I. Cameron, R. K. Kastelein, and A. P. Sobral, “Studies of Commerson’s dolphins, *Cephalorhynchus commersonii*, off Tierra del Fuego, 1976–1984, with a review of information on the species in the South Atlantic,” Report of the International Whaling Commission, vol. 9, pp. 3–70, 1988.
- [9] J. Stewman, K. Debure, S. Hale, and A. Russell, “Iterative 3-d pose correction and content-based image retrieval for dorsal fin recognition,” in *International Conference Image Analysis and Recognition*, pp. 648–660, Springer, 2006.
- [10] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second ed., 2004.
- [11] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (surf),” *Comput. Vis. Image Underst.*, vol. 110, pp. 346–359, June 2008.
- [12] M. Toews and W. Wells, “Sift-rank: Ordinal description for invariant feature correspondence,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 172–177, June 2009.
- [13] R. Szeliski, *Computer Vision: Algorithms and Applications*. New York, NY, USA: Springer-Verlag New York, Inc., 1st ed., 2010.
- [14] Y. Ke and R. Sukthankar, “Pca-sift: A more distinctive representation for local image descriptors,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR’04*, (Washington, DC, USA), pp. 506–513, IEEE Computer Society, 2004.
- [15] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, “Comparing icp variants on real-world data sets,” *Auton. Robots*, vol. 34, pp. 133–148, Apr. 2013.
- [16] S. Choi, T. Kim, and W. Yu, “Performance evaluation of ransac family,” in *Proc. BMVC*, 1997.
- [17] F. B. Neal and J. C. Russ, *Measuring shape*. CRC Press, 2012.

Interfaz Natural de Usuario para el Control de Robot Móvil con Gestos Faciales y Movimientos del Rostro Usando Cámara RGB

César Osimani¹, Martín Salamero², Carlos Bartó³, Marcos Lopez⁴

Centro de Investigación Aplicada y Desarrollo en Informática y Telecomunicaciones (CIADE-IT)
Universidad Blas Pascal (UBP) - Córdoba, Argentina

¹cosimani@ubp.edu.ar, ²msalamero@ubp.edu.ar, ³carlosbarto@gmail.com, ⁴marcoslopez@ubp.edu.ar

Resumen

Actualmente existen muchos esfuerzos por parte de grupos de investigación en acelerar el desarrollo de algoritmos y técnicas que faciliten la cotidianidad de las personas a través de una interacción más natural con los dispositivos. La disciplina que busca mejorar la experiencia del usuario en el manejo de computadoras se la conoce como Interacción Humano Computadora (HCI - Human Computer Interaction), y en algunos de sus campos se pueden encontrar las Interfaces Cerebro Computadora (BCI - Brain Computer Interface) que permiten una interacción mediante nuestro pensamiento, realizando la adquisición de las ondas cerebrales, que luego son procesadas para generar las acciones en el entorno. También se encuentran las Interfaces Naturales de Usuario (NUI - Natural User Interface) con las cuales se ofrece la posibilidad de control a través de gestos, posturas y movimientos con el cuerpo.

La utilización de recursos de visión artificial para reconocer los movimientos y gestos del rostro de las personas constituye un tema de interés para el desarrollo de las Interfaces Naturales de Usuario. El objetivo de este proyecto, además de contribuir con el avance de técnicas para el reconocimiento de gestos del rostro, es diseñar una interfaz de usuario compuesta por una pantalla de visualización y una cámara RGB estándar que permita controlar dispositivos a través de los movimientos de la cabeza y los gestos faciales, emulando gráficamente en la pantalla un conjunto de botones de control para facilitar al usuario reconocer con exactitud el comando que se está realizando. Este trabajo

también hace hincapié en la Interacción Humano Robot (HRI - Human RobotInteraction) para implementar las técnicas de visión artificial en el control del desplazamiento de un robot móvil.

Palabras clave: Visión Artificial; Interfaz Natural de Usuario; Reconocimiento de expresiones faciales; Control de robots

Contexto

La línea de investigación que se trata en este proyecto es el desarrollo de Interfaces Naturales de Usuario con el uso de técnicas de Visión Artificial. Se realiza en el marco de las actividades desarrolladas en el Centro de Investigación Aplicada y Desarrollo en Informática y Telecomunicaciones (CIADE-IT), dependiente de la Universidad Blas Pascal, específicamente en lo relativo al proyecto "Visión Artificial para la Interacción Natural Humano Computadora a través de los gestos del rostro" aprobado por el Ministerio de Ciencia y Tecnología del Gobierno de la Provincia de Córdoba en el llamado a presentación de proyectos 2015 para Grupos de Reciente Formación y Tutoría a equipos de investigación científica y tecnológica (GRFT) en la Resolución Nro.0045 del 12 de agosto de 2016.

Cabe destacar que se mantiene un proyecto en esta línea con investigadores de la Universidad de Almería (España) con quienes, a la fecha, se tiene la aceptación del artículo "Hand Posture Recognition with standard webcam for Natural Interaction" en el congreso WorldCist'17 (5th World Conference on Information Systems and Technologies).

Introducción

Para la creación de interfaces naturales contécnicas de visión artificial se requiere la elección de bibliotecas de programación para el procesamiento de las imágenes y el diseño de interfaces gráficas. Este grupo de trabajo tiene antecedentes en esta línea de investigación con el uso de las bibliotecas OpenCV (ver [1] y [2]) y Qt bajo lenguaje C++ (documentación en [3]). Se comienza adquiriendo las imágenes de video de una cámara RGB estándar. Cada una de las imágenes es analizada con uno de los métodos más utilizados, desarrollado por Viola & Jones [4], para la detección de rostros. Es un método estadístico que utiliza muestras de entrenamiento (imágenes con rostros e imágenes sin rostros) para extraer información que permite distinguir un rostro de aquello que lo no es. Este método también puede ser utilizado para la detección de ojos, bocas sonrientes o no, orejas, como también puede ser extendido para detectar otro tipo de objetos (aviones, autos, etc.). Para utilizar este método se realiza un entrenamiento con las imágenes de muestra, para así obtener un clasificador del objeto. Este tipo de clasificador entrega una región de interés (ROI - Region Of Interest) que es un rectángulo dentro de la imagen que contiene el objeto en cuestión.

Es importante indicar que un prototipo ya se encuentra desarrollado por este grupo de investigación, el cual permite controlar un menú de opciones a través de la orientación del rostro. El código fuente y un video demostrativo se encuentran en [5] y [6].

El control del menú se realiza con un clasificador de rostros (disponible en OpenCV) que identifica su posición dentro de la imagen y además reconoce su orientación a partir de un leve desplazamiento que sufre la región detectada cuando el rostro cambia su orientación. Esta característica se trata en [7].

OpenCV dispone de un clasificador para la sonrisa que genera un alto porcentaje de falsos positivos cuando se analiza una imagen que contiene diversos objetos. Sin embargo, da muy buenos resultados cuando la región de la imagen que se analiza es reducida y forma parte del rostro. Por este motivo, es altamente recomendable que en un primer paso se detecte el rostro, para luego dividir esta región de interés por la mitad (obteniendo una región

donde se encuentran los ojos y otra región con la boca). Por último, analizar la región inferior con el clasificador de la sonrisa para detectar si la persona está sonriendo o no. De esta forma, la sonrisa puede ser utilizada como acción de selección dentro de la interfaz natural de usuario. Para completar la interfaz natural de usuario con otra alternativa para realizar acción de selección, se puede optar por detectar la apertura de la boca bajando la mandíbula. Para ello, el uso de un clasificador que identifique la boca abierta es una opción. De esta forma, la interfaz quedaría preparada para interpretar lo siguiente:

- Cambios de orientación del rostro.
- Sonrisa y apertura de boca.

Entre los distintos usos que puede tener una interfaz natural de usuario podemos mencionar aquellos recursos tecnológicos para personas con problemas en su movilidad o que tengan limitaciones para la operación de aquellos dispositivos con los tradicionales mecanismos de control. Otro uso para este tipo de interfaz es el control de robots del tipo plataformas móviles, no sólo para el control de su desplazamiento sino también la posibilidad de controlar algún accesorio que pudiera estar montado sobre el robot, como por ejemplo, una cámara de video.

Líneas de Investigación y Desarrollo

Existen diversos trabajos en el estado del arte que estudian los mecanismos de control de robots a través de una interacción con el usuario por medio de la detección del rostro. En [8] se realiza la segmentación del rostro con la detección del color de la piel y un análisis geométrico para identificar la región del rostro. Luego se utiliza el algoritmo SURF (Speeded- Up Robust Features) para extraer descriptores visuales para por último comprobar la consistencia geométrica a través del método iterativo RANSAC (Random sample consensus) contra las imágenes de rostros de una base de datos.

El trabajo [9] utiliza el método de Viola &

Jones para detectar el rostro y embebe este algoritmo en un FPGA (Field Programmable Gate Array), logrando realizar el procesamiento de 30 cuadros por segundo. Esto lo hace apto para interfaces de comunicación con el usuario en sistemas embebidos montados en un robot.

Otros trabajos, tal como [10], también utilizan el método de Viola & Jones para detectar el rostro y es apto para ser ejecutado en un computador Raspberry Pi (creado por Raspberry Pi Foundation [11]).

Partiendo del prototipo [6] ya desarrollado por este grupo de investigación, el alcance del proyecto incluye lo siguiente:

- Incorporar detección de sonrisa y apertura de la boca para ejecutar acciones.
- Construcción de un robot móvil comandado de manera remota con los movimientos y gestos del rostro.
- Diseñar una interfaz gráfica de usuario en el que se visualicen botones de control.

El sistema completo consta de dos programas informáticos: uno para ser ejecutado en una laptop para comandar de manera remota al robot móvil, recibir las imágenes desde una cámara montada en el mismo y reconocer los movimientos y gestos del rostro; por otro lado, un programa para ser ejecutado en un computador Raspberry Pi para ocuparse de los motores y controlar la cámara que redirige las imágenes a la computadora remota.

Para el diseño de la interfaz gráfica se utilizará lenguaje C++ con la biblioteca Qt para visualizar, en la pantalla de la computadora, las imágenes que son recibidas desde el robot. La interfaz despliega botones de control, como se observa en la Figura 1. Estos botones pueden ser accionados con el mouse, con el teclado o a través de los gestos del rostro, formando con esta característica la Interfaz Natural de Usuario.



Figura 1. Botones de control del robot móvil y la cámara

Mediante los movimientos del rostro, se podrá controlar el desplazamiento del robot y accionar los botones u optar por el control del movimiento de la cámara, que dispondrá de dos grados de libertad (horizontal y vertical). Cabe destacar en este punto que, según la extensión de la ley de Fitts [12], el tiempo requerido para alcanzar un objetivo (en este caso, un botón) decrece cuando se incrementa el tamaño del mismo. Por ello se proponen botones de gran tamaño y en transparencia con el video a fin de reducir el tiempo de selección de los botones. En la interfaz se resaltaré un botón cuando el usuario oriente su cabeza y mirada hacia dicho botón. Para ejecutar la acción asociada al botón, se deberá abrir la boca. De esta forma, si el usuario se posiciona sobre el botón de avance y abre la boca, el robot avanzará mientras se mantenga la boca abierta. Si el usuario sonríe, se podrá alternar entre las siguientes tres posibilidades: controlar el desplazamiento del robot, controlar la orientación de la cámara o dejar de controlar el robot. Como se observa en la Figura 1, los botones de control cambiarán, como también lo hará el ícono mostrado en la esquina superior derecha, para que se pueda identificar sobre qué dispositivo se está realizando el control.

El robot móvil se confeccionará a partir del chasis provisto por la empresa Makerfire (Figura 2). Es una plataforma plana con cuatro motores de corriente continua. Al chasis se le incorporará la electrónica de control y se le suministrará energía mediante una batería de gel de 6 volts. Para permitir lograr que los motores giren en ambos sentidos, se utilizarán circuitos electrónicos Puente H, controlados con una placa Intel Galileo [13] que genera una señal tipo PWM proporcional a la velocidad de desplazamiento del móvil. El chasis también tendrá montado una cámara de video en un

soporte con dos servomotores (ver Figura 3). La placa Intel Galileo controla el movimiento de los dos servomotores para cambiar la elevación y el ángulo horizontal de la cámara. Adicionalmente, un sensor ultrasónico para medir proximidad colocado en el frente del robot actuará como un interruptor de emergencia cuando el robot se dirija hacia algún obstáculo, evitando la colisión.

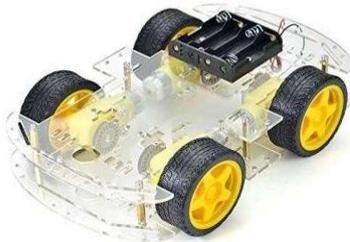


Figura 2. Chasis del robot Makerfire

La computadora Raspberry Pi montada en el robot dispone de pines GPIO [14] (General Purpose Input/Output) y a través de ellos se realizará la comunicación con la placa Intel Galileo. Para la comunicación inalámbrica desde el programa ejecutado en la laptop y el programa en la Raspberry se utilizará un punto de acceso de red inalámbrica basado en el estándar IEEE 802.11. La biblioteca Qt dispone de recursos para utilizar el protocolo TCP para el envío de datos, que posibilita tanto la comunicación para el control del robot como para la transmisión de las imágenes.

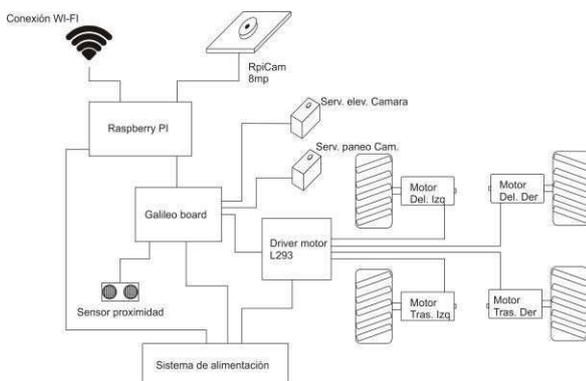


Figura 3. Diagrama en bloques del robot móvil

Resultados y Objetivos

Los objetivos de desarrollo se centran en el diseño de las herramientas para acercarnos a una interacción natural con las máquinas que nos rodean. El desarrollo de un robot móvil

controlado con el movimiento del rostro servirá como prueba de concepto acerca de la viabilidad y usabilidad de este tipo de tecnología para un usuario con movilidad reducida (o usuarios en general). Como se puede apreciar en [15], es posible su uso para el control de una silla de ruedas eléctrica.

Con el desarrollo de una interfaz natural, como la propuesta en este trabajo, facilitaría una mayor integración para personas que tienen dificultades en la comunicación oral. El desarrollo de un producto de apoyo que utilice técnicas de Comunicación Aumentativa y Alternativa, como por ejemplo ARASAAC [16], en donde el usuario, mediante pictogramas, genera una oración con la selección de un número reducido de imágenes, permitiendo expresarse mejorando su calidad de vida y autonomía personal.

Para lograr estos objetivos de desarrollo se realizarán las siguientes tareas, en tres momentos del proceso de diseño de la interfaz:

- **Antes de comenzar el diseño final:** Observación de distintos usuarios interactuando con el prototipo ya desarrollado [6] para conocer sus necesidades en el contexto real.
- **Durante la incorporación de detección de sonrisa y boca abierta:** Estas pruebas entregarán datos importantes para identificar una elaboración poco clara de la interfaz, que no se entiende y confunde, permitiendo realizar correcciones.
- **Con el producto final:** Las pruebas en esta etapa apuntarán a estudios pragmáticos aplicados al control de la interfaz para comandar el robot. Se recabarán datos para conocer su usabilidad, por ejemplo: conocer las opciones fáciles de comprender, en la que se desempeñan con más soltura, la que más agrada. Estas pruebas pueden conllevar a un rediseño completo del producto.

Se espera obtener resultados sobre la usabilidad de la interfaz tanto para usuarios en general como también para usuarios con movilidad reducida, sobre todo en sus miembros superiores. La Universidad Blas Pascal cuenta con la Comisión de Inclusión de Personas con Discapacidad que tiene por objeto impulsar actividades para el progresivo desarrollo de personas con dificultades que impiden su participación plena en la comunidad. Es importante destacar que se cuenta con el apoyo de esta comisión para realizar las pruebas con personas con movilidad reducida y con dificultades para su comunicación oral.

Formación de Recursos Humanos

El grupo se compone por cuatro investigadores con experiencias y roles distintos: el Ing. Carlos Bartó como asesor en metodología de la investigación, el Ing. César Osimani como director del proyecto y el Ing. Martín Salamero y el Lic. Marcos Lopez como investigadores en formación. Se cuenta con la colaboración de dos alumnos pasantes de Ing. Informática. Además, se asocia a la tesis "Reconocimiento de posturas de las manos para la Interacción Natural Humano Computadora en Ambientes Inteligentes con cámaras RGB" de César Osimani por el título Doctor en Ciencias de la Ingeniería de la FCEFYN de la UNC.

Financiamiento

La investigación en esta comunicación cuenta con el financiamiento para Grupos de Reciente Formación y Tutoría del Gobierno de Córdoba, y dos becas de investigación y una dedicación exclusiva a investigación por parte de la Universidad Blas Pascal.

Referencias

- [1]. "OpenCV - Open source Computer Vision", Online: <http://docs.opencv.org/trunk/index.html>, [Consulta: marzo de 2017].
- [2]. G Bradski, A. Kaehler, "Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library", O'Reilly Media, 2016.
- [3]. "Qt Library - Documentación", Disponible: <http://doc.qt.io>, [Consulta: marzo de 2017].
- [4]. P. Ithaya Rani, K. Muneeswaran, "Robust Real Time Face detection automatically from video sequence based on Haar features", "Communication and Network Technologies (ICCNT 2014)", Dec. 2014.
- [5]. "Menú controlado con orientación del rostro", <http://github.com/cosimani/menupencv>, 2015.
- [6]. "Video del menú controlado con el rostro", <https://www.youtube.com/watch?v=BUmQNs7GxPY>, 2015.
- [7]. C. Osimani y E. Kohmann, "Simple método para la Interacción Natural Humano Computadora con movimientos del rostro", CoNaIISI 2015, Bs. As. Arg., Nov. 2015.
- [8]. Shan An, Xin Ma, Rui Song, Yibin Li, "Face detection and recognition with SURF for human-robot interaction", 10th Asian Control Conference (ASCC 2015), June 2015.
- [9]. S.-S. Lee, S.-J. Jang, J. Kim, B. Choi, "A hardware architecture of face detection for human-robot interaction and its implementation", (ICCE-Asia), Oct. 2016.
- [10]. S. Fernandes, J. Bala, "Low Power Affordable and Efficient Face Detection in the Presence of Various Noises and Blurring Effects on a Single-Board Computer", 49th Annual Convention of the Computer Society of India (CSI), Volume 1. vol. 337, pp. 119-127, 2015.
- [11]. Raspberry Pi Foundation, About us, [online] Disponible:<https://www.raspberrypi.org/about>, [Consulta: marzo de 2017].
- [12]. R. A. Burnoa, B. Wub, R. Dohertya, H. Coletta, R. Elnaggar, "Applying Fitts' law to gesture based computer interactions", 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015), 2015.
- [13]. Placa Intel Galileo, [online] Disponible:<https://software.intel.com/es-es/iot/hardware/galileo>, [Consulta: marzo de 2017].
- [14]. S. Ahsan, "Peripheral control with a Raspberry Pi - An ultrasonic sensor", [online]. Disponible:<https://wiki.aalto.fi/download/attachments/84751400/burglaralarm.pdf>, [Consulta: marzo de 2017].
- [15]. M. A. Eid, N. Giakoumidis, A. El Saddik, "A Novel Eye-Gaze-Controlled Wheelchair System for Navigating Unknown Environments: Case Study With a Person With ALS", IEEE Access, Jan. 2016.
- [16]. F. C. Luque, E. B. López, "Símbolos pictográficos de ARASAAC: ¿son adecuados?". J. Navarro, M. T. Fernández, F. J. Soto, F. Tortosa (Coords.), (2012), Respuestas flexibles en contextos educativos diversos. ISBN: 978-84-616-0718-1.

Medición del Tamaño de Partículas de Minerales Mediante Procesamiento Digital de Imágenes

Mg. Ing. Carlos Gustavo Rodríguez Medina ¹, Dr. Ing. Oscar Daniel Chuk ²,
Ing. Regina Bertero, Ing. Pablo Trigo

Instituto de Investigaciones Mineras. Facultad de Ingeniería.
Universidad Nacional de San Juan

Av. Libertador Gral. San Martín 1109 oeste. San Juan
0264-4211700 (int. 285 ¹, int. 389 ^{1,2})

grodriguez@unsj.edu.ar ¹, dchuk@unsj.edu.ar ²

RESUMEN

La determinación de la granulometría resulta de mucha importancia en una amplia gama de industrias y constituye un parámetro crítico en la fabricación de diversos productos [1]. Por tal motivo es necesario la medición del tamaño de partículas en los procesos de molienda de minerales. La misma afecta a los productos y a los procesos [2] y además está relacionada con el consumo energético de los molinos intervinientes en dicha operación, como así también con la calidad del producto que se elabora, entre otros. Estas impactan directamente en los costos de producción [3]. Si bien existen diferentes métodos y tecnologías para determinar el tamaño de partículas de minerales (a través de técnicas manuales o por medio de dispositivos electrónicos), en el presente trabajo se expone una alternativa efectiva, de bajo costo y que tiene la finalidad de ser utilizada en línea con el proceso de molienda (sin la necesidad de estar tomando muestras para ser enviadas a laboratorios), mediante técnicas de procesamiento digital de imágenes. Se ha trabajado en el desarrollo de un algoritmo que permite la aplicación de técnicas de

procesamiento digital de imágenes para poder medir el tamaño de partículas microscópicas de minerales, que se encuentran en el rango que va de los 50 μm hasta los 800 μm aproximadamente. determinándose propiedades geométricas como así también medidas estadísticas de interés para el proceso.

Palabras clave: Procesamiento de imágenes, Tamaño de partículas. Algoritmo.

CONTEXTO

El desarrollo que se expone a continuación se lleva a cabo en el marco de un Proyecto de Investigación y Desarrollo para Jóvenes Investigadores de la provincia de San Juan, denominado PROJOVI, titulado "*Medición de propiedades geométricas de minerales a través de visión artificial con aplicación a la industria minera*". La misma es una Convocatoria de Proyectos recientemente lanzada y financiada (en su segunda edición) en forma conjunta por la Universidad Nacional de San Juan y el Gobierno Provincial orientada a la formación de grupos de investigadores jóvenes menores de 40 años.

Adicionalmente, los integrantes del mencionado proyecto conforman un equipo mayor, con una trayectoria de varios años en investigación y desarrollo, y en el que actualmente se están ejecutando proyectos correspondientes a convocatorias nacionales, tal como los PDTs (Proyectos de Desarrollo Tecnológico y Social), financiados por el Ministerio de Ciencia Tecnología e Innovación productiva (Mincyt), proyectos de Asistencia Exportadora "*Manuel Belgrano*", financiados por la Secretaría de Políticas Universitarias (SPU), y Proyectos internos de Investigación y Creación financiados por la Universidad Nacional de San Juan (proyectos CICITCA) y acreditados por dicha Universidad y el Ministerio de Educación.

1. INTRODUCCIÓN

La determinación de la granulometría en forma automática en la industria minera es uno de los problemas fundamentales en los procesos de control de molienda del mineral, especialmente cuando se procesan varias toneladas de mineral en forma continua, para así tener un mejor control sobre la operación óptima de carga de los molinos, además de poder cumplir con exigencias de mercados [8].

En particular, los minerales industriales (minerales No Metalíferos tales como calcita, bentonita, y otros) son usados como insumos para muchas industrias, tal como la de los plásticos, las pinturas, la siderurgia, materiales de construcción, etc.

Actualmente existen diversos métodos e instrumentos para medir el tamaño de partículas, dependiendo del tamaño de las mismas.

Los dispositivos para la medición de partículas secas en el rango de tamaños de interés más difundidos son los basados en difracción laser. El inconveniente que presentan estos equipos es que suelen ser de costos muy elevados, por lo cual solo grandes laboratorios de algunas empresas o de grandes centros de investigación pueden adquirirlos. Además los dispositivos

que aplican esta técnica requieren que se conozcan las propiedades ópticas de las partículas que componen la muestra, tal como el índice de refracción, entre otros [3].

De manera alternativa se pueden aplicar las técnicas de procesamiento digital de imágenes para la medición de partículas secas. Los sistemas de visión artificial se presentan como un campo emergente en la medición de tamaños y propiedades geométricas de partículas. Así lo manifiestan los artículos [4] y [5], y tesis doctorales como [6].

Recientemente se cuenta en el mercado con algunos aparatos para la medición del tamaño de partículas basados en el análisis de imágenes [7], pero algunos trabajan la muestra "*fuera de línea*" con el proceso (requieren tomar una muestra y llevarla luego a un laboratorio para ser analizada), y en otros el rango de medición es algo limitado (de 1 a 300 μm). Además, estos dispositivos también son de costos elevados.

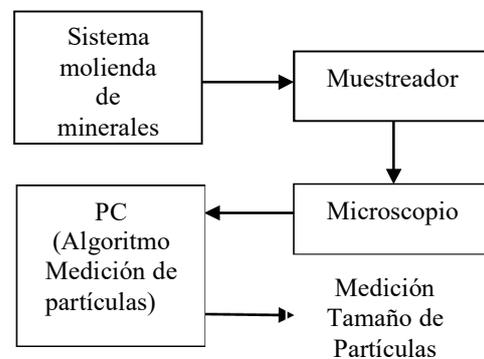


Figura 1. Sistema de visión artificial

El grupo de investigación se encuentra desarrollando un sistema de visión artificial el que va a permitir la medición de muestras "*on line*" con el proceso. Es decir, mediante un dispositivo muestreador inserto en los conductos de aire que transportan las partículas de mineral molido en un sistema de molienda, se toma una muestra y se captura la imagen a través de un microscopio USB. Esta imagen es tomada por una PC, la que a través de un algoritmo desarrollado específicamente mide el

tamaño y otras características de interés de la muestra. Los resultados podrán ser utilizados como entradas a un sistema de control automático de tamaño de partículas [3]. En la figura 1 se presenta un esquema del sistema de visión artificial mencionado.

El presente trabajo específicamente se limita a describir el algoritmo desarrollado para la determinación de tamaños de partículas y otras medidas estadísticas importantes para el proceso. El mismo ha sido escrito en lenguaje *M* de Matlab y permite analizar muestras de partículas secas de minerales no metalíferos en el rango de interés, y a un bajo costo.

Desarrollo del algoritmo

Los tamaños de partículas con los que se requiere trabajar se encuentran en el rango de los 50 a los 800 μm aproximadamente.

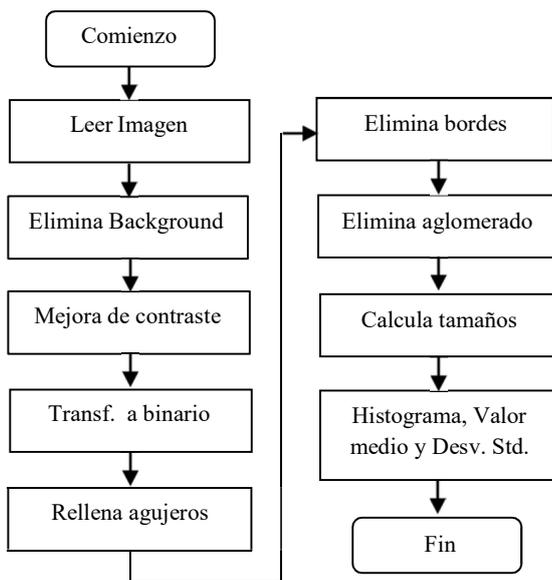


Figura 2. Diagrama de flujo del algoritmo desarrollado para medir tamaño de partículas de mineral.

Se captura la imagen a través del microscopio digital y se almacena en el disco de la PC para su procesamiento. En principio, se ha trabajado con la muestra esparcida sobre una superficie contrastante sobre la cual se toma la fotografía. Si se incorpora un dispositivo muestreador la

muestra puede ser analizada "on line" con el proceso, que en definitiva es lo que se desea realizar a través del sistema de Visión Artificial. Dado que el algoritmo ha sido implementado en Matlab, la manipulación de las imágenes es realizada en forma matricial. Por lo cual al leer la imagen, la misma se carga en memoria de forma matricial como una variable, formada por filas y columnas, cuya intersección determina la posición y el valor de cada pixel de la imagen (valor que varía entre 0 y 255).

La imagen capturada por el microscopio (figura 3.A) requiere de un procesamiento previo, con el fin de tener una imagen en condiciones para analizar y calcular el tamaño de las partículas. Como primer medida, se elimina el fondo de la imagen para descartar cualquier color y manchas que generen confusión con las partículas (figura 3.B). Luego se convierte la imagen a binario, es decir, mediante la determinación de un valor de umbral óptimo (valor de corte entre los grises más cercanos a los blancos o a los negros). Con el valor de umbral determinado se genera la imagen en blanco y negro presentada en la figura 3.C. Posteriormente el procesamiento continúa eliminando aquellas partículas que se hallan en el borde de la imagen, con la finalidad de que cuando se haga el cálculo de tamaño no se consideren aquellas partículas pegadas al borde y que pueden aparecer con un tamaño fraccionado, evitándose así introducir errores a los cálculos (figura 3.D).

También para descartar otro tipo de error, no se consideran aquellas partículas que se encuentran unidas unas con otras, denominadas usualmente como "aglomeradas". Para ello el algoritmo debe poder detectarlas, marcarlas y luego excluirlas del cálculo. En la figura 3.E. se han marcado con un color fucsia con el fin de que se puedan también distinguir visualmente con claridad, las que luego son eliminadas durante el procesamiento digital.

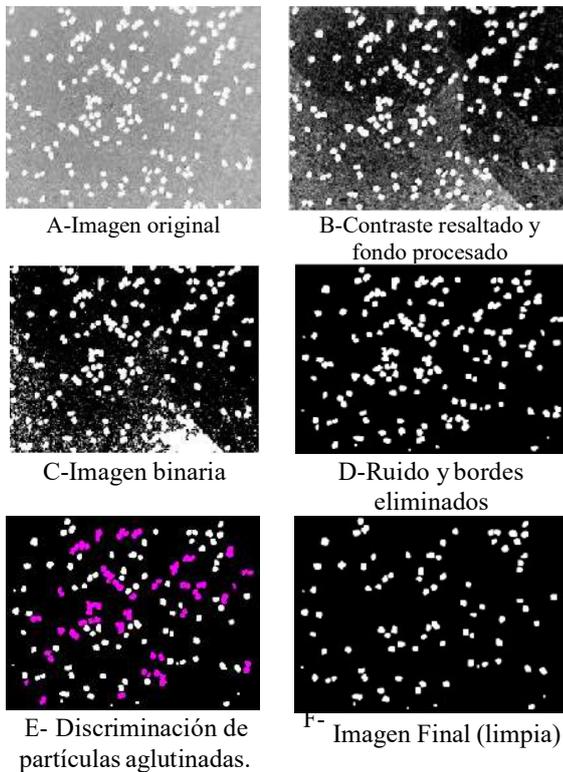


Figura 3. Procesamiento de imágenes a través del algoritmo para medir tamaño de partículas de mineral.

Se utiliza el cálculo definido como solidez S .

(Ecuación 1)

$$S = \frac{\text{Área del objeto}}{\text{Área de la envolvente convexa}}$$

Donde el *Área del objeto* es la cantidad de bits de color blanco acumulados en contacto entre sí dentro de un borde. El *Área de la envolvente convexa* es la cantidad de elementos binarios dentro del menor polígono convexo que rodea al objeto. La convexidad tiene una definición matemática concreta, pero se puede decir que es un polígono sin entradas cóncavas o bahías sobre sí mismo.

Para el caso del objeto A de la figura 4 se obtiene $S(A) = 0.8197$ mientras que para el objeto B – una partícula individual- el valor es $S(B) = 0.9622$. Estableciendo un valor de corte intermedio, típicamente $S = 0.9$, es posible discriminar objetos que deben ser incluidos en

las mediciones como partículas reales de los aglomerados que introducirían medidas erróneas.

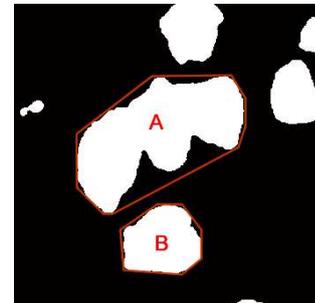


Figura 4. Objetos dentro de sus envolventes convexas.

En la figura 3.F. se visualiza la imagen final obtenida a través del algoritmo, luego de haber procesado la imagen original observada en la figura 3.A. La misma ya no contiene las partículas aglomeradas ni las pegadas a los bordes de la imagen a partir de la cual se determina el tamaño de las partículas.

El algoritmo determina el eje mayor y el eje menor de una elipse de área equivalente a la partícula, utilizando este último como tamaño equivalente. Dado que en el método patrón de los tamices, contra el cual se comparan los resultados, las partículas atraviesan los agujeros de los tamices a través de su eje menor.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el Laboratorio de Control Automático del Instituto de Investigaciones Mineras (Facultad de Ingeniería de la Universidad Nacional de San Juan) se realiza investigación y desarrollo respecto a la automatización de procesos de molienda de minerales hace casi dos décadas. Hace algunos años se viene desarrollando el Área de Procesamiento de Imágenes.

Actualmente se está trabajando en la medición de tamaños de partículas secas mediante un sistema de visión artificial que trabaje de manera *on-line* con el proceso de molienda y que permita el control automático del proceso. Otro de los ejes de investigación es la

caracterización de arenas especiales (forma, color, tamaños, clasificación del tipo de material, etc.) mediante el procesamiento de imágenes para ser utilizadas en la extracción de petróleo mediante el método de Fraking.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El algoritmo desarrollado, además de la determinación del tamaño, también calcula y presenta gráficamente el histograma correspondiente. De esta manera se , aprecian los tamaños y las frecuencias o cantidades de partículas con tamaños similares en la imagen analizada.

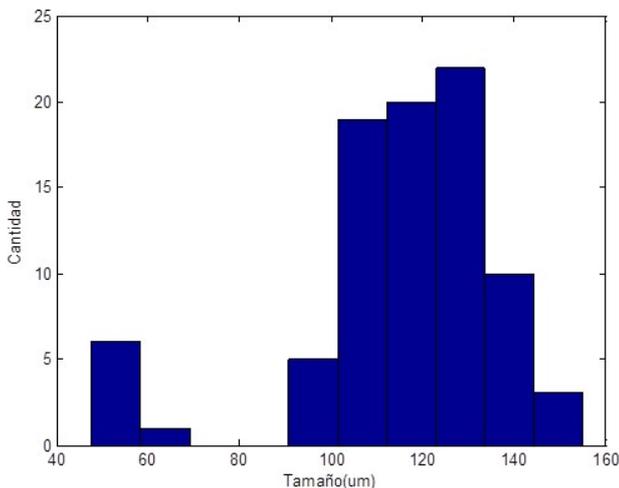


Figura 5. Histograma del tamaño de partículas de la muestra.

Tanto las imágenes de la figura 3 como el histograma de la figura 5, corresponden a una muestra de mineral molido de Calcita con tamaños de partículas entre los 106-150 µm, correspondientes a la clase granulométrica "100# +150#", de acuerdo a la tabla 1.

Esta tabla presenta una serie de muestras de diferentes clases granulométricas (clasificación de tamaños), determinados a través del método de los tamices, y las que se considerarán como muestras de medida patrón.

Al observar el histograma de la figura 5 se aprecia que la mayoría de las partículas se encuentran concentradas en dicho rango.

Clase	Tamaño (µm)
+20#	850
-20# +40#	850 - 420
-40# +70#	420 - 210
-70# +100#	210 - 150
-100# +150#	150 - 106
-150# +200#	106 - 75
-200# +325#	75 - 45
-325#	45

Tabla 1. Tamaños de muestras utilizadas.

Para esa misma prueba, el algoritmo determinó un *Valor Medio* de 114,75 µm y una *Desviación Estandar* de 22,26 µm. Dichos cálculos realizados también son comprobables gráficamente. Solo existe una pequeña cantidad de partículas medidas que están fuera del rango mencionado, lo cual puede deberse a ruido en la imagen analizada, contaminación de la muestra, y otras causas, lo que no resulta ser de importancia para los objetivos perseguidos.

Con el fin de validar los resultados obtenidos con la aplicación del algoritmo para la medición del tamaño de partículas, se han realizado diversas pruebas con diferentes *clases granulométricas* de mineral Calcita, de acuerdo a las indicadas en la tabla 1.

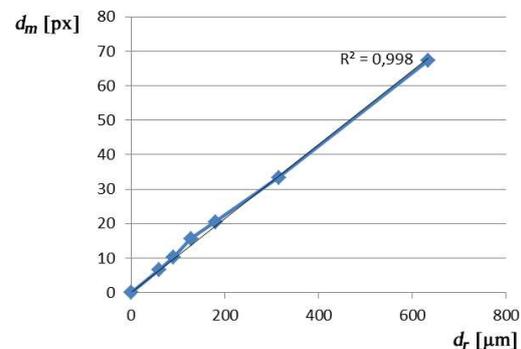


Figura 6. Tamaños de partículas reales vs. calculados

La figura 6 presenta una curva comparativa entre los tamaños reales " d_r ", de cada clase granulométrica usados como patrón [µm] y los diámetros equivalentes " d_m " calculados en pixeles [px]. Se puede observar una alta linealidad, lo que permite asegurar la validez del método tan solo con establecer una proporcionalidad entre ambas medidas. La calibración resulta un procedimiento sencillo.

Como conclusión, el algoritmo:

- Realiza la determinación de tamaños de partículas de la muestra contenida en cada imagen de manera confiable.

- Calcula y presenta el Histograma del tamaño de partículas de la muestra y medidas estadísticas de interés tal como el valor medio y la desviación estándar.

- Detecta y elimina la aglomeración de partículas.

- Integra un sistema de Visión Artificial de bajo costo que permitirá el análisis de partículas de mineral en forma "on line" con el proceso de molienda.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación se conforma por dos Ingenieros Electrónicos, con formación de posgrado de Doctorado y Maestría (cuya tesis fue defendida el año pasado), un Ingeniero en Minas, una Ingeniera Metalurgista cursando una Maestría en Metalurgia Extractiva, y un alumno recientemente egresado que está por comenzar su Doctorado en Procesamiento de Minerales.

En el seno del grupo, el año pasado un alumno de la carrera de grado de Ingeniería en Minas desarrolló una Beca de Investigación en la categoría Estudiantes avanzados y también llevó a cabo y presentó su tesina de grado.

El mencionado proyecto PROJOVI tiene como finalidad formar nuevos investigadores menores a 40 años, el que también se está ejecutando en el dicho grupo de investigación permitiendo sumar dos nuevos investigadores al equipo.

5. BIBLIOGRAFIA

[1] Sánchez M. A.; Nieto J.; Berrocal J. M. 2013. Un sistema de visión artificial para detectar y estimar el tamaño de rocas. Universidad de Salamanca.

[2] Malvern Instrument International. Tipos de instrumentos de medición de partículas actuales. 2017. <http://www.malvern.com/es/products/measurement-type/particle-size/>

[3] Rodriguez M., C. Gustavo; Chuk, O. Daniel; Bertero, Regina; Trigo, Pablo. Medición de propiedades geométricas de minerales a través de visión artificial con aplicación a la industria minera. Universidad Nacional de San Juan, Argentina. Informe de Avance de Proyecto PROJOVI. 2016.

[4] Igathinathane C.; Pordesimo L.O.; Columbus E.P.; Batchelor W.D.; Methuku S.R.; 2008. Shape identification and particles size distribution from basic shape parameters using Image computers and electronics in agriculture 63, 168–182.

[5] Kröner S., M.T. Doménech Carbó, 2013. Determination of minimum pixel resolution for shape analysis: Proposal of a new data validation method for computerized images. Powder Technology 245, 297–313.

[6] Andersson T, 2010. Estimating Particle Size Distributions Based on Machine Vision. Doctoral Thesis. Dept. of Computer Science and Electrical Engineering. Lulea University of Technology, Lulea, Sweden.

[7] Soto C. P.; Bollo M. C. Visión artificial en la minería con redes neuronales. Universidad Católica del Norte, Antofagasta, Chile.

[8] Corke P.; Roberts J.; Winstanley G. 1998. Vision Based Control for Mining Automation. IEEE Robotics and Automation Magazine. Vol 5, N 4. Pp: 44-49.

Modelos de Interacción y Aplicaciones en Realidad Virtual mediante Dispositivos Móviles

Matías Selzer^{1,2}, Elisabet Arriata², Leonardo Segovia², Nicolas Gazcón^{1,2},
Martín Larrea^{1,2}

¹Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica

²Departamento de Ciencias e Ingeniería en Ciencias de la Computación

Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

{matias.selzer, leonardo.segovia, ng, mll}@cs.uns.edu.ar, elisabet.arratia@gmail.com

RESUMEN

La Realidad Virtual (RV) se ha convertido, una vez más, en un tópico popular e interesante tanto en el ámbito de la investigación como en el campo comercial. Esta tendencia tiene su origen en el uso de dispositivos móviles como núcleo computacional y *displays* de RV. Tales dispositivos no están libres de limitaciones, tanto en el *software* como *hardware* que soportan. Esta línea de investigación tiene como objetivo analizar el impacto de los dispositivos móviles en los modelos de interacción de RV, y desarrollar nuevas aplicaciones de esta tecnología.

Palabras clave: *Realidad Virtual, Interacción Humano Computadora en Realidad Virtual, Realidad Virtual en Dispositivos Móviles*

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

Realidad Virtual

La RV es un entorno de escenas u objetos de apariencia real, generados mediante tecnología informática, que crea en el usuario la sensación de estar inmerso en él. Dicho entorno es contemplado por el usuario a través de un dispositivo conocido como gafas o casco de realidad virtual. Este puede ir acompañado de otros dispositivos, como guantes o trajes especiales, que permiten una mayor interacción con el entorno así como la percepción de diferentes estímulos que intensifican la sensación de realidad.

La aplicación de la RV, aunque centrada inicialmente en el terreno del entrenamiento y de los videojuegos, se ha extendido a muchos otros campos, como la medicina ([1, 2, 3]), la arqueología ([4, 5]), el entrenamiento militar ([6]) o diferentes tipos de simulaciones ([7]).

Abismos de Ejecución y Evaluación

Uno de los modelos más utilizados para el análisis de las interacciones y del nivel de inmersión que la RV ofrece es el modelo de Norman ([8]). Bajo tal modelo, el usuario piensa en un conjunto de acciones a realizar a medida que interactúa con el sistema. Mientras el plan se ejecuta, el usuario observa cómo la interfaz va respondiendo y, en base a estas respuestas,

decide cómo proseguir con su plan de acción. Esta manera de evaluar las interacciones es conocida como ciclo de ejecución-evaluación de Norman. Este modelo de evaluación introduce dos conceptos fundamentales:

- **Abismo de Ejecución:** es la distancia que hay entre las metas del usuario y los medios para lograrlas utilizando el sistema.
- **Abismo de Evaluación:** es la distancia que hay entre las respuestas del sistema y las metas del usuario.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Actualmente las líneas de investigación se focalizan en los distintos tipos de interacciones que se pueden producir entre el usuario y el entorno virtual, considerando ambientes virtuales generados desde dispositivos de bajo costo. Las interacciones son evaluadas mediante la aplicación del modelo de Norman, bajo el desarrollo de casos de aplicación.

Interacciones en RV: En las últimas décadas ha existido mucho interés en el desarrollo de sistemas de realidad virtual, especialmente en el ámbito de simuladores y en la medicina ([9, 10]).

En un principio, la única forma de interactuar con los ambientes virtuales era a través de la vista. Los usuarios sólo podían ver y apreciar el mundo virtual en el que se los situaba. Posteriormente se comenzó a incluir sonido, permitiendo que usuarios puedan oír distintos sonidos provenientes de los objetos virtuales. Luego, se comenzó a incluir el sentido del

tacto para que los usuarios puedan, por ejemplo, navegar a través del mundo virtual mediante el uso de dispositivos tangibles, como por ejemplo un teclado o un joystick. Se puede notar que cuantos más sentidos estén involucrados a la hora de interactuar con el mundo virtual, los usuarios sentirán que este mundo se asemeja cada vez más al mundo real. Se dice entonces que este mundo virtual es cada vez más inmersivo. Nuestra investigación analiza e investiga todas las posibles formas en que el usuario puede interactuar con el entorno virtual, y cómo este entorno virtual brinda una devolución al usuario, generando así un mayor nivel de realismo y una mayor inmersión.

RV sobre Dispositivos Móviles: Un sistema de RV funcional consiste al menos de algún tipo de visor, alguna computadora para procesar la aplicación y la ejecución el mundo virtual (CPU, memoria, SO, etc), y sensores para detectar datos del usuario, como por ejemplo la posición y la rotación de su cabeza.

Hace años, todos estos elementos resultaban muy costosos, pero actualmente los avances tecnológicos han logrado que los costos en su construcción se reduzcan. Sin embargo, este costo sigue siendo elevado para la mayoría de las personas, por lo que han surgido distintas alternativas para la construcción de estos sistemas. Una de ellas es mediante la utilización de dispositivos móviles, es decir, los *smartphones*, que disponen de todos los elementos necesarios anteriormente mencionados.

Nuestra línea de investigación consiste en la investigación y el desarrollo de aplicaciones de RV enfocadas

especialmente para funcionar en dispositivos móviles. Nos centramos principalmente en técnicas para reducir el costo computacional de los elementos gráficos y, en relación a la línea de investigación anterior, investigar sobre los distintos tipos de interacciones que se pueden lograr entre los usuarios y los dispositivos móviles, aprovechando al máximo las herramientas que nos proveen, como por ejemplo sus bastos sensores y sus posibilidades de intercomunicación.

Las tecnologías más importantes utilizadas en esta línea de investigación son:

- **Android:** es el Sistema Operativo que se utiliza en la mayoría de los dispositivos móviles de la actualidad, diseñado principalmente para dispositivos con *touchscreen* como *tablets* o *smartphones*. Su código fuente dispone licencia *Open Source*, lo que ha motivado a una comunidad de desarrolladores y entusiastas a realizar cientos de proyectos.
- **Arduino:** es una empresa que diseña microcontroladores para construir dispositivos digitales e interactivos que pueden comunicarse con el mundo exterior. Por esto, sus microcontroladores son más conocidos como Arduinos. En nuestros proyectos utilizamos Arduino UNO porque es uno de los microcontroladores más populares, poderosos y baratos de hoy en día.
- **Unity3d:** es un motor de videojuegos multiplataforma utilizado para desarrollar videojuegos para PC, consolas, dispositivos móviles, y páginas web, entre otros. Utilizamos este motor porque está disponible gratuitamente, es fácil de utilizar, y nos brinda la facilidad de crear aplicaciones para Android con un alto contenido gráfico.

3. RESULTADOS OBTENIDOS/ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales. Desde el punto de vista de la teoría de base en RV, se ha comenzado a realizar un relevamiento de los diferentes modelos de interacción y su implementación en ambientes virtuales generados por dispositivos móviles. En lo referido al desarrollo de aplicaciones de RV se han diseñado prototipos de recorridos virtuales con objetivos turísticos y educacionales, y ambientes virtuales para la medicina. Al mismo tiempo se desarrolló el trabajo “AnArU, a virtual reality framework for physical human interactions” el cual fue publicado en el CACIC 2015 y en una edición especial de *Journal of Computer Science & Technology* ([11, 12])

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo y concluidas relaciones con las líneas de investigación presentadas, así también como un proyecto en ejecución

Tesis Desarrollada

“Técnicas de deformación para objetos virtuales. El impacto entre vehículos como caso de estudio”, tesis de grado para la Licenciatura en Ciencias de la Computación. Alumno: Leonardo Segovia. Director: Martín Larrea.

Tesis en Desarrollo

“Realidad Virtual y el Tratamiento de Fobias”, tesis de grado para la Ingeniería en Computación. Alumna: Elizabet Arriata. Directores: Martín Larrea y Nicolás Gazcón.

“Interacción Humano Computadora en Ambientes Virtuales”, tesis de posgrado para el Magister en Ciencias de la Computación. Alumno: Matías Selzer. Director: Martín Larrea.

5. BIBLIOGRAFÍA

1. Freedman, S. A., Dayan, E., Kimelman, Y. B., Weissman, H., & Eitan, R. (2015). Early intervention for preventing posttraumatic stress disorder: an Internet-based virtual reality treatment. *European journal of psychotraumatology*, 6.
2. Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B. & Ressler, K. J. (2014). A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan War veterans. *American Journal of Psychiatry*.
3. Gorini, A., & Riva, G. (2014). Virtual reality in anxiety disorders: the past and the future. *Expert Review of Neurotherapeutics*.
4. Lynch, J., & Corrado, G. (2014). Arqueología virtual aplicada al sitio Villavil, Catamarca, Argentina Virtual Archaeology applied to the site Villavil, Catamarca, Argentina.
5. Gagne, R., Gouranton, V., Dumont, G., Chauffaut, A., & Arnaldi, B. (2014). Immersia, an open immersive infrastructure: doing archaeology in virtual reality. *Archeologia e Calcolatori, supplemento* 5, 1-10.
6. Carroll, J. M. (Ed.). (2003). *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann.
7. Schreuder, H. W., Persson, J. E., Wolswijk, R. G., Ihse, I., Schijven, M. P., & Verheijen, R. H. (2014). Validation of a novel virtual reality simulator for robotic surgery. *The Scientific World Journal*, 2014.
8. Armstrong, C. M., Reger, G. M., Edwards, J., Rizzo, A. A., Courtney, C. G., & Parsons, T. D. (2013). Validity of the Virtual Reality Stroop Task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology*, 35(2), 113-123.
9. Taffou, M., Ondřej, J., O'Sullivan, C., Warusfel, O., Dubal, S., & Viaud-Delmon, I. (2015, April). Auditory-visual virtual environment for the treatment of fear of crowds. In *Laval Virtual*.
10. Suied, C., Drettakis, G., Warusfel, O., & Viaud-Delmon, I. (2013). Auditory-visual virtual reality as a diagnostic and therapeutic tool for cynophobia. *Cyberpsychology, Behavior, and Social Networking*, 16(2), 145-152.
11. Selzer, M. N., & Larrea, M. L. (2015). AnArU, a virtual reality framework for physical human interactions. In *XXI Congreso Argentino de Ciencias de la Computación (Junín, 2015)*.
12. Selzer, M. N., & Larrea, M. L. (2015). AnArU, a virtual reality framework for physical human interactions. *Journal of Computer Science & Technology*; 2015, vol. 15, no.2, 50-54.

Visualización Basada en Semántica

Martín Larrea^{1,2,3}, Silvia Castro^{1,2,3}, Dana Urribarri^{1,2,3,4}, Sebastián Escarza^{1,2}

¹Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab)

²Departamento de Ciencias e Ingeniería en Ciencias de la Computación (DCIC)

³Instituto de Ciencias e Ingeniería de la Computación (ICIC),

⁴Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

{mll, smc, dku, se}@cs.uns.edu.ar

RESUMEN

Esta línea de investigación tiene como objetivo crear un modelo de visualización que considere la semántica de los datos, del contexto y de las etapas del “Modelo Unificado de Visualización” (MUV) para poder asistir al usuario en el seteo de los parámetros de la visualización logrando así una mejor representación visual. Este modelo deberá formalizar el modelo de referencia mediante una Ontología de Visualización, incorporar una Ontología de Datos e incorporar una capa de inferencia que le permita derivar nueva información, a partir de la ya conocida. Esta línea también busca introducir el uso de semántica y razonamiento semántico en diferentes aspectos del proceso de visualización.

Palabras clave: *Ontología, Representación Formal, Visualización basada en semántica, Visualización.*

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

Actualmente existe un gran número de modelos de referencia en Visualización ([1,2,3,4,5]) que, de diversas maneras, han identificado y delineado los principales componentes y procesos que sufren los datos para ser visualizados. En particular, en el seno del VyGLab hemos estado desarrollando el Modelo Unificado de Visualización (MUV) ([6]), un modelo que constituye un marco conceptual de referencia en términos del cual ubicar los procesos y estados intermedios de los datos y definir las interacciones explícitamente. Si bien estos esfuerzos por compilar, caracterizar y clasificar los aspectos más relevantes del área han sido y siguen siendo fructíferos, aún no existe un consenso definitivo que permita consolidar una teoría de base en Visualización. La incorporación de una semántica bien definida en el proceso de Visualización ya ha sido identificada como una necesidad ([7,8]) en el sentido que permitiría formalizar dicho proceso, estableciendo un vocabulario común que permita a los usuarios plantear sus requerimientos, y a los diseñadores de herramientas de visualización, expresar las transformaciones de los datos desde su obtención hasta la construcción de la vista, las operaciones que deben proveerse y las formas de interacción posibles entre los usuarios y la visualización. Para lograr una formalización consensuada del proceso de Visualización, es

preciso describir axiomáticamente el marco de referencia provisto por los modelos mencionados anteriormente. En este contexto, las ontologías surgen como una herramienta natural para estos fines. Un ejemplo de los beneficios de la formalización del proceso de visualización se encuentra en la formalización de los tipos de datos, es decir una Ontología de Datos. Contar con una ontología que clasifique los conjuntos de datos a visualizar es una guía que asiste a la hora de elegir la técnica de visualización apropiada para determinado conjunto de datos. En la literatura es posible encontrar taxonomías de datos, una versión menos formal de la ontología, pero no son presentadas desde un punto de vista de visualización. Por otro lado, el desafío de una visualización es encontrar una metáfora visual que permita entender y percibir en forma efectiva un conjunto de datos. Para este propósito es útil contar con cuantificaciones de diferentes aspectos de cada técnica de visualización. Una visualización debe proveer también un conjunto de interacciones a partir de las cuales el usuario explorará el conjunto de datos con una mínima carga cognitiva. La tecnología computacional actual permite la exploración de grandes conjuntos de información. Por un lado, esta situación es extremadamente útil pero la creciente cantidad de información genera una sobrecarga cognitiva. Mientras que el poder computacional ha crecido en forma exponencial, la habilidad para interactuar con dichos datos solo se ha incrementado en forma lineal. Hoy en día, una gran variedad de usuarios acceden, extraen y muestran información que está distribuida sobre diferentes puntos, con diferentes tipos, formas y contenidos. En muchos casos, el usuario debe tener un control activo sobre el proceso

de visualización pero, aún en este caso, es difícil obtener una visualización efectiva. Es común que la información que se desea representar no tenga una manifestación visual obvia, ante esta situación el proceso de mapeo del conjunto de datos a la vista puede llegar a ser no trivial ([9,10]).

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este trabajo presenta dos ejes de investigación las cuales se entrelazan y tienen varios puntos en común. Tales ejes son:

1. “Formalización del proceso de Visualización mediante el uso de Ontologías”.
2. “Visualización Basada en Semántica”.

Formalización del proceso de Visualización mediante el uso de Ontologías. Este eje es una continuación directa de la línea de investigación presentada en ediciones anteriores de este Workshop ([11]), en la cual se plantea la necesidad de definir una Ontología de Visualización y de brindar soporte a dicha especificación formal a través de una plataforma de software que facilite la integración de los diversos componentes. Este trabajo integra el uso de ontologías en Visualización. Las ontologías, si bien se originaron hace tiempo en el área de Representación del Conocimiento, han cobrado nueva relevancia con el surgimiento de proyectos como la Web Semántica. A raíz de ello, numerosos estándares han sido definidos. Especificaciones como RDF, RDF Schema y OWL (el Lenguaje de Ontologías de la Web) permiten definir vocabularios estructurados que modelan jerarquías de herencia entre conceptos, relaciones, restricciones y reglas, que permiten un tratamiento más riguroso del conocimiento al

poseer una semántica formal asociada. En cuanto a la integración de semántica en visualización, aún no se cuenta con abordajes sistemáticos al problema y se identifican dos enfoques principales. Por un lado, se utilizan representaciones formales para mejorar la integración, consulta y descripción de los datos del usuario como una forma de enriquecer la visualización, y por el otro, trabajos más relacionados con el que proponemos utilizan descripciones semánticas, aunque en principio algo limitadas, del proceso de visualización en sí mismo para ayudar al usuario en la definición y configuración del mismo y en la selección de una técnica de visualización apropiada.

Visualización Basada en Semántica. Debido a que el objetivo de una visualización es lograr una representación que ayude al usuario a interpretar un conjunto de datos y comunicar su significado, es importante controlar el mapeo de las dimensiones físicas a las perceptuales; un usuario inexperto podría utilizar un mapeo incorrecto afectando negativamente la visualización resultante. Una estrategia para mejorar esta situación es guiar al usuario en la selección de los diferentes parámetros involucrados en la visualización a través de información semántica.

3. RESULTADOS OBTENIDOS/ESPERADOS

A nivel de la primera línea de investigación presentada, la transformación de los objetivos en resultados sigue las etapas delineadas en [11]. Al momento de esta publicación se cuenta con una definición parcial de la Ontología de Visualización. Concretamente, la arquitectura general de la ontología ya ha sido definida así como los principales elementos de la misma. Adicionalmente,

también se ha definido un conjunto de conceptos destinados a caracterizar el espacio de datos del usuario y la representación visual, y se continúa trabajando en la definición de los conceptos ontológicos necesarios para describir el proceso de visualización como una red de etapas interconectadas, así como en las definiciones necesarias para especificar las interacciones entre el usuario y la visualización. Sobre la segunda línea de investigación, se ha logrado incluir la semántica de los datos en el mapeo de datos a colores en una representación visual ([12,13,14]). Gracias a este trabajo, esta etapa del proceso de visualización ya no requiere de la participación del usuario; las decisiones de qué color usar para cada dato se determinan a partir de un razonador semántico. El próximo objetivo es integrar métricas al proceso de decisión semántica.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo relacionadas con las líneas de investigación presentadas:

- Tesis Doctoral. Sebastián Escarza. Tema: Ontologías de Visualización. Dirección: Dra. Silvia Castro.
- Tesis Magister. Georgina Inés Cerúsico. Tema: Taxonomías de Técnicas de Visualización para la Formalización del Proceso de Visualización. Dirección: Dra. Silvia Castro y Dr. Martín Larrea.
- Tesis Magister. Alejandra Elizabeth Herrera. Tema: Coordenadas Paralelas. Visualización e Interacciones. Dirección: Dra. Silvia Castro.
- Tesis Magister. César Escobal Blanco. Tema: La Reducción de Datos en las

Técnicas de Visualización. Dirección:
Dra. Silvia Castro.

5. BIBLIOGRAFÍA

- [1] E. H. Chi. A taxonomy of visualization techniques using the data state reference model. In Proceedings of the IEEE Symposium on Information Visualization (InfoVis'00), page 6975. IEEE Computer Society Press, 2000.
- [2] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. IEEE Symposium on Visual Languages, 0:336, 1996.
- [3] K. Brodlie and N. M. Noor. Visualization Notations, Models and Taxonomies. pages 207–212, Bangor, United Kingdom, 2007. Eurographics Association.
- [4] Brunetti, Josep Maria, et al. "Formal linked data visualization model." Proceedings of International Conference on Information Integration and Web-based Applications & Services. ACM, 2013.
- [5] Ma, Kwan-Liu. A Pervasive Parallel Processing Framework for Data Visualization and Analysis at Extreme Scale. No. DOE--UCD5373. Univ. of California, Davis, CA (United States), 2017.
- [6] Martig S., Castro S., Fillotrani P., Estévez E., Un Modelo Unificado de Visualización, Proceedings, pp. 881-892, 9º Congreso Argentino de Ciencias de la Computación. 6 al 10 de Octubre de 2003. La Plata. Argentina.
- [7] [DBD*05] D. J. Duke, K. W. Brodlie, D. A. Duce, and I. Herman. Do you see what I mean? IEEE. Computer Graphics and Applications, 25(3):6–9, 2005
- [8] M. Chen, D. Ebert, H. Hagen, R. S. Laramee, R. van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. IEEE Comput. Graph. Appl., 29(1):12–19, 2009.
- [9] Carpendale, M. S. T. 2001. Considering Visual Variables as a Basis for Information Visualization. Technical Report. University of Calgary, Department of Computer Science
- [10] Gu, Yi, et al. "Visualization and recommendation of large image collections toward effective sensemaking." Information Visualization 16.1 (2017): 21-47.
- [11] S. Escarza, S. Castro, S. Martig. Formalización del Proceso de Visualización mediante el uso de Ontologías. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 261-265. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6
- [12] M. Larrea, S. Martig and S. Castro. Semantics-based Color Assignment in Visualization. Journal of Computer Science & Technology. Vol. 10 - No. 1 – April 2010 - ISSN 1666-6038.
- [13] M. Larrea, S. Martig and S. Castro. Formalización del Proceso de Visualización Basada en Semántica. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 270-274. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6.
- [14] Escarza, S., Larrea, M., Urribarri, D., Castro, S., Martig, S. "Integrating Semantics into the Visualization Process". Dagstuhl Follow-Up Seminar Book "Scientific Visualization: Interactions, Features Metaphors" (2011), pp. 92-102. ISBN: 978-3-939897-26-2.

Visualización y Realidad Aumentada en el Campo de las Ciencias Geológicas

Luján Ganuza¹, Juan Manuel Trippel Nagel^{1,3}, Nicolás Gazcón^{1,3}, Silvia Castro¹, Ernesto Bjerg^{2,3}, Florencia Gargiulo^{2,3}, Gabriela Ferracutti^{2,3}, Krešimir Matković⁴ y Eduard Gröller⁵

¹ Lab. de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur

² INGEOSUR y Departamento de Geología, Universidad Nacional del Sur

³ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

⁴ VRVis Research Center, Vienna, Austria

⁵ Technische Universität Wien, Institut für Computergraphik und Algorithmen, Vienna, Austria

{mlg, juan.trippel, nfg, smc}@cs.uns.edu.ar, ebjerg@ingeosur-conicet.gob.ar, {mfgargiulo, gferrac}@uns.edu.ar, Matkovic@vrvis.at, groeller@cg.tuwien.ac.at

Resumen

En el campo de las Ciencias Geológicas, un desafío importante consiste en encontrar una representación adecuada de una gran cantidad de datos de distintos tipos, que abarcan análisis de muestras minerales, datos topográficos, proyecciones cartográficas, datos geofísicos, entre otros. La exploración y análisis de estos datos requiere un soporte visual adecuado. Por otra parte, es importante la asistencia de dispositivos móviles al momento de adquirir estos datos en el trabajo de campo.

En esta línea de investigación se está trabajando en el diseño y la generación de visualizaciones y de sistemas que las soporten con el objetivo de asistir al geólogo en varias de sus tareas habituales. Por un lado, una sub-línea se refiere a la visualización de datos geológicos que provean un soporte adecuado para la exploración eficiente de distintos conjuntos de datos de micro-química. Por otro, se trabaja en incorporar visualizaciones que asistan al geólogo en la obtención de datos geológicos en el trabajo de campo; esto se lleva a cabo mediante el uso de tecnología de Realidad Aumentada y el uso de dispositivos móviles.

Palabras clave: *Visualización de datos geológicos, Representación de composiciones minerales, generación de*

contornos, Realidad Aumentada, Dispositivos Móviles.

Contexto

La línea de investigación presentada está inserta en el proyecto “Análisis Visual de Grandes Conjuntos de Datos” (24/N037), dirigido por la Doctora Silvia Castro. El proyecto es financiado por la Secretaría General de Ciencia y Tecnología de la Universidad Nacional del Sur; y acreditado por la Universidad Nacional del Sur, Bahía Blanca.

Introducción

En los distintos campos de aplicación referidos a las Ciencias Naturales, el crecimiento vertiginoso de la cantidad de información genera volúmenes de datos cada vez más grandes y difíciles de comprender y analizar sin un soporte visual adecuado.

Es en estos casos donde de la visualización contribuye significativamente a la exploración y entendimiento de estos conjuntos de datos, siempre y cuando se cuente con un soporte adecuado [B*16]

[C93][C*99][K03][K*17][M98][NS79].

En el campo de las Ciencias Geológicas un desafío importante consiste en encontrar una representación adecuada de una gran cantidad de datos. Los tipos de los datos a explorar abarcan análisis de muestras minerales, datos topográficos, proyecciones cartográficas, datos geofísicos, entre otros. La exploración y análisis de estos datos requiere un soporte visual adecuado.

Otro desafío de importancia, es la asistencia al geólogo en el trabajo de campo. Es común en su desempeño la utilización de dispositivos como GPS, brújulas, clinómetros o magnetómetros. Sin embargo, sería de gran utilidad contar con la asistencia y las posibilidades que otorgan dispositivos computacionales tales como los actuales dispositivos móviles. Gracias a las facilidades que brinda la Realidad Aumentada (RA) al permitir aumentar el mundo que nos rodea, es posible asistir al geólogo en su trabajo de campo. Tareas como marcar puntos de interés en un mapa interactivo o la visualización *in situ* de la información referida, por ejemplo, a estructuras tectónicas podrían ser realizadas de manera interactiva durante el trabajo de campo mediante el uso de dispositivos móviles y RA.

Esta línea de investigación propone el estudio e implementación de dos sistemas de visualización interactivos de datos geológicos, que provean un soporte adecuado para la exploración, análisis y síntesis eficiente de los datos.

Líneas de Investigación y Desarrollo

En el contexto de esta línea de investigación se están desarrollando los siguientes trabajos en paralelo:

1. Visualización Aplicada a la Categorización Semi-Automática de Composiciones Minerales.
2. Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.

Visualización Aplicada a la Categorización Semi-Automática de Composiciones Minerales.

Este trabajo se lleva a cabo con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS), el INGEOSUR CCT-CONICET, Bahía Blanca, Argentina, y el VRVis Research Center y la TU Wien, Viena, Austria.

Este trabajo se enfoca en la visualización de un conjunto de datos geológicos, en particular, el conjunto de minerales que integran el grupo de los Espinelos. Este grupo de minerales resulta un excelente candidato a ser explorado y visualizado ya que es considerado un importante indicador petrogenético, proveyendo información vital en lo referido al ambiente tectónico de las rocas presentes en determinada área en el contexto de la tectónica global [BR01] [CD97] [L91] [R94].

En 2001, Barnes y Roeder [BR01] compilaron en una base de datos más de 26.000 análisis de Espinelos correspondientes a rocas ígneas y metamórficas. En base a estos análisis delinearon contornos para un conjunto de campos composicionales característicos (o patrones). Los geólogos suelen utilizar estos contornos para estimar el ambiente tectónico donde una muestra de Espinelos se podría haber formado.

Esta tarea es propensa a errores e involucra una tediosa comparación manual de diagramas superpuestos.

El objetivo de este trabajo de investigación es desarrollar una herramienta de Visualización que no solo permita visualizar y explorar conjuntos de datos correspondientes a composiciones minerales pertenecientes al grupo de los Espinelos, sino que también ofrezca técnicas que permitan categorizar

semiautomáticamente e interactivamente un conjunto arbitrario de Espinelos en función de los contornos de Barnes y Roeder.

Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.

Este trabajo es llevado a cabo por becarios doctorales y postdoctorales del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (UNS) con asistencia del INGEOSUR CCT-CONICET. Una de las actividades que realizan habitualmente los geólogos consiste en trabajos de campo.

Estas tareas se efectúan en ambientes exteriores, y en muchos casos en entornos que pueden ser hostiles y en los que sólo se tiene comunicación a través de los satélites.

La RA aplicada a las ciencias geológicas busca asistir en la práctica de dichas tareas, simplificando y complementando el uso de herramientas propias de la actividad como lo son las cartas geológicas, las brújulas y los mapas. Para esto se incorpora a la visión de la realidad, información que pueda resultar de utilidad para el geólogo como lo es, por ejemplo, información de formaciones geológicas u objetos virtuales que podrían indicar puntos de interés del geólogo. Las interacciones que se pueden lograr con la visualización de datos en tiempo real en el mismo trabajo de campo pueden contribuir efectivamente a esta disciplina, aportando soluciones y resultados en base al entendimiento de los datos y a sus relaciones.

Debido a la naturaleza móvil y en entornos abiertos, el subsistema de *tracking* es de suma importancia para asistir correctamente al geólogo [Vea12]. En este sentido los sistemas de navegación (GPS, GLONASS) permiten la obtención de la ubicación del sistema en cualquier punto del globo con una precisión lo suficientemente buena para satisfacer los requerimientos de tracking de posición. Por otro lado, los sensores inerciales y magnéticos, provistos en la mayoría de los dispositivos móviles modernos (*Tablets* o *SmartPhones*), proveen el mecanismo de obtención de la orientación.

Resultados y Objetivos

Visualización Aplicada a la Categorización Semiautomática de Composiciones Minerale.

Siguiendo la línea de este trabajo, en 2012, se presentó la aplicación *SpinelViz* [G*12] que consiste en un visor interactivo 3D para graficar y explorar Espinelos y que permite visualizar diferentes conjuntos de datos a la vez. Posteriormente, se desarrolló el *Spinel Explorer* [G*14] que integra los gráficos específicos dedicados a la exploración de los Espinelos con otros gráficos convencionales en un marco de análisis visual interactivo.

En 2015, se presentó una extensión del *Spinel Explorer* [G*15] que incorpora técnicas que permiten categorizar semiautomáticamente e interactivamente un conjunto arbitrario de Espinelos en función de los contornos de Barnes y Roeder. El sistema integra completamente los contornos de Barnes y Roeder y permite visualizarlos y contrastarlos con los datos

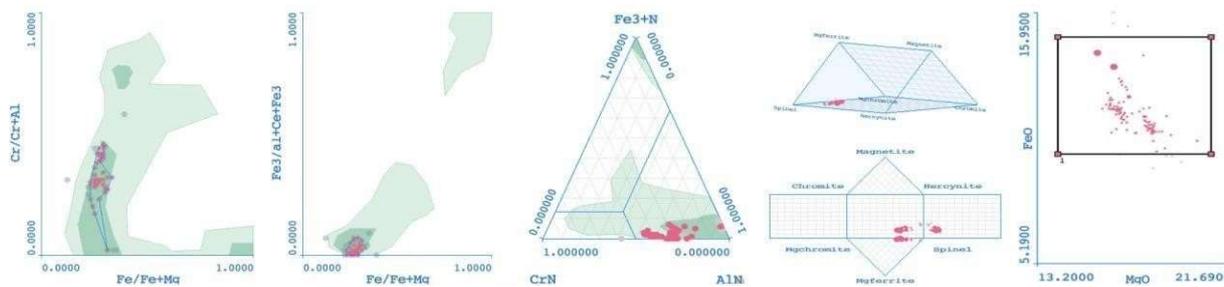


Figura 1: Captura de pantalla de una sesión de análisis en el Spinel Explorer [G*15]

del usuario de manera instantánea. Además, el sistema contrasta automáticamente los datos del usuario con todos los contornos de Barnes y Roeder y arroja una lista con los contornos que mejor se adaptan a los datos (ver Figura 1). Además, se implementaron interacciones específicas que permiten al usuario categorizar un conjunto de datos en un tiempo considerablemente menor al utilizado para esta tarea en el flujo de trabajo tradicional.

Actualmente se están incorporando al sistema nuevas funcionalidades que permiten generar contornos a partir de los datos del usuario y comparar estos nuevos contornos con los contornos de Barnes y Roeder [G*17].

Visualización de Datos Geológicos mediante RA y dispositivos móviles en el trabajo de campo.

En esta línea se está comenzando con el desarrollo de un *framework* destinado a la visualización *in situ* de información geológica. Este soporta la inclusión de terrenos en 3D que puedan ser superpuestos a la vista del mundo real [Fra08, Fed16]. De esta manera se ofrecerán soluciones a los requerimientos básicos de un sistema de RA, esto es, los subsistemas de tracking y registración, visualización e interacción.

Actualmente se logró unificar la visualización de una superficie 3D, generada a partir de un mapa de altura, con la vista del terreno real (ver Figura 2). Para el subsistema de tracking se utilizaron la tecnología GPS y GLONASS para determinar la posición y una fusión de sensores inerciales y magnéticos para la obtención de la orientación.

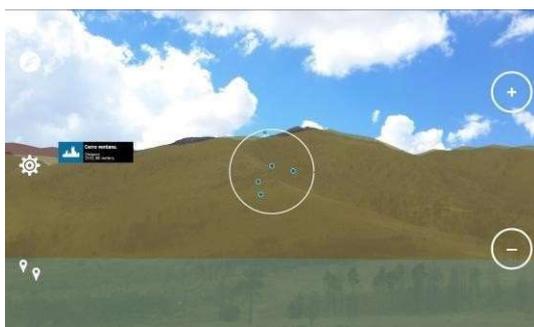


Figura 2: Visualización del terreno 3D y de un punto de interés.

En cuanto a la visualización del terreno se desarrolló una librería basada en OpenGL para aprovechar las capacidades provistas por las GPU integradas en los dispositivos móviles.

El objetivo es utilizar este *framework* para desarrollar una aplicación que asista al geólogo en el trabajo de campo. Pudiendo almacenar puntos de interés georeferenciados con la posibilidad de obtener información del contexto.

De esta manera se aportará una solución al problema de la ubicación relativa de los distintos puntos de interés respecto a la ubicación actual del usuario en el mundo.

Con estas nuevas funcionalidades el geólogo podría agilizar el trabajo de campo procesando datos in-situ, sin necesidad de realizarlo en un trabajo posterior de oficina.

Formación de Recursos Humanos

En lo concerniente a la formación de recursos humanos se detallan las tesis concluidas y en desarrollo relacionadas con la línea de investigación presentada:

Tesis en Desarrollo

- Tesis Doctoral. María Luján Ganuza. Tema: *Servicios Web en Visualización*. Dirección: Dra. Silvia Castro.
- Tesis Doctoral. Juan Manuel Trippel Nagel. Tema: *Realidad Aumentada Móvil en Exteriores para Visualización de datos Geológicos*. Dirección: Dra. Silvia Castro, Dr. Ernesto Bjerg.

Referencias

[B*16] Byška, J., Le Muzic, M., Gröller, M. E., Viola, I., & Kozlikova, B. (2016). AnimoAminoMiner: Exploration of Protein Tunnels and their Properties in Molecular Dynamics. *IEEE transactions on visualization and computer graphics*, 22(1), 747-756.

- [BR01] Barnes, S. J.; Roeder, P.L. 2001. The Range of Spinel Compositions in Terrestrial Mafic and Ultramafic Rocks. *Journal of Petrology*, vol. 42, number 12, pp: 2279-2302.
- [C93] Cleveland, W. S., 1993. *Visualizing Data*. Hobart Press. New Jersey, United States of America.
- [C*99] Card, S., Mackinlay, J., Shneiderman, B., *Readings in Information Visualization – Using Vision to Think*, M. K., 1999.
- [Fed16] Fedorov R., Frajberg D., Fraternali P. A Framework for Outdoor Mobile Augmented Reality and Its Application to Mountain Peak Detection. *AVR (1) 2016*: 281-301.
- [G*09] Ganuza, M. L.; Castro, S. M.; Martig, S. R.; Ferracutti, G.; Bjerg, E.; 2009. Mineral Compositions Visualization Implementig the Spinel Prism. *Congreso Argentino de Ciencias de la Computación Proceedings. CACIC 2009*, ISBN 978-897-24068-4-1, pp. 576-585. Jujuy, Argentina.
- [G*12] Ganuza, M. L., et al., SpinelViz: An interactive 3D application for visualizing spinel group minerals. *Computers & Geosciences*, Volume 48, November 2012, Pages 50–56
- [G*14] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S., Bjerg, E., Gröller, E., Matković, K., *The Spinel Explorer - Interactive Visual Analysis of Spinel Group Minerals*, *IEEE Trans. on Visualization and Computer Graphics* 20, 12 (2014), 1913-1922.
- [G*15] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S., Bjerg, E., Gröller, E., Matković, K., *Interactive Semi-Automatic Categorization for Spinel Group Minerals*, *IEEE VAST 2015*: 197-198.
- [G*17] Ganuza, M.L., Ferracutti, G., Gargiulo, M.F., Castro, S., Bjerg, E., Gröller, E., Matković, K., *Interactive Visual Categorization of Spinel Group Minerals. Enviado al SCCG2017 (33rd Spring Conference on Computer Graphics). En evaluación.*
- [K03] Koutek, M., 2003. *Scientific Visualization in Virtual Reality: Interaction Techniques and Application Development*. Computer Graphics & CAD/CAM group, Faculty of Information Technology and Systems (ITS), Delft University of Technology (TU Delft).
- [K*17] Krone, M., Frieß, F., Scharnowski, K., Reina, G., Fademrecht, S., Kulschewski, T. Ertl, T. 2017. *Molecular Surface Maps*. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 701-710.
- [L91] Lindsley, D. H. Ed. 1991. *Oxide Minerals: petrologic and magnetic significance*. Mineralogical Society of America, Departament of Geological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. vol. 25, p. 509.
- [M98] McCormick, B. H., 1998. *Visualization in scientific computing*. SIGBIO News. ACM, Vol.10, pp. 15-21.
- [NS79] Nielson, G. M.; Shriver, B.; Rosenblum, Lawrence. 1979. *Visualization in Scientific Computing*. IEEE Computer Society Press. United States of America.
- [R94] Roeder, P. L. 1994. Chromite: from the fiery rain of chondrules to the Kilauea Iki lava lake. *Canadian Mineralogist* 32, pp. 729-746.
- [Vea12] E. Veas, R. Grasset, E. Kruijff, D. Schmalstieg. *Extended Overview Techniques for Outdoor Augmented Reality*, *IEEE Trans. on Visualization and Computer Graphics (Proceedings Virtual Reality 2012)*, Vol. 18, pp. 1-12, April 2012.
- [Fra08] L. Frauciel, J. Vairon, P. Nehlig, P.Thierry, I. Zendjebil, F. Ababsa. *Outdoor Augmented Reality as a tool for bringing 3D geology to the field: the RAXENV project.*

Ingeniería de Software

Análisis de Enfoques de Aplicaciones para Dispositivos Móviles

Pablo Thomas, Lisandro Delia, Leonardo Corbalan, Germán Cáseres, Nicolás Galdamez, Alfonso Cuitiño, Juan Fernandez Sosa, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

526 e/ 10 y 11 La Plata Buenos Aires

{pthomas, ldelia, corbalan, gcaseres, ngaldamez, acuitino, jfernandez, ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con aspectos de Ingeniería de Software, orientados al análisis de los diversos enfoques de desarrollo de Aplicaciones para Dispositivos Móviles.

Palabras claves: Dispositivo Móvil - Plataformas para Dispositivos Móviles - Aplicaciones Nativas- Aplicaciones Híbridas – Aplicaciones Interpretadas – Aplicaciones por Compilación Cruzada

Contexto

Esta línea de Investigación forma parte del Proyecto (2014-2018) “*Tecnologías para Sistemas de Software Distribuidos. Calidad en Sistemas y Procesos. Escenarios Educativos mediados por TICS*”, en particular del subproyecto *Ingeniería de Software en el Desarrollo de Aplicaciones para Dispositivos Móviles* del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

Introducción

La computación móvil puede definirse como un entorno de cómputo con movilidad física. Un usuario debe ser capaz de acceder a datos, información u otros objetos lógicos desde cualquier dispositivo en cualquier red, al mismo tiempo que va moviéndose, cambiando su locación geográfica.

El desarrollo de aplicaciones para dispositivos móviles plantea nuevos desafíos originados en las características únicas de esta actividad. La necesidad de tratar con diversos estándares, protocolos y tecnologías de red; las capacidades limitadas, aunque en constante crecimiento, de los dispositivos; las restricciones de tiempo impuestas por un mercado altamente dinámico y la existencia de distintas plataformas de hardware y software son sólo algunas de las dificultades a las que se enfrentan los desarrolladores en esta área.

Las aplicaciones se generan en un entorno dinámico e incierto. En su mayoría se trata de aplicaciones pequeñas, no críticas, destinadas a un gran número de usuarios finales que son liberadas en versiones rápidas para poder satisfacer las demandas del mercado.

Todas las particularidades previamente mencionadas hacen que el desarrollo de software para dispositivos móviles difiera considerablemente del tradicional. Ello

conduce a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software como disciplina, acompañando este proceso de desarrollo tecnológico.

La existencia de una enorme competencia en el mercado de las aplicaciones móviles, promovida en parte por la proliferación de dispositivos móviles y el número de negocios que están migrando sus servicios a Internet, revela la importancia que implica para las empresas el posicionamiento en este mercado. Para conseguirlo, es necesario reducir al máximo el tiempo de desarrollo de las aplicaciones y al mismo tiempo hacer que éstas se ejecuten en el mayor número de dispositivos posible.

Este propósito se ve obstaculizado por la excesiva fragmentación de hardware y software existente, originada por el alto número de dispositivos distintos, con sus propios sistemas operativos y plataformas de desarrollo.

La alternativa más costosa para hacer frente a esta problemática consiste en el desarrollo nativo de las aplicaciones en cada una de las plataformas existentes, utilizando el entorno de desarrollo integrado (IDE por sus siglas en inglés), el lenguaje de programación y las herramientas propias de cada plataforma.

Las aplicaciones nativas así desarrolladas poseen un conjunto de características ventajosas entre las que sobresalen el acceso a todas las capacidades del dispositivo (cámara, GPS, acelerómetro y agenda, entre otras), el alto rendimiento, la posibilidad de trabajar sin acceso a Internet y de correr en segundo plano notificando al usuario sólo en caso de requerir su atención. Estas aplicaciones pueden distribuirse a través de las tiendas en línea correspondientes. Sin embargo, el precio de todas estas ventajas es alto: no es posible reusar el código fuente entre plataformas diferentes, el esfuerzo se multiplica y se elevan los costos de desarrollo, actualización y distribución de nuevas versiones.

El desarrollo multiplataforma se contrapone al nativo y se centra en la reutilización de código. Se procura entonces optimizar la relación costo/beneficio compartiendo la

misma codificación entre las versiones para las distintas plataformas. La construcción de aplicaciones Web Móviles constituye un ejemplo representativo de este enfoque. Estas aplicaciones se diseñan para correr dentro de un navegador, se desarrollan con tecnología web bien conocidas (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo; su puesta en marcha es rápida y sencilla.

Las desventajas de las aplicaciones Web Móviles recaen sobre su rendimiento. Los tiempos de respuesta se dilatan afectados por la interacción cliente-servidor y las restricciones de seguridad impuestas a la ejecución de código por medio del navegador limitan el acceso a todas las capacidades del dispositivo. Además, al no poseer el *look and feel* de las aplicaciones nativas, resultan menos atractivas para el usuario final.

Las aplicaciones híbridas constituyen otro tipo de desarrollo multiplataforma basado en tecnologías web (HTML, Javascript y CSS) pero que, a diferencia de las anteriores, no son ejecutadas por un navegador. En su lugar, corren en un contenedor web especial con mayor acceso a las capacidades del dispositivo a través de una API específica.

Las aplicaciones híbridas permiten la reutilización de código en las distintas plataformas, el acceso al hardware del dispositivo, y la distribución a través de las tiendas de aplicaciones. Sin embargo conservan algunas de las desventajas de las aplicaciones Web Móviles: la utilización de componentes no nativos en la interfaz perjudica la experiencia de usuario, y la ejecución se ve ralentizada por la carga asociada al contenedor web.

Otro tipo de aplicación multiplataforma lo constituyen las aplicaciones interpretadas, las cuales son traducidas en su mayor parte a código nativo, mientras que un resto se interpreta en ejecución. Se implementan de forma independiente de las plataformas utilizando diversas tecnologías y lenguajes, tales como Java, Ruby y XML, entre otros.

La obtención de interfaces nativas constituye una de las principales ventajas de

este tipo de aplicaciones, y la dependencia total con el entorno de desarrollo el obstáculo más notable. Appcelerator Titanium es el entorno de desarrollo más popular.

Finalmente, las aplicaciones generadas por compilación cruzada también constituyen un tipo de desarrollo multiplataforma. Estas aplicaciones se compilan de manera nativa creando una versión específica de alto rendimiento para cada plataforma destino. Ejemplos de entornos de desarrollo para generar aplicaciones por compilación cruzada son Applause, Embarcadero Delphi XE6 y Xamarin.

Líneas de Investigación y Desarrollo

- Enfoques de desarrollo de Aplicaciones para Dispositivos Móviles
- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Nativas.
- Aplicaciones Web Móviles.
- Aplicaciones Móviles Híbridas (PhoneGap, Sencha Touch, Ionic).
- Aplicaciones Móviles Interpretadas (Appcelerator Titanium, Native Script).
- Aplicaciones Móviles generadas por compilación cruzada (Xamarin, Applause, Embarcadero Delphi XE6).
- Análisis y estudio comparativo de rendimiento, consumo de energía, tamaño de software, entre otros, en los distintos enfoques de Aplicaciones Móviles.
- Experiencia de usuario en Aplicaciones Móviles generadas con distintos enfoques de desarrollo.

Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros involucrados en esta línea de investigación.
- Utilizar y comparar el uso de diversos frameworks para el desarrollo de aplicaciones móviles nativas y multiplataforma.
- Estudiar el uso de bluetooth en los diferentes enfoques de desarrollo.
- Analizar y comparar el impacto que distintos frameworks de desarrollo multiplataforma tienen sobre el rendimiento de las aplicaciones móviles construidas con ellos. Enfocar este análisis sobre la velocidad de procesamiento, el acceso a disco, el consumo de batería, tamaño de aplicación, administración de imágenes y otras capacidades; considerando además el tipo, gama y sistema operativo del dispositivo.
- Cuantificar las diferencias de rendimiento entre las aplicaciones multiplataforma y las aplicaciones nativas.
- Se ha realizado el estudio comparativo entre diferentes tipos de desarrollo de aplicaciones móviles a partir de la implementación de diversas aplicaciones con idéntica funcionalidad, que permiten interactuar con una plataforma de e-learning
- En relación con el análisis de rendimiento, consumo de batería y gestión de imágenes se han realizado un conjunto de pruebas. Los escenarios de prueba diseñados incluyeron los sistemas operativos Android y iOS ejecutándose sobre diversos dispositivos smartphones y tablets. Se ha medido el impacto de los parámetros de evaluación estudiados en los frameworks de desarrollo multiplataforma Apache Cordova, Native Scripts, Ionic, Titanium, y Xamarin.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de

Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

1. Hayes, I. S. *Just Enough Wireless Computing*. Prentice Hall Professional Technical Reference. 2002. ISBN:0130994618
2. Spyros Xanthopoulos, Stelios Xinogalos, *A Comparative Analysis of Cross-platform Development Approaches for Mobile Applications*, BCI' 2013, Greece
3. Tracy, K.W., *Mobile Application Development Experiences on Apple's iOS and Android OS*, Potentials, IEEE, 2012.
4. Delia L., Galdamez N., Thomas P, Pesado P., *Un Análisis Experimental de Tipo de Aplicaciones para Dispositivos Móviles*, CACIC 2013, CAECE Mar del Plata, Octubre 2013
5. <http://www.appcelerator.com/>
6. <https://github.com/applause/applause>
7. <https://www.embarcadero.com/es/products/delphi>
8. <http://xamarin.com/>
9. Abrahamsson, P. *Mobile software development -the business opportunity of today*. Proceedings of the International Conference on Software Development, (pp. 20-23). 2005. Reykjavik.
10. <http://devgirl.org/2012/12/04/easy-phonegap-push-notifications-with-pushwoosh/>
11. <http://phonegap.com/>
12. <http://jquerymobile.com/>
13. <http://backbonejs.org/>
14. Digital Possibilities. Mobile Development Frameworks Overview <http://digital-possibilities.com/mobile-development-frameworks-overview/>
15. Markus Falk. Mobile Frameworks Comparison Chart, <http://www.markus-falk.com/mobile-frameworks-comparison-chart/>
16. Anup Kumar y Bin Xie, *Handbook of Mobile Systems Applications and Services*. Editorial CRS Press, ISBN 978-1-4398-0152-9, Año 2012.
17. Sambasivan, D.; John, N.; Udayakumar, S.; Gupta, R., *Generic framework for mobile application development*, Internet (AH-ICI), 2011 Second Asian Himalayas International Conference on Computing & Processing (Hardware/Software).
18. Choi, Y.; Yang, J.-S.; Jeong, J., *Application framework for multi platform mobile application software development*, Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on Computing & Processing (Hardware/Software)
19. Anthony Wasserman , Carnegie Mellon Silicon Valley, *Software Engineering Issues for Mobile Application Development*, 2º Workshop on Software Engineering for Mobile Application Development, MobiCASE '11, Santa Monica, California, USA, October 2011.
20. Delia L., Galdamez N., Corbalan L., Thomas P, Pesado P., *Multi-Platform Mobile Application Development Analysis*, IEEE Ninth International Conference on Research Challenges in Information Science, May 2015, Athens, Greece, ISBN 978-1-4673-6630-4
21. Delia L., Galdamez N., Corbalan L., Thomas P, Pesado P., *Un Análisis comparativo de rendimiento en Aplicaciones Móviles Multiplataforma*, CACIC 2015, UNNOBA Junín, Octubre 2015.

Análisis de la Usabilidad de los Sitios Web de Universidades: Elaboración de un Conjunto de Buenas Prácticas Aplicables en su Diseño y Construcción

Olariaga, Sandra¹; Ligorria, Laura²; Di Gionantonio, Alejandra³;
Colacioppo, Nicolás⁴; Scandolo, Ivan⁵; Peralta, Lorena⁶;
Savorgnano, Marcela⁷; Reyna, Ramiro⁸; Manera, Roxana⁹

Laboratorio de Investigación de Software (Lis)
Departamento de Ingeniería en Sistemas de Información
Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC)
Maestro Marcelo Lopez y Cruz Roja Argentina – Ciudad Universitaria – Córdoba, Argentina

(1) solariaga@yahoo.com; (2) liuniversidad@gmail.com; (3) ing.alejandrag@gmail.com; (4) nicolas_colacioppo@hotmail.com; (5) ivanscand@gmail.com; (6) peralta.lorena.d@gmail.com; (7) marales511@outlook.com; (8) ramiroreyna90@gmail.com; (9) r_manera@hotmail.com

Resumen

Dentro de la Ingeniería de Software una de las disciplinas que la conforman es el Testing de Software, el cual se ocupa de la validación y verificación del mismo. Dentro de este campo se encuentra el testing de sitios web, en el cual se evalúan distintos aspectos de las páginas web a fin de determinar los aspectos favorables y desfavorables. Particularmente en el presente proyecto se procederá a testear la interfaz de usuario de sitios web universitarios, desde el momento en el que la página esté publicada, dicho proyecto será realizado en el Laboratorio de Investigación de software del Departamento de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba. Este trabajo tendrá una incidencia directa en los usuarios de sitios web universitarios, ya que como resultado del proceso de investigación se generará un conjunto de buenas prácticas en la construcción de portales universitarios, lo cual facilitará el acceso a los usuarios a dichos sitios web universitarios.

Palabras Clave: usabilidad, sitios web, evaluación

Contexto

El presente proyecto está siendo llevado a cabo en el Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de la Universidad Tecnológica Nacional Facultad Regional Córdoba. (U.T.N.-F.R.C.).

El Código del proyecto es UTN4082 dentro del área de Ciencias de la Computación e Informática, radicado dentro de la Secretaria de Ciencia y Tecnología de la Universidad Tecnológica Nacional Facultad Regional Córdoba. Fecha de inicio: 01/05/2016 Fecha de finalización: 30/04/2018.

1 Introducción

En los últimos años se han desarrollado una gran cantidad de recursos web, para la divulgación de información, comercio, educación e incluso interacciones personales, ya que les ofrece la posibilidad de contar con diferentes niveles de información dependiendo de las necesidades del usuario. En este entorno web cabe destacar la importancia de la usabilidad en la interacción cada vez más

creciente entre el hombre y la máquina.

"La usabilidad trata sobre el comportamiento humano; reconoce que el humano es emotivo, no está interesado en poner demasiado esfuerzo en algo, y generalmente prefiere las cosas que son fáciles de hacer contra las que son difíciles de hacer." [1].

Usabilidad es un término adaptado de la palabra en inglés "usability", para indicar que algo se puede usar.

Según Jakob Nielsen [2], la usabilidad es un atributo de calidad. Evalúa qué tan fáciles de usar son las interfaces de usuario.

Si a esto lo trasladamos a Internet, se puede decir que un sitio web es usable si un usuario, al finalizar la navegación, se ha manifestado satisfecho por el cumplimiento de sus objetivos iniciales. [3]

En referencia a esto, la norma ISO 9241-11, expresa que la usabilidad se refiere al alcance en el que un producto puede ser utilizado por usuarios específicos para alcanzar metas específicas con efectividad, eficiencia y satisfacción en un contexto específico de uso [4].

Por otro lado, se debe tener en cuenta que no hay que hacer pensar al usuario en su proceso de interacción con una página web, sino que ésta sea tan clara, obvia y fácil de entender, que implique la acción directa y no la pérdida de tiempo en su intento de exploración y comprensión de lo que se le ofrece por pantalla [5].

En relación a lo anterior, y teniendo en cuenta la diversidad de actividades y acciones que un usuario puede y debe realizar al interactuar con

un sitio web, particularmente si éste pertenece al ámbito educativo, es que se hace necesario su validación y verificación en cada paso de su diseño y construcción, ya que hoy en día, estos sitios son la principal fuente de información de las actividades y ofertas educativas que las universidades e instituciones nos ofrecen, y debemos interactuar con ellos al estar insertos en nuestras actividades académicas.

2 Líneas de Investigación, Desarrollo e Innovación

El proyecto se inscribe dentro de los lineamientos de investigación en innovaciones curriculares en Educación Superior Universitaria dentro del área de Sistemas e interfaces adaptables en la intervención humano-computadora.

Para el desarrollo de este trabajo de investigación se aplicará el método empírico-analítico, que se basa en la experimentación y en la lógica empírica, junto a la observación de sitios web universitarios y análisis de usabilidad de los mismos.

Primer paso: Realizar una investigación exploratoria acerca de los estándares y normas establecidas sobre la usabilidad aplicada a los sitios web educativos.

Segundo paso: Analizar y definir el perfil del usuario a ser considerado, estableciendo las características inherentes al mismo. Definir también los aspectos deseables de usabilidad para los sitios web universitarios.

Tercer paso: Seleccionar los sitios web universitarios nacionales a ser evaluados.

Cuarto paso: Realizar un análisis exhaustivo a fin de determinar y definir las herramientas más

adecuadas, para llevar a cabo la evaluación de la usabilidad de los sitios web universitarios nacionales.

Quinto paso: Se procederá a la evaluación de los sitios web universitarios seleccionados con la herramienta definida desde el perfil de usuario establecido.

Sexto paso: Elaborar los informes correspondientes con las conclusiones a las que se ha arribado.

Ya hemos definido el perfil del usuario con el que vamos a trabajar y hemos realizado un relevamiento de las normas y estándares existentes. También hemos seleccionado los sitios web universitarios nacionales a ser evaluados.

3 Resultados y Objetivos

Objetivo Principal

Elaborar un conjunto de recomendaciones de buenas prácticas al diseñar una página web universitaria, a partir de la evaluación de la usabilidad de sitios web de universidades nacionales que ya están publicadas en internet por las instituciones correspondientes, teniendo en cuenta el punto de vista del usuario.

Objetivos específicos

* Investigar los estándares y normas establecidas a fin de obtener los parámetros a considerar, para realizar la evaluación de la usabilidad de los sitios web universitarios

* Definir el perfil del usuario, es decir quién va a ser el sujeto desde el punto de vista del que va a ser realizada la evaluación, es decir, establecer las características que tendrá el

sujeto que va a hacer uso del sitio web y cuyo punto de vista será objeto importante de esta evaluación.

* Determinar las herramientas a ser utilizadas y los sitios web universitarios a ser evaluados.

* Llevar a cabo la evaluación de la interacción de los usuarios con las páginas universitarias seleccionadas a fin de determinar si cumplen con los parámetros de usabilidad que se establezcan, tomando como base las medidas de usabilidad propuestas por Jacob Nielsen. (Heurísticas) [2].

* Elaborar informes acerca de cuáles son los aspectos a ser considerados a fin de que la página web de los sitios universitarios ofrezca un nivel óptimo de usabilidad.

Nuestro trabajo se viene desarrollando a través de varias tareas, algunas de ellas se han realizado simultáneamente, dando el punto de partida a otras. En consecuencia, ya se han efectuado tareas recopilación de información sobre normas y estándares, lo cual permitió obtener los parámetros a considerar, para realizar la evaluación de la usabilidad de los sitios web universitarios. También se ha definido el perfil del usuario, es decir, se han establecido las características que tendrá el sujeto que va a hacer uso del sitio web y cuyo punto de vista será objeto importante de estudio. De igual manera hemos procedido a seleccionar las universidades nacionales a ser evaluadas.

Actualmente el equipo de docentes abocados al presente proyecto se encuentra realizando una selección de recursos, técnicas y estrategias para determinar y confeccionar la herramienta más adecuada para ser utilizada en este proyecto. Para ello hemos definido dimensiones y subdimensiones, las que se muestran a

continuación. A partir de ellas se desarrollará la intervención con los usuarios de los sitios a evaluar para conocer sus puntos de vista, interacción, opinión y valoración de los mismos y extraer a partir de ello una serie de conclusiones que nos permitirá elaborar el compendio de buenas prácticas que constituye el objetivo principal del proyecto.

Dimensiones y subdimensiones a analizar

Dimensión	Subdimensión	Items
Aspectos generales	Claridad	El sitio web tiene una dirección clara? Muestra cuáles son los contenidos principales que ofrece el sitio web? El tamaño de la letra es el adecuado? Es posible modificar el tamaño de la letra desde el sitio web? Hay un botón para imprimir la página web?
	Sencillez	Es simple de comprender la estructura del sitio web?
	Versatilidad	Contiene los links suficientes?
	Medios de transporte	Presenta mapa de la ubicación geográfica? Brinda información clara acerca de su ubicación geográfica y cómo llegar?
Identidad	Claridad	Puede leer y entender claramente el nombre de la institución? Brinda información acerca de las autoridades que representan la institución?
	Medios de contacto	Se especifican dirección de mail y números de teléfonos de la institución?
	Logotipo	Es representativo de la institución?
Orientación para Potenciales ingresantes	Becas	Se brinda información acerca de becas y jornadas de información sobre el ingreso a la institución.?
	Tramites administrativos	Explica los requisitos necesarios para inscribirse.?
	Volumen de información	Se ha evitado la sobrecarga informativa.?
	Tutorías	Se menciona si existen tutorías?
Navegabilidad	Mapa del sitio	Marca el punto donde se encuentra dentro del sitio web?
	Sentido de avance	Podemos avanzar y retroceder dentro de la página sin usar los botones del navegador? Muestra los enlaces visitados y aquellos que no ha visitado aún? ¿Se ha sentido perdido dentro de la estructura de la página web.?
	Sencillez	Existen elementos dentro del sitio web que puedan distraerlo? Hay que dar demasiados clicks para llegar a obtener la información deseada.?
	Búsqueda	El resultado de la búsqueda es útil? Existe un botón de búsqueda dentro del sitio web?
	Manejo de errores	Hay enlaces que no lleven a ningún lado.?
	Multimedia	Tamaño de imágenes
Tipo de imágenes	Son claras y sencillas? El uso de imágenes o animaciones proporciona algún tipo de valor añadido? Se han incluido animaciones cíclicas?	

4 Contribuciones al avance científico, tecnológico, transferencia al medio

Nuestro trabajo de investigación aportará una serie de recomendaciones claras y precisas que actúen como un compendio de buenas prácticas a seguir al momento de realizar un diseño de

sitios web universitarios centrado en el usuario. Lo cual posibilitará detectar las debilidades y fortalezas que los mismos presentan, en relación con las variables de información sobre las universidades, brindando una orientación adecuada para ingresantes en el manejo del sitio.

Permitirá también, identificar el nivel de cobertura alcanzado en cada uno de los parámetros de los sitios web estudiados.

También será de utilidad a los fines de ayudar a la modificación del diseño e implementación de los sitios web universitarios, mediante la aplicación del análisis basado en los aspectos definidos y su complementación con los test de usuarios obtenidos, ya que algunas interfaces de sitios web universitarios son complejas en su utilización, a gran cantidad de usuarios les sería dificultoso navegar e interactuar adecuadamente con estos sitios.

5 Contribuciones a la formación de recursos humanos

El desarrollo del Proyecto que engloba este trabajo, tendrá un alto impacto en la formación de RRHH en el área de desarrollo web mediante la determinación de pautas para lograr la eficiencia y la optimización de la usabilidad del diseño de los sitios web universitarios. Los investigadores que recién inician forjarán sus primeras herramientas en I+D lo que les permitirá ampliar conocimientos en el área en cuestión además de incrementar sus antecedentes académicos. El proyecto contribuirá también a consolidar los conocimientos de dos integrantes del grupo, uno que está realizando la Diplomatura en Testing; y a fortalecer la formación de un becario, estudiante de Ingeniería en Sistemas de Información. A su vez induce a la interacción de trabajo en conjunto de investigadores de

diferentes ámbitos y departamentos académicos de la Facultad Regional Córdoba tales como el Departamento de Ingeniería en Sistemas de Información y el Departamento de Ciencias Básicas que unirán sus esfuerzos logrando un intercambio de experiencias y conocimientos que enriquecerán su saber.

6 Bibliografía

- [1] McQuillen, D., “Taking Usability Offline”, *Darwin Magazine*, June 2003.
- [2] Nielsen, J., Loranger, H., Gallud Jurado, E., Usabilidad. Prioridad en el diseño Web. Anaya Multimedia-Anaya Interactiva, 1ª ed. Madrid, 2006.
- [3] Morales Martínez, M., Analítica web para empresas. Arte, ingenio y anticipación. Editorial UOC, 1º ed. Barcelona, 2010.
- [4] Martínez de la Teja, G. “Usabilidad y Accesibilidad en WEB”. Sociedad de Ergonomistas de México. Universidad. Fecha de consulta: 01-08-2016. Disponible en <http://www.semec.org.mx/archivos/6-11.pdf>
- [5] Krug, S., No me hagas pensar: Una aproximación a la usabilidad en la web, 2ª ed. Madrid: Pearson Educación, 2006.

Revista Edutec-e Revista Electrónica de Tecnología Educativa Nro. 37 Septiembre 2011. Evaluación de la usabilidad de un sitio web educativo y de promoción de la salud en el contexto universitario.

Análisis del Modelo WQM para Métricas de Usabilidad Web

Jesús Aguirre, Daniel Riesco, Germán Montejano

Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
jaguirre,driesco,gmonte@unsl.edu.ar

Resumen

La usabilidad es un aspecto muy importante a considerar durante el desarrollo de software, por lo tanto es necesario evaluarla para realizar las modificaciones pertinentes y lograr que el software cumpla con las recomendaciones del estándar ISO-IEC 9126. Los Sitios Web (SW) deben determinar criterios de usabilidad en el momento de analizar, diseñar, desarrollar, administrar y elaborar contenidos para la Web. De esta manera, las métricas analizadas pueden garantizar un nivel de usabilidad con calidad para el producto final. El presente trabajo recopila métricas del modelo de calidad Web WQM y analiza los atributos de calidad de las mismas para que los potenciales clientes accedan al SW y puedan satisfacer las necesidades del usuario: efectividad, seguridad, capacidad de aprendizaje, eficiencia y memorabilidad. Con esto se busca elaborar un referente para guiar a los desarrolladores Web en la medición de la usabilidad a partir de valores propuestos por estándares Web o por diferentes autores del diseño Web.

Palabras clave: usabilidad Web, modelo de calidad, modelo WQM, estándares, analítica web

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación denominado: “Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad”, de la Facultad de Ciencias Físico Matemáticas y Naturales- Universidad Nacional de San Luis (Proyecto PROICO N.º 03-1516- Director: Daniel Riesco). Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos/empresas que contribuyen a su financiamiento.

Introducción

El objetivo de la Ingeniería Web (IW) es proveer de metodologías para el desarrollo de SW con calidad del producto y del proceso. Entre las características técnicas de calidad más

relevantes se encuentran: usabilidad, funcionalidad, confiabilidad y eficiencia [1].

La evaluación de la calidad de un producto de software tiene un rol fundamental, ya sea durante su Proceso Desarrollo de Software (PDS) o en la selección de un producto ya desarrollado. El estándar ISO/IEC 9126-1 identifica 6 atributos clave de la calidad: funcionalidad, confiabilidad, usabilidad, eficiencia, facilidad de recibir mantenimiento y portabilidad [2]. La norma ISO/IEC 9126-1 define a la usabilidad como uno de los atributos de calidad de software más relevante. En ella establece a la usabilidad como “la capacidad en que un producto de software puede ser entendido, aprendido y usado por determinados usuarios bajo ciertas condiciones en un contexto de uso específico” [2].

Como la Usabilidad Web (UW) es uno de los factores de calidad que más se persiguen, existe en la actualidad diferentes técnicas para evaluarla en forma empírica y no por el sentido común o experiencia de los desarrolladores Web. La UW es un aspecto de calidad que creció a través de las experiencias de usuario y especialistas en diseño visual, entre otros factores. Muchos desarrolladores no se preocupan para que los potenciales clientes accedan a la Web con mayor facilidad y accesibilidad, ya que no creen que desarrollar SW con recomendaciones o pautas estandarizadas beneficia a la organización y a los clientes, porque suponen invertir más tiempo y recursos [3, 4].

Un estándar Web es un conjunto de reglas normalizadas definidas por organizaciones independientes que desarrollan especificaciones técnicas y directrices que deben ser cumplidos por un servicio, producto o proceso. El objetivo principal es el de crear una Web

interoperable, accesible y eficiente. A través del uso de estas tecnologías abiertas y de uso libres, se pretende lograr una evolución de la Web con alta calidad [3].

Los modelos de calidad del software, proporcionan una base para la especificación de requisitos de calidad y para la evaluación de la misma. Son un conjunto de buenas prácticas utilizadas por las organizaciones en la mejora continua del producto y para asegurar la calidad. En general, los modelos descomponen la calidad jerárquicamente en una serie de características y subcaracterísticas tomando como base la norma ISO/IEC 9126-1. El Modelo de Calidad Web (del inglés Web Quality Model) distingue tres dimensiones: características de calidad (funcionalidad, confiabilidad, usabilidad, eficiencia, portabilidad y mantenibilidad), procesos de ciclo de vida (desarrollo, operación, mantenimiento, esfuerzo y reuso) y características Web (contenido, presentación y navegación) (ver figura 1) [5]. El modelo WQM presenta un 48% de métricas de UW, este alto número confirma la tendencia de satisfacer al usuario desarrollando SW usables con calidad.

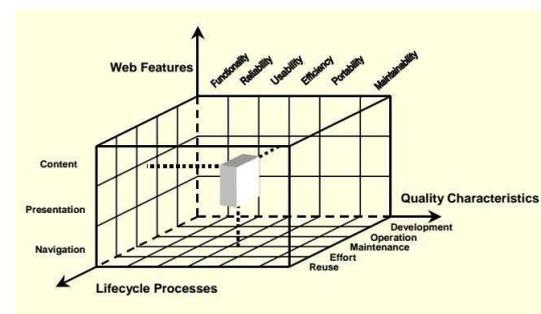


Figura 1: El modelo WQ

La Analítica Web (AW) es una disciplina creciente que permite tomar decisiones objetivas y estratégicas basadas en información relevante. La AW

es esencial porque nos revela información sobre identificar las preferencias de los usuarios, lo que realizan los clientes en el SW, mejorar el posicionamiento en los buscadores, hacer un diseño web más usable, optimizar las ventas de los productos en línea, entre otros factores [6, 7].

Para medir la UW es necesario definir un conjunto de métricas adecuadas, para luego utilizar la información que revelan y realizar los cambios pertinentes para modificar las fallas detectadas. Generalmente, la UW se la ha evaluado con diferentes técnicas y métodos en etapas finales del PDS, lo que implica altos costos en la modificación de las fallas detectadas. Un visión más actual, es evaluar la UW durante las etapas de análisis, desarrollo e implementación del SW.

A partir del modelo tridimensional de calidad WQM, se analizaron las métricas con atributos de UW, se investigaron los valores sugeridos por los estándares Web [1, 2] y por diferentes autores para métricas de texto, formato de texto, legibilidad lingüística, enlaces, imágenes, formato de páginas web, accesibilidad, diseño web, y navegación [8, 9, 10, 11].

Líneas de Investigación, Desarrollo e Innovación

El grupo de trabajo investiga, básicamente en los siguientes campos relacionados a la UW:

- clasificación de las métricas del modelo WQM en 9 categorías: texto, formato de texto, legibilidad lingüística, enlaces, imágenes, formato de páginas web, accesibilidad, diseño web, y navegación [5],

- Análisis de métricas y definición de criterios para evaluar los atributos de la UW según las recomendaciones o normas estandarizadas [1,2],
- investigar valores sugeridos por diferentes autores de la temática para las diferentes categorías [8, 9, 10, 11].

Resultados y Objetivos

La presente línea de investigación está orientada a profundizar en la medición de la UW a través de la realización de las siguientes tareas :

- estudio y análisis de diferentes métricas para la evaluación de usabilidad en PDS Web,
- análisis de las normas y recomendaciones estandarizadas para incorporar en la UW,
- análisis de sugerencias de diferentes autores para integrar a la UW.

A futuro, se pretende llevar a cabo las siguientes acciones de investigación:

- actualizar el modelo WQM con métricas que incorporen los cambios tecnológicos actuales,
- adaptar e incorporar métricas para evaluar UW en diseños responsivos para múltiples dispositivos,
- identificar métricas para evaluar la UW en entornos MDA,
- expresar las anteriores métricas en OCL.

Formación de Recursos Humanos

En el marco del presente proyecto se presenta información de las tesis dirigidas por integrantes del proyecto,

finalizadas o en las últimas etapas de su elaboración durante el año 2016, según el siguiente detalle: 15 tesis de maestría y doctorado en colaboración con distintas universidades internacionales: la Universidad Politécnica de Valencia, la Universidad Castilla-La Mancha, la Universidad Minho, la Universidad Estatal de Minas Gerais (proyecto financiado por la SPU y CAPES).

Referencias

- [1] Presman, R., “Ingeniería del Software: un enfoque práctico”, McGRAW-HILL INTERAMERICANA EDITORES, 7 Ed., 2010.
- [2] ISO/IEC 9126-1, "Software Engineering -Product Quality - Part 1", 2001.
- [3] World Wide Web Consortium, W3C, <http://www.w3c.es/estandares>
- [4] Web Accessibility Initiative, WAI, <https://www.w3.org/WAI/>
- [5] Calero, C., Ruiz, J., & Piattini, M., “Classifying web metrics using the web quality model”, 2005, Online Information Review, 29.
- [6] Krug S.; “No me hagas pensar. Una aproximación a la usabilidad en la Web”, segunda edición, Prentice Hall, 2005
- [7] Nielsen J., Pernice K., Whitenon K., “How People Read on the Web. The Eyetracking Evidence”, Nielsen Norman Group
- [8] Felder L., “Writing for the Web. Creating Compelling Web Content Using Words, Pictures and Sound”, New Rider , 2012.
- [9] Lynch P., & Horton S., “Web Style Guide. Foundations of User Experience Design”, 4th edition, Yale University Press, 2016.
- [10] Redish J., “Letting Go of the Words: Writing Web Content that Works”, Morgan Kaufmann, 2012
- [11] Lynch P., Horton S., Web Style Guide. Foundations of User Experience Design, 4th edition, Yale University Press, 2016

Aplicación de la Lógica Difusa en la Evaluación de Modelos Conceptuales de Procesos de Negocio

Carlos Salgado, Mario Peralta, Daniel Riesco, Lorena Baigorria, Germán Montejano,

Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

e-mail: {csalgado, mperalta, driesco, flbaigor, gmonte}@unsl.edu.ar

Resumen

Los complejos problemas empresariales, caracterizados por la presencia de conflictos de valores, requieren la inclusión de algún tipo de proceso de toma de decisiones para tratar las múltiples y frecuentes perspectivas opuestas en los distintos modelos de procesos de negocio. El modelado de procesos de negocio presenta una visión global de la organización que permite entender mejor la dinámica de la empresa y las relaciones que se dan en su interior y con su entorno. El modelado del negocio es la técnica por excelencia para alinear los desarrollos con las metas y objetivos de las organizaciones. Los modelos cumplen un rol fundamental en la especificación de los Procesos de Negocio (PN). Por ello, es de vital importancia la calidad de los mismos para que ayuden a mejorar el desempeño y evolución de la organización. Acorde a ello, se propone un método, basado en la Lógica Difusa (LD), para evaluar modelos conceptuales de PN. El método propuesto permite que un grupo de evaluadores emita su opinión de forma independiente con términos lingüísticos comprensibles para cada usuario/actor que interviene en el proceso de evaluación/toma de decisiones. Ello posibilita la interacción de forma más comprensible entre los distintos miembros que intervienen en el modelado de los PN.

Palabras clave: Lógica Difusa, Procesos de Negocio, Modelado de

Procesos de Negocio, Modelos Conceptuales.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Introducción

El desarrollo de modelos de PN constituye una de las tareas claves en las primeras etapas del ciclo de vida de los PN. Los modelos son utilizados como medios para que los participantes puedan entender fácilmente los procesos que representan. Además, son empleados como punto de partida a la hora de realizar cambios y adaptaciones de los PN a las nuevas necesidades de las empresas. Por ello, es un factor primordial que estos modelos sean de alta calidad.

Al hablar de calidad en el modelado conceptual, se debe distinguir entre la

calidad del producto y la calidad del proceso [1]. Moody en [2], establece que es fundamental que toda propuesta de evaluación adhiera a estándares aceptados y aplicados y propone que deberían ser consistente con las normas de calidad ISO 9000 [3], e ISO/IEC 9126 [4], ya que un modelo conceptual es un tipo particular de producto. En este trabajo se consideró la norma actual de producto de software: ISO 25000 [5]. En base a ella, se tuvieron en cuenta las características externas, principalmente la entendibilidad, la mantenibilidad y la modificabilidad de los modelos, e internas, principalmente acoplamiento, cohesión y tamaño de los modelos.

La complejidad de un modelo conceptual puede estar influenciada por los elementos que lo componen (tareas, subprocesos, participantes, eventos, etc.). Por lo tanto, no es aconsejable definir una medida general para su complejidad [6]. Rolón en [7] propone un conjunto de medidas para la calidad de modelos conceptuales de PN desarrollados en BPMN. Estas medidas se basan en la propuesta de García Rubio de medidas para la calidad de proceso software [8]. Desde este punto de vista, tener métodos que permitan medir la calidad de dichos modelos será de gran ayuda en cuanto a la administración, difusión y mantenimiento de los PN que ellos representan. Desde este punto de vista, el proceso de evaluación de requerimientos de calidad de los modelos conceptuales de PN es de suma importancia. Por lo tanto, será de gran utilidad contar con un método cuantitativo para la evaluación y comparación de las características deseables de todo modelo que se apoye en los principios y prácticas de la ingeniería de software. En este sentido, en [9], presentamos un método para la evaluación de modelos de PN sin importar su representación. El método propuesto en [9] permite evaluar las principales características de calidad que

se considera que todo modelo de PN debe satisfacer. Sin embargo, y como argumentan Huang y Wu en [10], “la condición difusa de los pensamientos humanos puede ejercer cierta influencia en la respuesta de los expertos sobre sus preferencias con respecto a los factores, criterios y alternativas, y de sus cálculos y juicios subjetivos”. Respecto del modelado de PN, esta condición difusa del pensamiento humano, ejerce una influencia muy importante en cuanto a la percepción de las reglas del negocio, y por ende en la definición y construcción de los modelos que representan dichas reglas. Desde este punto de vista, el método no permite evaluar con certeza estas características difusas del razonamiento que pueden llevar a ambigüedades en los modelos.

Muchos problemas de decisión se sitúan en un entorno complejo e implican sistemas contradictorios de criterios, incertidumbre e información imprecisa. Se han propuesto numerosos métodos para resolver problemas de criterios cuando la información disponible es precisa. Sin embargo, la incertidumbre y la incertidumbre inherentes a la estructura de la información hacen que el modelo matemático riguroso resulte adecuado para resolver problemas multicriterios con información imprecisa [11, 12, 13, 14].

La utilización de la LD en la evaluación de los modelos de PN, permite evaluarlos considerando dicha condición difusa, lo que dará una mejor valoración de la calidad de los modelos, y su aproximación a la realidad que representan.

Líneas de Investigación, Desarrollo e Innovación

En base a lo expresado previamente, en esta línea de investigación se trabaja en el estudio, definición y uso de distintos métodos de evaluación para diversos sistemas del mundo real. En

particular, se propone un método para evaluar modelos conceptuales de PN basado en la LD. El objetivo es brindar un medio que ayude en la toma de decisión a la hora de evaluar la calidad de los modelos de PN.

El método propuesto brinda ciertas ventajas al permitir hacer un control difuso. Un sistema de LD utiliza cualquier tipo de información y la procesa de manera similar que el pensamiento humano; por ello, los sistemas de LD son adecuados para tratar información cualitativa, inexacta e incierta. Permiten, además, tratar con procesos complejos, lo que la hace una alternativa interesante para modelar problemas de toma de decisiones.

El control difuso permite operar con conceptos vagos o ambiguos propios del razonamiento humano cualitativo, fundado sobre un soporte matemático que permite extraer conclusiones cuantitativas a partir de un conjunto de observaciones (premisas) y reglas cualitativas (base de conocimiento).

Cuando se cuenta con información imprecisa e insuficiente, usar instrumentos estadísticos no es suficiente para obtener resultados significativos. La LD surge precisamente para tratar con este tipo de problemas y lograr darles una solución óptima. De esta forma, una combinación entre un sistema de LD y la experiencia o conocimiento que tienen los encargados de tomar las decisiones es una excelente manera de obtener buenos resultados [15].

En el proceso de desarrollo de modelos de PN, muchas veces la información acerca de las reglas de negocio que se deben representar y modelar, suelen ser imprecisas o insuficientes, lo que lleva a modelos imprecisos. El uso de la LD en la evaluación de dichos modelos, permitirá, a través de los mecanismos que provee dicha lógica, evaluar aquellas situaciones imprecisas, e incluso

ambiguas, producidas en la construcción de los modelos.

Resultados Obtenidos y Objetivos

En base a lo expresado, se definió un método para el análisis y evaluación de los modelos conceptuales de PN basado en la LD. El método propuesto se divide en 5 fases bien diferenciadas, comenzando con el establecimiento de los requerimientos de calidad a evaluar. Para hacer este estudio se tienen en cuenta las buenas prácticas de modelado de PN. Las buenas prácticas consideradas para el trabajo son tan amplias para modelado independiente del lenguaje, como también especifica un lenguaje como BPMN.

En base a las buenas prácticas que surgen de estudios científicos, expertos del dominio y personal técnico de las organizaciones, se definen funciones de pertenencia que permitirán tomar decisiones en cuanto a la entendibilidad y mantenibilidad de los modelos de PN.

Para cada una de las variables que se desean estudiar en los modelos se define una función de pertenencia. Para definir dichas funciones se tuvo en cuenta que existen buenas prácticas generales que trascienden los lenguajes en las que sean aplicadas y que tienden a mejorar la comprensión de los modelos, así como a reducir los errores que se deriven del modelado. Por ejemplo, 7PMG: *Seven Process Modeling Guidelines* [16], son guías de modelado definidas a partir de evidencia empírica en el modelado de procesos.

El uso del estándar BPMN proporciona a las organizaciones la capacidad de comprender y comunicar sus PN internos de manera estándar en una notación gráfica. Sin embargo, el uso de dicho estándar, no garantiza que los procesos se modelen de forma clara y eficaz. La forma en que los modeladores interpretan las condiciones de negocio y

cómo definen su estructura, es crucial para asegurar que se entienden correctamente. Por ello, y en base a estas guías y buenas prácticas, se definen las reglas borrosas que especifican el vínculo entre las variables de entrada y salida del sistema.

La interpretación de una regla si-entonces involucra dos pasos: (1) evaluar el antecedente mediante la aplicación de cualquier operador difuso y (2) implicar o aplicar el resultado del antecedente al consecuente. Esto se hace evaluando la función de pertenencia. Es decir, se trata de evaluar la activación de una regla en función del grado de cumplimiento del antecedente. Para realizar dicha tarea se usan operadores de composición de conjuntos difusos y se aplica un sistema de inferencia. En esta etapa del proceso, las salidas de cada una de las reglas se combinan para obtener un único conjunto difuso. Las entradas del proceso de agregación son las funciones de pertenencia truncadas obtenidas de la etapa de inferencia para cada una de las n-reglas.

Las operaciones básicas realizadas con conjuntos difusos se llevan a cabo mediante la aplicación de algún operador binario clasificado como T-normas (para operaciones de intersección) o S-normas (para operaciones de unión) [17, 18, 19, 20, 21]. En esta etapa se obtiene un valor nítido o concreto a partir del conjunto difuso de salida, el cual proporciona la solución del sistema planteado.

En la fase final del método, se debe realizar un análisis y comparación de los resultados obtenidos en la evaluación de los modelos respecto de las preferencias de los usuarios, obtenidas en la aplicación del método. Además, se debe documentar el proceso de evaluación y los resultados obtenidos, de manera que dicha documentación sirva como referencia e historial de la evolución de los modelos de proceso de negocio estudiados en futuras evaluaciones de dichos modelos. Esta documentación

puede servir como punto de referencia y comparación a la hora de evaluar nuevos modelos y procesos de negocio. Esta fase trata con actividades de análisis y comparación de las preferencias de calidad y los resultados obtenidos. A partir de las metas establecidas y el punto de vista de los interesados en los modelos y procesos de negocio a evaluar, esta etapa culmina con las conclusiones y recomendaciones del caso.

Esta etapa es una de las actividades más relevantes del método. Por ello, es de suma utilidad tener la información recopilada durante la aplicación del método volcada en estructuras y representaciones que sean claras de leer e interpretar. Desde esta perspectiva, se propone un formulario tipo que debería llenarse una vez realizada la evaluación de los modelos. Dicho formulario permite, entre otras cosas, tener presente qué funciones de pertenencia se utilizaron; si fueron definidas por el grupo evaluador o si se utilizaron otras definidas y almacenadas en un repositorio previamente. Además, se registran datos de los modelos, de los evaluadores, y si existen evaluaciones previas se incluye una referencia a ellas.

Se ha evaluado a través del método propuesto los mismos modelos que en [9]. Se comenzó por el proceso de compra y ventas de la empresa. Luego se realizó una comparación de los resultados. El análisis de los formularios de evaluación de ambos métodos permite visualizar que el método propuesto utilizando LD presenta un lenguaje adecuado y entendible para los distintos usuarios/actores que intervienen en el uso de los procesos de negocio.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación se presentó un trabajo final de Especialización en Ingeniería de

Software, [22], y una tesis de Maestría en ingeniería de Software referente al Modelado de Procesos de Negocio [23]. Actualmente, se está trabajando en una tesis doctoral en la que se combina el análisis de calidad de modelos de PN con la LD. Además, se está desarrollando una tesis de maestría en lo referente, como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación.

Referencias

- [1] M. Piattini, F. Ó. García Rubio, and I. Caballero, *Calidad de Sistemas Informáticos*: Alfaomega-RA-MA, 2007.
- [2] D. Moody, "Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions," *Data & Knowledge Engineering. Elsevier B.V.* 243–276, 2005.
- [3] ISO, "ISO Standard 9000-2000: Quality Management Systems: Fundamentals and Vocabulary, International Standards Organisation (ISO)." 2000.
- [4] ISO/IEC, "ISO/IEC Standard 9126: Software Product Quality, International Standards Organisation (ISO), International Electrotechnical Commission (IEC)," 2001.
- [5] ISO, "ISO/IEC 25000," <http://iso25000.com/index.php/en/>.
- [6] Fenton, "Software Measurement: A Necessary Scientific Basis," *IEEE Transactions on Software Engineering.* 20(3), pp. 199-206, 1994.
- [7] E. Rolon, F. Ruiz, F. Ó. García Rubio, and M. Piattini, "Aplicación de Métricas Software en la Evaluación de Modelos de Procesos de Negocio," *Revista Electrónica de la Sociedad Chilena de Ciencia de la Computación*, 2005.
- [8] F. Ó. García Rubio, "FMESP: Marco de Trabajo Integrado para el Modelado y la Medición de los Procesos Software," Doctoral, Departamento de Informática, U.C.L.M. Ciudad Real. España, 2004.
- [9] N. Debnath, C. Salgado, M. Peralta, M. Berón, D. Riesco, and G. Montejano, "MEBPCM: A Method for Evaluating Business Process Conceptual Models. A Study Case.," presented at the Ninth International Conference on Information Technology: New Generations (ITNG), Las Vegas, Nevada, USA, 2012.
- [10] L.-C. Huang and R. Y.-H. Wu, "Applying fuzzy analytic hierarchy process in the managerial talent assessment model – an empirical study in Taiwan's semiconductor industry.," *International Journal of technology Management.* 105-130, 2005.
- [11] X. Wang and E. Triantaphyllou, "Ranking irregularities when evaluating alter-natives by using some ELECTRE methods.," *Omega*, pp. 36:45–63, 2008.
- [12] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning.," *Information Sciences*, pp. 199–249(I), 301–357(II). , 1975.
- [13] R. Bellman and L. Zadeh, "Decision making in a fuzzy environment.," *Management Science*, 17B(4):141–64. 1970.
- [14] H. Zimmermann, "Fuzzy set theory and its applications.," *2nd Edition. Boston: Kluwer Academic Publishers*, 1991.
- [15] B. Kosko, "Pensamiento borroso: la nueva ciencia de la lógica borrosa," *Barcelona: Crítica.*, 1995.
- [16] J. Mendling, H. Reijers, and W. van der Aalst, "Seven Process Modeling Guidelines," *Information & Software Technology*, vol. 52(2), pp. 127-136, 2010.
- [17] A. Kaufmann and J. Gil Aluja, "Las matemáticas del azar y la incertidumbre: elementos básicos para su aplicación en economía," *Madrid: Ramón Areces*, 1990.
- [18] E. Trillas, "Conjuntos Borrosos," *Madrid: Vicens Vives.*, 1980.
- [19] J. Jang, E. Mizutani, and C. Sun, "Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence," *New York: Prentice Hall.*, 1997.
- [20] A. Kulkarni, "Computer vision and fuzzyneural systems," *New York: Prentice Hall.*, 2001.
- [21] N. Kasabov, "Foundations of neural networks, fuzzy systems and knowledge engineering," *Cambridge: The MIT Press*, 1998.
- [22] C. Salgado, "Tesis de Especialización en Ingeniería de Software: El Modelado de Procesos de Negocio: Aplicando LSP para la Evaluación de Lenguajes de Modelado de Procesos de Negocio". FCFMyN, UNSL, 2010.
- [23] C. Salgado, "Tesis de Maestría en Ingeniería de Software: MEMPN: Método para la Evaluación de Modelos Conceptuales de Procesos de Negocio ", Facultad de Ciencias Físico-Matemáticas y Naturales., Universidad Nacional de San Luis, San Luis, Argentina., 2013.

Aproximaciones para el Desarrollo Multiplataforma y Mantenimiento de Aplicaciones Móviles

Marta C. Fennema, Susana I. Herrera, Rosa A. Palavecino, Paola D. Budán, Federico Rosenzvaig, Pablo J. Najjar Ruiz, Álvaro J. Carranza, Emmanuel Saavedra

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
{rosypgg, sherrera, pbudan, frosenzvaig}@unse.edu.ar, najjarpablo@yahoo.com.ar

Resumen

Esta línea de investigación está abocada al análisis de herramientas de generación automática de aplicaciones nativas para múltiples Sistemas Operativos móviles. Vinculado a este análisis se desean definir métricas que permitan evaluar la capacidad de mantenimiento de aplicaciones móviles obtenidas con herramientas de generación automática. Asimismo, se destinarán los esfuerzos de investigación a definir un método ágil apropiado para el desarrollo de aplicaciones móviles nativas y multiplataforma.

Palabras clave: aplicaciones móviles, herramientas de generación automática de aplicaciones, métodos de desarrollo.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado “*Computación Móvil: desarrollo de aplicaciones y análisis forense*”. El mismo se ejecutará en el período 2017-2018, y es una continuación del proyecto sobre Computación Móvil, desarrollado entre los años 2012 y 2016 en la Universidad Nacional de Santiago del Estero (UNSE) [7].

El equipo de investigación del proyecto está conformado por docentes del Instituto de Investigaciones en Informática de la Facultad de Ciencias Exactas de la UNSE. Además, el proyecto cuenta con el asesoramiento de

investigadores de Institutos de Investigación de la Facultad de Informática de la Universidad Nacional de La Plata, y de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta.

Este proyecto está financiado por el Consejo de Ciencia y Técnica de la UNSE, durante el período Enero 2017- Diciembre 2018.

2 Introducción

El desarrollo de una aplicación móvil implica considerar distintos factores [2], entre los cuales se pueden mencionar: la Arquitectura de la Aplicación (considerado el más relevante), el Modelo de Contexto, el Modelo de Usuario, la Representación del Espacio, la Dinámica de la Aplicación, la Representación de los Puntos de Interés (PI), el Modelo de Dominio, y los Mecanismos de Censado. Estos conceptos varían en función de la arquitectura de la aplicación y viceversa; por ello, la consideración de la arquitectura de la aplicación es el factor más relevante.

Con respecto a esto, las aplicaciones móviles pueden desarrollarse con arquitecturas de tres capas denominadas presentación, aplicación, y datos; que se pueden alojar en la web [17] o en diferentes ubicaciones [4]. Para el desarrollo de esta línea de investigación se seguirá la clasificación de

arquitecturas propuesta por Challiol [2], quien sostiene que la arquitectura de una aplicación móvil puede ser: cliente, híbrida o web.

Por otro lado, hasta el año 2015 el SO más utilizado por los usuarios de aplicaciones móviles fue Android StatCounter [16]. Actualmente, se observa que un 25% de la población utiliza otros SO móviles. Considerando esta variación y el hecho que este proyecto de investigación aplica sus esfuerzos en el m-learning, donde se requieren aplicaciones para diversos SO, se precisa trabajar en la producción multiplataforma. En este campo, han surgido diversas herramientas que permiten el desarrollo híbrido de aplicaciones; que son esencialmente aplicaciones web ejecutándose en un navegador Shell en una aplicación que posee acceso a las capas de la plataforma nativa. Las aplicaciones híbridas tienen muchos beneficios sobre las aplicaciones nativas puras, específicamente en términos de soporte de plataformas, velocidad de desarrollo y acceso a código de terceros. Las herramientas más conocidas son las siguientes:

- *Apache Cordova*. Es un framework de desarrollo móvil de código abierto que permite utilizar las tecnologías estándar web para el desarrollo multiplataforma, evitando el lenguaje de desarrollo nativo de cada una de las plataformas móviles [19].
- *PhoneGap*. Es un framework que permite crear aplicaciones multiplataforma utilizando HTML y JavaScript. Es una distribución gratuita y de licencia libre de Apache Cordova [12]. Posibilita la creación de aplicaciones híbridas mediante la utilización de una serie de APIs que permiten controlar características de los dispositivos como el acelerómetro, la cámara, las notificaciones o la geolocalización

[5]. Phonegap permite conectar un dispositivo móvil con una aplicación web mediante dos componentes principales: Webview y Plugins.

- *Ionic*. Es una herramienta, gratuita y open source, para el desarrollo de aplicaciones híbridas basadas en HTML5, CSS y JS. Está construido con Sass y optimizado con AngularJS [10]. Está orientado al desarrollo de aplicaciones móviles híbridas. Debido a que Ionic es un marco de trabajo HTML5 necesita de un envoltorio nativo como Cordova o Phonegap para ejecutarse como una aplicación nativa.

Si bien el uso de estos *frameworks* es creciente en la actualidad, las aplicaciones híbridas no son efectivas cuando se trabaja con acceso al hardware de dispositivos, como por ejemplo cámaras, gps, acelerómetro, brújula, lo cual es necesario en aplicaciones con RA. Es por ello que se investigará acerca de *frameworks* que generen aplicaciones multiplataformas nativas.

Recientemente, ha surgido Xamarin, una plataforma de desarrollo que permite obtener aplicaciones de código nativo, multiplataforma, en C#, para iOS, Android y WindowsPhone [6]. Se explorarán también herramientas abiertas y la posibilidad de desarrollarlas o mejorar las existentes.

En cuanto a métodos de desarrollo, algunos autores sostienen que las aplicaciones móviles pueden desarrollarse tanto con métodos predictivos, como *Waterfall*, como con métodos ágiles, como el Desarrollo Rápido de Aplicaciones, basado en la prototipación [18]. Sin embargo, los métodos que se destacan en la literatura corresponden a métodos ágiles. Entre ellos, sobresale Mobile-D adaptable a diferentes dominios o contextos, se caracteriza por propiciar desarrollos de

mediana envergadura en muy breves períodos de tiempo. Para ello, propone iteraciones divididas en cinco fases compuestas por etapas, en donde cada fase tiene un momento de planificación, de trabajo y de lanzamiento [1].

Spartaru [15] propone un conjunto de mejoras al método Mobile-D, incorporando a las fases, un estudio en categorías de aplicaciones de móviles, los paradigmas de desarrollo, la inclusión del usuario final en el ciclo de vida del método, pruebas de rendimiento de componentes y la adopción de principales líneas de producto de software. Rahimian et.al. [13] presentan una metodología híbrida ágil para el desarrollo de aplicaciones móviles basada en aproximaciones de ingeniería, más específicamente, una metodología basada en riesgos, en donde se le da un rol fundamental a la priorización de requerimientos para el desarrollo de las iteraciones.

En síntesis, esta línea de investigación está abocada a indagar sobre cómo desarrollar aplicaciones móviles multiplataforma de calidad, que se adapten a la demanda de los equipos 4G, y a comparar los distintos métodos y herramientas existentes para optimizarlo.

3 Líneas de investigación y desarrollo

La principal línea de investigación de esta propuesta consiste en el estudio de métodos, técnicas y herramientas que permitan el desarrollo de sistemas móviles multiplataforma; lo cual significa indagar sobre:

- a) Herramientas para la generación automática de aplicaciones móviles nativas que interactúan con hardware del dispositivo,
- b) Métodos ágiles para el desarrollo de aplicaciones móviles multiplataforma.

En esta propuesta se pretende investigar sobre plataformas o entornos de desarrollo de aplicaciones móviles que permite generar aplicaciones nativas que puedan ser optimizadas mediante la modificación del código.

En cuanto al dominio de aplicación, se continuará trabajando en Educación y Salud, desarrollando aplicaciones para m-learning y para rehabilitación de personas con discapacidad cerebral. Se incorpora en esta nueva etapa el dominio de la informática forense.

4 Objetivos y resultados

El objetivo general que guía esta investigación aplicada es:

Contribuir al progreso del campo de la Computación Móvil mediante la optimización del desarrollo y mantenimiento de aplicaciones móviles nativas multiplataforma.

Los objetivos específicos son:

- a) Analizar herramientas (libres y propietarias) de generación automática de aplicaciones nativas para múltiples SO móviles.
- b) Definir métricas para evaluar la capacidad de mantenimiento de aplicaciones móviles con herramientas de generación automática.
- c) Comparar herramientas para el desarrollo de aplicaciones móviles con RA.
- d) Definir un método ágil apropiado para el desarrollo de aplicaciones móviles nativas y multiplataforma, basadas en generación automática de código.
- e) Determinar la calidad de aplicaciones móviles construidas con las diversas herramientas analizadas, usando

prototipos construidos con métodos ágiles y las métricas definidas.

Esta línea de investigación está incluida en un proyecto en estado inicial, por lo cual no se tienen resultados preliminares a la fecha.

Sin embargo, como se mencionó anteriormente, es continuación del proyecto de investigación sobre Computación Móvil 2012-2016 [7], algunos de cuyos principales resultados en esta línea de investigación son:

- Herrera, S. (2016). Tesis doctoral en Ciencias Informáticas (UNLP) referida a un marco para el análisis y evaluación de experiencias de m-learning. Directora: Cecilia Sanz. Co-directora: Cristina Fennema.
- Rocabado, S.; Sanchez, E.; Herrera, S.; Cadena, C. (2016). Eficiencia energética en dispositivos móviles para facilitar su uso en zonas rurales aisladas [14].
- Córdoba, Melisa et al (2016). Sistema Alternativo de Comunicación para Niños con Parálisis Cerebral Infantil. Artículo enmarcado en una tesis final de grado que se encuentra en estado avanzado de desarrollo [3].
- Rocabado, S. H. (2014). Caso de estudio de comunicaciones seguras sobre redes móviles ad-hoc. Tesis de finalización de Maestría en Redes de Datos. Facultad de Informática, Universidad Nacional de La Plata.
- Najar (2013) El diseño de modelos de arquitecturas móviles, tanto web como híbridas, orientadas a servicios. Este trabajo final de grado se enfatizó en el desarrollo de aplicaciones Android, dado que hasta el año 2015 este SO predominaba en el mercado en Argentina.

- Herrera, S., Najar, P., Palavecino, R., Goñi, J. (2013) Evaluación de la calidad en aplicaciones móviles [8].
- Herrera, S.; Najar Ruiz, P.; Contreras, N.; Fennema, C.; Lara, C. (2013). Marco Sistémico para el Desarrollo de Aplicaciones de m-Turismo [9].

5 Formación de recursos humanos

La Directora del proyecto pertenece a la Universidad del Norte Santo Tomás de Aquino. La Codirectora y el resto de los integrantes, al Instituto de Investigación en Informática y Sistemas de Información de la UNSE. El proyecto cuenta con asesores del Instituto de Investigación de la Facultad de Informática de la Universidad Nacional de La Plata, y de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta.

En esta línea de investigación se desempeñan cuatro investigadores de la UNSE y tres estudiantes avanzados de la Licenciatura en Sistemas de Información.

Se prevé realizar cuatro trabajos final de grado y un trabajo final de posgrado. Este último para la Especialización de Enseñanza en Tecnología.

REFERENCIAS

1. Abrahamsson, P., Hanhineva, A., Hulkko, H., Ihme, T., Jääliñoja, J., Korkala, M., ... & Salo, O. (2004). Mobile-D: an agile approach for mobile application development. In Companion to the 19th annual ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications (pp. 174-175). ACM.
2. Challiol, C. (2012). *Apuntes de Curso de Posgrado sobre Computación Móvil. Curso dictado en Universidad Nacional de Santiago del Estero. Secretaría de Extensión de la Facultad de Ciencias Exactas y Tecnologías.*
3. Córdoba, Melisa N. V.; Najar, Pablo; Budan Paola Daniela. (2016). Sistema

- Alternativo de Comunicación para Niños con Parálisis Cerebral Infantil. Argentina. San Salvador de Jujuy. 2016. Libro. Artículo Completo. Jornada. XI Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. CODINOA
4. Emmanouilidis, C., Koutsiamanis, R.A., Tasidou, (2013) A. *Mobile guides: Taxonomy of architectures, context awareness, technologies and applications*. Journal of Network and Computer Applications. Volume 36, Issue 1, Pages 103–125. Elsevier, January.
 5. Genbetadev. (2016). <http://www.genbetadev.com/frameworks/phonegap>. Fecha de acceso: 13/08/2016.
 6. Hermes, D. (2015). Xamarin Mobile Application Development: Cross-Platform C# and Xamarin.Forms Fundamentals. Ed. Apress. ISBN 978-1484202159.
 7. Herrera, S. I., Najar Ruíz, P. J., Rocabado Moreno, S. H., Fénema, M. C., & Cianferoni, M. C. (2013). Optimización de la Calidad de los Sistemas Móviles. In XV Workshop de Investigadores en Ciencias de la Computación.
 8. Herrera, S., Najar, P., Palavecino, R., Goñi, J. (2013) Evaluación de la calidad en aplicaciones móviles. Investigaciones en Facultades de Ingeniería del NOA. Santiago del Estero: Universidad Nacional de Santiago del Estero. vol. n°. p103 - 104. issn 1853-7871.
 9. Herrera, S.; Najar Ruiz, P.; Contreras, N.; Fennema, C.; Lara, C. (2013). Marco Sistémico para el Desarrollo de Aplicaciones de m-Turismo. Brasil. Palmas, Tocantins. Libro. Artículo Completo. Congreso. Congresso Brasileiro de Sistemas: pensar a compreensão de sistemas. Universidade Federal do Tocantins.
 10. Ionic. (2016). Disponible en: <http://ionicframework.com/docs/guide/preface.html>. Fecha de acceso: 13/08/2016.
 11. Najar, P. J. (2013). Prototipo de sistema móvil para e-turismo. Trabajo final de Licenciatura en Sistemas de Información. Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero.
 12. PhoneGap. (2013). A guid to building cross-platform apps using the W3C standards based Cordova/PhoneGap framework. Packt Publishing.
 13. Rahimian, V., & Ramsin, R. (2008). Designing an agile methodology for mobile software development: A hybrid method engineering approach. In 2008 Second International Conference on Research Challenges in Information Science (pp. 337-342). IEEE
 14. Rocabado, S.; Sanchez, E.; Herrera, S.; Cadena, C. (2016). Eficiencia energética en dispositivos móviles para facilitar su uso en zonas rurales aisladas. Argentina. Concordia. Libro. Artículo Breve. Workshop. XVIII Workshop de Investigadores en Ciencias de la Computación. RedUNCI.
 15. Spataru, A. C. (2010). Agile development methods for mobile applications. Master of Science Thesis submitted to Computer Science School of Informatics, University of Edinburgh.
 16. StatCounter Global Stats (2015). Ranking de los 8 Sistemas Operativos Móviles más usados en Argentina. Disponible en <http://gs.statcounter.com/#mobile_os-AR-monthly-201401-201501-bar>. Fecha de consulta: 18/04/2015.
 17. Talukder, A.K., Ahmed, H., Yavagal, R. (2010) Mobile Computing: Technology, Applications, and Service Creation. 2° Edición. McGraw-Hill communications engineering series. ISBN 9780070144576.
 18. Vique, R. R. (2012). Métodos para el desarrollo de aplicaciones móviles. PID_00176755.
 19. Wargo, J. (2015). Apache Cordova 4 Programming. Addison-Wesley Professional.

Asistencia Dirigida por Ontologías al Diseño Arquitectónico de Videojuegos

Facundo Gonzalez Gola¹, Alejandro Sánchez^{1,2}, Germán Montejano^{1,3}

¹Universidad Nacional de San Luis
Ejercito de Los Andes 950, San Luis, Argentina - Tel: +54 266 4520300
gonzalezgolafacundo@gmail.com, {asanchez,gmonte}@unsl.edu.ar

²Universidad Nacional de la Patagonia San Juan Bosco
Bvd Brown 2051 Puerto Madryn, Tel. +54-280-4883585

³Universidad Nacional de la Pampa
Av. Uruguay 151, (6300) Santa Rosa, La Pampa, Argentina, Tel.: +54-2954-245220 Int. 7125

Resumen

El diseño arquitectónico de videojuegos es complejo. Es normal que los videojuegos se proyecten para ser ejecutados sobre una amplia variedad de dispositivos, sistemas operativos, e interpretes. A su vez, siempre se utilizan componentes de software – llamados motores – que proveen un número de funcionalidades básicas, pero plantean limitaciones sobre la plataforma y el entorno de desarrollo. Esto se combina con requerimientos de performance, y en muchos casos de escalabilidad, naturales en este dominio. Por otro lado, los equipos de desarrollo son típicamente multidisciplinares, lo que dificulta la comunicación de decisiones de diseño.

Esta línea plantea la investigación y desarrollo de ontologías para dirigir el diseño arquitectónico de videojuegos. Estas ontologías explicitarán y alinearán las conceptualizaciones utilizadas por las disciplinas involucradas, facilitarán la integración de datos sobre los artefactos relacionados al diseño, y proveerán una teoría lógica para guiar parte del diseño

arquitectónico y explicar inconsistencias en el mismo.

Palabras clave: ontología, videojuegos, diseño arquitectónico.

Contexto

Esta línea de investigación, desarrollo e innovación se desprende del proyecto “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad” de la facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis.

Introducción

El mercado mundial de videojuegos se encuentra en constante crecimiento. La ESA¹, asociación creadora del evento más importante a nivel mundial de video juegos (el Electronic Entertainment Expo o E3), reporta en (Essential facts about the computer and video game industry 2016) un gran crecimiento del mercado Norte Americano con respecto al año anterior. En 2016 el 63% de la población Norte Americana ha jugado al menos 3 horas o más por semana, significando un

¹ Entertainment Software Association

mercado de un total de 23,5 billones de dólares.

Sin embargo, muchos proyectos de desarrollo de videojuegos han fracasado y continúan haciéndolo en la actualidad. Considere por ejemplo el caso del juego ET, de la empresa Atari (Dingman 2014). El juego fue sacado de producción y llevó a la empresa a la bancarrota. Otro caso es el juego Duke Nukem Forever, que demoró 14 años en salir al mercado por problemas de diseño (Schiesel 2011). Peor aún es el caso del juego No Man's Sky, donde sus desarrolladores fallaron en cumplir con las características prometidas y debieron afrontar procesos legales (BBC-Tech 2016). Casos como estos se reportan frecuentemente en la actualidad, con mayor proporción en pequeñas y medianas empresas.

Es posible atribuir muchos de los fracasos en proyectos de desarrollo de videojuegos al diseño arquitectónico de los mismos. La complejidad en esta actividad viene dada principalmente por la gran variedad de plataformas sobre las que se pretenden liberar los juegos y sus limitaciones físicas, los componentes de software que se deben utilizar y sus condicionamientos sobre las plataformas, los típicos requerimientos de performance y escalabilidad, y la frecuente composición multidisciplinaria de los equipos de desarrollo.

El mercado demanda que los videojuegos sean liberados sobre una amplia gama de dispositivos que va desde consolas (computadoras específicamente diseñadas para juegos), pasando por computadoras hogareñas hasta teléfonos móviles. Estos dispositivos disponen de su software de base sobre el cual deben ejecutar los programas, tales como sistemas operativos y máquinas virtuales,

provistos bajo diferentes licencias, y disponibles en diferentes versiones. Tanto los dispositivos como su software evolucionan rápidamente.

Existen componentes de software – llamados motores de juegos – que pueden facilitar el desarrollo al proveer un número de funcionalidades tales como la renderización de imágenes, la emisión de sonidos, la captura de eventos generados por el usuario, la administración de memoria, o la ejecución de scripts bajo determinadas condiciones de disparo. Dichos motores son desarrollados para determinados dispositivos asumiendo la disponibilidad de ciertas capacidades y software de base.

A su vez, el desarrollo de videojuegos típicamente involucra expertos de una variedad de dominios, como ser artes gráficas, musicalización, y narración. Estos expertos cuentan con una formación muy diferente a los expertos IT y en consecuencia una conceptualización distinta para realizar el trabajo. La comunicación puede resultar poco eficaz y sea difícil inconsistencias en requerimientos y decisiones de diseño arquitectónicas.

Fallar en la detección de inconsistencias en el diseño arquitectónico de un videojuego pueden condenar al fracaso el proyecto. Por ejemplo, el juego requiere una cantidad de memoria mínima que algunas de las plataformas sobre las que se liberará no poseen. Otro ejemplo es utilizar para un juego de música una máquina virtual que permite la portabilidad del mismo, pero impide la ejecución simultánea de sonidos. Si el equipo no dispone de un arquitecto con suficiente conocimiento de la máquina virtual, este problema no se

detectaría hasta que el primer prototipo fuera probado.

Los casos como los mencionados pueden ser prevenidos si se dispusiera de un servicio que respondiera preguntas como las que siguen debajo.

- A. ¿Es consistente un conjunto dado de decisiones arquitectónicas?
- B. ¿Cuáles son limitaciones de memoria y procesamiento impuestas por las plataformas elegidas?
- C. ¿Qué componentes/software de base se ajusta a los requerimientos del video juego proyectado?

Existen antecedentes del uso de ontologías (Gruber 1993) para dirigir aplicaciones que respondan a preguntas como estas en el contexto del desarrollo de software. Una ontología es una representación formal y compartida de una conceptualización. Este enfoque resulta más flexible que el de una aplicación que responda directamente, ya que una ontología es una representación explícita de las teorías subyacentes que puede ser modificada para ajustarse a cambios, como por ejemplo, los originados por nuevas tecnologías en consolas o dispositivos móviles.

En (Pan et al. 2012), la expresión “desarrollo de software dirigido por ontologías” o ODS (por las siglas de su nombre en inglés: ontology-driven software development) es usado para referirse a la aplicación de tecnologías de ontologías (Baader et al. 2010) para facilitar los procesos de desarrollo de software, y específicamente, la práctica de desarrollo de software dirigido por modelos o MDS (de las siglas de model-driven software development). Los autores

exploran la aplicación al contexto del desarrollo de software de las capacidades de las ontologías para chequear consistencia y proveer guía.

Estos enfoques, en general, implican ontologías de distintos tipos (Guarino 1998). Se suele requerir el desarrollo/reutilización de ontologías de dominio, y el desarrollo de ontologías de aplicación. En ambos casos, las ontologías pueden incorporar entidades de ontologías de alto nivel. Las relaciones entre estas ontologías se desarrollan conforme a la disciplina de correspondencia entre ontologías (Euzenat and Shvaiko 2013).

A su vez, existen metodologías de desarrollo que permiten la participación, supervisión, y evaluación de los productos por parte de expertos del dominio. Esto es particularmente importante de cara al objetivo de que se consideren las conceptualizaciones de expertos de una variedad de dominios. Entre dichas metodologías podemos citar la propuesta en (Grüninger and Fox 1995), o en (Uschold and King 1995), o Kactus (Bernaras, Laresgoiti, and Corera 1996), o Methontology (Gómez-Pérez, Fernández-López, and De Vicente 1996), Ontology (Staab et al. 2001), Terminae (Corcho, Fernández-López, and Gómez-Pérez 2003), o REFSENO (Tautz and Wangenheim 1998),

El concepto de ontología y sus fundamentos y herramientas se han aplicado en el desarrollo de videojuegos. Por ejemplo, en (Llansó García 2014; Llansó et al. 2011) se propone una metodología para definir de manera formal y ágil el dominio del juego utilizando análisis formal de conceptos o FCA (por su nombre en inglés formal concept analysis). En (Leon Z. and

Sanchez 2010) utilizan UML y OCL para bosquejar una ontología para videojuegos para dispositivos móviles. El trabajo de (Trindade 2015) se enfoca en el desarrollo de una ontología para identificar requerimientos técnicos en el diseño de juegos.

Líneas de investigación, desarrollo e innovación

Esta línea de I-D-I investiga y desarrolla ontologías para dirigir el diseño arquitectónico de videojuegos. La línea se encuentra en etapa inicial. Se investigará la conceptualización para la toma de decisiones de diseño arquitectónico, se desarrollarán ontologías que formalicen este conocimiento y permitan luego el desarrollo de aplicaciones que puedan guiar al diseñador y detectar inconsistencias en especificaciones arquitecturas.

Objetivos y resultados esperados

El objetivo es especificar de manera explícita y formal el conocimiento necesario para decisiones de diseño arquitectónico en el desarrollo de videojuegos. Las ontologías abordarán los aspectos fundamentales del dominio de videojuegos y las teorías para responder a preguntas como las planteadas en los ítems A, B y C de la sección Introducción. Se busca proyectar aplicaciones dirigidas por dichas ontologías que asistan al diseñador en las decisiones de diseño y en la detección de inconsistencias. Se espera obtener:

- Ontologías sobre videojuegos especificando conceptos tales como dispositivos, software de base, motores de videojuegos, y las dependencias entre estos.

- Ontologías especificando una teoría lógica para responder a preguntas de diseño como las indicadas en los puntos A, B y C.
- Una aplicación prototípica que resulte una respuesta innovadora a los desafíos de diseño arquitectónico en el ámbito de los videojuegos.

Formación de Recursos Humanos

Esta línea de I-D-I continúa con la formación en investigación de alumnos de grado y posgrado de la UNSL. Está previsto que uno de los autores desarrolle en la línea su Tesis de Maestría en Ingeniería de Software, así como también dos alumnos de grado sus tesinas. Es importante destacar el apoyo académico y científico para los temas desarrollados en la formación de recursos humanos calificados, ya que son carreras de posgrado categorizadas "A" para los desarrollos de las tesis de posgrado y carreras de grado acreditadas por 6 años para los desarrollos de las tesinas.

Referencias

Baader, Franz et al. (2010). *The Description Logic Handbook: Theory, Implementation and Applications*. 2nd. New York, NY, USA: Cambridge University Press.

BBC-Tech (2016). No Man's Sky investigated over 'misleading' adverts.

Bernaras, A., I. Laresgoiti, and J. Corera (1996). "Building and Reusing Ontologies for Electrical Network Applications". In: *Proceedings of the 12th European Conference on Artificial*

Intelligence (ECAI'96). Ed. by Wolfgang Wahlster. Budapest, Hungary: John Wiley and Sons, pp. 298–302.

Corcho, Oscar, Mariano Fernández-López, and Asunción Gómez-Pérez (2003). “Methodologies, Tools and Languages for Building Ontologies: Where is Their Meeting Point?” In: *Data Knowl. Eng.* 46.1, pp. 41–64.

Dingman, Hayden (2014). Atari: Game Over film review: Unearthing the El Dorado of video games. PC World.

Essential facts about the computer and video game industry (2016). Tech. rep. Entertainment Software Association.

Euzenat, Jérôme and Pavel Shvaiko (2013). *Ontology matching*. 2nd. Heidelberg (DE): Springer-Verlag.

Gómez-Pérez, Asunción, Mariano Fernández-López, and A.J De Vicente (1996). “Towards a Method to Conceptualize Domain Ontologies”. In: *ECAI-96 Workshop on Ontological Engineering*. ECAI-96 Workshop Proceedings.

Gruber, Thomas R. (1993). “A Translation Approach to Portable Ontology Specifications”. In: *Knowl. Acquis.* 5.2, pp. 199–220.

Grüninger, M. and M. Fox (1995). “Methodology for the Design and Evaluation of Ontologies”. In: *IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing*, April 13, 1995.

Guarino, Nicola (1998). *Formal Ontology in Information Systems: Proceedings of*

the 1st International Conference June 6-8, 1998, Trento, Italy. 1st. Amsterdam, The Netherlands, The Netherlands: IOS Press.

Leon Z., Annie C. and Abraham L. Sanchez (2010). “An Ontology for Mobile Video Games”. In: *2010 Ninth Mexican International Conference on Artificial Intelligence*, pp. 154–159.

Llansó García, David (2014). “Metodología ontológica para el desarrollo de videojuegos”. PhD thesis. Universidad Complutense de Madrid.

Pan, Jeff Z. et al. (2012). *Ontology-Driven Software Development*. Springer Publishing Company, Incorporated.

Schiesel, Seth (2011). 14 Years of Waiting Have Come to an End. NY Times.

Staab, Steffen et al. (2001). “Knowledge Processes and Ontologies”. In: *IEEE Intelligent Systems* 16.1, pp. 26–34.

Tautz, Carsten and Christiane Gresse von Wangenheim (1998). *REFSENO: A representation formalism for software engineering ontologies*. Tech. rep. 15. Fraunhofer IESE.

Trindade, Glauco Ofranti (2015). “Video game development ontology”. PhD thesis. Universidade Federal do Rio de Janeiro.

Uschold, Mike and Martin King (1995). “Towards a methodology for building ontologies”. In: *Workshop on Basic Ontological Issues in Knowledge Sharing*.

Búsqueda de Estrategias para la Clasificación del Contenido en Foros Técnicos de Discusión

Nadina Martínez Carod, Gabriela Aranda, Alejandra Cechich,
Valeria Zoratto, Carina Noda, Mauro Sagripanti

Grupo de Investigación en Ingeniería de Software del Comahue (GIISCO)
<http://giisco.uncoma.edu.ar>

Facultad de Informática. Universidad Nacional del Comahue
Buenos Aires 1400, (8300) Neuquén
Contacto: {nadina.martinez, gabriela.aranda, alejandra.cechich}@fi.uncoma.edu.ar

RESUMEN

En las últimas décadas la disciplina Information Retrieval ha avanzado considerablemente. Esto se debe gran parte a que las organizaciones actuales hacen cada vez más esfuerzos para reutilizar el conocimiento, definiendo estrategias para tener catalogadas y reutilizar soluciones ya probadas [7].

Por el otro lado la evolución de la Web trajo consigo distintas herramientas de trabajo colaborativo. Dentro de dichas herramientas, los foros de discusión son ampliamente utilizados para plantear problemas, expresar sugerencias, así como para intercambiar conocimientos y experiencias. Dentro de un foro de discusión, un usuario de la comunidad puede realizar una pregunta, y el resto de los miembros de dicha comunidad pueden responder proponiendo soluciones al problema planteado. De esta manera, mediante el uso de esta herramienta, se genera un volumen de información bastante importante, que puede ser utilizado como fuente de conocimiento para un sistema de recuperación de información.

El objetivo fundamental de nuestro proyecto es definir una herramienta que, a partir de información existente en hilos de discusión de foros técnicos, la clasifique y establezca un orden entre soluciones posibles para problemas recurrentes del área de programación.

CONTEXTO

La línea de investigación presentada se denomina “Reuso de Conocimientos en Foros de Discusión II” y forma parte del programa “Desarrollo de Software Basado en Reuso – Parte II”, con período de vigencia 2017-2020.

El programa mencionado extiende el programa “Desarrollo de Software Basado en Reuso” realizado durante el período 2013-2016.

1. INTRODUCCION

La disciplina que se encarga de recuperar información (Information Retrieval) surge en la década de 1950 [12], ante la necesidad de procesar y reutilizar la información almacenada en grandes volúmenes. Desde ese momento, este campo ha madurado y han ido surgiendo valiosos aportes en distintas ramas de investigación. Por ejemplo, algunos proyectos se han enfocado en utilizar la información almacenada en documentos específicos, mientras que otros han desarrollado técnicas para generación automática de tesauros (lista de sinónimos, en conjunto con lista de antónimos, etc.) para su uso en consultas. En general el proceso de recupero de información comienza con la consulta de un usuario al sistema. Las respuestas a dicha consulta poseen diferentes grados de relevancia, y para organizarlas, se determina un ranking el cual evalúa el grado de respuesta a una consulta.

Si bien el conocimiento en la Web se encuentra diseminado en distintos tipos de aplicaciones, los foros de discusión en particular se caracterizan por ser herramientas colaborativas con grandes volúmenes de información, accesibles a la comunidad en general como fuente de consulta (la gran mayoría de los foros cumplen estas características). En estas herramientas se intercambia conocimiento constantemente. Esas son las razones por las cuales nos

enfocamos en los foros de discusión, y específicamente en aquellos que tratan temas técnicos.

En general, la mayoría de los métodos automáticos de IR se basan en analizar la ocurrencia de palabras en los documentos, lo que produce listas de palabras fuertemente relacionadas. El principal problema detectado es que no todas las palabras relacionadas con una palabra de consulta son significativas en el contexto de la consulta. Este es un aspecto fundamental considerado en el proyecto.

Dado que existen en la Web muchos foros de discusión sobre la misma temática, se pueden hallar preguntas y respuestas similares diseminadas en varios de ellos, por lo que generalmente es necesario navegar por varios hilos hasta dar con una solución correcta. Incluso muchas veces es necesario considerar características de calidad para evaluar soluciones [1][3][8], desafío que se intenta lograr al dar las posibles soluciones.

Existen varias propuestas de reuso de conocimiento disponible en foros de discusión, como [2] que implementa un sistema recomendador que busca y propone mensajes con contenido similar, en [4], los mensajes se clasifican de acuerdo a una jerarquía de temas preestablecida. El enfoque de Nicoletti [17] clasifica los mensajes acorde a una jerarquía de temas obtenido de Wikipedia. También existen propuestas de generación de algoritmos de ranking basados en la calidad de los atributos, como [11].

Bajo este prisma, nuestro proyecto tiene como objetivo principal favorecer el reuso de la información contenida en conversaciones existentes en la Web, con el valor agregado de un análisis de calidad de dichas fuentes de información. Además, se ha experimentado tanto con la aplicación de algoritmos de análisis de lenguaje natural como de aprendizaje automático, y se está analizando la aplicación de sentiment analysis para mejorar las búsquedas. El análisis de lenguaje natural en foros de discusión permite analizar el tipo de fragmento dentro de un hilo de discusión, como en [10]. Por este motivo, nuestro proyecto está orientado a determinar un ranking de soluciones posibles, y cada línea dentro del proyecto se enfoca en esta acción desde diferentes aspectos.

2. LINEAS DE INVESTIGACION Y DESARROLLO

El proyecto de investigación se denomina “Reuso de Conocimientos en Foros de Discusión – Parte II” y está enmarcado dentro del Programa de Investigación “Desarrollo de Software Basado en Reuso – Parte II”, con período de vigencia 2017-2020.

El programa mencionado extiende la investigación realizada durante el programa denominado “Desarrollo de Software Basado en Reuso”, realizado en el período 2013-2016. Respecto a este proyecto en particular, el objetivo es extender los estudios realizados sobre reuso de conocimiento en foros de discusión técnicos, incorporando la definición de métodos y algoritmos de recomendación para la asistencia inteligente a usuarios en la búsqueda de soluciones a preguntas recurrentes. Por otra parte, el programa está conformado por otros dos subproyectos que profundizan en las temáticas de Reuso Orientado al Dominio y Reuso Orientado a Servicios.

Dicho programa está desarrollado por el Grupo de Ingeniería de Software de la Universidad Nacional del Comahue, (GIISCo), formado por docentes y estudiantes de la Facultad de Informática de la Universidad Nacional del Comahue, junto con asesoría y colaboración de otras universidades. En particular, este proyecto es desarrollado en colaboración con la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires. Aunque el objetivo del Grupo GIISCo es brindar soporte en investigación y transferencia de tópicos relacionados con la Ingeniería de Software, el proyecto también involucra a docentes pertenecientes a otras áreas de la Facultad, como Programación y Teoría de la Computación, lo que permite abordar la investigación desde ópticas diferentes, enriqueciendo el desarrollo con un trabajo conjunto y colaborativo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Como antecedentes de este proyecto de investigación, en el año 2013 se presentó un modelo de calidad para foros de discusión en

base a modelos de datos y de información en la Web y en estándares para la calidad de datos software [9]. La validación de la selección de atributos y sub-atributos de dicho modelo se realizó mediante encuestas [13]. Además, durante 2014 se implementó una versión preliminar de una herramienta que permite la recuperación de información desde foros de discusión técnicos y su análisis mediante un conjunto preliminar de métricas de calidad, a partir de las cuales se propone un ranking de soluciones posibles para una pregunta. Dicha herramienta fue aplicada en varios casos de estudio con hilos de discusión reales y sus resultados están presentados en una tesis de licenciatura con fecha de defensa a realizarse próximamente.

Además, durante 2015 y 2016 se avanzó en el análisis de casos de estudio más amplios, a partir de una cadena de búsqueda y en el estudio del orden esperado confrontado al orden obtenido por medio de las herramientas de análisis de texto [15][16]. Para ello se utilizó la herramienta Lucene, con mecanismos personalizados para establecer stopwords propias del dominio. Actualmente, se está avanzando en el uso de bases de datos léxicas (como WordNet [24]), en combinación con otras estrategias de Recuperación de Información. Esta línea de investigación está siendo desarrollada como parte de otra tesina que evaluará los resultados obtenidos al aplicar distintas funciones de las bases de datos léxicas [25] en la búsqueda de mensajes relacionados a una pregunta particular. Durante el desarrollo de esta tesina se establecerán nuevos corpus de hilos de discusión reales sobre los cuales aplicar las técnicas seleccionadas implementando una estrategia de validación empírica.

Por otra lado se está evaluando la aplicación de técnicas de Data Mining y de modelos de aprendizaje automático supervisados y no supervisados [18][19] como así también técnicas y herramientas disponibles de PLN [21] que puedan ser combinadas con las de aprendizaje automático para facilitar la incorporación de contenido sintáctico y/o semántico en la creación de ejemplos de forma automática. Este es el objetivo de una tesina de licenciatura que está comenzando a desarrollarse.

Otro enfoque que está comenzando a estudiarse es el papel que juegan los distintos usuarios activos dentro de un foro (los que participan compartiendo opiniones y experiencias). Dicho análisis se está realizando para incorporarlo como posible mejora del recomendador de hilos de foros de discusión deseado. Bajo esta premisa, se han estudiado diferentes propuestas [20] [23] [22] y se está trabajando en una tesis en curso, a partir de una estrategia empírica basada en la observación de hilos de discusión reales obtenidos de la web.

4. FORMACION DE RECURSOS HUMANOS

El proyecto avanza en la línea del proyecto comenzado en 2013, el cual tenía como objetivo definir un modelo de calidad a partir de información contenida en foros de discusión técnicos.

El proyecto actualmente se encuentra conformado por un grupo de docentes, asesores y alumnos desarrollándose en las áreas de Ingeniería en Sistemas, Programación y Teoría de la Computación, trabajando en forma colaborativa e interdisciplinaria.

Las personas que colaboran, asesoran y forman parte del proyecto son:

La conformación del equipo con docentes de distintos departamentos, sumado a la asesoría externa mencionada, permite el trabajo cooperativo de un grupo interdisciplinario. Además, la incorporación de estudiantes de la Facultad amplia los posibles tipos de desarrollo relacionados a la temática del proyecto.

5. BIBLIOGRAFIA

1. ISO/IEC 25012:2008, Software product Quality Requirements and Evaluation (SQuaRE): Data quality model. 2008.
2. W. Chen, R. Persen (2009), "A Recommender System for Collaborative Knowledge".
3. C. Calero, A. Caro, M. Piattini (2008), "An Applicable Data Quality Model for Web Portal Data Consumers", World Wide Web, vol. 11, no. 4, pp. 465-484.
4. D. Helic, N. Scerbakov (2003), "Reusing Discussion Forums as Learning Resources in WBT Systems".
5. I. Rafique et al(2012), "Information Quality Evaluation Framework: Extending ISO 25012

- Data Quality Model”, *International Journal of Computer and Information Sciences*, vol.6.
6. R. Wang, D. M. Strong (1996), “Beyond accuracy: What data quality means to data consumers”, *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-33.
 7. Smith y Duffy (2001), Re-using knowledge: why, what and where. En *Proceedings de 2001 International Conference on Engineering Design*, Glasgow.
 8. P. Di Maio (2009), Toward Pragmatic Dimensions of Knowledge Reuse and Learning on the Web. *Proceedings of I-KNOW’09 and I-SEMANTICS’09*, Graz, Austria.
 9. G. Aranda, N. Martínez Carod, P. Faraci, A. Cechich. *Hacia un framework de evaluación de calidad de información en foros de discusión técnicos*. ASSE 2013,
 10. A. Tigelaar, R. Op Den Akker and D. Hiemstra, *Automatic summarisation of discussion fora*, *Natural Language Engineering*, ISSN 1469-8110, Vol 16, Issue 02, pp. 161-192, 2010.
 11. H. Kuna, et al. , *Generación de un Algoritmo de Ranking para Documentos Científicos del Área de las Ciencias de la Computación*, , CACIC 2013, XIX pp. 787-796, 2013.
 12. Singhal,. *Modern information retrieval: A brief overview*.IEEE Data Eng. Bull., 2001, vol. 24, no 4, p. 35-43
 13. N.Martínez Carod et al. *Análisis de la información presente en foros de discusión técnicos*. In CACIC 2013, pp. 847- 856, 2013.
 14. G. Aranda, N. Martínez-Carod, S. Roger, P. Faraci, and A. Cechich. *Una herramienta para el análisis de hilos de discusión técnicos*. In CACIC 2014, pages 803 - 812, Oct. 2014.
 15. V. Zoratto, G. Aranda, S. Roger, A. Cechich, *Análisis de estrategias para clasificar contenidos en foros de discusión: Un caso de estudio* ASSE 2015, pp. 176-190.
 16. V. Zoratto, G. Aranda, S. Roger, A. Cechich, *Analyzing Discussion Forums Threads About Java Programming Language Usage*, *Electronic Journal of SADIO*, 2016 .ISSN (versión online): 1514-6774. En revisión. Publicación estimada Noviembre 2016.
 17. M. Nicoletti, S. Schiafino, and D. Godoy. *Mining interests for user profiling in electronic conversations*. *Expert Syst. Appl.* , 40(2):638-645, Feb. 2013.
 18. I. Witten, E. Frank and M. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier. 2011
 19. Bing Liu. *Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data*. Springer. 2008
 20. M. Lui and T. Baldwin. *Classifying user forum participants: Separating the gurus from the hacks, and other tales of the internet*. In *Proceedings of Australasian Language Technology Association Workshop* , pages 49-57, 2010.
 21. C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA, USA, 1999.
 22. T. Hecking, I. Chounta, and H. U. Hoppe. *Investigating social and semantic user roles in MOOC discussion forums*. In LAK, pages 198-207. ACM, 2016.
 23. S. Bhatia and P. Mitra. *Classifying user messages for managing web forum data*. In Z. G. Ives and Y. Velegrakis, editors, *WebDB* , pages 13-18, 2012
 24. G. A. Miller, R. Beckwith, C. D. Fellbaum, D. Gross, K. Miller. 1990. *WordNet: An online lexical database*. *Int. J. Lexicograph.* 3, 4, pp. 235–244.
 25. A. Gangemi, R. Navigli, P. Velardi. *The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet*, In *Proc. of ODBASE 2003*, Catania, Sicily (Italy), 2003, pp. 820–838.
 26. R. Navigli, S. P. Ponzetto. *BabelNet: Building a Very Large Multilingual Semantic Network*. *Proc. of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, Uppsala, Sweden, July 11–16, 2010, pp. 216–225.

Clasificación y Evaluación de Métricas de Mantenibilidad Aplicables a Productos de Software Libre

José M. Ruiz, Cristian D. Pacifico, Martín M. Pérez

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos
 Av. Tavella 1424, (E3202KAC) Concordia, Entre Ríos, Argentina – Tel: (0345) 423-1400
 {ruiz,cripac,marper}@fcad.uner.edu.ar

RESUMEN

Un producto software es “*software libre*” si cumple cierta forma de licenciamiento. Esta forma contempla determinadas libertades que, en teoría, puede ejercer el usuario. De estos derechos, uno significativamente importante es la libertad de *estudiar cómo funciona el programa y adaptarlo a sus necesidades*. Para ejercer esta libertad, no basta con que la licencia exija la disponibilidad del código fuente, sino que también se requiere que el software posea características internas que faciliten la legibilidad y modificación del código.

La “*Mantenibilidad*”¹ (*Maintainability*) es la facilidad con la que se modifica, mejora y/o adapta un producto software. Esta característica es identificada y definida por normas de calidad ampliamente aceptadas, que recomiendan establecer métricas para su evaluación.

La línea de I+D presentada aquí, pretende evaluar y clasificar métricas para la “*mantenibilidad*” aplicables a productos software libre y desarrollar herramientas que las implementen, con el objetivo de favorecer la comprensión y modificación del código.

Palabras clave: Software Libre, métricas, mantenibilidad, calidad.

CONTEXTO

La presente investigación surge en el marco del proyecto de investigación y desarrollo **PID-UNER 7049** “*Guías para aplicación de Normas de Calidad para los procesos de Ingeniería de Software en productos desarrollados con Lenguajes de Programación Open Source*”. El Proyecto de I+D es llevado a cabo en la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos. Esta unidad académica, mediante varias acciones llevadas en conjunto con cámaras empresariales, empresas y desarrolladores, apoya y fomenta el desarrollo e implementación de Software Libre. Este proyecto de I+D se establece como ámbito científico-académico para el desarrollo y redacción de guías y normas para certificación de productos y procesos de Software Libre y Tecnologías Abiertas.

1. INTRODUCCIÓN

Software Libre

Un producto software es considerado “*software libre*” si cumple cierta forma de licenciamiento. Esta forma o “*licencia*”

¹El término original en inglés es “*Maintainability*”. La traducción al español de la norma ISO/IEC 25010 es el término “*mantenibilidad*”, y en la norma ISO/IEC 9126 se utiliza la frase “*facilidad de mantenimiento*”. El presente trabajo utiliza la traducción adoptada por la ISO/IEC 25010.

contempla determinadas libertades que, en teoría, puede gozar el usuario. Estas libertades y derechos son establecidas por elección manifiesta de su autor, en tanto que el producto puede ser copiado, estudiado, modificado, utilizado libremente con cualquier fin y redistribuido con o sin cambios o mejoras [1].

Estas libertades son formalizadas en una licencia que adopta un manifiesto de Software Libre. Una licencia es aquella autorización formal con carácter contractual que un autor de un software da a un interesado para ejercer “*actos de explotación legales*” [2].

El manifiesto más utilizado en proyectos de este tipo es la **Licencia Pública General** de proyecto GNU, o **GNU GPL** (General Public License). La licencia GPL estipula los criterios que se tienen que cumplir para que un programa sea considerado libre. En concreto las cuatro libertades que definen al Software Libre son [3]:

- La libertad de ejecutar el programa para cualquier propósito.
- La libertad de estudiar cómo funciona el programa y de adaptarlo a sus necesidades.
- La libertad de redistribuir copias, para que pueda ayudar al prójimo.
- La libertad de mejorar el programa y poner las mejoras a disposición del público, para que toda la comunidad se beneficie.

Con estas libertades, los usuarios (personas, organizaciones, compañías) controlan el programa/producto y lo que éste hace. Estas libertades son derechos, no obligaciones, cada persona puede elegir no usarlas, pero también puede elegir usar todas ellas.

Cabe destacar que aceptar las libertades del

Software Libre no excluye de su uso comercial. Un programa libre debe estar disponible para uso comercial, desarrollo comercial y distribución comercial. El desarrollo comercial del Software Libre ha dejado de ser inusual [4] [5].

Si una empresa es productora de productos de software libre (o los patrocina), busca lograr los beneficios inherentes al modelo de ciclo de vida propuesto por la filosofía del *open source*: lograr que su producto sea ampliamente usado, recibir reportes de errores y modificaciones de su producto y establecer líneas de negocios de servicios de personalización, consultorías y capacitación para uso del producto [6] [7].

En este contexto, la libertad de modificar el código fuente no se limita al derecho del usuario, sino a que es ejercicio fundamental para que el proyecto de software libre perdure y se desarrolle. Por eso, no basta con cumplir con la exigencia de disponibilidad del código fuente inferida en la licencia, sino que también hace falta que el producto posea calidades internas que favorezcan y faciliten la legibilidad y modificación del código.

Mantenibilidad y Mantenimiento del Software

El estudio y la definición de los factores o características que afectan la calidad de un software es una de las áreas más desarrolladas en la Ingeniería de Software [8]. De hecho, se han establecido estándares para fijar las normas de calidad.

Pressman [8] describe la calidad del software como un “*proceso eficaz de software que se aplica de manera que crea un producto útil que proporciona valor medible a quienes lo producen y a quienes lo utilizan*”.

La publicación de la norma ISO 9126 [9], en el año 1991, constituye un hito en la definición de estándares de calidad de producto software. Luego, en el año 2001,

fue reemplazado por dos estándares relacionados: el ISO/IEC 9126 [10], que describe las particularidades de un modelo de calidad del producto software; y el estándar ISO/IEC 14598 [11], que abordaba el proceso de evaluación de productos software.

El modelo más actual está representado por la familia de normas ISO/IEC 25000 [12], que es el resultado de la evolución de las normas citadas más arriba. La ISO/IEC 25000 conocida como SQuaRE (*System and Software Quality Requirements and Evaluation*), es una familia de normas que tiene por objetivo la creación de un marco de trabajo común para evaluar la calidad del producto software.

El *modelo de calidad de producto* definido por la ISO/IEC 25010 está compuesto por ocho características de calidad: adecuación funcional, eficiencia de desempeño, compatibilidad, usabilidad, fiabilidad, seguridad, **mantenibilidad** y portabilidad [13].

El IEEE² define **mantenibilidad** como: “*La facilidad con la que un sistema o componente software puede ser modificado para corregir fallos, mejorar su funcionamiento u otros atributos o adaptarse a cambios en el entorno*” [14]. Esta definición está conectada directamente con la definición del IEEE para **mantenimiento** del software: “*es el proceso de modificar un componente o sistema software después de su entrega para corregir fallos, mejorar su funcionamiento u otros atributos o adaptarlo a cambios en el entorno*”.

Según el modelo de la ISO/IEC 25000, “**Mantenibilidad**” (*Maintainability*) o **Facilidad de Mantenimiento** es la capacidad del producto software para ser modificado efectiva y eficientemente, debido a

necesidades evolutivas, correctivas o perfectivas. Esta característica se subdivide a su vez en cinco subcaracterísticas aplicables a un producto software: Modularidad, Reusabilidad, Capacidad para analizarse, Capacidad para modificarse, y Capacidad para probarse.

Estudios previos señalan al mantenimiento posterior como la fase que más recursos requiere a lo largo del ciclo de vida del producto, dos veces superior a los costos de desarrollo [8]. A pesar que el modelo SQuaRE no especifica un conjunto de métricas para evaluar la mantenibilidad de productos software, sí establece las características deseables de dichas mediciones. Tienen que ser objetivas, independientes y reproducibles, expresadas por medio de escalas válidas y suficientemente precisas para apoyar comparaciones fiables.

La **mantenibilidad** debe establecerse como objetivo tanto en las fases iniciales del ciclo de vida, para reducir las posteriores necesidades de mantenimiento, como durante las fases de mantenimiento, para reducir los efectos laterales [15]. Existen unos pocos factores que afectan directamente la **mantenibilidad** de un producto. Los tres más significativos son:

Proceso de desarrollo: la **mantenibilidad** debe formar parte integral del proceso de desarrollo del software y ser uno de sus principios rectores.

Documentación: En múltiples ocasiones, ni la documentación ni las especificaciones de diseño están disponibles y, por tanto, los tiempos y costos de comprensión, corrección y/o modificación se incrementan.

Comprensión de Programas: La causa básica de los altos costos del mantenimiento es la presencia de obstáculos a la

²IEEE: Institute of Electrical and Electronics Engineers.

comprensión del funcionamiento del producto.

Para abordar el concepto de métrica en ingeniería de software es necesario mencionar el concepto de medición, según Fenton: “*La medición es el proceso mediante el cual se asignan números o símbolos a los atributos de las entidades en el mundo real, de manera que se les define de acuerdo con reglas claramente determinadas*” [16]. El IEEE define como métrica a “*una medida cuantitativa del grado en que un sistema, componente o proceso posee un atributo dado*” [14].

Existe la necesidad de medir y de controlar la complejidad del software, y si bien un solo valor de esta métrica de calidad es difícil de derivar, es posible desarrollar medidas de diferentes atributos internos de programa (modularidad efectiva, independencia funcional, entre otros). Estas medidas y las métricas derivadas de ellas brindan un mecanismo consistente y objetivo para valorar la calidad [8].

Se han propuesto métricas diseñadas explícitamente para mantenibilidad que deberían afectar las actividades de mantenimiento. El IEEE. 982.1 [17] sugiere un índice de madurez de software (IMS) que proporcione un indicio de la estabilidad de un producto de software (con base en cambios que ocurran para cada liberación del producto).

Otros estudios [18, 19, 20] han puesto de manifiesto la necesidad de contar con mediciones de la mantenibilidad del producto, en especial si éste se trata de software libre.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo pretende formalizar procedimientos y normas de aplicabilidad para realizar mediciones orientadas a la mantenibilidad de productos de software libre, permitiendo que las métricas obtenidas ajusten la calidad interna del código fuente con el objetivo de facilitar y asegurar su libertad de modificar el producto.

Puntualmente, este trabajo propone realizar las siguientes tareas:

1. Evaluar de las principales normas y buenas prácticas referidas a la **mantenibilidad** de software, en especial los casos de estudios aplicados a software.
2. Clasificar los patrones de medición estudiados, realizando una valoración individual de su impacto en la aplicación de las libertades de Software Libre.
3. Diseñar un núcleo de mediciones/métricas recomendables para aplicar sobre **mantenibilidad**, definiendo su alcance y aplicación,
4. Desarrollar herramientas para la recolección de métricas de **mantenibilidad** en base al diseño anterior.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los objetivos de este trabajo son:

- Lograr un estudio comparativo de mediciones de mantenibilidad teniendo en cuenta el impacto en productos de software libre.
- Obtener una guía para poder aplicar métricas de mantenibilidad a proyectos de software libre.
- Incorporar la guía y herramientas

desarrolladas como guías de aplicaciones de normas de calidad específicas para software libre.

4. FORMACIÓN DE RECURSOS HUMANOS

Este trabajo forma parte de 1 (un) proyecto de tesis de la Maestría en Sistemas de Información. Adicionalmente se prevé contemplar, para los detalles de implementación de herramientas, la realización de 1 (un) Trabajo Final –tesina de grado– para la Licenciatura en Sistemas, ambas carreras dictadas en de la Facultad de Ciencias de la Administración de la UNER

BIBLIOGRAFÍA

- [1] R. M. Stallman, *Free Software, Free Society: Selected Essays*, Boston, Massachusetts, MA: GNU Press, 2002.
- [2] F. y. F. H. da Rosa, *Guía práctica sobre software libre: su selección y aplicación local en América Latina y el Caribe*, Montevideo - Uruguay: UNESCO, 2007.
- [3] Fundación del Software Libre de Europa, [En línea]. Available: <https://fsfe.org/about/basics/freesoftware.es.html>.
- [4] T. Wasserman, *Building a Business on Open Source Software*, 2009.
- [5] P. y. A. N. Cooper, «Free and open source software business and sustainability models. OSS Watch,» 2014. [En línea]. Available: <http://oss-watch.ac.uk/resources/businessandsustainability>.
- [6] M. Ferris, «Open source code and business models: More than just a license. OpenSource.com,» 2014. [En línea]. Available: <http://opensource.com/business/13/5/open-source-your-code>.
- [7] L. Ibañez, «Open source economic model: Sell the license or charge a consulting fee? OpenSource.com,» 2013. [En línea]. Available: <http://opensource.com/education/13/2/open-source-economic-model>.
- [8] R. S. Pressman, *Ingeniería del Software. Un enfoque práctico*. 7 Ed. ISBN: 978-607-15-0314-5, 2005 ed., MCGRAW-HILL, 2010.
- [9] ISO 9126, «Software Quality Model,» International Organization for Standardization, Geneva, Switzerland, 1991.
- [10] ISO/IEC 9126, «Software engineering - Product quality» International Organization for Standardization, Geneva, Switzerland, 2001.
- [11] ISO/IEC 14598-6, «Software engineering - Product evaluation -,» International Organization for Standardization, Geneva, Switzerland, 2001.
- [12] ISO/IEC 25000, «Systems and software engineering – Systems and software product Quality Requirements and Evaluation (SQuaRE). Guide to SQuaRE,» International Organization for Standardization, Geneva, Switzerland, 2014.
- [13] ISO/IEC 25010, «Systems and software engineering – Systems and software product Quality Requirements and Evaluation (SQuaRE) – System and software quality models.,» International Organization for Standardization, Geneva, Switzerland, 2014.
- [14] Engineers, Institute of Electrical and Electronics, «610.12 - IEEE Standard Glossaries of Software Engineering Terminology,» IEEE Computer Society Press, New York, NY. USA, 1990.
- [15] B. W. Boehm, J. R. Brown y M. Lipow, *Quantitative evaluation of software quality*, Los Alamitos, CA, USA: IEEE Computer Society Press, 1976.
- [16] N. E. Fenton y S. L. Pfleeger, *Software Metrics: A Rigorous and Practical Approach*, Boston, MA, USA: PWS Publishing Co., 1997.
- [17] Engineers, Institute of Electrical and Electronics, «982.1 - IEEE Standard Dictionary of Measures of the Software Aspects of Dependability,» IEEE Computer Society Press, New York, NY. USA, 1988.
- [18] Y. & X. B. Zhou, *Predicting the maintainability of open source software using design metrics*, In Wuhan Univ.: J. Nat Sci., 2008, pp. 13:p 14-20.
- [19] S. R. S. a. K. C. L. Yu, «Measuring the maintainability of open-source software,» de *International Symposium on Empirical Software Engineering*, 2005.
- [20] B. G. J. H. H. E. S. M. Smit, «Maintainability and Source Code Conventions: An Analysis of Open Source Projects,» University of Alberta, 2011.

Coordinación de Dispositivos en Ambientes Ubicuos Mediante Coreografías

Germán Montejano^{1,2}; Oscar Testa²; Rubén Pizarro²; Darío Segovia²; Oscar Dieste³; Efraín R. Fonseca C.⁴

¹Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-2652-424027 – Int. 251

gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

²Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-425166 – Int. 28

[otesta, ruben]@exactas.unlpam.edu.ar

³Facultad de Informática
Universidad Politécnica de Madrid
Campus Montegancedo – (28660) Boadilla del Monte – Madrid – España
Tel.: +34 91 336 5011
odieste@fi.upm.es

⁴Departamento de Ciencias de la Computación
Universidad de las Fuerzas Armadas ESPE
erfonseca@espe.edu.ec

Resumen

Actualmente nos encontramos involucrados en ambientes donde los dispositivos ubicuos forman parte de nuestra vida cotidiana y de nuestras tareas diarias. En casi todos los casos, los dispositivos ubicuos no proporcionan servicios de forma aislada, sino que deben cooperar con otros dispositivos. La necesidad de desarrollar sistemas donde una multiplicidad de dispositivos ubicuos se coordinen entre ellos para lograr un fin no es sólo un problema académico, responde también a necesidades de la industria. La computación orientada a servicios proporciona mecanismos para la composición de servicios que permiten construir sistemas de negocio complejos y aplicaciones a partir de una gran cantidad de servicios heterogéneos, simples y distribuidos. Las similitudes entre la composición de servicios y la coordinación de dispositivos ubicuos es sorprendente. Si cada dispositivo ubicuo en un ambiente pervasivo es proveedor, o consumidor de un servicio, la coordinación de dispositivos se ajusta con la composición de servicios en ambientes distribuidos.

Los mecanismos de composición establecidos para servicios web no son directamente aplicables. La utilización de un estándar ya existente en ambientes de internet puede facilitar la adopción de la propuesta en la práctica.

Nuestra propuesta es adaptar y aplicar las especificaciones de coreografías actualmente existentes en SOA para la coordinación de servicios proporcionados por dispositivos ubicuos.

Palabras clave: sistemas ubicuos, SOA, servicios, composición de servicios.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software, Conceptos, Métodos y Herramientas en un Contexto de “Ingeniería de Software en Evolución” – Facultad de Ciencias Físico- Matemáticas y Naturales, Universidad Nacional de San Luis y en el Proyecto de Investigación: Ingeniería de Software: composición de servicios en ambientes ubicuos – Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa.

Las líneas aquí presentadas actualmente forman parte de las bases de un anteproyecto de tesis doctoral y de una tesis de maestría, ambas en ingeniería de software por la Universidad Nacional de San Luis.

Introducción

En la actualidad, las actividades cotidianas del hombre se han hecho dependientes de una gran cantidad de dispositivos electrónicos tales como: ordenadores personales, ordenadores portátiles, teléfonos móviles, PDAs, tabletas, sensores de muchas y diversas utilidades, entre otros; los cuales logran comunicarse entre sí gracias a diversos protocolos de comunicación inalámbrica, redes de celulares, redes de área local (LAN), redes de área extensa (WAN), Bluetooth, etc. Estamos en la presencia de nuevos dispositivos de comunicación, lo que conlleva un nuevo escenario social, donde la interacción permanente con estos elementos es ineludible. Por ejemplo, la cantidad de móviles existentes en el mercado se aproxima a la cantidad de habitantes mundiales, según un informe de la Unión Internacional de Comunicaciones. En este informe, se estima que hasta finales del año 2014 hubo casi 7.000 millones de suscripciones de telefonía celular, lo que corresponde al 96 % de la población global; es decir, la cantidad de usuarios de telefonía móvil se acerca al número de personas que viven en el planeta [1]. Los avances de las comunicaciones entre dispositivos ha permitido que estos sean generadores y consumidores de servicios al mismo tiempo, es decir, de acuerdo a las capacidades del dispositivo puede no solo obtener, sino también ofrecer a otros equipos sus funciones y así cooperar entre ellos. La tendencia actual es hacia los ambientes ubicuos, los cuales se caracterizan por estar poblados de numerosos dispositivos que, gracias a la integración extrema de los elementos electrónicos, son invisibles al usuario y están en permanente rastreo de la actividad humana [2].

Dispositivos ubicuos son todos aquellos dispositivos que pueden existir en todas partes,

es decir, son dispositivos electrónicos que tienen capacidad de procesamiento y comunicación y pueden ser encontrados en lugares diversos de la vida cotidiana.

La computación ubicua es un desarrollo tecnológico que intenta que las computadoras no se perciban en el entorno como objetos diferenciados, y que la utilización por parte de los seres humanos sea lo más transparente y cómoda posible, facilitando de esta manera la integración en la vida cotidiana. Desde hace varios años los dispositivos ubicuos han ganado importancia y presencia en la vida cotidiana de las personas, debido principalmente a que: poseen distintos tipos de sensores (posicionamiento, proximidad, luminosidad, temperatura, etc.), facilitan la conectividad incluso en áreas con poca señal o acceso a las redes, permiten la convergencia tecnológica (computo, medios, telefonía, etc) y brindan acceso a servicios de distinta índole (mapas, ayudas, etc).

Por composición entendemos la forma en que se pueden combinar o enlazar un número indeterminado de dispositivos para llevar adelante una tarea determinada. En ambientes ubicuos, la composición de dispositivos, presenta nuevos desafíos tales como: la heterogeneidad (ya sea por la diversidad de dispositivos involucrados, como por la presencia de dispositivos de varios fabricantes), las contingencias de los dispositivos y la personalización de los mismos (por ej. provisión de servicios de acuerdo a las preferencias del usuario). Dado que los dispositivos en donde los servicios son ejecutados poseen limitaciones de recursos (ej. poca memoria y batería), se deben hacer consideraciones especiales respecto a la eficiencia y rendimiento de la composición de servicios [3].

La composición en este tipo de ambientes implica que los dispositivos deben dialogar entre ellos para poder compartir los servicios que ofrecen con la finalidad de obtener un servicio con valor agregado, o bien para abordar la solución de una problemática particular, como podría ser la seguridad de un hogar, o la seguridad vial, por mencionar

algunos ejemplos.

Si bien hoy en día podemos decir que distintos sensores o dispositivos se pueden comunicar entre ellos, compartiendo de alguna manera sus servicios, generalmente lo realizan a partir de protocolos propietarios y sin seguir definiciones estándares, provocando que otros componentes de otros proveedores (o incluso de los mismos) no puedan ser utilizados. Esto obviamente representa una importante limitación en la composición de dispositivos ubicuos. Adicionalmente la composición de

dispositivos ubicuos presenta un nuevo desafío. Los mecanismos de composición en ambientes masivos, necesitan hacer frente a las distintas contingencias que pueden ocurrir con estos dispositivos. Los dispositivos ubicuos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas variantes hacen que la composición de dispositivos¹ ubicuos se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

Líneas de Investigación y Desarrollo

La computación orientada a servicios, y en particular los servicios web en ambiente de internet, proporcionan mecanismos para la composición de servicios. Dichos mecanismos, como las orquestaciones, son aspectos bien conocidos de la computación orientada a servicios que permiten construir sistemas de negocio complejos y aplicaciones a partir de una gran cantidad de servicios heterogéneos, simples y distribuidos. Podría pensarse que son aplicables a ambientes ubicuos. Sin embargo, en contextos como puede ser la Internet de las

Cosas (IoT) donde los servicios son dinámicos, móviles, menos fiables y dependientes del dispositivo, los mecanismos de composición establecidos para servicios web no es directamente aplicable [8].

Adicionalmente la composición de múltiples dispositivos ubicuos presenta nuevos desafíos que no son compatibles con la composición de servicios web. En particular, los mecanismos de composición en ambientes masivos como lo es el de dispositivos móviles, necesita hacer frente las distintas contingencias que pueden ocurrir con estos elementos, así como también contemplar la heterogeneidad de los mismos.

Estos dispositivos tienen distintas limitantes como son la cantidad de memoria disponible, la durabilidad de la batería, la disponibilidad de acuerdo a la red del lugar donde se encuentre en un momento determinado. En ambientes ubicuos, la disponibilidad y confiabilidad de los dispositivos no puede ser garantizada. Todas estas dificultades hacen que la composición de dispositivos se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

Finalmente existen distintos proyectos en la actualidad donde se intenta integrar sensores y dispositivos ubicuos a la vida cotidiana. Específicamente podemos mencionar la domótica, donde varios dispositivos y sensores deben actuar en coordinación para prevenir un incidente de seguridad (ya sea por robo o por incendio) en nuestros hogares. Sin embargo, existen áreas de aplicación más relevantes.

En la industria, existe lo que se llama Industria 4.0 [9], donde lo que se intenta es integrar dentro de una planta fabril la intercomunicación de todos los dispositivos que componen la cadena de producción con el fin de que coordinen entre ellos las tareas a realizar en base a los tiempos a cumplir, stocks disponibles, demanda en línea de los productos, etc. Otra área donde los dispositivos ubicuos están ganando importancia es la automotriz, donde los esfuerzos se enfocan en que distintos sensores monitoreen funciones vitales del conductor

¹ Si bien los autores se refieren a la composición de servicios, se hace dentro de un contexto de dispositivos ubicuos, lo cual a los fines de este trabajo se puede interpretar como composición de dispositivos, haciendo que la terminología para este caso particular sea más adecuada.

(como es el caso de presión arterial, pulsaciones, etc) y en caso de que detecten anomalías actúen en conjunto con otros dispositivos del vehículo para evitar accidentes. Es claro que en este punto se hace necesaria una mayor investigación y desarrollo de tecnologías que permitan solucionar en todo o en parte estos desafíos planteados, haciendo foco en la composición de distintos dispositivos de una manera abierta y estándar.

Los mecanismos de composición en ambientes ubicuos como los dispositivos móviles, necesitan hacer frente a las distintas contingencias que pueden ocurrir con estos dispositivos, así como también contemplar la heterogeneidad de los mismos. La heterogeneidad no sólo se refiere a la existencia de dispositivos de distintos modelos, sistemas operativos y fabricantes, sino también a los mecanismos de comunicación e interacción que poseen los mismos, en algunos casos propietarios, provocando que la interacción y coordinación entre ellos representa un desafío de enorme magnitud. Estos dispositivos tienen, a su vez, limitantes adicionales como son la cantidad de memoria disponible, la durabilidad de la batería o la conectividad de acuerdo a la red del lugar donde se encuentre en un momento determinado. Todas estas dificultades hacen que la composición de servicios incluyendo dispositivos móviles se transforme en un área de investigación muy importante donde los avances no han sido claros al día de hoy[3].

A medida que los dispositivos ubicuos son menos potentes (ej: cámaras de seguridad, sensores, etiquetas RFID, etc), la disponibilidad y confiabilidad de los mismos no puede ser garantizada. En este tipo de ambientes, mecanismos automáticos y dinámicos son necesarios para la composición de dispositivos, ya que de esta forma se puede compensar la falta de disponibilidad de un dispositivo en un momento determinado[8].

Resultados y Objetivos

Por lo expuesto, vemos que existe un campo de trabajo importante en el desarrollo de composición de servicios en ambientes ubicuos, más precisamente en la coreografía de servicios, la cual no es abordada en los estudios previos de la materia.

Por todo esto, nuestra propuesta es poder adaptar y aplicar las especificaciones actualmente existentes en SOA para la coordinación de servicios disponibles en ambientes pervasivos a través de la utilización de dispositivos ubicuos, más concretamente, el objetivo de esta investigación es:

Definir un mecanismo de coordinación de dispositivos ubicuos que garantice su interoperabilidad independientemente del modelo y fabricante del mismo; utilizando los estándares de SOA y de coreografías para la composición de servicios.

Debemos destacar que la aplicación de los conceptos de SOA a dispositivos ubicuos no consiste en una mera traslación de los conceptos de un ambiente a otro, sino que será necesario para ello extender las especificaciones de SOA existentes de modo que se adapten a las circunstancias particulares de los sistemas ubicuos. Asimismo se deberá mantener total compatibilidad con las especificaciones relacionadas a SOA y coreografía de servicios existentes.

Hasta el momento, en la investigación hemos realizado el análisis del estado del arte, que ha dado las bases para poder enfocar claramente el resto de las etapas de la investigación. Se está trabajando actualmente en la simulación de un ambiente pervasivo, donde se pueda llevar adelante una coreografía conformada a través del estándar de WS-CDL. En este punto nos encontramos tratando de resolver de qué manera se puede saber a través de la especificación de la coreografía, en qué lugar exacto se encuentra el dispositivo para realizar la tarea que le fue encomendada. Para ello se están analizando y estudiando otros paradigmas de programación distribuida para poder ver de qué manera se maneja este

concepto, como lo es JADE.

Podemos, entonces, concluir que la composición y coordinación de dispositivos ubicuos en ambientes pervasivos está sometido a varios desafíos. Por otro lado hemos visto que la arquitectura SOA posee la potencialidad necesaria para responder a los desafíos que mencionamos, con el agregado de que nos brinda la ventaja de tratarse de una tecnología conocida y afianzada en el ambiente de sistemas distribuidos, con estandarizaciones de mucha índole y en los campos más importantes de la materia y que a su vez cuenta con una amplia base de lenguajes, frameworks y herramientas disponibles en el mercado, de utilización directa.

Nuestra investigación, a través del objetivo planteado, permite trasladar progresivamente los conceptos de SOA a dispositivos ubicuos en ambientes pervasivos, permitiendo además conocer y estudiar las ventajas e inconvenientes de la aplicación de la arquitectura de servicios en dichos ambientes.

Formación de Recursos Humanos

Además de los resultados obtenidos/esperados en el punto 3, se espera como resultado en la formación de recursos humanos, la continuación de esta misma línea de proyecto como tesis doctoral de alguno(s) de los investigadores. También se espera lograr una mayor interrelación con la Universidad de Minas Gerais con la que se cuenta con un convenio con tal objetivo como parte de él. Se espera avanzar también en un convenio de colaboración con la Universidad Politécnica de Madrid para la aplicación de las metodologías aquí presentadas en los proyectos de Ingeniería de Software Empírica. Adicionalmente, se espera que otras tesis de Maestría, así como tesinas de Licenciatura surjan a partir de los logros obtenidos en la presente línea de investigación.

Bibliografía

- [1] U. I. d. T. (UIT), “Unión Internacional de Telecomunicaciones.” <https://www.itu.int/net/pressoffice/pressreleases/2014/23-es.aspx>, 10 2015.
- [2] M. Weiser, “Hot topics-ubiquitous computing,” *Computer*, vol. 26, pp. 71–72, Oct 1993.
- [3] Q. Z. Sheng, X. Qiao, A. V. Vasilakos, C. Szabo, S. Bourne, and X. Xu, “Web services composition: A decade’s overview,” *Information Sciences*, vol. 280, no. 0, pp. 218–238, 2014.
- [4] M. Viroli, “On competitive self-composition in pervasive services,” *Science of Computer Programming*, vol. 78, no. 5, pp. 556–568, 2013. Special section: Principles and Practice of Programming in Java 2009/2010 & Special section: Self-Organizing Coordination.
- [5] S. W. Loke, “Supporting ubiquitous sensor-cloudlets and context-cloudlets: Programming compositions of context-aware systems for mobile users,” *Future Generation Computer Systems*, vol. 28, no. 4, pp. 619–632, 2012.
- [6] F. Palmieri, “Scalable service discovery in ubiquitous and pervasive computing architectures: A percolation-driven approach,” *Future Generation Computer Systems*, vol. 29, no. 3, pp. 693–703, 2013. Special Section: Recent Developments in High Performance Computing and Security.
- [7] S. Najar, M. K. Pinheiro, and C. Souveyet, “A New Approach for Service Discovery and Prediction on Pervasive Information System,” *Procedia Computer Science*, vol. 32, pp. 421–428, 2014. The 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014).
- [8] G. Cassar, P. Barnaghi, W. Wang, S. De, and K. Moessner, “Composition of services in pervasive environments: A Divide and Conquer approach,” in *Computers and Communications (ISCC)*, 2013 IEEE Symposium on, pp. 000226–000232, July 2013.
- [9] Wikipedia, “Industria 4.0 — Wikipedia, La enciclopedia libre,” 2016. [Internet; descargado 4-noviembre-2016].
- [10] H.-I. Yang, R. Bose, A. (Sumi) Helal, J. Xia, and C. Chang, “Fault-Resilient Pervasive Service Composition,” in *Advanced Intelligent Environments* (A. D. Kameas, V. Callagan, H. Hagra, M. Weber, and W. Minker, eds.), pp. 195–223, Springer US, 2009.

Definición de una Arquitectura de Procesos Utilizando la Metodología BPTrends para la Aplicación del Ciclo de Vida BPM

Marisa Perez[†], Juan Pablo Ferreyra[†], Claudia Verino[†], Diego Cocconi[†]

Departamento de Ingeniería en Sistemas de Información / Facultad Regional San Francisco / Universidad Tecnológica Nacional (UTN)
Av. de la Universidad 501, 2400, San Francisco, Córdoba, Argentina, (03564) 431019 / 435402
[†]{mperez, jpferreyra, cverino, dcocconi}@sanfrancisco.utn.edu.ar

Resumen

La gestión de procesos de negocio (BPM) ha ganado mucha popularidad en la actualidad. Se fundamenta en la representación explícita de los procesos de negocio para posibilitar la aplicación de un ciclo de mejora continua en el que intervienen las fases de diseño y análisis, configuración, ejecución y evaluación. El propósito de este trabajo es remarcar la importancia de la fase de diseño y análisis, como entrada a las demás fases, y fundamentalmente a la fase de evaluación, para obtener valiosa realimentación que será utilizada en la mejora de procesos. El principal objetivo de la línea de investigación propuesta es la definición de una arquitectura de procesos que revele la dinámica del área donde se va a aplicar e identifique defectos del modo de trabajo actual. Una vez alcanzado este objetivo se generará la configuración adecuada para la ejecución de los procesos y se realizará el correspondiente monitoreo, obteniendo información para proponer una nueva arquitectura que resolverá las falencias identificadas en la actual y será propuesta como mejora dentro del proyecto de investigación del cual surgió este estudio.

Palabras clave: gestión de procesos de negocio, mejora continua, arquitectura de procesos, BPTrends, sistemas de información.

Contexto

El presente trabajo forma parte del proyecto de investigación I+D UTN 4090 “*Optimización organizacional basada en un modelo de gestión por procesos en la Secretaría Extensión y Cultura de la Universidad Tecnológica Nacional, Facultad Regional San Francisco*”. El mismo está homologado como proyecto de investigación y desarrollo de la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional. En el marco de dicho proyecto se propone promover BPM para la mejora del desempeño de las organizaciones, orientándose a la gestión por procesos y específicamente para fomentar BPM en la institución con el fin de mejorar su eficiencia y rendimiento, y lograr la alineación de los procesos de negocio con sus metas y estrategias. Se ha comenzado por evaluar su aplicación en la Secretaría de Extensión y Cultura de la Facultad Regional San Francisco de la UTN.

1. Introducción

La *gestión de procesos de negocio* (BPM, del inglés *Business Process Management*) ha ganado mucha popularidad en la actualidad. BPM permite a las organizaciones administrar y optimizar sus procesos de negocio, de manera que los mismos evolucionen y conduzcan a una optimización organizacional

[1-2]. Además, BPM no se trata sólo de mejorar la forma en que se llevan a cabo actividades individuales, sino de la gestión de cadenas enteras de eventos, actividades y decisiones que agregan valor a la organización y sus clientes. Estas cadenas de eventos, actividades y decisiones se denominan *procesos de negocio* [3]. Un proceso de negocio consiste en un conjunto de actividades que se llevan a cabo de manera coordinada en un entorno organizacional [4].

En particular, BPM ha surgido de la evolución y el desarrollo de las arquitecturas de software y, por otro lado, de la administración de negocios. Esta última disciplina ha contribuido principalmente a través de dos grandes factores: (1) las *cadena de valor*, vistas como la descomposición funcional de los procesos de una organización para analizar su contribución a los objetivos de la misma; y (2) la *orientación por procesos* como forma de organizar las actividades. Por eso es importante conocer y tener en cuenta que BPM tiene una fuerte raíz en conceptos que no son exclusivos del software; la base administrativa permite obtener un mejor entendimiento de la forma en que la organización trabaja y cómo sus procesos aportan a sus objetivos. Una organización puede alcanzar sus objetivos de negocio de una manera eficiente y eficaz sólo si las personas y otros recursos organizacionales, como los sistemas de información, son alineados a los objetivos organizacionales. Cada proceso de negocio contribuye a alcanzar uno o más objetivos de la organización [4].

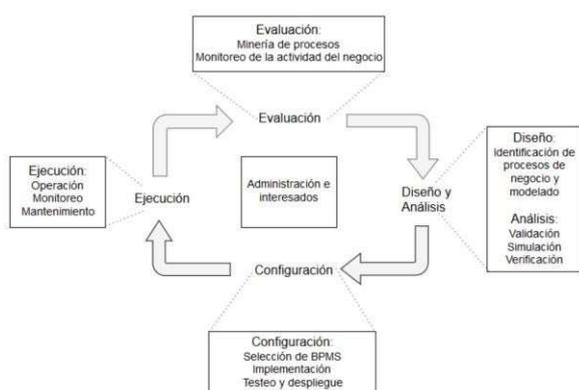


Fig. 1: Ciclo de vida de BPM [4].

BPM permite aplicar un ciclo de mejora continua que se conoce como *ciclo de vida de BPM* (Fig. 1), en el que intervienen las fases de *diseño y análisis, configuración, ejecución y evaluación*. En la fase de diseño y análisis se identifican los procesos de negocio de la organización, lo cual implica un análisis del estado actual de los procesos con el objetivo de detectar problemas existentes e identificar oportunidades de mejora para el (re)diseño de los mismos. Durante la fase de configuración, se especifican aspectos necesarios y se configuran para que los modelos de procesos puedan ser interpretados por un *sistema de gestión de procesos de negocio* (BPMS, del inglés *Business Process Management System*). En la fase de ejecución, el BPMS permite la ejecución de los procesos configurados. Finalmente, en la fase de evaluación se analiza el resultado de la ejecución para identificar problemas y aspectos que puedan ser mejorados [5].

El propósito de este trabajo es remarcar la importancia de la fase de diseño y análisis, ya que durante la misma se puede obtener y mejorar una *arquitectura de procesos* que sirve de entrada a las demás fases y al finalizar cada iteración, puede ser optimizada sobre la base de los resultados obtenidos en las otras etapas. Una *arquitectura de procesos* (o *mapa de procesos*) define “qué” hace y hará en el futuro la organización. La arquitectura de procesos es una colección de procesos de negocio y sus interdependencias entre ellos [6]; estos procesos deberían estar alineados a los objetivos de la organización. Las metodologías modernas reconocen la necesidad de trabajar a varios niveles, pero interconectados entre ellos, como en la pirámide que propone la metodología BPTrends (Fig. 2) [7].

La *metodología BPTrends* (del inglés, *BPTrends Process Change Methodology*) provee un modo de obtener la arquitectura de procesos. Esta metodología abarca dos métodos complementarios (Fig. 3): uno para el desarrollo de la arquitectura de procesos (*Metodología de la arquitectura de procesos de negocio BPTrends*) y otro para proyectos de rediseño de procesos de negocio

(Metodología de rediseño de procesos de negocio) [8].

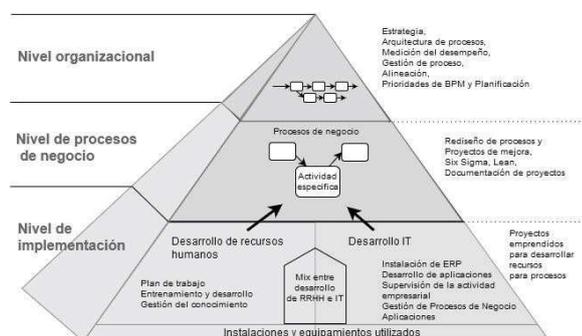


Fig. 2: La pirámide BPTrends [7].

La metodología de la arquitectura de procesos de negocio BPTrends comprende varias fases, que pueden apreciarse en la figura: *entender el contexto organizacional, modelar los procesos organizacionales, definir medidas de desempeño, definir gobernabilidad de los procesos, alinear la arquitectura*. En el presente trabajo el foco estará puesto en las primeras tres fases, las cuales centran su atención en el desarrollo de la arquitectura de procesos.

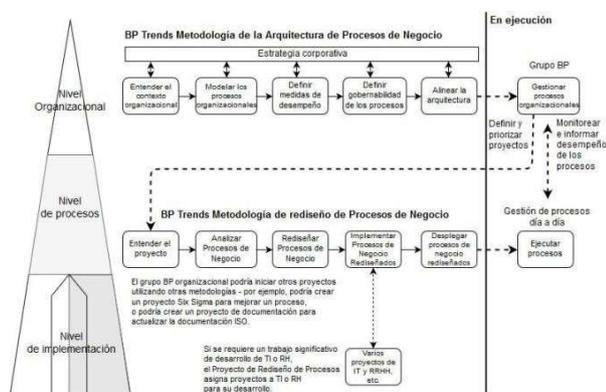


Fig. 3: Metodología BPTrends [8].

En la primera fase (*entender el contexto organizacional*) se intenta entender a la organización como un todo; comienza con un análisis de la organización para definir la estrategia, los objetivos y las relaciones claves, refinando el entendimiento de la organización y sus *stakeholders*, incluyendo clientes, proveedores, distribuidores y entidades gubernamentales. Las cadenas de valor son definidas, consiguiendo además un conocimiento sólido sobre los objetivos y

metas de cada una de ellas. La segunda fase (*modelar los procesos organizacionales*) se centra en definir la arquitectura de procesos de negocio de cada cadena de valor. Dependiendo de las necesidades de la organización, los recursos, políticas y reglas de negocio, recursos tecnológicos, recursos humanos, etc. podrán ser alineados a los procesos en la arquitectura. Finalmente, la tercera fase (*definir medidas de desempeño*) se centra en tomar la arquitectura y para cada proceso, definir cómo será monitoreado y medido; de este modo, se establecen los *indicadores clave de rendimiento* (KPI, del inglés *key performance indicators*), que son indicadores de alto nivel mediante los cuales los ejecutivos intentan monitorear y asegurar que los objetivos, estrategias e iniciativas relacionadas son obtenidas [8].

A modo de caso de estudio, se aplicará el enfoque descrito a la Secretaría de Extensión y Cultura de la Facultad Regional San Francisco. Una vez probada la efectividad del mismo, se propone y sugiere aplicarlo a otras áreas de la Facultad para optimizarlas y lograr la visión global de la institución completa.

2. Líneas de Investigación, Desarrollo e Innovación

Siguiendo la línea de investigación correspondiente a lo expuesto en este trabajo, se llevarán a cabo actividades relacionadas con las siguientes áreas temáticas:

- Gestión de procesos de negocio.
- Mejora de procesos.
- Metodologías de análisis y diseño de procesos.

3. Objetivos y Resultados Esperados

El principal objetivo de la línea de investigación propuesta es la definición de una arquitectura de procesos que revele la dinámica del área donde se va a aplicar e identifique defectos del modo de trabajo actual (cuellos de botella, retrabajos, etc.). Dicha arquitectura será el punto de partida

para comenzar con las fases de configuración y ejecución del ciclo de vida BPM, y será mejorada a partir de la información obtenida en la fase de evaluación.

Las organizaciones generalmente muestran su estructura mediante *organigramas*, en los que aparecen los diferentes grupos o departamentos, quién los gestiona y a quién deben reportar; si bien esta información es útil, hay información sobre la organización que el organigrama no brinda: los clientes, los productos o servicios ofrecidos, de dónde vienen los recursos, cuáles son o cómo se utilizan. Tampoco es posible saber cómo fluye el trabajo de una actividad a otra en el desarrollo de los productos o servicios. Las personas se ven influenciadas por este modelo vertical y tienden a pensar en términos del mismo, enfocándose en quién reporta a quién y estableciendo objetivos para cada departamento en forma independiente del resto. Para pensar en las organizaciones como un todo, las personas deben aprender a visualizarlas con modelos que permitan conocer cómo tales organizaciones funcionan en realidad, en términos de cadenas de valor en lugar de pensar en términos de divisiones, departamentos o su propia unidad funcional [8]. La arquitectura de procesos puede brindar esa vista global buscada.

Para alcanzar este objetivo, durante la fase de diseño y análisis del ciclo BPM, habrá que descubrir y documentar los procesos que son ejecutados, construir una taxonomía con los mismos e identificar sus debilidades. Surgirá entonces un modelo que documentará el estado actual de cada proceso (*cómo es hoy*, del inglés *As-Is*), que puede ser expresado en BPMN (del inglés, *Business Process Modeling Notation*), la notación estándar propuesta por OMG (2013) y ampliamente utilizada para modelar procesos de negocio [2]. Se pretende identificar objetivos, metas, responsables, entradas, salidas, métricas y actividades de cada proceso y, si es posible, cuantificar los problemas asociados a los mismos haciendo uso de los KPI; se priorizarán los problemas identificados en términos de impacto y esfuerzo necesario para resolverlos. A continuación, se considerarán

propuestas de mejora en aquellos aspectos que se crean necesarios; se identifican los cambios en los procesos que ayudarán a hacer frente a los problemas encontrados. Luego se realizará la mejora a los procesos pertinentes, modelando y documentando los procesos optimizados (modelo *cómo debería ser*, del inglés *To-Be*).

Durante la fase de configuración se procederá a obtener el esquema adecuado para la ejecución de los procesos mediante el modelado usando la notación BPMN. Para la generación de los procesos ejecutables se está evaluando el BPMS a utilizar, dentro de las alternativas conocidas (*Bizagi*, *Signavio*, *Bonitasoft*), valorando especialmente las que cuenten con versión libre (*opensource*), pues puede resultar útil la posibilidad de crear y/o ajustar conectores para integrar con otras aplicaciones.

En la fase de ejecución se pretende realizar el monitoreo adecuado de los procesos, para lo cual se utilizan las métricas definidas en la arquitectura de procesos (KPI). Se espera que a través de la herramienta seleccionada se puedan obtener datos útiles para la última fase del ciclo de vida, que implica la evaluación de los resultados para producir mejoras. A tal fin, se está evaluando utilizar técnicas de *minería de procesos* [9], lo cual permitirá mejorar la arquitectura realizando los ajustes necesarios.

Debido a la naturaleza iterativa que conlleva implícitamente la mejora continua de los procesos, necesaria para lograr la plena optimización organizacional se pretenden estructurar los resultados de la última fase de forma que sirvan como entrada para una nueva iteración.

4. Formación de Recursos Humanos

El grupo de esta línea de investigación está conformado por docentes y alumnos de la carrera de Ingeniería en Sistemas de Información. Todos los docentes se desempeñan en cátedras del área Sistemas de Información. Además tres de ellos se encuentran en la etapa de desarrollo de sus tesis de maestría (dos en Ingeniería en

Sistemas de Información y uno en Ingeniería de Software); otro está comenzando su tesis de doctorado (mención Ingeniería en Sistemas de Información) y la restante integrante está cursando una Diplomatura en BPM. Todos ellos con temas altamente vinculados al área del proyecto. Además un integrante ya ha presentado su trabajo final de Especialización en Ingeniería en Sistemas de Información en un área referida al tema. Como iniciativa del grupo, se prevé la capacitación y formación de recursos humanos, que contempla las siguientes actividades:

- Participación en cursos de actualización y posgrado en el área de estudio (procesos de negocio).
- Intercambio de ideas y conocimientos con docentes-investigadores de otras Facultades que trabajen en el área (Santa Fe, Mendoza, Rosario).
- Ofrecer cursos o charlas de capacitación para contribuir a la formación de otros docentes en la temática, para afianzar conocimientos de los alumnos y para extender en la comunidad científico-tecnológica, educativa y empresarial los conocimientos teóricos y prácticos logrados a medida que avanza el proyecto.
- Transferencia de tecnologías de procesos a otras áreas de la Facultad y a la industria local.
- Dirección y asesoramiento sobre el área a interesados de la industria local.
- Incorporar a las cátedras del tronco integrador de la carrera de Ingeniería en Sistemas de Información (Sistemas y Organizaciones, Análisis de Sistemas, Diseño de Sistemas y Administración de Recursos) contenidos sobre la gestión de procesos de negocio utilizando la experiencia y los conocimientos obtenidos del presente proyecto.
- Involucrar alumnos y graduados de la carrera de Ingeniería en Sistemas de Información en la realización de actividades del presente proyecto e incentivarlos para acercarse a propuestas de becas para investigación.

5. Referencias

- [1] Duipmans, E. F., Pires, L. F. y da Silva Santos, L. O. B. (2014). "A transformation-based approach to business process management in the cloud". *Journal of grid computing*, 12(2), 191-219.
- [2] Ferreyra, J. P. (2016). "Gestión de Procesos de Negocio basada en Computación en la Nube". Trabajo Final Integrador de Especialización. Santa Fe.
- [3] Dumas, M., La Rosa, M., Mendling, J. y Reijers, H. A. (2013). "Fundamentals of business process management" (Vol. 1, p. 2). Heidelberg: Springer.
- [4] Weske, M. (2012). "Business process management: concepts, languages, architectures", 2da. ed. Springer Publishing Company, Incorporated.
- [5] Van Der Aalst, W. M., Ter Hofstede, A. H. y Weske, M. (2003, Junio). "Business process management: A survey". En: *International conference on business process management* (pp. 1-12). Springer Berlin Heidelberg.
- [6] Eid-Sabbagh, R. H., Dijkman, R., y Weske, M. (2012, Septiembre). "Business process architecture: use and correctness". En: *International Conference on Business Process Management* (pp. 65-81). Springer Berlin Heidelberg.
- [7] Burlton, R. T. (2015). "Delivering business strategy through process management". En: *Handbook on Business Process Management 2* (pp. 45-78). Springer Berlin Heidelberg.
- [8] Harmon, P. (2014). "Business process change", 3era. ed. Morgan Kaufmann.
- [9] van der Aalst, W. M. P. (2010). "Process Mining. Discovery, Conformance and Enhancement of Business Processes". Springer.

Derivación de Casos de Prueba a Partir de Escenarios

Gladys Kaplan^{1,2}, Jorge Doorn^{2,3}, Walter Panessi¹, Claudia Ortiz¹, Eugenia Céspedes¹

¹Departamento de Ciencias Básicas, Universidad Nacional de Luján (UNLu)

²Departamento de Ingeniería e Innovación Tecnológica, Universidad Nacional de La Matanza (UNLaM)

³Escuela de Informática, Universidad Nacional del Oeste (UNO)

(gkaplan@unlam.edu.ar, jdoorn@exa.unicen.edu.ar, wpanessi@unlu.edu.ar, cortiz@unlu.edu.ar, eugeniacespedes@outlook.com)

Resumen

Como se describió en un artículo precedente [Kaplan 15], este proyecto tiene por objetivo concebir un mecanismo para generar tempranamente los casos de prueba partiendo del conocimiento obtenido en la etapa de Ingeniería de Requisitos (IR), particularmente del proceso de requisitos basado en escenarios [Leite 97] [Leite 04]. En dicha oportunidad se presentó la propuesta general del proyecto, mientras que en el presente artículo se profundiza la relación *escenarios futuros – casos de prueba* (EF-CP) y se describe un mecanismo de derivación automática. En muchos casos, la etapa de prueba es la variable de ajuste cuando los tiempos y el presupuesto de un proyecto están en crisis. Lograr que las pruebas, total o parcialmente, se realicen después de cada modificación asegura la calidad de los requisitos del software. Esto fuerza la necesidad de una creación, actualización y ejecución de los CP automática, con el objetivo de probar el software tantas veces como sea necesario. Que además permitan la mínima distracción de las etapas en desarrollo y a

un bajo costo es indispensable para que no se transformen en un elemento de perturbación. Cabe destacar que, al ser una actividad transversal en el proceso de construcción del software, estos CP evolucionan durante el diseño ya que es cuando se conoce en qué componente de software se alojará cada requisito (software allocation). Este nuevo conocimiento permite generar CP de caja blanca con un alto nivel de detalle.

Palabras clave: ingeniería de requisitos, casos de prueba funcionales, escenarios futuros.

Contexto

Este proyecto es una continuación de otros realizados en diferentes Universidades. Desde el año 1995 a la fecha se ha estudiado el proceso de requisitos basado en escenarios [Leite 04], utilizado en el presente artículo, donde se estableció una estrategia para realizar un proceso de requisitos a partir de una necesidad del usuario-cliente hasta obtener un documento de requisitos de software de alta calidad. En este proceso, se describen los modelos con sus

procesos de construcción. Los modelos utilizados son el Léxico Extendido del Lenguaje (LEL), los Escenarios (EA y EF), la Ficha de Información Extemporánea (FiE), el LEL de Requisitos (LEL_R) y la Especificación de Requisitos del Software (ERS).

Introducción

Como se mencionó en la sección anterior, el presente proyecto basa su estudio en el proceso de requisitos [Leite 04], el cual crea un glosario denominando LEL [Leite 90] [Hadam 08] del dominio de la aplicación. Durante la construcción del LEL y de un conjunto de escenarios actuales (EA) que representan situaciones observables, aparece la Información Extemporánea (FiE) que registra conocimiento que no puede ser representado en el modelo en construcción en ese momento. Con el LEL, los EA y las FiE se construye un conjunto de escenarios futuros [Doorn 02] que representan situaciones no observables del contexto futuro. Estos escenarios se validan y negocian hasta obtener la versión final que representan los procesos del negocio para cuando el sistema de software se encuentre en ejecución. Es de estos escenarios futuros de donde se extraen los requisitos del software.

Las pruebas construidas tempranamente en el proceso de construcción del software son del tipo funcional ya que verifican QUÉ se debe probar dejando para un momento posterior el CÓMO se debe hacer. Es así que las pruebas funcionales son del tipo caja negra,

mientras que las pruebas durante el diseño y del código son del tipo caja blanca.

Existen diversos trabajos relacionados con la generación de casos de prueba durante la etapa de IR a partir de un modelo existente. Entre los más conocidos están la generación de CP a partir de Casos de Uso [Riebisch 03] [Palacio 09] [Odalys 06] [Correa 11] [Gutiérrez 06], a partir de algoritmos metaheurísticos [Blanco 9] utilizando el lenguaje de transformación QVT, a partir de diagramas de secuencia extendidos [Lamancha 09], a partir de diagramas de actividad [Mingsong 09] [Linzhang 04] [Boghdady 12] [Boghdady 11], a partir de diagramas de secuencia [Javed 07], a partir de diagramas de estado [Badri 09], a partir de UML [Abbors 10] [Correa 12] [Wendland 11] [Salem 11], a partir de especificación de Requisitos [Hasling 08] entre otros.

Las pruebas son una etapa del proceso de construcción del software que determina, junto a las actividades de validación, la calidad del nuevo sistema de software a construir. Paradójicamente tienden a minimizarse cuando los tiempos o los costos de un proyecto entran en crisis debido a una mala estimación, por una pobre administración del proyecto o por la presencia de más cambios que los esperados. Este último caso, el de muchos cambios en los requisitos, aumenta considerablemente el costo del producto en construcción debido a la necesidad de actualizar constantemente los CP y como consecuencia, repetir todas las pruebas

para asegurar la consistencia de todo el software. Para evitar que las pruebas sean una actividad regulable según las necesidades del proyecto y no del producto, es que se hace indispensable automatizar la generación de los CP encontrando algún mecanismo de prueba lo más automático posible.

A lo largo del proyecto se espera probar la hipótesis de trabajo subyacente que considera que la automatización de la generación de los CP permite sostener en el tiempo la calidad del producto a construir.

Líneas de Investigación, Desarrollo e Innovación

Para alcanzar el objetivo final del proyecto se determinaron diferentes hitos. En una primera etapa se analizaron los CP funcionales y su relación con los modelos de requisitos. En la segunda etapa del proyecto se construyó una plantilla para registrar los casos de prueba. En la etapa que se presenta en este artículo se generó un proceso de derivación desde los EF a la plantilla de CP. El proceso para realizarlos es que por cada EF se abstraen los objetivos funcionales a validar y se generan los CP necesarios. La construcción de escenarios propone que cada EF debe satisfacer un objetivo y sus subobjetivos. Dado que un CP puede satisfacer un subobjetivo que esté incluido en varios escenarios y como un objetivo de un escenario puede requerir varios CP para probarlo, la relación entre EF y CP es de M:N.

Luego, se espera realizar la comprobación

de los casos de prueba con los casos de aceptación para analizar la posibilidad de automatizar el proceso.

Como línea complementaria de investigación se analiza el impacto de las jerarquías conceptuales de los léxicos representados en los glosarios de requisitos, específicamente en el LEL, y se realizó una inspección para asegurar su inclusión en el glosario.

Resultados y Objetivos

Como se mencionó [Kaplan 15] se cuenta con aproximadamente 150 casos de estudio para probar el proceso. Se han tomado estos casos para realizar las derivaciones y refinar el proceso.

Los casos de prueba se utilizan para verificar los requisitos de software. Es justamente desde la ERS, desde donde se pueden reconocer las trazas hacia los EF que le dieron origen, con el objetivo de identificar el conocimiento contextual de los requisitos del software y generar así los CP, tomando solo dicha información. Los casos realizados han demostrado que la información del EF es suficiente para generar un CP. Esto nos permitió definir la plantilla de CP donde se deriva casi automáticamente la información del EF. Se analizó que los casos de aceptación son un buen mecanismo para validar el CP semi automáticamente.

Se espera analizar la posibilidad de que los CP no existan, sino que en el momento de su recuperación se construyan automáticamente. Esto haría que estén siempre actualizados los CP ante cambios en los requisitos. Se

estudiará la posibilidad de automatizar la generación de los CP evitando la intervención humana y beneficiando así que las pruebas puedan estar siempre actualizadas.

Formación de Recursos Humanos

Se planifica la finalización de la carrera de grado Licenciatura en Sistemas de Información de la alumna Eugenia Cespedes (30 materias aprobadas) y del alumno Julián Massolo (25 materias aprobadas) en UNLu. También la presentación de la tesis de Maestría en Ingeniería de Software de Claudia Ortiz y Walter Panessi en UNLP y la finalización de la tesis doctoral de Gladys Kaplan y de David Petrocelli en UNLP.

Referencias

- [Abbers 10] Abbers, F., Backlund, A. ; Truscan, D., “MATERA - An Integrated Framework for Model-Based Testing”, Engineering of Computer Based Systems (ECBS), 17th IEEE International Conference 2010
- [Badri 09] Mourad Badri, Linda Badri and Maxime Bourque-Fortin, “Automated State-Based Unit Testing for Aspect Oriented Programs A Supporting Framework”, Journal of Object Technology. Vol. 8, No. 3, May/June 2009.
- [Blanco 06] Raquel Blanco, Eugenia Díaz, Javier Tuya,” Generación automática de casos de prueba mediante búsqueda dispersa ”, Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.2, No. 1, 2006.
- [Boghdady 12] P. N. Boghdady, Nagwa L. Badr, M. A. Hashim, Mohamed F. Tolba, “Test Cases Automatic Generator (TCAG): A Prototype”, First International Conference, AMLTA 2012, Cairo, Egypt, December 8-10, 2012.
- [Boghdady 11] Boghdady, P.N. Badr, N.L. ; Hashim, M.A. ; Tolba, M.F., “An enhanced test case generation technique based on activity diagrams”, Computer Engineering & Systems (ICES), 2011
- [Correa 11] Natalia Correa, Roxana Giandini, “Casos de Prueba del Sistema Generados en el Contexto MDD/MDT”, LIFIA- Laboratorio de Investigación y Formación en Informática Avanzada, Universidad Nacional de La Plata, 40 JAIIO, ASSE 2011.
- [Correa 12] Natalia Correa, Roxana Giandini (2012). Casos de Prueba del Sistema Generados en el Contexto MDD/MDT. 41 JAIIO - ASSE 2012 - ISSN: 1850-2792 - Pág. 91-105
- [Doorn 02] Doorn J., Hadad G., Kaplan G. (2002) Comprendiendo el Universo de Discurso Futuro, WER’02 - Workshop on Requirements Engineering, Valencia, Spain.
- [Gutiérrez 06] Javier J. Gutiérrez, María J. Escalona, Manuel Mejías y Jesús Torres, Hacia una propuesta de pruebas tempranas del sistema, XV Jornadas de Ingeniería del Software y Bases de Datos. JISBD 2006.
- [Hadad 08] Hadad G.D.S., Doorn J.H., Kaplan G.N. Creating Software System Context Glossaries, In: Mehdi Khosrow-Pour (ed) Encyclopedia of Information Science and Technology. IGI Global, Information Science Reference, Hershey, PA, USA, ISBN: 978-1-60566-026-4, 2nd edn, Vol. II. 2008.
- [Hasling 08] Bill Hasling, Helmut Goetz, Klaus Beetz. Model Based Testing of

- System Requirements using UML Use Case Models. International Conference on Software Testing, Verification, and Validation. 2008
- [Javed 07] Javed, A.Z. “Automated Generation of Test Cases Using Model-Driven Architecture”, *Automation of Software Test*, 2007.
- [Kaplan 15] Gladys Kaplan, Jorge Doorn, Walter Panessi, Claudia Ortiz, Eugenia Cespedes, Julian Massolo, David Petrocelli, “Generación semi automática de casos de prueba a partir de escenarios”, WICC 2015.
- [Lamancha 09] Beatriz Pérez Lamancha, Macario Polo, “Generación automática de casos de prueba para líneas de producto de software”, *Revista Española de Innovación, Calidad e Ingeniería del Software*, Vol.5, No. 2, 2009.
- [Leite 97] Leite, J.C.S.P., Rossi, G., Balaguer, F., Maiorana, V., Kaplan, G., Hadad, G., Oliveros, A., “Enhancing a Requirements Baseline with Scenarios”, *Requirements Engineering Journal*, Vol.2, N° 4, 1997.
- [Leite 90] Leite J.C.S.P., Franco, A.P.M., (1990) “O Uso de Hipertexto na Elicitação de Linguagens da Aplicação”, *Anais de IV Simpósio Brasileiro de Engenharia de Software*, SBC.
- [Leite 04] Leite, J.C.S.P., Doorn, J.H., “Perspectives on Software Requirements: An introduction” en el libro “Perspectives on Software Requirements”, Kluwer Academic Publishers, EEUU, ISBN: 1-4020-7625-8, Capítulo 1, 2004.
- [Linzhang 04] Wang Linzhang, Yuan Jiesong ; Yu Xiaofeng ; Hu Jun ; Li Xuandong; Zheng Guoliang, “Generating test cases from UML activity diagram based on Gray-box method”, *Software Engineering Conference*, 2004. 11th Asia-Pacific
- [Mingsong 09] Chen Mingsong, Qiu Xiaokang, Li Xuandong, “Automatic Test Case Generation for UML Activity Diagrams”, *Journal of Object Technology*. Vol. 8, No. 3, May/June 2009.
- [Odalys 06] Jordán Enriquez, Odalys; Vázquez Ruiz, Orelvis, “GENERACIÓN DE CASOS DE PRUEBA A PARTIR DE CASOS DE USO EN LAS PRUEBAS DE SOFTWARE”, *Ingeniería Industrial*, vol. XXVII, núm. 1, 2006, pp. 7-10 Instituto Superior Politécnico José Antonio Echeverría La Habana, Cuba ISSN: 0258-5960, 2006.
- [Palacio 09] Liliana González Palacio, “MÉTODO PARA GENERAR CASOS DE PRUEBA FUNCIONAL EN EL DESARROLLO DE SOFTWARE”, *Revista Ingenierías Universidad de Medellín*, vol. 8, No. 15 especial, pp. 29-36 - ISSN 1692-3324, 150 p. Medellín, Colombia, 2009
- [Riebisch 03] M. Riebisch, I. Philippow, and M. Götze, “UML-Based Statistical Test Case Generation,” in *Revised Papers from the International Conference NetObjectDays on Objects, Components, Architectures, Services, and Applications for a Networked World*, 2003, pp. 394-411.
- [Salem 11] Yasmine Ibrahim Salem y Riham Hassan. 2011. Requirement-Based Test Case Generation and Prioritization. 978-1- 61284- 185-4/111 - 2011 IEEE.
- [Wendland 11] Marc-Florian Wendland, Ina Schieferdecker and Alain Vouffo-Feudjio..Requirements-driven testing with behavior trees. Fourth International Conference on Software Testing, Verification and Validation Workshops. 2011.

Desarrollo de Aplicaciones Móviles 3D

Pablo Thomas, Federico Cristina, Sebastián Dapoto, Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

50 y 120 La Plata Buenos Aires

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

526 e/ 10 y 11 La Plata Buenos Aires

{pthomas, fcristina, sdapoto, ppesado}@lidi.info.unlp.edu.ar

Resumen

Se presenta una línea de investigación y desarrollo, que tiene por objeto estudiar temas relacionados con aspectos de Ingeniería de Software, orientados al desarrollo de aplicaciones móviles sobre diversas plataformas y entornos operativos. En particular, se pone el énfasis en el desarrollo de herramientas para dispositivos móviles que utilicen un entorno tridimensional (3D).

Palabras claves: Dispositivos Móviles - Aplicaciones 3D - Aplicaciones Multiplataforma - M-Learning

Contexto

Esta línea de Investigación forma parte del proyecto (2014-2018) “*Tecnologías para Sistemas de Software Distribuidos. Calidad en Sistemas y Procesos. Escenarios Educativos mediados por TICS*”, en particular del subproyecto *Ingeniería de Software en el Desarrollo de Aplicaciones para Dispositivos Móviles* del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación de la Nación.

Hay cooperación con Universidades de Argentina y se está trabajando con Universidades de Europa en proyectos financiados por el Ministerio de Ciencia y Tecnología de España y la AECID.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado.

Introducción

En la actualidad los dispositivos móviles forman parte de la vida cotidiana, tanto en el ámbito personal como en el profesional. Estos dispositivos son cada vez más sofisticados y su evolución tecnológica permite ejecutar aplicaciones cada vez más complejas y con exigentes requerimientos de hardware.

Debido a esta evolución, existen además una gran cantidad de librerías y frameworks que permiten desarrollar aplicaciones tridimensionales para dispositivos móviles.

Cada uno de estos frameworks de desarrollo 3D posee diferentes características que los hacen adecuados para distintos tipos y magnitudes de proyectos. La elección de un framework en particular se puede basar en diversos criterios tales como su comunidad de usuarios, los lenguajes de codificación que utiliza, la calidad de los gráficos 3D resultantes, los tutoriales y documentación existentes, los requerimientos de hardware, la facilidad de instalación, de aprendizaje y de uso, entre otros.

Las herramientas desarrolladas con estos frameworks son visualmente más agradables y atraen más fuertemente a sus usuarios, dado que los entornos tridimensionales son más cercanos a la realidad que los bidimensionales, y permiten involucrarse de forma más activa.

Por otro lado, el ámbito educativo debe adaptarse a los cambios y nuevas formas de aprendizaje. M-learning (mobile learning) plantea métodos modernos de apoyo al proceso de aprendizaje mediante el uso de dispositivos móviles.

Las aplicaciones móviles 3D son una herramienta ideal para atraer a los alumnos. Sin embargo, muchos de los potenciales usuarios de estas aplicaciones educativas pueden no disponer de dispositivos de última generación, por lo que resulta de vital importancia realizar un análisis en profundidad de los parámetros que inciden en la performance final de una aplicación 3D.

Líneas de Investigación y Desarrollo

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Mobile Learning
- Aplicaciones Móviles Multiplataforma
- Aplicaciones Móviles 3D
- Frameworks para el desarrollo de Aplicaciones Móviles 3D
- Performance de aplicaciones móviles 3D

Resultados esperados/obtenidos

Los resultados esperados/obtenidos se pueden resumir en:

- Avanzar en la capacitación continua de los miembros de la línea de investigación.
- Avanzar en el aprendizaje de frameworks que permiten desarrollar aplicaciones 3D multiplataforma, particularmente para dispositivos móviles.
- Avanzar en el desarrollo de aplicaciones educativas, teniendo como finalidad enriquecer las experiencias interactivas y motivar el aprendizaje mediante su uso.
- Avanzar en el análisis de performance de uno o más frameworks 3D con el fin de

detectar los puntos críticos y mejorar la ejecución de las aplicaciones 3D.

- Se han evaluado diferentes herramientas para el desarrollo de aplicaciones móviles 3D. En particular se analizaron los frameworks Unity y Unreal Engine, notando que difieren considerablemente en algunas de sus características.
- Se ha desarrollado el prototipo móvil R-Info3D, similar a la aplicación de escritorio R-Info, que puede ser utilizada desde distintas plataformas móviles. El entorno tridimensional de R-Info3D mejora la experiencia del alumno y permite una mejor visualización de las tareas que realiza el robot desde diferentes puntos de vista. Figura 1.



Figura 1. R-Info3D.

- Se ha construido un modelo virtual 3D del edificio de la Facultad de Informática de la UNLP.
- Se ha desarrollado el prototipo móvil InfoUNLP3D que sirve de guía para los estudiantes en sus primeras experiencias dentro de la facultad. Figura 2.

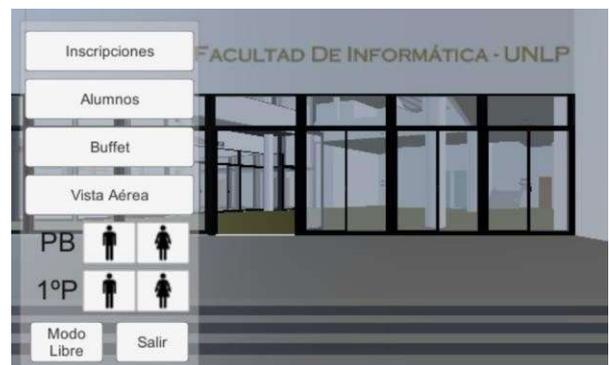


Figura 2. InfoUNLP3D.

- Se ha desarrollado un prototipo de análisis de performance de aplicaciones 3D, que estudia la degradación de la ejecución de dichas aplicaciones a medida que se incrementa ciertos factores de relevancia como lo son la cantidad de objetos simples y complejos, objetos con y sin textura, luces y sombras, y sistemas de partículas. Figuras 3, 4 y 5.



Figura 3. Prototipo de Performance. Objeto complejo: edificio de la Facultad de Informática.

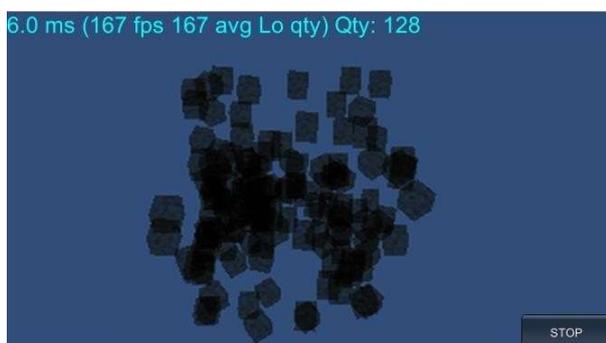


Figura 4. Prototipo de Performance. Objetos con textura: transparencia.

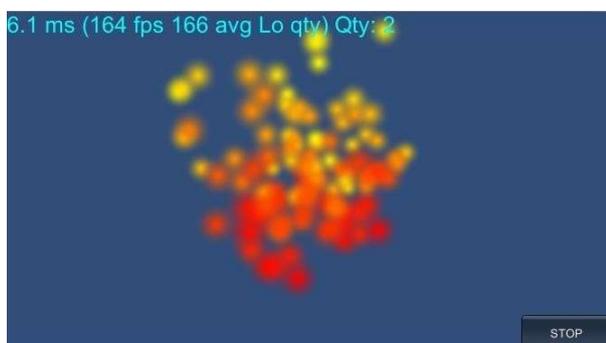


Figura 5. Prototipo de Performance. Sistema de partículas.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Grado y Tesis de Postgrado en la Facultad de Informática, y Becarios III-LIDI en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

Referencias

- Cristina, F.; Dapoto, S.; Thomas, P.; Pesado, P. "InfoUNLP3D: An interactive experience for freshman students". XXII Congreso Argentino de Ciencias de la Computación CACIC 2016. San Luis, Argentina. Octubre 2016.
- Cristina, F.; Dapoto, S.; Thomas, P.; Pesado, P. "Prototipo móvil 3D para el aprendizaje de algoritmos básicos". XXI Congreso Argentino de Ciencias de la Computación CACIC 2015. Junín, Argentina. Octubre 2015. ISBN: 978-987-3724-37-4.
- Cristina, F.; Dapoto, S.; Thomas, P.; Pesado, P. "A simplified multiplatform communication framework for mobile applications". IEEE International Conference on Computer Engineering & Systems (ICCES). December 2014. ISBN 978-1-4799-6593-9.
- De Giusti Armando, De Giusti Laura, Leibovich Fabiana, Sanchez Mariano, Rodriguez Eguren Sebastian. "Entorno interactivo multirrobot para el aprendizaje de conceptos de Concurrencia y Paralelismo". TE&ET 2014. Chilecito.
- De Giusti Laura, Leibovich Fabiana, Sánchez Mariano, Rodriguez Eguren Sebastian, Chichizola Franco, Naiouf Marcelo, De Giusti Armando. "Herramienta interactiva para la enseñanza temprana de Concurrencia y Paralelismo: un caso de estudio". CACIC 2014. San Justo. ISBN: 978-987-3806-05- 6.
- De Giusti Laura, Leibovich Fabiana, Sánchez Mariano, Chichizola Franco, Naiouf Marcelo, De Giusti Armando. "Desafíos y herramientas para la enseñanza temprana de Concurrencia y

- Paralelismo". CACIC 2013. Mar del Plata. ISBN: 978-987-23963-1-2.
7. De Giusti Armando, Frati Emmanuel, Sanchez Mariano, De Giusti Laura. "LIDI Multi Robot Environment: Support software for concurrency learning in CS1". CTS 2012. ISBN: 978-1-4673- 1380-3.
 8. ICETI 2011. "Web based Interactive 3D Learning Objects for Learning Management Systems". Stefan Hesse, Stefan Gumhold. <https://www.inf.tu-dresden.de/content/institutes/smt/cg/publications/paper/ICETI2011.pdf>.
 9. Kantel E., Tovar G., Serrano A."Diseño de un Entorno Colaborativo Móvil para Apoyo al Aprendizaje a través de Dispositivos Móviles de Tercera Generación." IEEE-RITA 5, no. 4 (2010): 146-151.
 10. 2009 LA Web Congress. "Interacting with 3D Learning Objects". Rosa Paredes, J. Alfredo Sánchez, Liliana Rojas, Daniel Strazzulla, Ronel Martínez-Teutle. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5341602>.
 11. Unity 3D Homepage: <https://unity3d.com/>.
 12. Unreal Engine Homepage: <https://www.unrealengine.com/>.

Desarrollo e Implementación de un Modelo Basado en Estándares para el Logro y Evaluación de la Usabilidad en Aplicaciones Informáticas

CASTRO Marcelo, SÁNCHEZ RIVERO David, VARGAS Alejandro, ARAGÓN Fabiana, REINOSO Elizabeth, APARICIO María, FARFÁN José, CÁNDIDO Andrea, CASTRO Daniel, CAZÓN Liliana & ZAPANA, José

Investigación + Desarrollo en Gobierno Electrónico / Facultad de Ingeniería / Universidad Nacional de Jujuy

Av. Italia y Av. Martiarena / S. S. de Jujuy / Provincia de Jujuy
Tel. 388-4221591

vdsanchezrivero@fi.unju.edu.ar, mcastro@fi.unju.edu.ar, lavargas@fi.unju.edu.ar, fraragon@fi.unju.edu.ar, edrreinoso@fi.unju.edu.ar, mcaparicio@fi.unju.edu.ar, jhfarfan@fi.unju.edu.ar, agcandido@fhycs.unju.edu.ar, ddcastro@fi.unju.edu.ar, lbcazon@fce.unju.edu.ar, jvzapana@fi.unju.edu.ar

RESUMEN

El presente proyecto de investigación tiene por objeto diseñar, desarrollar e implementar un modelo basado en estándares internacionales relacionados con la calidad: ISO 9241, ISO/IEC 9126-1, ISO/IEC 9126-4, ISO/IEC 25010, WCAG 2.0 e ISO/IEC 25062, para el logro y evaluación de la Usabilidad en aplicaciones informáticas teniendo en cuenta el alcance, la dimensión y la relación con otros términos de calidad en las interfaces de usuario; accesibilidad, calidad en uso y experiencia de usuario(UX).

Además se estudiarán y se incorporarán para el diseño del modelo, los conceptos sobre Ingeniería de la usabilidad, Diseño centrado en el uso, patrones de diseño de interfaces de usuario y Desarrollo dirigido por modelos (MDD). En relación a la evaluación de la Usabilidad se estudiarán los distintos métodos existentes, haciendo hincapié en las recomendaciones del estándar ISO/IEC 25062:2006, además se utilizarán metodologías y herramientas relacionadas con la evaluación de la accesibilidad.

Para probar el modelo desarrollado se creará un Laboratorio de Usabilidad que dependerá del Gabinete de Herramientas de Software, ubicado en el centro de cómputos de la Facultad de Ingeniería de la UNJu. El Laboratorio de Usabilidad proveerá servicios a distintas entidades públicas y privadas que

deseen evaluar la usabilidad de sus productos informáticos.

Palabras clave: Usabilidad, Accesibilidad, Diseño Centrado en el Usuario (DCU), Experiencia de usuario(UX), Calidad en uso, Evaluación de software.

CONTEXTO

El proyecto “Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas” es una continuación e incorpora conceptos y experiencias de los proyectos desarrollados por nuestro equipo de investigación denominados, “Tic’s: Automatización y Estandarización del Proceso de Gobierno Electrónico”, entre los años 2007 al 2009, “Desarrollo de una herramienta para automatizar el proceso de Gobierno Electrónico”, cuyo periodo de trabajo comprendió los años 2010 al 2012 y “Desarrollo de un modelo basado en servicios digitales comunes reutilizables para Gobierno Electrónico” desde el año 2013 al 2015. Estos constituyeron la base conceptual para elaborar un modelo que permitiera, basado en estándares logro y evaluación de la Usabilidad en aplicaciones informáticas.

El proyecto se desarrolla en la Facultad de Ingeniería de la Universidad Nacional de Jujuy, posee código D-0142 y posee Categoría “A”. Además se encuentra subsidiado por la Secretaría de Ciencia y Técnica y Estudios Regionales, dependiente de la U.N.Ju.

1.INTRODUCCIÓN

En el modelo para la evaluación de la usabilidad basado en estándares que se propone en el proyecto de investigación, se analizarán las diferencias y relaciones existentes entre los conceptos de usabilidad, diseño centrado en el usuario, facilidad de uso, accesibilidad, ergonomía, experiencia de usuario, calidad en uso y usabilidad universal.

La Usabilidad significa que la gente que utilice un producto pueda realizar rápida y fácilmente sus tareas. [1]

La usabilidad realmente significa tener la certeza de que algo funciona bien, es decir, que cualquier persona con conocimientos o habilidades medias (incluso por debajo de la media) puede navegar por un sitio web, usar un avión o abrir una puerta, sin que cualquier labor que lleve a cabo conlleve su frustración. [2]

La Usabilidad universal es la posibilidad y grado con el que usuarios diversos, con tecnologías diferentes e intereses también distintos pueden acceder a la información y a los servicios.[3]

Es importante tener presente que la usabilidad no es algo simple o una propiedad que ofrezca una única cara relacionada con la interfaz de usuario ofrecida. La usabilidad tiene múltiples componentes e involucra cinco atributos: aprendizaje, eficiencia, memorización, tratamiento de errores y satisfacción. [4]

La accesibilidad Web significa que personas con algún tipo de discapacidad van a poder hacer uso de la Web. En concreto, al hablar de accesibilidad Web se está haciendo referencia a un diseño Web que va a permitir que estas personas puedan percibir, entender, navegar e interactuar e interactuar con la Web, aportando a su vez contenidos. [5]

La ergonomía es el conjunto de conocimientos de carácter multidisciplinar aplicados para la adecuación de los productos, sistemas y entornos artificiales a las necesidades, limitaciones y características de sus usuarios, optimizando la eficacia, seguridad y bienestar. [6]

La experiencia de Usuario (UX) es la respuesta a cómo se sienten los usuarios después de interactuar con un sistema. [7] [8]

[9]

La calidad en uso, es un concepto que surge desde la ingeniería del software y viene a salvar las diferencias entre ésta y la interacción persona-ordenador (HCI).

1.1. Estándares Internacionales

Para la característica de usabilidad los estándares internacionales plantean la calidad del producto software y la calidad de la interacción, es decir calidad del producto y calidad del proceso. A su vez las disciplinas que trabajan estrechamente con la usabilidad, son la Ingeniería del Software y la Interacción Hombre Máquina (HCI).

En el caso de la Ingeniería del software se preocupa por la calidad del producto software a través de la evaluación de características tales como la funcionalidad, la Fiabilidad, la eficiencia, la mantenibilidad, la portabilidad, y de la Usabilidad. La Interacción Hombre Máquina, en cambio, se preocupa por la calidad del proceso de interacción, verificando la accesibilidad, la experiencia de usuario y la usabilidad.

Los estándares internacionales trabajan con diversos factores de calidad asociados tanto al producto software como al proceso de interacción que ofrecen los mismos. En la tabla 1, se puede apreciar la relación entre los distintos estándares internacionales y los factores de calidad que consideran.

Estándar	Factores de calidad considerados
ISO9241-11	Usabilidad
ISO 9241-210	Usabilidad, UX y DCU
ISO 9241-171	Accesibilidad
ISO/IEC9126-1	Funcionalidad, eficiencia, fiabilidad, usabilidad, portabilidad y mantenimiento
ISO/IEC 9126-4	Calidad en uso
ISO/IEC 25010	Calidad de un producto software y de la interacción de dicho producto

Tabla 1 – Relación entre Estándar y Factores de calidad
El estándar ISO 9241-11 [10] establece que la usabilidad es efectividad, eficiencia y

satisfacción. Efectividad en el sentido que el usuario es capaz, utilizando el producto software, de lograr sus objetivos. Eficiencia ya que logra sus objetivos y lo hace destinando los recursos necesarios (tiempo, aprendizaje, etc.). Satisfacción del usuario en el uso del producto software tornándose placentero, confortable y útil.

En el estándar ISO 9241-210 [11] se definen los conceptos y las técnicas asociadas a Usabilidad, Experiencia de Usuario (UX) y Diseño Centrado en el Usuario (DCU).

En la norma ISO 9241-171 [12] se identifican un conjunto de requisitos relacionados con el logro de la accesibilidad. Brinda información consistente con documentos de la W3C relacionados con accesibilidad en la Web (WCAG 2.0.)

La ISO/IEC 9126-1 [13] establece las características, y sub-características que se deben tener en cuenta para evaluación de la calidad interna y externa de un producto software.

El estándar ISO/IEC 9126-4 [14] introduce el concepto de calidad en uso, relacionado con la calidad de la interacción. Además incorpora métricas para su estimación y anexos relacionados con la elaboración de informes sobre la calidad en uso/usabilidad de un producto software.

El ISO/IEC 25010 [15] sustituye a la ISO/IEC 9126-1, y sigue considerando la usabilidad como producto software. Mantiene algunos criterios tradicionales relacionados con usabilidad, pero agrega criterios adicionales como el riesgo y la adecuación al contexto.

1.2. Evaluación de la Usabilidad

Existen técnicas que dan soporte a las actividades de evaluación de interfaces de usuario, por ejemplo las descritas por el estándar ISO 9241-210 [11], éstas incluyen las actividades de planificar el proceso de diseño, entender y especificar el contexto de uso, identificar y especificar los requisitos del usuario, diseñar e implementar la interface de usuario.

Algunos métodos de evaluación de la usabilidad y soporte al DCU [16] [17] se pueden observar en la tabla 4.

Nombre del método	Breve descripción
Observación de usuarios	Recogerde manera sistemática información sobre el comportamiento y las prestaciones de los usuarios en un contexto específico mientras realizan su actividad.
Medición de prestaciones	Colección de medidas con la intención de conocerel impacto de los aspectos de usabilidad.
Análisis de incidentes críticos	Recolecciónde eventos específicos (positivos y negativos).
Cuestionarios	Métodos de evaluación indirectaque consiguen opiniones de los usuarios sobre la interfaz de usuario utilizando cuestionarios predefinidos.
Entrevistas	Similar a los cuestionarios con mayor flexibilidad y utilizando una interacción cara a cara.
Thinkingaloud	Involucrar a usuariosy pedirles que verbalicen sus opiniones, creencias, dudas, etc. mientras interactúan con el sistema.
Evaluación y diseño colaborativo	Métodos que permiten a diferentes tipos de usuarios colaborar enla evaluación y diseño de sistemas.
Métodos creativos	Métodos que persiguenla elicitaciónde nuevos productos con nuevas características.
Métodos basados en documentos	Examen de documentos existentes por especialistas para formarseun juicio profesional del sistema.
Desarrollos basados en modelos	Utilización de representacionesabstractas del producto evaluado para permitir la predicción de las prestaciones de los usuarios.
Evaluación experta	Evaluación basada en el conocimiento de expertos.
Evaluación automática	Evaluación dirigida por algoritmos centrados en criterios de usabilidad y conocimiento ergonómico con los que diagnosticar deficiencias.

Tabla 4 - Métodos de evaluación de la usabilidad y soporte al DCU

Para el desarrollo del modelo de evaluación a generar se propone trabajar con los estándares descriptos precedentemente y en

especial con la norma ISO/IEC 25062 [18] que define el CIF (Common Industry Format) como estándar a utilizar para confeccionar informes de pruebas de usabilidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El equipo de trabajo se encuentra desarrollando actividades de investigación y desarrollo, fundamentalmente en el área de la Ingeniería de software y específicamente en la formalización del proceso de gobierno electrónico, Ingeniería Web, Calidad del Software y Arquitectura dirigida por Modelos (MDA) a partir de los proyectos: “TIC’s: automatización y estandarización del proceso de Gobierno Electrónico”, “Diseño y Desarrollo de una herramienta para automatizar el proceso de Gobierno Electrónico”, “Desarrollo de un modelo basado en MDA para gobierno electrónico” y “Desarrollo de un modelo basado en servicios digitales comunes reutilizables para Gobierno Electrónico”.

En el nuevo desafío que presenta el proyecto “Desarrollo e implementación de un modelo basado en estándares para el logro y evaluación de la Usabilidad en aplicaciones informáticas”, se trabaja sobre ejes disciplinares relacionados a la utilización de estándares internacionales en aspectos relacionados a la usabilidad.

En la actualidad la demanda de aplicaciones informáticas requiere el desarrollo de interfaces de usuario de calidad. Para ello los diseñadores y los desarrolladores deben considerar múltiples requisitos, pero fundamentalmente la usabilidad del producto. En consecuencia se hace necesario abordarla como un proceso ingenieril para lograr los mejores resultados y poder evaluarla.

A través de estudios preliminares realizados por el grupo de investigación en los proyectos ejecutados desde al año 2007 y descriptos anteriormente, se detectó que las interfaces de usuario de diversas aplicaciones informáticas, particularmente las desarrolladas en ambientes Web tanto en el sector público como en el empresarial, no fueron concebidas

y evaluadas, teniendo en cuenta estándares internacionales relacionados a la usabilidad del producto software.

Sumado a lo expuesto precedentemente, se puede decir que organizaciones gubernamentales y no gubernamentales relacionadas a brindar ayuda y servicios a personas con capacidades diferentes, no cuentan con instalaciones especiales y procesos para realizar pruebas de accesibilidad a aplicaciones informáticas, que abarcan a interfaces de usuario de software basados en la Web, como a aplicaciones del tipo HCI.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se espera que el modelo a desarrollar en el presente proyecto de investigación como el laboratorio a montar en el ámbito de la Facultad de Ingeniería de la UNJu, tendrán un rol importante en la Provincia y en la región, brindando servicios a una gran parte de la sociedad, que va desde empresas dedicadas al desarrollo del software, pasando por compañías que desarrollan sus propias aplicaciones; hasta organismos públicos que desarrollan software para brindar servicios al ciudadano e instituciones no gubernamentales que permiten el acceso a las nuevas tecnologías de la información y comunicación a personas con capacidades diferentes, tales como el Programa para personas con discapacidad de la Universidad Nacional de Jujuy (UnjuProDis), el Instituto Helen Keller, creado para la atención de niños sordos y el Centro de Rehabilitación Integral de la Asociación de Protección al Paralítico Cerebral (Appace).

Por lo expuesto, el proyecto de investigación aspira satisfacer demandas relacionadas con aspectos técnicos informáticos, como la usabilidad y la accesibilidad de un producto software, pero también a aspectos sociales brindando servicios a instituciones que trabajan con personas con capacidades diferentes.

Los resultados obtenidos hasta el momento, se encuentran plasmados en los trabajos presentados en diversos eventos

científicos entre los que podemos citar “Evaluando la calidad de uso en una aplicación web de gobierno electrónico” y “Evaluando la usabilidad de Galaxy Conqueror con ISO 25062”. Los citados trabajos fueron desarrollados en el año 2016.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto de investigación se encuentra conformado por once docentes investigadores, siete de los cuales se encuentran categorizados y cuatro en proceso de categorización. Además cuenta con cuatro alumnos que están siendo iniciados en tareas de investigación y actividades de desarrollo.

Por otra parte el desarrollo de las tareas de investigación, ha generado en los años 2012, tres anteproyectos de tesis en la Maestría en Ingeniería de Software, de la Universidad Nacional de San Luis pertenecientes a integrantes del equipo de trabajo.

Además se realizó la presentación y aprobación de una propuesta técnica de Tesis Doctoral para cursar el Doctorado en Ciencias Informáticas en la Facultad de Informática de la Universidad Nacional de la Plata, que se encuentra en etapa de elaboración final de la tesis doctoral.

Simultáneamente algunos miembros del equipo de trabajo, se encuentran dirigiendo cuatro proyectos finales de carrera de grado, pertenecientes a alumnos de Ingeniería Informática, de la Facultad de Ingeniería de la Universidad Nacional de Jujuy.

5. BIBLIOGRAFÍA

- [1] Dumas, Joseph S. and Janice C. Redish, A Practical Guide to Usability Testing, Rev. Edition, Intellect Books, 1999.
- [2] Krug, Steve, Don't Make Me Think: A Common Sense Approach to the Web, 2nd Edition, New Riders Publishing, 2005.
- [3] Shneiderman, Ben, Universal Usability, ACM 43(5), Pag. 84-91, 2000.
- [4] Nielsen, Jakob, The Usability Engineering Life Cycle, IEEE Computer 25(3), Pag. 12-22, 1992.
- [5] Accesibilidad, WorldWide Web Consortium, W3C. <http://www.w3c.es>
- [6] ISO/TR16982, Ergonomics of human-system interaction -Usability methods supporting human-centred design, 2002.
- [7] Norman, Donald, The Design of Everyday Things, Basic Books, 2002.
- [8] Chong Law Effie Lai, Roto, Virpi, Vermeeren Arnold, Kort Joke, Hassenzahl Marc, Towards a shared definition of user experience. CHI Extended Abstracts, Pag. 2395-2398, 2008.
- [9] User experience design, http://semanticstudios.com/user_experience_design/
- [10] ISO 9241-11, Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability, 1998.
- [11] ISO 9241-210, Ergonomics of human-system interaction, Part 210: Human-centred design for interactive systems, 2010.
- [12] ISO/DIS 9241-171, Ergonomics of human-system interaction, Guidance on software accessibility, 2006.
- [13] ISO/IEC 9126-1, Software engineering - Product quality -Part 1: Quality model, 2001.
- [14] ISO/IEC TR 9126-4, Software engineering -Product quality -Part 4: Quality in use metrics, 2004.
- [15] ISO/IEC 25010, Systems and software engineering -Systems and software Quality - Requirements and Evaluation (SQuaRE) - System and software quality models, 2010.
- [16] Métodos de evaluación de la usabilidad. <http://www.usabilitynet.org>.
- [17] MAUSE, Towards the MAuration of Information Technology USability Evaluation, [http://www.cost294.org\(digital library\)](http://www.cost294.org(digital_library)).
- [18] ISO/IEC 25062, Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) -Common Industry Format (CIF) for usability test reports, 2006.

Elicitación del Vocabulario del Contexto Ayudada por Mapas Conceptuales

Graciela D. S. Hadad^{1,2}, Alberto Sebastián¹, Jorge H. Doorn^{2,3}, Ezequiel Robledo², Daniela Raffo¹

¹Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano

²Escuela de Informática, Universidad Nacional del Oeste

³DIIT, Universidad Nacional de la Matanza

graciela.hadad@comunidad.ub.edu.ar, alberto.sebastian@comunidad.ub.edu.ar, jdoorn@exa.unicen.edu.ar, ezeroble@hotmail.com.ar, daniela.raffo@comunidad.ub.edu.ar

RESUMEN

La Ingeniería de Requisitos ha promovido una gran variedad de métodos, técnicas y herramientas para la producción de requisitos del software, sin embargo, se siguen observando deficiencias en los modelos generados. Frecuentemente, estos modelos se escriben en lenguaje natural dada su facilidad de comprensión por parte de todos los involucrados, aunque conllevan defectos inherentes a su esencia: la ambigüedad del texto. Asimismo, en muchos casos es notoria la falta de completitud de dichos modelos. Esta deficiencia es de más difícil detección en forma temprana, y suele ser muy dañina cuando se arrastra a subsiguientes artefactos del software. Las omisiones en los modelos no provienen, en general, de una mala construcción del propio modelo, sino que se originan con mayor frecuencia en la elicitación. Es por ello que se propone focalizar la elicitación desde un punto de vista cognitivo, estudiando las posibilidades que puedan brindar algunas técnicas provenientes de la Gestión de Conocimiento. La propuesta se centra en la elicitación del vocabulario utilizado en un contexto de aplicación, el que es modelado mediante el Léxico Extendido del Lenguaje.

Palabras clave: Ingeniería de Requisitos, Gestión de Conocimiento, Elicitación, Completitud de Modelos, Ambigüedad.

CONTEXTO

La propuesta que se presenta es parte de los proyectos de investigación “Gestión de la calidad de un modelo léxico en el proceso de requisitos” de la Universidad de Belgrano y “Tratamiento de los factores situacionales y la completitud en la ingeniería de requisitos” de la Universidad Nacional del Oeste.

1. INTRODUCCIÓN

Ackoff [1] ha señalado que a menudo se falla más veces por resolver el problema incorrecto que por obtener una solución deficiente al problema correcto. Esto apunta básicamente a la dificultad en establecer los requisitos apropiados para el software a construir, lo que queda en evidencia por la cantidad de proyectos de software de gran envergadura que han fracasado debido principalmente a una pobre producción de requisitos [2] [3] [4] [5].

El proceso de producción de requisitos consiste en capturar información de un contexto de aplicación y definir lo que luego debe ser construido [6]. Para facilitar estas actividades, suelen utilizarse modelos escritos en lenguaje natural [7] [8] [9], pues son un medio propicio para interactuar con los clientes, aunque presentan algunas debilidades que afectan su calidad. La ambigüedad suele ser el primer inconveniente [10], siendo otro el poder establecer el grado de completitud alcanzado [11] [12]. Dado que la completitud es una propiedad casi imposible de lograr en problemas complejos, se suele establecer como meta lograr un nivel aceptable de completitud [13].

Un modelo que ha sido ampliamente difundido en la Ingeniería de Requisitos es el Léxico Extendido del Lenguaje (LEL) [14], debido a que es un glosario con particularidades que fortalecen la comunicación entre todos los involucrados a lo largo del desarrollo del software, ya que permite definir con precisión los términos que se utilizan en el contexto del problema. Además, sirve de soporte para construir y usar modelos de requisitos [9] [15] [16]. Es por ello que ha sido objeto de varios estudios referidos a la completitud en proyectos de investigación previos [11] [17] [18]. Estos estudios se centraron en estimar el tamaño del LEL mediante métodos estadísticos, concluyendo que dicho modelo presentaba un alto número de omisiones. En relación a la falta de completitud, un trabajo experimental sobre elicitación directa de requisitos menciona la posible existencia de una “zona ciega” para detectarlos [19].

Según Nonaka y Takeuchi [20], existen dos tipos bien definidos de conocimiento: explícito y tácito. Es explícito si puede ser transferido de un individuo a otro usando algún tipo de sistema de comunicación formal. El conocimiento tácito es aquel conocimiento poco codificado y profundamente arraigado en la experiencia y los modelos mentales. Para adquirir este conocimiento en las organizaciones, dichos autores sugieren enfocarse en percepciones, impresiones e intuición a través del uso de metáforas, imágenes o experiencias.

Las actividades de elicitación y modelado en la Ingeniería de Requisitos están fuertemente relacionadas con la Gestión del Conocimiento [21]. El propósito de la Gestión de Conocimiento es la transmisión de conocimiento desde el lugar donde se genera hasta el lugar donde se va a aplicar en el marco de una organización [22]. Constituye un proceso cíclico mediante el cual se captura, analiza, organiza, almacena, difunde y crea nuevo conocimiento, para mejorar procesos o resolver problemas en cualquier ámbito [23]. Por lo que la Gestión del Conocimiento puede ayudar a la Ingeniería de Requisitos y viceversa. Una de las técnicas más destacadas

para gestionar el conocimiento organizacional son los mapas conceptuales [24]. Estos son un recurso esquemático para representar un conjunto de significados conceptuales incluidos en una estructura de proposiciones [25]. La teoría del aprendizaje significativo [26] presupone que la estructura cognitiva está organizada de forma jerárquica, es decir, que las proposiciones y conceptos más generales y menos específicos incluyen a las proposiciones y conceptos menos generales y menos inclusivos. De esta forma, los mapas conceptuales seguirían la misma organización que la estructura cognitiva. Es por ello que uno de sus usos más frecuentes es como herramienta educativa, permitiendo aprender diferentes dominios y materiales [27]. En la Ingeniería de Requisitos existe una fase de aprendizaje para comprender el problema a resolver, por lo que se supone que los mapas conceptuales podrían colaborar en la elicitación. Sin embargo, debe notarse que una estructura cognitiva organizada jerárquicamente según [26], está en oposición con lo mencionado en [20] respecto al conocimiento tácito, por lo que la misma debe ser construida como parte del proceso de elicitación y modelado.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

El proceso de requisitos propuesto por Leite et al. [13] se basa en la creación y uso de modelos en lenguaje para lograr un mayor involucramiento de los clientes. Este proceso utiliza el enfoque “elicitación guiada por el modelo” en cada una de sus etapas [28], es decir, la estrategia consiste en capturar la información necesaria en el contexto de aplicación para construir el/los modelo/s pertinentes a cada etapa: i) comprender el vocabulario del contexto de aplicación, creando el LEL; ii) comprender el comportamiento en el contexto observable, creando un conjunto de Escenarios Actuales; iii) definir en colaboración con los clientes y usuarios el contexto del software a construir, creando un conjunto de Escenarios Futuros; y iv) explicitar los requisitos, creando el

documento de Especificación de Requisitos del Software.

Entonces, el primer modelo que se construye siguiendo este proceso de requisitos es el LEL [14], con el fin de obtener una mejora en la comprensión del vocabulario que se utiliza en el contexto del problema. Es por ello que este modelo inicial debe tener la mayor calidad posible, ya que los restantes modelos harán uso de la terminología definida en él y algunos de ellos pueden ser derivados del propio LEL [13]. Se han realizado varios estudios sobre este modelo, centrados principalmente en su verificación mediante diversas variantes de inspección [29] [30] y en el diseño de heurísticas más precisas para su modelado [31] [32].

Mediante verificaciones y validaciones sistemáticas en el proceso de requisitos se pueden lograr mejoras sustanciales en los modelos elaborados, al poder identificar y corregir diversos defectos. Sin embargo, aquel conocimiento no adquirido y, por ende, no representado en modelos, continuará siendo ignorado luego de dichas actividades. Para lograr mejorar la calidad del LEL promoviendo una mayor completitud del mismo, es que se debería disponer de mecanismos que colaboren en la elicitación. Se propone utilizar para ello la técnica de mapas conceptuales [25] como instrumento de apoyo a la elicitación.

Esta técnica de mapas conceptuales ya ha sido usada en un proyecto de investigación previo, con el fin de dar soporte a una variante de inspección del modelo LEL [30], dando resultados relativamente aceptables en cuanto a cantidad de defectos detectados y tiempo insumido [33]. Esta variante consiste en construir un mapa conceptual por cada término definido en el LEL, donde las oraciones en cada definición del término son proposiciones en el mapa conceptual. A partir de la lectura sistemática de cada mapa se detectan cierto tipo de defectos, principalmente ambigüedades y omisiones. Aunque dentro de estas últimas, el proceso de inspección puede sugerir términos candidatos del LEL, no es posible establecer la falta de información en la definición de términos ni la

falta de otros términos, que escapen al conjunto de texto proporcionado por el propio modelo LEL.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo del trabajo de investigación en curso es definir heurísticas más precisas en la adquisición de la información necesaria para elaborar el modelo LEL en la etapa inicial del proceso de requisitos descripto, de manera tal de alcanzar un mayor nivel de completitud.

En tal sentido, se realizará un estudio sobre la elicitación de conocimiento en el contexto del problema, utilizando técnicas tradicionales de recolección de hechos [34], frente al uso de mapas conceptuales como técnica de la Gestión del Conocimiento, de manera tal de construir el modelo LEL en base a estos mapas. Se analizarán las diferencias entre los modelos léxicos construidos siguiendo las heurísticas establecidas en [13] [31] aplicando técnicas de elicitación tradicionales, y los mismos modelos construidos utilizando los mapas conceptuales en la elicitación. Se estudiará también si los mapas conceptuales efectivamente pueden colaborar en la adquisición de conocimiento tácito, o deberán complementarse con otros instrumentos, tales como los mapas mentales [35].

En una primera aproximación al uso de mapas conceptuales como soporte a la elicitación, se ha construido un mapa en el contexto del alquiler de bicicletas en una ciudad (ver Figura 1). A partir del mapa conceptual, se elaboró una versión inicial del modelo LEL, sin seguir ninguna heurística a priori. La Figura 2 presenta dos términos del LEL, definidos en base al mapa conceptual de la Figura 1. Mediante este ejercicio preliminar se observó que disponer del mapa conceptual facilitó detectar términos candidatos al LEL, al identificar visualmente los conceptos ubicados en el mapa. A través de las relaciones entre conceptos del mapa, se pudieron describir varios de los *impactos* de cada término (componente del término donde se define su connotación en el contexto bajo estudio). También la visualización de las relaciones entre conceptos en el mapa

conceptual incentivó que en la definición de cada término se utilizaran otros términos del LEL, contribuyendo esto a cumplir con los principios pilares del LEL [13]: *circularidad*

(maximizar el uso de términos en la definición de otros términos) y *vocabulario mínimo* (minimizar el uso de términos externos al LEL).

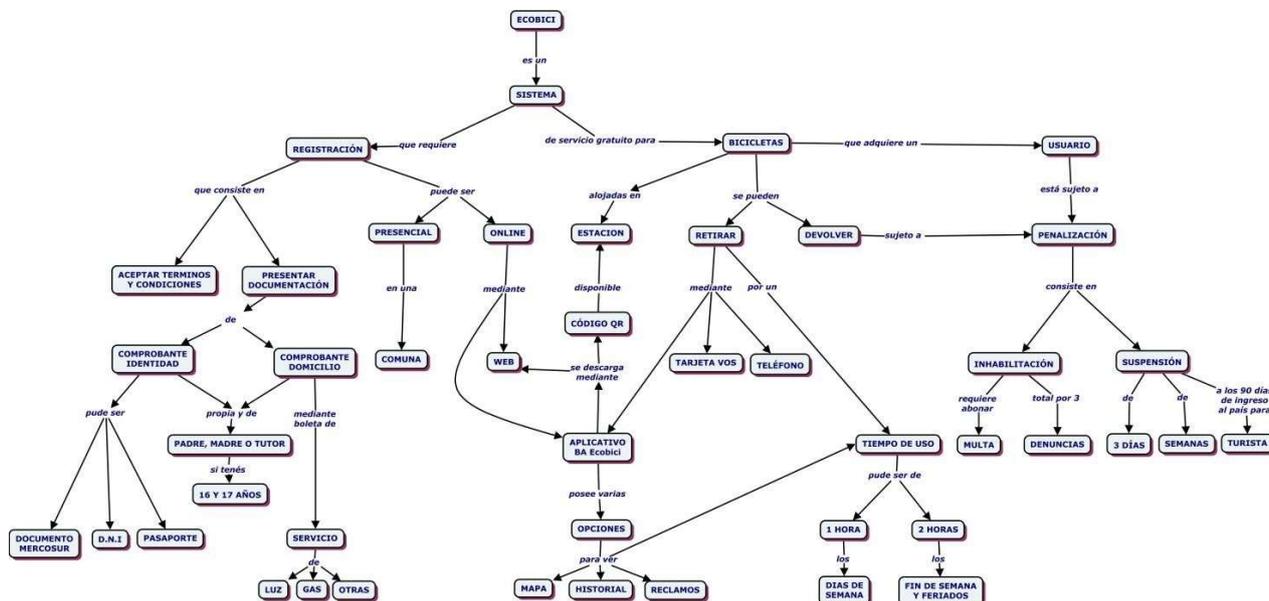


Figura 1. Mapa conceptual del sistema de alquiler de bicicletas

<p>Nombre: PROCESO DE REGISTRO</p> <p>Tipo: Verbo</p> <p>Noción:</p> <ul style="list-style-type: none"> Es la actividad que realiza el <u>usuario</u> para acceder al <u>sistema Ecobici</u>. Se realiza en forma presencial en una <u>comuna</u>. Se realiza en forma online mediante el sitio Web de la Ciudad o la <u>aplicación BA Ecobici</u>. <p>Impacto:</p> <ul style="list-style-type: none"> El <u>usuario</u> debe presentar la <u>documentación</u>. El <u>usuario</u> debe aceptar los <u>términos y condiciones</u>. Se valida la <u>documentación</u>. Se confirma el registro en el <u>sistema Ecobici</u>. Se habilita la <u>tarjeta VOS</u> para utilizar el <u>sistema Ecobici</u>, si el <u>usuario</u> ya la posee.
<p>Nombre: USUARIO</p> <p>Tipo: Sujeto</p> <p>Noción:</p> <ul style="list-style-type: none"> Es una persona que está registrada en el <u>sistema Ecobici</u>. <p>Impacto:</p> <ul style="list-style-type: none"> Realiza el <u>proceso de registro</u>. Puede <u>retirar la bicicleta</u>. Puede <u>devolver la bicicleta</u>. Puede ser <u>suspendido</u> o <u>inhabilitado</u>.

Figura 2. Términos del LEL del sistema alquiler de bicicletas

lo largo de todo el proceso de desarrollo de software y del cual pueden derivarse otros modelos.

4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto de la Universidad de Belgrano participan tres investigadores, dos de ellos en formación, mientras que en el tema de completitud dentro del proyecto de la Universidad Nacional del Oeste participa un investigador con dos alumnos becarios.

El año pasado, el Lic. Alberto Sebastián aprobó su Trabajo Final de Especialización “Verificación de un modelo Léxico Extendido del Lenguaje con Mapas Conceptuales” en la UCA, el que está directamente relacionado con el trabajo de investigación en curso.

5. BIBLIOGRAFÍA

[1] Ackoff, R. (1974) Redesigning the Future: Systems Approach to Societal Problems. John Willey & sons. ISBN: 978-0471002963.

[2] Finkelstein, A., Dowell, J. (1996) A comedy of Errors: The London Ambulance Service Case Study. 8th Int Workshop on Software Specification and Design, IEEE Computer Society Press.

[3] Breitman, K., Leite, J.C.S.P., Finkelstein, A. (1999) The World’s a Stage: A Survey on Requirements

En base al trabajo propuesto, se espera mejorar la calidad del modelo Léxico Extendido del Lenguaje, el cual es utilizado a

- Engineering using a Real-Life Case Study. *Journal of the Brazilian Computer Society*, 6(1):13-37.
- [4] Interim Report on the August 14, 2003 Blackout, New York Independent System Operator (NYISO), Enero2004, <http://www.hks.harvard.edu/hepg/Papers/NYISO.blackout.report.8.Jan.04.pdf>
- [5] Audit of the Status of the Federal Bureau of Investigation's Sentinel Program, U.S. Department of Justice Office of the Inspector General Audit Division, Audit Report 14-31, Sep. 2014, <https://oig.justice.gov/reports/2014/a1431.pdf>
- [6] Nuseibeh, B., Easterbrook, S. (2000) Requirements Engineering: A Roadmap. *Future of SE Track 2000*, pp.35-46
- [7] Alexander, I., Maiden, N. (2004) Scenarios, Stories, Use Cases. *Through the Systems Development Life-Cycle*. John Wiley & Sons.
- [8] Seyff, N., Maiden, N., Karlsen, K., Lockerbie, J., et al. (2009) Exploring how to use scenarios to discover requirements. *Requirements Engineering*, 14(2):91-111.
- [9] Antonelli, L., Rossi, G., Leite, J.C.S.P., Oliveros, A. (2012) Deriving requirements specifications from the application domain language captured by Language Extended Lexicon. 15th Workshop on Requirements Engineering.
- [10] Berry, D.M., Kamsties, E. (2004) Ambiguity in Requirements Specification. En: Leite & Doorn (eds.) *Perspectives on Software Requirements*, pp.7-44. Kluwer Academic Publishers.
- [11] Doorn, J.H., Ridao, M. (2003) Compleitud de Glosarios: Un Estudio Experimental. 6th Workshop on Requirements Engineering.
- [12] Ridao, M., Doorn, J. (2006) Estimación de Compleitud en Modelos de Requisitos Basados en Lenguaje Natural. 9th Workshop on Requirements Engineering, ISSN: 1413-9014, pp. 151-158.
- [13] Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridao, M.N. (2004) Defining System Context using Scenarios. En: *Perspectives on Software Requirements*, Kluwer Academic Publishers. ISBN:1-4020-7625-8,cap.8,pp.169-199.
- [14] Leite, J.C., Franco, A.(1993) A Strategy for Conceptual Model Acquisition. *RE'93*, pp.243-246.
- [15] Leonardi, M.C., Ridao, M.N., Mauco, M.V., Felice, L. (2015) A Natural Language Requirements Engineering Approach for MDA. *International Journal of Computer Science, Engineering and Applications*, 5(1):1-18.
- [16] Mighetti, J., Hadad, G. (2016) Uso de un Léxico y Escenarios para Mitigar Amenazas a Requisitos en el Desarrollo Global de Software, 19th Workshop on Requirements Engineering, pp. 407-420.
- [17] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2012) Un abordaje al problema de completitud en requisitos de software. XVIII Congreso Argentino de Ciencias de la Computación. pp. 827-836.
- [18] Hadad, G.D.S., Litvak, C.S., Doorn, J.H. (2014) Problemas y Soluciones en la Compleitud de Modelos en Lenguaje Natural. II Congreso Argentino de Ingeniería. ISBN:978-987-1662-51-7.
- [19] Martínez, S., Oliveros, A., Zuñiga, J., Corbo, S., Forradelas, P. (2014) Aprendizaje de la elicitación y especificación de requerimientos. XX Congreso Argentino de Ciencias de la Computación.
- [20] Nonaka, I., Takeuchi, H. (1995) *The Knowledge creating company. How japanese companies create the dynamics of innovation*. Oxford University Press.
- [21] Hinojosa, C., Raura, G., Fonseca, C., Dieste, O. (2015) *La Gestión del Conocimiento Aplicada en la Ingeniería de Requisitos: Un Caso de Estudio en Ecuador*. 18th Workshop on Requirements Engineering.
- [22] Belly, P. (2014) *Emprender el camino de la Gestión de conocimiento*. Temas Grupo Editorial, 1ra. ed., Buenos Aires. ISBN: 978-987-1826-80-3.
- [23] Gallego, D., Ongallo, C. (2004) *Conocimiento y Gestión*, Pearson Prentice-Hall.
- [24] García Martínez, R. (1994) *Adquisición de Conocimiento*. En Abecasis, S. y Heras, C. *Metodología de la Investigación*. Nueva Librería. ISBN 950-9088-65-x.
- [25] Novak, J., Cañas, A. (2008) *The Theory Underlying Concept Maps and How to Construct and Use Them*. Technical Report, Florida Institute for Human and Machine Cognition.
- [26] Ausubel, D., Novak, J., Hanesian, H. (1978) *Educational Psychology: A Cognitive View*, 2ª ed. New York: Holt, Rinehart & Winston.
- [27] Novak, J., Gowin, D.B. (1988) *Aprendiendo a aprender*. Ediciones Martínez Roca, Barcelona.
- [28] Kaplan, G., Doorn, J.H. (2015) *Advanced & Delayed Information in Requirements Engineering*. En: *Encyclopedia of Information Science and Technology*, Third Edition. IGI Global.
- [29] Kaplan, G., Hadad, G., Doorn, J. Leite, J. (2000) *Inspección del Léxico Extendido del Lenguaje*. 3rd Workshop on Requirements Engineering, pp.70-91.
- [30] Sebastián, A., Hadad, G.D.S. (2015) *Mejoras a un Modelo Léxico mediante Mapas Conceptuales*. XXI Congreso Argentino de Ciencia de la Computación, pp. 526-535. ISBN: 978-987-3724-37-4.
- [31] Hadad, G., Doorn, J., Kaplan, G. (2009) *Creating Software System Context Glossaries*. En: *Encyclopedia of Information Science and Technology*, 2nd Edition, pp.789-794. IGI Global.
- [32] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2014) *Heurísticas para el modelado de requisitos escritos en lenguaje natural*. XX Congreso Argentino de Ciencias de la Computación, pp. 682-691.
- [33] Sebastián, A., Hadad, G.D.S. (2016) *Experimento Controlado en la Inspección de un Léxico mediante Mapas Conceptuales*. III Congreso Argentino de Ingeniería, Chaco.
- [34] Carrizo, D., Dieste, O., Juristo, N. (2008) *Study of elicitation techniques adequacy*. 11th Workshop on Requirements Engineering, pp.104-114.
- [35] Buzan, T. (2002) *El libro de los mapas mentales*. Ediciones Urano. ISBN-13: 978-8479531461

Especificaciones Formales Tempranas del Comportamiento de Sistemas de Software

Fernando Asteasuain -Manuel Dubinsky – Federico Díaz – Juan Lagostena

Contacto: fasteasuain@undav.edu.ar

Ing. en Informática – Dpto. Tecnología y Administración – Universidad Nacional de Avellaneda

Resumen

El objetivo general del presente proyecto es facilitar la especificación formal del comportamiento de artefactos de software. La especificación formal del comportamiento esperado de artefactos de software ha sido identificada como uno de los mayores obstáculos para el desarrollo de software basado en modelos, y para la transferencia de técnicas de validación y verificación formal como model checking.

La comunidad en Ingeniería de Software cree que una posible solución debe poder combinar varios condimentos: deben permitir la descripción parcial del comportamiento, la facilidad para especificar sistemas abiertos, se deben usar sintaxis simples y conocidas, deben basarse en notaciones con suficiente poder expresivo y deben soportar heterogeneidad (combinando declaratividad para estar cerca de la especificación de los requerimientos y notaciones operacionales basadas en autómatas o formalismos similares).

El objetivo específico de esta investigación es el desarrollo de un nuevo lenguaje declarativo con el suficiente poder expresivo para especificar el comportamiento de sistemas abiertos, y la capacidad para operacionalizar las especificaciones. En el mismo se combinará la posibilidad de describir comportamiento parcial, permitiendo el modelado incremental, junto con la posibilidad de especificar el comportamiento de artefactos describiendo su interacción con un ambiente o contexto externo, habilitando las especificaciones de sistemas abiertos.

Palabras claves: *Lenguaje Declarativo, Modelado de Comportamiento, Especificación de Comportamiento.*

Contexto

El presente proyecto se encuentra enmarcado dentro del proyecto UNDAV CYT denominado “Especificaciones formales tempranas del comportamiento de sistemas de software”, dirigido por el Dr. Fernando Asteasuain. El grupo de investigación está conformado por los autores del presente trabajo. El proyecto tiene una duración de dos años, y actualmente está en ejecución el segundo año del proyecto. El proyecto está financiado 100% por la Universidad Nacional de Avellaneda.

1. Introducción

El desarrollo de software libre de errores representa el mundo ideal para la Ingeniería de Software, y todas las herramientas y técnicas buscan desde hace años acercarse cada vez más a ese objetivo de máxima. En la actualidad se estima que buena parte del tiempo que insume el desarrollo de software se lo lleva el proceso de detección y corrección de errores o “bugs” [8]. Diversas técnicas se han elaborado desde la comunidad de Ingeniería de Software para intentar atacar este problema. Entre ellas, el proceso de Testing ha ido creciendo en importancia y alcance, desde los clásicos tests de caja blanca, negra, de unidad, regresión, hasta métodos más complejos como la generación automática de casos de test [10] o el testing de mutación [11], entre otros. Pese a todo el crecimiento de esta área, el proceso de testing es por naturaleza incompleto, ya que no garantiza la ausencia de errores. La comunidad busca entonces apoyo en áreas como la validación y verificación formal de software, y el desarrollo de software basado en modelos [27]. Entre estas técnicas se destaca Model Checking [1], que en pocas palabras busca

establecer de manera automática si un modelo dado de un sistema satisface o no reglas que describen su comportamiento.

Sin embargo, existen obstáculos para poder transferir estas herramientas al mundo industrial del desarrollo de software [2-4]. Uno de los más importantes es la dificultad para completar el ciclo de la especificación de propiedades y de modelos: los lenguajes formales usados para la especificación requieren usuarios expertos, y las especificaciones obtenidas en ocasiones no son concisas, y resultan en ocasiones difíciles de manipular, entender, comunicar y validar [3, 4, 12]. La especificación de propiedades es un proceso que incluye dos pasos: 1) escribir la propiedad en un lenguaje formal y 2) revisar, validar esa propiedad para asegurarse que está describiendo el comportamiento que se tiene en mente. La mayoría de las aproximaciones (por ejemplo, basadas en lógicas temporales como LTL o notaciones basadas en autómatas) requieren usuarios expertos para poder escribir las propiedades, lo cual representa un importante obstáculo a la hora de adoptar técnicas formales de verificación [2,4,12]. Adicionalmente, estas aproximaciones no proveen al usuario de instrumentos que faciliten la validación de las propiedades: es una tarea compleja revisar propiedades, compararlas, modificarlas. Todos estos son aspectos claves a la hora de explorar y entender las implicancias de una propiedad.

Otro punto mencionado como un problema para los lenguajes de especificación es su poder expresivo. Muchos autores han señalado problema de expresividad en lógicas temporales como LTL [18,21] y se han propuesto extensiones como ser la inclusión de autómatas de Büchi en lógicas temporales [21]. Sin embargo, estas extensiones vuelven demasiado complejos los lenguajes, debilitándose su poder de uso.

Uno de los intentos para atacar estos problemas resulta del lenguaje declarativo Feather Weight Visual Scenarios (FVS) [12,17,28]. Uno de los principales aportes de FVS fue facilitar la tarea de la descripción y validación de propiedades, proveyendo un lenguaje gráfico y declarativo,

donde se obtienen especificaciones simples y concisas. Además, FVS cuenta con un alto poder expresivo, capaz incluso de expresar propiedades Ω -regulares.

Sin embargo, estos y otros avances [6-7,19-23] resultan insuficientes para llevar a cabo por completo la transferencia al mundo industrial. Por ejemplo, FVS está limitado a lógicas temporales razonando únicamente sobre eventos en sistemas cerrados. En este sentido, la comunidad en Ingeniería de Software cree que una posible solución debe poder combinar varios condimentos: deben permitir la descripción parcial del comportamiento [15,26], la facilidad para especificar sistemas abiertos [26,35], se deben usar sintaxis simples y conocidas [23, 25], deben basarse en notaciones con suficiente poder expresivo [12,18,20,21], y deben soportar heterogeneidad (combinando declaratividad para estar cerca de la especificación de los requerimientos [12] y notaciones operacionales basadas en autómatas o formalismos similares [4]).

Dado este contexto, la línea de investigación detrás del presente proyecto es generar una sinergia entre tres líneas de investigación: descripción declarativa de propiedades, especificaciones parciales y síntesis de comportamiento para sistemas abiertos a partir de objetivos.

Sistemas abiertos implica por un lado el diseño y manejo de algoritmos de control en Ingeniería de Software, y por otro, la posibilidad de contemplar el comportamiento esperado del ambiente, visión común dentro de la Ingeniería de Requerimientos [24]. Para mencionar un pequeño ejemplo de interacción con un ambiente, basta considerar el siguiente caso basado en el modelado de un sistema para el manejo de las luces interiores de un auto [16]. En el mismo, se busca evitar que la batería se descargue cuando funciona en un modo “caro”, que ocurre cuando las luces interiores se encienden y el auto está apagado. Para cumplir con este requerimiento la solución recae en pedir que el auto se encienda (y así la batería se recargue) entre dos funcionamientos consecutivos de la batería en

modo “caro”. Sin embargo, que el auto se encienda o no está fuera del control del sistema de luces, ya que forma parte de un comportamiento controlado por el ambiente o entorno. Este es sólo un pequeño ejemplo de la necesidad de extender FVS para especificar y sintetizar comportamiento en sistemas abiertos.

La síntesis de comportamiento ha sido atacada metodológicamente y algorítmicamente desde la síntesis y generación automática de controladores [5,9,14]. Se proveen mecanismos para diferenciar y distinguir el comportamiento que un componente puede hacer y controlar, y aquellas cosas que están fuera de su alcance: el comportamiento que proviene del entorno o ambiente que interactúa con el sistema. Sin embargo, existen limitaciones desde el punto de vista del lenguaje de especificación. Estas aproximaciones están fuertemente basadas en lógicas temporales (LTL) y algunas extensiones como fluents [3]. La expresividad de estos lenguajes ya ha sido desafiada por la comunidad, por lo que existe la necesidad de basarse en lenguajes de especificación con mayor poder expresivo. Asimismo, en estas técnicas no es posible determinar el controlador más general cuando se evalúan propiedades de tipo “liveness”, lo cual constituye una limitación conceptual de estas herramientas.

Por otro lado, existen otras aproximaciones basadas en especificaciones parciales que también han atacado el problema de expresividad en los lenguajes de especificación [6,15]. Estas aproximaciones se alejan de las estructuras lineales para enfocarse en estructuras “branching time”. Una característica de esta generación de especificaciones parciales es que posibilitan la descripción, manipulación y cómputo de posibles implementaciones que satisfacen los requerimientos planteados. En este contexto, es posible también el uso de MTS para tomar decisiones sobre la especificación operacional que más se ajuste a los requerimientos, particularmente en aquellas situaciones donde existe más de una implementación posible y se

debe guiar al usuario para llegar a una única solución. Si bien el aporte de las estructuras “branching time” es importante a la hora de especificar comportamiento, existen limitaciones a la hora de trasladarlas a sistemas abiertos, donde se describen propiedades a ser satisfechas por un ambiente o entorno. En estas estructuras se describe el comportamiento de un componente, pero sin hacer distinción entre lo que el componente puede hacer y lo que puede controlar.

2. Líneas de Investigación y Desarrollo

En el presente trabajo se están explorando las siguientes líneas de investigación:

- Análisis de expresividad en notaciones operacionales declarativas.
- Combinar distintas fuentes de especificación buscando obtener modelos operacionales y ejecutables.
- Aplicar el lenguaje desarrollado en casos de estudio de relevancia industrial como protocolos de comunicación y otros relacionados a los dominios de aplicación.
- Publicar los resultados parciales y finales en conferencias y revistas relativas al área de investigación
- Analizar especificaciones parciales basadas en estructuras "branching time".
- Diseñar extensiones a especificaciones parciales para manejar sistemas abiertos.
- Analizar técnicas de síntesis de controladores actuales.
- Potenciar lenguajes de especificación para síntesis de controladores.
- Generar extensión del lenguaje FVS en base a las extensiones detectadas para especificaciones parciales y síntesis de controladores.
- Desarrollar un nuevo lenguaje declarativo en base a las extensiones propuestas

- Generación y desarrollo de herramientas de software que den soporte a lo investigado.
- Modelado de comportamiento en arquitecturas de software.

3. Resultados Obtenidos/Esperados

En esta primera etapa los resultados alcanzados fueron:

- Publicaciones en congresos nacionales (JAIIO, CONAISSI).

- Combinar distintas fuentes de especificación: La traducción de escenarios y reglas FVS a autómatas de Buchi permitió combinar especificaciones declarativas realizadas en FVS con otras operacionales como aquellas basadas en notaciones de autómatas. Esto le provee a todo el entorno FVS una gran versatilidad ya que el comportamiento puede ser razonado y analizado desde requerimientos expresados en diversas fuentes y variantes. Esto es fundamental para continuar profundizando los casos de estudio, en particular haciendo énfasis en protocolos de comunicación.

- Potenciar crecimiento líneas de investigación y desarrollo dentro de la UNDAV: La difusión de los resultados logrados en el transcurso del año por parte de los docentes miembros del grupo permitió tanto el acercamiento de otros docentes como así de estudiantes de la carrera. En este sentido la investigación del proyecto permitió consolidar temas concretos para la realización de tesis por parte de los estudiantes, así como también la posibilidad de que realicen la PPS (Práctica Profesional Supervisada) dentro del proyecto de investigación. De la misma manera, también está contemplada una fuerte colaboración con el Laboratorio de Software Libre de la carrera de Ingeniería en Informática de la UNDAV, tanto para la obtención de casos de estudio como para el desarrollo de software.

- Aplicar el lenguaje desarrollado en casos de estudio de relevancia industrial como protocolos de comunicación y otros relacionados a los dominios de aplicación: Se

trabajó en entornos de arquitectura de software, modelando interacciones complejas de comunicación entre componentes. La diversidad en el tipo de comunicaciones incluye comunicación entre artefactos de software, entre artefactos de software y hardware, así como también la interacción con capas intermedias de middleware como sensores.

- Desarrollar un nuevo lenguaje declarativo en base a las extensiones propuestas. Como primer paso se decidió estudiar la aplicabilidad de FVS para relevar y razonar sobre comportamiento arquitectónico. El avance actual de la investigación sobre este dominio permitió concluir que no fue necesario extender FVS de manera significativa en esta primera etapa, ya que fue lo suficientemente expresivo como para poder modelar comportamiento arquitectónico de relevancia.

- Construcción de herramientas de software: En esta primera etapa se desarrollaron herramientas de software que dan soporte al lenguaje FVS. Se desarrolló la herramienta GTxFVS, que da la posibilidad al ingeniero de Software de trabajar en un entorno moderno para la especificación declarativa de comportamiento. De la misma manera, se implementó una herramienta basada en nociones orientadas a aspectos para la relevación dinámica de arquitecturas de software.

Los resultados esperados incluyen:

- Consolidación del lenguaje FVS en el dominio de arquitecturas de software.

- Inclusión de nociones de especificación parcial de comportamiento y lógicas de tipo "branching".

- Síntesis de controladores bajo el contexto de sistemas abiertos.

- Publicaciones en revistas y congresos nacionales e internacionales.

- Dirección de tesis de licenciatura y supervisión de prácticas profesionales (PPS).

4. Formación de Recursos Humanos

Para este segundo año del proyecto están contempladas la incorporación de estudiantes avanzados para realizar tareas iniciales de investigación y en especial la realización de tesis de licenciatura. De la misma manera, se planea incorporar estudiantes para que realicen la PPS (Práctica Profesional Supervisada) dentro del proyecto de investigación. En particular, durante el primer cuatrimestre de 2017 un estudiante avanzado de la carrera, el estudiante Eric Loza, realizará su PPS desarrollando una herramienta de software dentro del marco del proyecto.

5. Bibliografía

- [1] Clarke, E.M. and Grumberg, O. and Peled, D. Model Checking. MIT Press. 1999
- [2] M. Dwyer, G. Avrunin, and J. Corbett. Patterns in property specifications for finite-state verification. In Proceedings ICSE, volume 99, 1999.
- [3] D. Giannakopoulou and J. Magee. Fluent model checking for event-based systems. In Proceedings of the 9th European software engineering conference, page 266. ACM, 2003.
- [4] R. Smith, G. Avrunin, L. Clarke, and L. Osterweil. Propel: An approach supporting property elucidation. In ICSE, volume 24, pages 11-21, 2002.
- [5] N. Piterman, A. Pnueli, and Y. Sa'ar. Synthesis of reactive (1) designs. Lecture notes in computer science, vol. 3855, pp. 364-380, 2006.
- [6] K. G. Larsen and B. Thomsen. A modal process logic. In LICS, pp 203–210. IEEE Computer Society, 1988.
- [7] G. S. Walia and J. C. Carver. A systematic literature review to identify and classify software requirement errors. Inf. Softw. Technol., 51(7):1087-1109, 2009.
- [8] Amalinda Post, Igor Menzel, Jochen Hoenicke, Andreas Podelski: Automotive behavioral requirements expressed in a specification pattern system: a case study at BOSCH. Requir. Eng. 17(1): 19-33 ,2012.
- [9] B. Jobstmann, S. Galler, M. Weiglhofer, and R. Bloem. Anzu: a tool for property synthesis, in Proceedings of the 19th CAV'07, pp. 258-262, Springer-Verlag, 2007.
- [10] Cohen, D.M. and Dalal, S.R. and Parelius, J. and Patton, G.C. The combinatorial design approach to automatic test generation. In IEEE Software Journal, volume 13, number 5, pages 83-88, 1996.
- [11] Offutt, A.J. and Pan, J. Automatically detecting equivalent mutants and infeasible paths. In Software Testing, Verification and Reliability Journal, volume 7, number 3, pages 165-192, 1997.
- [12] Fernando Asteasuain, Víctor Braberman. Specification Patterns can be formal and still easy, SEKE page 430-436 – 201.
- [13] A. Alfonso, V. Braberman, N. Kicillof, and A. Olivero. Visual timed event scenarios. In 26th ICSE'04, pages 168-177, 2004.
- [14] Nicolas D'Ippolito, Víctor Braberman, Nir Piterman, Sebastian Uchitel, Synthesising Non-Anomalous Event-Based Controllers for Liveness Goals. CM TRANSACTIONS ON SOFTWARE ENGINEERING AND METHODOLOGY, Volume in press - 2012
- [15] German Sibay, Víctor Braberman, Sebastian Uchitel, Jeff Kramer, Distribution of Modal Transition Systems, 18th International Symposium on Formal Methods FM 2012, Volume in press – 2012.
- [16] N. Noda and T. Kishi. An aspect-oriented modeling mechanism based on state diagrams. In 9th International Workshop on AOM, 2006.
- [17] F. Asteasuain, V. Braberman. “Declaratively building behavior by means of scenario clauses”. Requirements Engineering Journal. ISSN: 0947-3602. Diciembre 2015. DOI: 10.1007/s00766-015-0242-2
- [18] Z. Wu. On the expressive power of qltl. In Proceedings of the 4th international conference on Theoretical aspects of computing, pages 467-481. Springer-Verlag, 2007.
- [19] Kupferman, O. and Vardi, M. Module checking. In Computer Aided Verification, 75-86, 1996. Springer.
- [20] Shoham Ben-David, Marsha Chechik, Arie Gurfinkel, Sebastián Uchitel: CSSL: a logic for specifying conditional scenarios. SIGSOFT FSE 2011: 37-47
- [21] Ahmed Bouajjani, Yassine Lakhnech, Sergio Yovine: Model-Checking for Extended Timed Temporal Logics. FTRTFT 1996:306-326.
- [22] I. Krka, Y. Brun, G. Edwards, and N. Medvidovic. Synthesizing partial component-level behavior models from system specifications. In ESEC/FSE '09, pages 305–314. ACM, 2009.
- [23] Matthew B. Dwyer, George S. Avrunin, James C. Corbett: Patterns in Property Specifications for Finite-State Verification. ICSE 1999: 411-420
- [24] A. V. Lamsweerde. Goal-oriented requirements engineering: A guided tour. In RE'01 –
- [25] Margus Veanes, Colin Campbell, Wolfgang Grieskamp, Wolfram Schulte, Nikolai Tillmann, Lev Nachmanson: Model-Based Testing of Object-Oriented Reactive Systems with Spec Explorer. Formal Methods and Testing '08: 39-76
- [26] David Harel: Statecharts: A Visual Formalism for Complex Systems. Sci. Comput. Program. 8(3): 231-274 (1987)
- [27] DALAL, Siddhartha R., et al. Model-based testing in practice. En Proceedings of the 21st ICSE ACM, 1999. p. 285-294.
- [28] F. Asteasuain, V. Braberman “Specification patterns: formal and easy”. IJSEKE ISSN: 0218-1940. Vol. 25, No. 4 (2015) 669–700, DOI: 10.1142/S0218194015500060.

Evaluación de la Accesibilidad Web. Caso de Estudio: Sitios Web de la UNCA

Chayle, Carolina I., Herrera, Claudia M., Barrera, Maria A., Pauletto, Ana C. y Blanco, Sergio D.

Departamento de Informática/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
Maximio Victoria N° 55 - C.P: 4700 - San Fernando del Valle de Catamarca
Telefono: 03833- 435112 – int 168
cchayle@gmail.com, herrera.claudia.mabel@gmail.com

Resumen

En la sociedad del conocimiento se intensifica la comunicación de información a través de entornos Web. La accesibilidad Web como criterio de calidad en los sistemas informáticos y su interés a nivel internacional día a día cobra más relevancia, dada su implementación en diversos espacios socio-económicos-culturales.

Cabe señalar que existen legislaciones a nivel internacional, mientras que en la Argentina en el año 2010 se sancionó la Ley N° 26.653 de “Accesibilidad de la Información en las Páginas Web”. La accesibilidad Web se entiende como la posibilidad de que la información de la página Web, pueda ser comprendida y consultada por personas con discapacidad y por usuarios, que posean diversas configuraciones en su equipamiento o en sus programas.

También los sitios educativos, deberían ofrecer políticas que mejoren el acceso Web, como así mismo, cumplir normas de diseño universal, habilitando el uso de la información a la mayor cantidad posible de usuarios sin límites arbitrarios.

En este trabajo se pretende realizar una Evaluación de la Accesibilidad Web en los sitios Web de la Universidad Nacional de Catamarca (UNCa), de acuerdo con las recomendaciones realizadas en las Normas de Accesibilidad Web 2.0 propuestas por la ONTI, con el objeto de recomendar acciones a seguir en el diseño de los mismos para mejorar su calidad permitiendo de esta manera el acceso a todos sin restricciones.

PALABRAS CLAVES: Accesibilidad Web, UNCA.

Contexto

Se presentan los avances vinculados a la accesibilidad Web enmarcados en el proyecto “Evaluación de la accesibilidad Web. Caso de Estudio: Sitio Web de la UNCA”, perteneciente a los Proyectos de Investigación y Desarrollo de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA (2016-2017)

Introducción

Internet, por ser una herramienta global

de consulta desde hace ya un tiempo, es uno de los recursos más importantes utilizados para el intercambio de información. Su uso para todos los niveles educativos se producen por la facilidad de difusión, intercambio y socialización de los conocimientos. La WWW (World Wide Web) utiliza una nueva tecnología que permite acceder a los recursos de Internet de una manera sencilla en un ambiente interactivo donde la información es fácil de encontrar y acceder, a través de ella se ha generado una nueva cultura, la cybercultura, con su modo de pensar, de hablar y de sentir. Este flujo de información global permite a los usuarios interactuar con grupo de personas dispersas por el mundo. Permite salvar obstáculos de acceso físico, geográfico, de tiempo y de restricciones horarias. Pero, aquellas personas que sufren problemas sensoriales (problemas de visión, oído), cognitivos (dificultades en la lectura o la comprensión), o motrices (incapacidad para manejar total o parcialmente el mouse o el teclado), que podrían aprovechar al máximo las potencialidades de Internet, se enfrentan a una barrera digital que les imposibilita el acceso a esta importante herramienta. La accesibilidad Web se ha convertido en un problema común, sobre el que se han centrado varios organismos mundiales. Una de las más importantes es la desarrollada por la W3C (Consortio World Wide Web), a través de la Iniciativa para la Accesibilidad a la Web o WAI (Web Accessibility Initiative) en 1999 que propone por primera vez las recomendaciones denominadas “Guías de Accesibilidad al Contenido Web” (WCAG) y desde entonces soporta la evolución y el desarrollo de materiales y especificaciones para contribuir a la concientización, comprensión e implementación de la Accesibilidad en la Web. Uno de los temas objeto de estudio

por parte de diversos organismos mundiales es la accesibilidad web, entre las principales acciones se mencionan las abordadas por W3C, ISO, Fundación Sidar, Centro de Investigación y Desarrollo de Adaptaciones Tiflotécnicas (CIDAT), promovido por ONCE. En torno al estándar WCAG varios países han desarrollado su propia legislación de Accesibilidad. Ejemplos de estas legislaciones existentes son Estados Unidos; Reino Unido; la Legislación Italiana; en el 2009 el gobierno australiano, aprobó la Estrategia Nacional de Transición de Accesibilidad Web (Web Accessibility National Transition Strategy), entre otras. En el año 2010 en Argentina se sanciona, la Ley 26.653 de Accesibilidad de la Información en las Páginas Web. En el año 2011, se aprueba la Guía de Accesibilidad 1.0 para Sitios Web del Sector Público, como parte integrante de los Estándares Tecnológicos para la Administración Pública Nacional (ETAP). La accesibilidad Web se entiende como la posibilidad de que la información de la página Web, pueda ser comprendida y consultada por personas con discapacidad y por usuarios, que posean diversas configuraciones en su equipamiento o en sus programas. El cumplimiento de la ley favorece, a personas con discapacidad, y también a otros grupos de usuarios en condiciones desfavorables, como ser: personas con dificultades por el envejecimiento, por baja iluminación o espacios reducidos; personas con limitaciones tecnológicas o diferencias culturales, y quienes sean inseguros o inexpertos en el uso de dispositivos electrónicos. Los sitios educativos deberían impartir políticas que mejoren el acceso Web, como así mismo, cumplimentar normas de diseño universal, habilitando el uso de la información a la mayor audiencia posible sin límites de arbitrarios. Debido a su

gran relevancia, es importante que se logre la igualdad de oportunidades para su acceso. Es uno de los temas principales en los organismos interesados en la accesibilidad, ya que las limitaciones y el mal uso por parte de los diseñadores de las tecnologías imperantes de publicación Web están dando lugar a situaciones de imposibilidad de acceso a la información por muchas personas, especialmente para aquellas que poseen algún tipo de discapacidad, agravando así la denominada infoexclusión o brecha digital, que supone la discriminación de una parte importante del total de usuarios. Existen estándares internacionales para la accesibilidad web, los cuales ayudan a los usuarios con o sin discapacidad a mejorar la accesibilidad y lograr la independencia del dispositivo de acceso. Es responsabilidad de los medios que emiten la información asegurar su accesibilidad a las personas. En este sentido, los sitios educativos, deberían ofrecer políticas que mejoren el acceso Web, como así también, cumplir normas de diseño universal, habilitando el uso de la información a la mayor cantidad posible de usuarios sin ningún tipo de restricciones. En este sentido, se reconocen publicaciones de trabajos de investigación en universidades como ser "Evaluación de la accesibilidad de páginas web de universidades españolas y extranjeras incluidas en rankings universitarios internacionales", la principal novedad de este trabajo además de aplicar la nueva versión WCAG 2.0, es que incluye la evaluación simultánea de universidades españolas y extranjeras; lo cual permite la comparación de los resultados en ambos ámbitos o extranjeras, otra publicación es "Accesibilidad de Portales Web Universitarios", en este estudio se dan a conocer las principales barreras de accesibilidad detectadas en el análisis de

de páginas de Universidades españolas, en relación con las pautas vigentes WCAG 1.0.

Líneas de Investigación, Desarrollo e Innovación

El presente proyecto es un investigación aplicada en la cual se realizarán estudios de carácter descriptivo y comparativo, se analizarán los sitios Web de la Universidad Nacional de Catamarca, con el fin de determinar el grado de adhesión a la normativa vigente sobre accesibilidad establecidos en la Guía de Accesibilidad para Sitios Web del Sector Público Nacional, recomendada por la ONTI, de acuerdo con lo establecido en la Ley 26.653 de Accesibilidad de la Información en las Páginas Web.

Con respecto a la obtención de la información se prevé recolección de muestras mediante la observación directa a los sitios que pertenezcan a la UNCA. Las metodologías, técnicas y herramientas serán exhaustivamente analizadas en una fase posterior al estudio comparativo y se validarán sus resultados para elegir la que brinde un adecuado nivel de accesibilidad. Se utilizará un validador de accesibilidad automático o servicio en línea, y se complementarán con evaluación manual.

Resultados y Objetivos

Como resultado se espera obtener un diagnóstico del estado de situación de los sitios web de la UNCA, de acuerdo con las recomendaciones realizadas en las Normas de Accesibilidad Web 2.0 propuestas por la ONTI. Los resultados obtenidos permitirán dar recomendaciones a los desarrolladores de las páginas Web de la UNCA a la hora de diseñar éstas y así poder ayudar a los

usuarios con o sin discapacidad a mejorar la accesibilidad y lograr la independencia del dispositivo de acceso; y de esta forma acercar a la universidad al cumplimiento de la Ley 26.653.

Los objetivos del proyecto son:

- Evaluar la accesibilidad de los sitios Web de la Universidad Nacional de Catamarca, para obtener un nivel de calidad en todas sus páginas Web.
- Analizar requerimientos de Accesibilidad Web.
- Estudiar las herramientas conceptuales y prácticas del diseño para los sitios.
- Estudiar herramientas automáticas utilizadas para la medición de accesibilidad.
- Seleccionar la herramienta que mejor se adapte para el caso de estudio.
- Recomendar acciones a seguir en el diseño de los sitios de la UNCA.
- Acercar a la universidad al cumplimiento de la Ley Nacional N° 26.653 de Accesibilidad de la Información en las Páginas Web.

Formación de Recursos Humanos

El proyecto está integrado por dos profesores adjuntos, un docente auxiliar, y un alumno. El mismo tiene como objetivo la formación de recursos humanos a través de:

- Participación en eventos científicos regionales, nacionales e internacionales de la especialidad, como congresos, simposios, seminarios y cursos, por ello el programa de capacitación y formación de recursos humanos, contempla las siguientes actividades:
- Participación en cursos de actualización y posgrado en el área de estudio.
- Participación de los integrantes del proyecto en la dirección,

asesoramiento y evaluación de tesis de grado de la carrera de Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA.

- Desarrollo de tesis de posgrados por parte de los docentes integrantes.

En cuanto a la formación específica del equipo de docentes investigadores, se destaca que los mismos se encuentran abocados a actividades de capacitación y estudios de posgrado, para la elaboración de 2 (dos) tesis correspondiente a la carrera Maestría en Ingeniería del Software de la Universidad Nacional de San Luis.

Referencias

- OFICINA ESPAÑOLA, 2008 Oficina Española (2008) Word Wide Web - Guía Breve de Accesibilidad Web, Disponible en:
<http://www.w3c.es/divulgacion/guiasbrev es/accesibilidad>
- US Government Electronic and Information Technology Accessibility Standards, 2003.
 PUBLICLY AVAILABLE SPECIFICATION: A Guide to Good Practice in Commissioning Accessible Websites, 2006
- STANCA LAW, 2004 Torres R., S. Accesibilidad en la Web para las Personas con Discapacidad Mucho Dicho Pero Muy Poco Hecho. Revista Scientia Et Technica, ISSN 0122-1701.
www.redalyc.org/articulo.oa?id=84920503060
- SECRETARIA DE GABINETE, 2011. Jefatura de Gabinete de ministros, Bs. As. Técnicas para las pautas de accesibilidad. (2011).
http://www.jefatura.gob.ar/tecnicas-para-las-pautas-deaccesibilidad_p86.
- LEY DE ACCESIBILIDAD DE LA INFORMACIÓN EN PÁGINAS WEB,

2010. Ley 26653 de Accesibilidad de la Información en Páginas Web. Argentina. 2010.

<http://www.infoleg.gov.ar/infolegInternet/anexos/175000-179999/175694/norma.htm>

JGM, 2014 JGM. Jefatura de Gabinete de ministros, Bs. As. Técnicas para las pautas de accesibilidad. (2011).

http://www.jefatura.gob.ar/tecnicas-para-las-pautas-deaccesibilidad_p86

DISCAPNET, 2010. Accesibilidad de Portales web Universitarios [en línea].

Madrid:Discapnet.

http://www.discapnet.es/Castellano/areastematicas/Accesibilidad/Observatorio_infoaccesibilidad/informesInfoaccesibilidad/Paginas/AccesibilidaddePortalesWebUniversitarios.aspx

RIBERA, M.; Térmens, M.; Frías, A. (2009). La accesibilidad de las webs de las universidades españolas. Balance 2001-2006. Revista Española de Documentación Científica, vol. 32(3), 66-88.

HILERA, J.R.; Fernández, L.; Suárez, E.; Vilar, E.T. (2013). Evaluación de la accesibilidad de páginas web de universidades españolas y extranjeras incluidas en rankings universitarios internacionales. Revista Española de Documentación Científica, 36(1):e004. doi:<http://dx.doi.org/10.3989/redc.2013.1.91>

LEY 26653 de Accesibilidad de la Información en Páginas Web. Argentina. 2010.<http://www.infoleg.gov.ar/infolegInternet/anexos/175000-179999/175694/norma.htm>

EVALUACIÓN DE LA ACCESIBILIDAD DE PÁGINAS WEB DE UNIVERSIDADES ESPAÑOLAS Y EXTRANJERAS incluidas en rankings universitarios internacionales - <http://redc.revistas.csic.es/index.php/redc/article/viewFile/774/863>.

ONTI. Oficina Nacional de Tecnologías

de la Información.
www.jgm.gov.ar/sgp/paginas.dhtml?pagina=27

Evaluación de la Usabilidad en Sitios Web Móviles Adaptativos

Rocío A Rodríguez, Isabel B Marko, Pablo M Vera, Gabriela Y Vallés, Gaspar Acevedo Zain

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
 Departamento de Ingeniería e Investigaciones Tecnológicas
 Universidad Nacional de La Matanza
 Florencio Varela 1903, San Justo, Provincia de Buenos Aires
 {rocio.rodriguez, imarko, pvera, gvalles, gacevedo} @unlam.edu.ar

RESUMEN

Si bien los dispositivos móviles se encuentran actualmente insertos en la sociedad de forma masiva, aún en la web no se cuenta con demasiados sitios realizados específicamente para estos dispositivos. Siendo una práctica más habitual el desarrollar un solo sitio web, el cual pueda adaptarse a los distintos dispositivos, cambiando estratégicamente la forma en la que se visualizan sus elementos. Es por ello que en el presente proyecto se pone el foco en la calidad de los sitios web adaptativos, notándose deficiencias comunes en los mismos a fin de planificar la construcción de métricas que permitan por una parte evaluar los mismos y por otra parte considerar, al momento de construirlos, parámetros que ayuden generar sitios web adaptativos de mayor calidad.

Palabras clave: Diseño Web Adaptativo, Sitios Web, Dispositivos Móviles, Usabilidad

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al grupo de investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza (UNLaM). El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D

alumnos de la universidad de las carreras de Ingeniería.

Cabe destacar que el presente proyecto tiene como antecedente un proyecto previo desarrollado en la UNLaM el cual estuvo basado en evaluación de sitios web móviles. El proyecto actual permitirá continuar con esta línea de investigación, permitiendo profundizar en los sitios web adaptativos.

1. INTRODUCCIÓN

Si bien tener un sitio web móvil permite establecer los contenidos que se van a priorizar, disminuir la cantidad de información visual que carece de importancia, ofreciendo una solución ágil para un usuario en movimiento; son pocas las instituciones que cuentan con una solución web de escritorio y otra móvil; siendo el diseño adaptativo (responsive design) una primera iniciativa significativa para mejorar la visualización de los sitios en pantallas reducidas. El diseño adaptativo es una buena solución para poder ajustar los controles de una página web a distintos tamaños de pantalla.

Actualmente existe una clara tendencia en el crecimiento del uso de los dispositivos móviles para conectarse a internet. Cada vez menos personas esperan llegar a su casa y sentarse frente a una computadora para consultar los emails, leer alguna noticia o consultar alguna página particular. La disponibilidad y practicidad de los dispositivos móviles hacen que sean la

primera fuente de referencia para estos casos. Sin embargo, al querer ingresar a muchos sitios web desde los dispositivos móviles surgen diversos problemas ya que la mayoría de los sitios no fueron diseñados y estructurados pensando en dispositivos móviles sino para ser visualizados desde una computadora. Por ello se propone el enfoque mobile first (pensar primero en el usuario móvil), “es pensar en el móvil como punto de partida, esto obliga a concentrarse en lo esencial de un producto y a hacer foco solo en lo que tiene sentido para este dispositivo” [1].

Un relevamiento realizado previamente por el equipo de investigación (cuyos resultados han sido publicados en una revista indexada [2]), muestra que sobre 135 municipios que tiene la Provincia de Buenos Aires a lo que se le suma la Ciudad Autónoma de Buenos Aires, dando un total de 136 municipios; tan sólo 5 de ellos tiene sitio web móvil (3,68%), sin embargo 40 de ellos tienen un diseño web adaptativo (29,41%). Con lo cual puede notarse que esta práctica de realizar diseños adaptativos es cada vez más común. Cabe destacar que se trata del mismo sitio web pero con visualizaciones distintas que permiten que se adapte a las características de pantalla de los distintos dispositivos.

La web móvil tiene diversos problemas, algunos de ellos que parecen ser heredados de las páginas web tradicionales esto se debe a que no se ha puesto el foco en los usuarios móviles. Entre esos inconvenientes es posible mencionar:

- (1) Peso de la página;
- (2) Exceso de imágenes;
- (3) Imágenes no preparadas para dispositivos móviles;
- (4) Contenido no priorizado;
- (5) Páginas demasiado grandes, lo que obliga a realizar zoom y scroll en ambas direcciones.

Para lograr evitar desarrollar distintas versiones de un sitio web y hacer que el mismo se visualice correctamente tanto en un dispositivo móvil como en una computadora de escritorio nace el diseño adaptativo.

Es importante comprender que el diseño adaptativo no es una versión diferente del sitio web, sino que es el mismo sitio que se va adaptando y acomodando sus elementos para una mejor visualización en el dispositivo de acceso. Con lo cual no cuenta con todas las ventajas que tendría la construcción de un sitio móvil.

El diseño adaptativo se base en la utilización de hojas de estilo de nivel 3 con CSS Media Queries [3] lo que permite aplicar distintos estilos según sea el tamaño de la pantalla en la cual se está visualizando el sitio web. De esta forma una misma página se podrá ver “correctamente” indistintamente del tamaño en la que se esté visualizando.

“El diseño web adaptativo utiliza hojas de estilo en cascada complejas, para modificar de forma fluida la apariencia de un sitio web dependiendo del ancho de la ventana de visualización del dispositivo en uso” [4].

La figura 1 muestra como varía la visualización de un mismo sitio web al cambiar el tamaño de la pantalla. En la misma pueden diferenciarse tres medidas:

- tamaño grande (A) que es adecuado por ejemplo para la visualización en una computadora de escritorio,
- tamaño mediano (B) que se adapta por ejemplo para ser visualizada en una tablet
- tamaño pequeño (C) adecuada para un dispositivo con pantalla reducida como por ejemplo un teléfono celular.

Sin embargo, la mayoría de los sitios web implementados utilizando esta tecnología no lo hacen pensando en el usuario móvil y presuponen que por el solo hecho de emplearla ya están creando un sitio amigable para dichos usuarios.



Figura 1. Visualización de un sitio adaptativo en tres tamaños de pantalla

Tal como se mencionó previamente, el desarrollo adaptativo se basa en la utilización de CSS Media Queries [3] aplicando distintos estilos según el tamaño de visualización de la página. No obstante, existen diversos frameworks con estilos pre-diseñados que se pueden aplicar rápidamente para la construcción de sitios web adaptativos. Estos estilos generalmente están complementados con archivos JavaScript que agregan funcionalidad a ciertos controles, haciéndolos dinámicos e incluso permitiendo además animar parte de la interfaz gráfica. Entre los frameworks existentes pueden mencionarse MATERIALIZECSS [5], BOOSTRAP [6], FOUNDATION [7], SKELETON [8]. Sin embargo, al utilizarlos puede notarse que la mayoría de estos frameworks no están correctamente diseñados ya que se basan en unidades absolutas (píxeles) para determinar los distintos tamaños de pantalla cuando lo correcto sería utilizar unidades relativas (em o porcentajes) para asegurar una correcta visualización en pantallas pequeñas más allá de la resolución que posean. “Para que un sitio sea usable cuando se ve en una variedad

de tamaños de pantalla, un cierto grado de flexibilidad debe ser incorporada en el diseño. Div, fuentes, botones, y todos los demás elementos deben ser capaces de crecer y contraerse sobre la base del tamaño de la pantalla. En el diseño web adaptativo, esto se logra mediante el abandono de los píxeles, en favor de porcentajes o unidades relativas llamadas ems” [9].

Un mal uso del diseño adaptativo incorporará diversos errores, entre ellos: (1) Páginas muy pesadas; (2) Gran cantidad de imágenes; (3) Imágenes no preparadas para dispositivos móviles; (4) Contenido no priorizado; (5) Páginas demasiado extensas... Siendo los mismos problemas mencionados anteriormente al visualizar la web tradicional en un dispositivo móvil, solucionando únicamente el no tener que realizar scroll en dos direcciones ó zoom para visualizar una parte de la pantalla.

Es por ello que resulta sumamente importante poder construir métricas que permitan evaluar y ofrecer reportes que permitan mejorar los sitios web adaptativos.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Análisis de las soluciones realizadas con frameworks existentes
- Definición de principios y pautas de usabilidad
- Construcción de Métricas para evaluar sitios web adaptativos

3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha comenzado a trabajar en la definición de pautas para el diseño adaptativo, las cuales se muestran en la figura 2 y fueron explicadas en una publicación realizada en un congreso académico [10].



Figura 2. Pautas para el Diseño Adaptativo con el Enfoque Mobile First

Teniendo estas pautas como base, se continúa profundizando en la temática del presente proyecto. Dado que este proyecto tiene su inicio en el presente año aún no hay resultados formales para reportar.

Se espera como contribución científica obtener un conjunto de métricas que permitan relevar sitios web adaptativos detectando sus deficiencias.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por 7 personas.

- 4 Docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 3 alumnos de ingeniería quienes se están formando en actividades de Investigación y Desarrollo los cuales cuentan con becas asignadas

Vinculado con esta temática se encuentra en realización una tesis de doctorado en la Universidad Nacional de La Plata.

5. BIBLIOGRAFIA

- [1] Cuello J.; Vittone J. “Diseñando Apps móviles”. Primera edición: junio de 2013.
- [2] Rodríguez, R., Vera, P., Marko, I., Alderete C., Conca A. “El Gobierno Electrónico y la Implementación de las TIC para Brindar Nuevos Canales de Comunicación. Revista Latinoamericana de Ingeniería de Software, 3(5), 2015, pp. 187-196. ISSN 2314-2642
- [3] W3C Media Queries. 2012. <https://www.w3.org/TR/css3-mediaqueries/>
- [4] Walsh, T. A., McMinn, P., & Kapfhammer, G. M. “Automatic Detection of Potential Layout Faults Following Changes to Responsive Web Pages”. In Automated Software Engineering (ASE), 2015 30th IEEE/ACM International Conference on (pp. 709-714). IEEE.
- [5] Wang A, Chang A, Mark A, Louie K, “Carnegie Mellon University” <http://materializecss.com/>

- [6] Otto M., Thornton J. “Bootstrap”. 2010.
<http://getbootstrap.com/>
- [7] Zurb. “Foundation - The most advanced responsive front-end framework in the world”. 2011. <http://foundation.zurb.com/>
- [8] Gamache D. “Skeleton”. 2014.
<http://getskeleton.com/>
- [9] Guerra, Y. M., González, R. C., Febles, J. P. “Diseño web adaptativo para la plataforma educativa ZERA”. Revista Cubana de Ciencias Informáticas, 10(2), pp. 100-115. 2016.
- [10] Vera P, Rodríguez R, Martínez M. R, “Guía metodológica para el desarrollo de sitios web móviles adaptativos con el enfoque Mobile First”. Congreso Nacional de Ingeniería Informática y Sistemas de Información. 2016

Evaluación de Transformaciones de Modelos

Corina Abdelahad, Enrique Miranda, Daniel Riesco

Departamento de Informática

Universidad Nacional de San Luis

Ejército de los Andes 950 – San Luis – Capital – Argentina

C.P.: 5700

Tel.: +54-0266- 4520300 – Int. 2102

[cabdelah, eamiranda, driesco]@unsl.edu.ar

Resumen

La Ingeniería Dirigida por Modelos está en auge desde hace algunos años y junto con ella la especificación de transformaciones se ha vuelto una de las tareas más importantes llevando a la creación de muchos lenguajes de transformación de modelos. ATL y QVT son algunos de los lenguajes de transformación más utilizados en la actualidad. Teniendo en cuenta el incremento en la utilización de estos tipos de lenguajes, es importante comenzar a considerar aspectos relacionados con la compresión, reutilización y mantenimiento de las transformaciones especificadas en estos lenguajes. Las estrategias para evaluar y medir transformaciones ayudan al Ingeniero de Software en las tareas de compresión, reutilización y posterior mantenimiento en las mismas.

El presente trabajo propone una línea de investigación que tiene como principal lineamiento el estudio, análisis y definición de frameworks de evaluación en el contexto de transformación de modelos.

Palabras claves: Métricas, Lenguajes de Transformación de Modelos, Framework de Evaluación, Mantenimiento.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de software de Calidad – Facultad de Ciencias Físico,

Matemáticas y Naturales, Universidad Nacional de San Luis. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

Introducción

La Ingeniería Dirigida por Modelos (Model-Driven Engineering-MDE) [1, 2] ofrece un escenario ideal para fortalecer el papel de la trazabilidad en el desarrollo de software. MDE propone un proceso de desarrollo de software en el cual la clave son los modelos y las transformaciones entre ellos. En este proceso, el software se desarrolla construyendo uno o más modelos, y a través de un refinamiento, mediante transformaciones, se generan otros modelos o código ejecutable. La transformación de modelos se refiere al proceso de transformación (relaciones y mapeo) de elementos de un modelo a elementos correspondientes de otro modelo. Los modelos están definidos por sus metamodelos, los cuales son los encargados de brindar la sintaxis para la construcción de los mismos. Es decir, un metamodelo define el lenguaje con el cual se construyen modelos. Una transformación de modelos posee como entrada un modelo acorde a un determinado metamodelo y produce como salida otro modelo acorde también a un determinado metamodelo. Más precisamente, y de acuerdo con la definición de Kleppe et al [3]:

“Una transformación es la generación automática de un modelo destino a partir de un modelo origen de acuerdo a una definición

de transformación".

Especificar transformaciones se ha vuelto una de las tareas más importantes en MDE. Esto se ve reflejado en la creación de muchos lenguajes de transformación de modelos como por ejemplo ATL y QVT[5, 7]. Debido al importante crecimiento en la utilización de dichos lenguajes, es imprescindible comenzar a contemplar tareas de mantenimiento y evolución en dichas transformaciones. En este contexto, resulta sumamente útil comenzar a contemplar y medir el esfuerzo que conlleva la comprensión, reutilización y mantenimiento [4] en las transformaciones. La definición de métricas en conjunto con su evaluación brindan información valiosa para el ingeniero de software a la hora de realizar estas actividades.

Líneas de Investigación, Desarrollo e Innovación

En general, al momento de medir y evaluar cualquier tipo de artefacto (ya sea, un programa, una técnica, un proyecto, etc.) de forma sistemática, es necesario definir una estrategia que contemple un conjunto de principios, métodos, técnicas y herramientas que permitan especificar, definir y recolectar métricas e indicadores y sus respectivos valores [9]. Además, con el objetivo de llevar a cabo el análisis y el proceso de toma de decisión, es necesario asegurar que las medidas y los valores de los indicadores sean repetibles y comparables. Por lo tanto, es mandatorio guardar no solo la información de la medición, sino también ciertos metadatos como procedimiento de medición, escala, tipo de escala, modelo de indicador elemental, niveles de aceptabilidad, entre otros.

Una métrica es la especificación de un proceso de medición que transforma un atributo de una entidad en una medida. Un indicador elemental es la especificación de un

proceso de evaluación, el cual recibe como entrada los valores correspondiente a una métrica en particular y produce un valor indicador (es decir, información contextual). Sin embargo en los distintos trabajos relacionados con la temática, en general, se le ha dado poca relevancia a la especificación de este tipo de conceptos. Así mismo, se puede observar una falta de consenso entre los diferentes estándares y manuscritos relativos a la medición y evaluación.

Esta línea de investigación está orientada al estudio, profundización y desarrollo de métricas y criterios los cuales pueden ayudar al ingeniero de software en la medición y evaluación de la comprensión, mantenimiento y reutilización de las transformaciones.

Resultados y Objetivos

A continuación se mencionan los resultados obtenidos hasta el momento dentro del marco de la línea de investigación planteada.

En primera instancia, se especificó un conjunto de métricas en el cual se ha utilizado un enfoque similar al propuesto por Olsina et. Al [6, 9]. En dichos trabajos, los autores proponen modelos y frameworks de calidad en conjunto con estrategias de evaluación que apuntan a un enfoque integrado para medir y evaluar diferentes áreas de las nuevas aplicaciones que se utilizan en la web.

Primeramente se definió un conjunto de atributos (o requerimientos) agrupados en categorías, subcategorías, etc. De esta manera se obtiene una estructura de árbol que pasa a ser el elemento principal del proceso de medición y evaluación. Para esta instancia de la investigación se desarrolló el árbol exhibido en la figura 1. El mismo ha sido desarrollado con el principal objetivo de medir y evaluar transformaciones en lenguaje QVT-R, desde el punto de vista del mantenimiento de software [10].

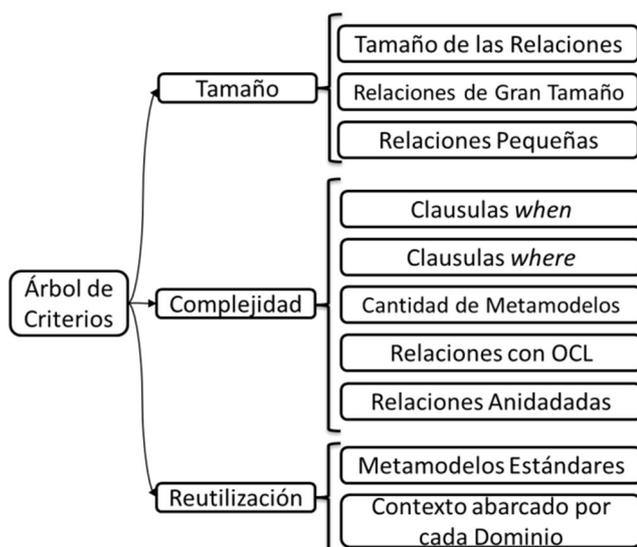


Figura 1: Árbol de Características y Atributos para evaluar transformaciones

El mismo está compuesto de 3 categorías de más alto nivel las cuales, a su vez, contienen atributos atómicos. Dichas categorías son *Tamaño*, *Complejidad* y *Reutilización*. En la categoría *Tamaño* se ha definido 3 atributos que pretenden medir la dificultad que conlleva la comprensión y mantenimiento de la transformación bajo estudio, desde el punto de vista del tamaño que presente la misma. Por otra parte, también se ha incluido la categoría *Complejidad*. La misma está compuesta de atributos que apuntan a medir desde ciertos aspectos sintácticos, la complejidad de una transformación. La última categoría se denomina *Reutilización*. Esta posee dos atributos que pretenden reflejar cuan reutilizable es el código que compone la transformación.

Una vez definido el árbol de atributos, es necesario establecer las métricas que permitirán cuantificar dichos atributos. Para diseñar una métrica, se debe definir el método de medición y el procedimiento de cómputo, en conjunto con la escala en donde se refleja la medición. El tipo de la escala depende de la naturaleza de las relaciones entre los valores de la misma; es decir, algunos aspectos que se pueden ver reflejados por medio de ciertas preguntas, como: los valores en la escala, ¿mantienen el orden y/o las distancias entre categorías? ¿Existe el cero (ausencia de

atributo medido)? entre otras. Entre los tipos de escalas más utilizadas en Ingeniería de Software se encuentran: nominal, ordinal, intervalos, porcentaje y absolutas. Esto es importante, ya que cada tipo de escala, determina el uso de operaciones matemáticas y técnicas estadísticas adecuadas para analizar la información.

Por motivos de extensión de este artículo se especifica un ejemplo de definición de la métrica: *Relaciones con OCL* [8](Categoría: *Complejidad*).

Atributo: Relaciones con OCL.

Definición: La métrica está orientada a medir la proporción de relaciones que usan expresiones OCL, sobre el total. Las relaciones que poseen expresiones OCL dificultan el entendimiento por parte del ingeniero de software, ya que el mismo no sólo debe comprender la semántica de las sentencias en el lenguaje de transformación, sino también la semántica de OCL. Claramente, si una transformación posee muchas relaciones con OCL, dificultará el proceso de comprensión y mantenimiento de la misma.

Métrica Indirecta: Proporción de Relaciones con Expresiones OCL (PROCL).

Objetivo: determinar la cantidad de relaciones de la transformación que poseen expresiones OCL en proporción a la cantidad total de relaciones.

Método de Cálculo:

$$P\ ROCL = (\#RelOCL / \#TotalRel) * 100$$

Escala Numérica: Continua.

Tipo de Valor: Real.

Tipo de Escala: Proporción.

Métricas Relacionadas:

- 1) Número Total de Relaciones con Expresiones OCL (#RelOCL);
- 2) Número Total de Relaciones (#TotalRel).

El conjunto de métricas definidas en la etapa previa será útil si se definen los indicadores adecuados en el proceso de evaluación. Un indicador permite especificar como calcular e interpretar los atributos especificados en el árbol en la Figura 1. A continuación se define el indicador para el

atributo *Relaciones con OCL*.

Atributo: Relaciones con OCL.

Indicador Elemental:

Nombre: Nivel de Desempeño en Relaciones con OCL (D_PROCL).

Modelo Elemental:

Especificación: el mapeo es de la siguiente manera:

$$D_PROCL = \begin{cases} 100 - PROCL & PROCL \leq 100 \\ 0 & PROCL \geq 100 \end{cases}$$

Donde $PROCL$ es la métrica indirecta definida anteriormente.

Criterios de Decisión:

Criterio 1: Insatisfactorio;

Rango: $0 \leq D_PROCL \leq 40$;

Criterio 2: Marginal;

Rango: $40 < D_PROCL \leq 60$;

Criterio 3: Satisfactorio;

Rango: $60 < D_PROCL \leq 100$;

Escala Numérica: Continua.

Tipo de Valor: Real.

Tipo de Escala: Proporción.

Finalmente, mientras el proceso de medición produce datos, el de evaluación retorna información contextual respecto de lo que se está evaluando. El resultado de este análisis se refleja en distintos tipos de reportes, documentos, sugerencias, de acuerdo al objeto de estudio. De esta manera se puede proveer un punto de partida para intervenir los puntos débiles que se detecten en el conjunto de requerimientos de calidad.

Utilizando el enfoque presentado previamente, se realizó un caso de estudio en donde se evalúa una transformación definida para la empresa especializada en optimización multi-objetivo ESTECO [11]. Dicha transformación permite la conversión de muchos flujos de trabajo ingenieriles definidos en el formato propietario de ESTECO al estándar de procesos de negocio BPMN2. Las reglas de transformación, han sido validadas experimentalmente aplicándolas en la empresa mencionada anteriormente.

Para este estudio, se han utilizado 3

intervalos de aceptabilidad en una escala de porcentaje para todos los indicadores elementales. Los valores pueden caer en los intervalos: i) Insatisfactorio, indicado con color rojo; ii) Marginal, en color amarillo; o iii) Satisfactorio determinado por el color verde.

Característica / Atributo	I.E %
Tamaño	
Tamaño de las Relaciones	0
Relaciones Grandes	78
Relaciones Pequeñas	86
Complejidad	
Cláusulas when	61
Cláusulas where	81
Cantidad de Metamodelos	14
Relaciones con OCL	100
Relaciones Anidadas	80
Reutilización	
Metamodelos Estándares	0
Contexto Abarcado por cada Dominio	65

Tabla 1. Indicadores Elementales para la transformación analizada. Lo valores están especificados en %.

Como resultado del análisis del caso de estudio se puede concluir los siguientes aspectos: 1) desde el punto de vista del *Tamaño*, las relaciones poseen, en promedio, más líneas de código que las estimadas para una relación grande, sin embargo la transformación posee pocas relaciones “pequeñas” y “extensas”; 2) desde un enfoque más general, la *Complejidad* de la transformación no muestra muchos atributos como *Insatisfactorios*, sin embargo, se puede notar que hace uso de muchos metamodelos, aspecto que desfavorece el proceso de comprensión; 3) con respecto a la *Reutilización*, es posible destacar lo siguiente: i) la transformación posee un metamodelo no estándar y esto perjudica su reutilización en otro contexto por parte de los ingenieros de software; ii) el contexto del dominio, teniendo en cuenta todos los metamodelos, que debe tener el ingeniero de software para comprender la transformación es *Marginal*, es decir, no está dentro de lo óptimo pero tampoco se considera voluminoso.

Dentro de los objetivos planteados a corto plazo se pretende: i) expandir el árbol de atributos abarcando otras características que reflejen aspectos del mantenimiento no contemplados en esta instancia de la investigación; ii) desarrollar estrategias de extracción automática de métricas en conjunto con la integración en un entorno de desarrollo como Eclipse o Netbeans; iii) integrar frameworks de evaluación de transformaciones de modelos con Ingeniería de Requerimientos.

Formación de Recursos Humanos

Las tareas realizadas en esta línea de investigación sirve como base para el desarrollo de tesis de posgrado, ya sea de doctorado o maestrías en Ingeniería de Software y desarrollo de trabajos finales de la carrera Licenciatura en Ciencias de la Computación e Ingeniería en Informática de la Universidad Nacional de San Luis, en el marco del Proyecto de Investigación.

Referencias

- [1] J.-M. Favre, J. Estublier, and M. Blay. Beyond MDA: Model Driven Engineering (L'Ingénierie Dirigée par les Modeles: audéla du MDA) Edition Hezmes-Lavoisier. Technical report, ISBN 2-7462-1213-7. Février, 2006.
- [2] M. Vo lter, T. Stahl, J. Bettin, A. Haase, S. Helsen, and K. Czarnecki. Model-Driven Software Development: Technology. Engineering, Management. Wiley, 5:6, 2006.
- [3] A. G. Kleppe, J. Warmer, and W. Bast. MDA Explained: The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
- [4] M. Petrenko, V. Rajlich, and V. R. Partial Domain Comprehension in Software Evolution and Maintenance. International Conference on Program Comprehension (ICPC08), pages 13–21, 2008.
- [5] QVT: <http://www.omg.org/spec/QVT/1.1>, 2017.
- [6] P. Becher, P. Lew, and L. Olsina. Specifying Process Views for a Measurement, Evaluation, and Improvement Strategy. *Advances in Software Engineering*, 2012:2, 2012.
- [7] Jouault, F., Kurtev, I.: Transforming models with ATL. In: *Satellite Events at the MoDELS 2005 Conference*, pp. 128–138. Springer, Berlin (2006)
- [8] OCL: <http://www.omg.org/spec/OCL/>
- [9] L. Olsina, P. Lew, A. Dieser, and B. Rivera. Updating Quality Models for Evaluating New Generation Web Applications. *Journal of Web Engineering*, 11(3):209–246, 2012.
- [10] C. Abdelahad, E. Miranda, y N. Perez. Medición y Evaluación de Transformaciones a Nivel Metamodelo: un Enfoque Orientado al Mantenimiento. Congreso Nacional de Ingeniería Informática y Sistemas de Información – CoNaIISI'14: pag. 1083-1094.
- [11] ESTECO. <http://www.esteco.com>,

Extendiendo Transformaciones MDA con Metamodelo de Patrones de Diseño

Luis Roqué Fourcade*, Liliana Arakaki*, Daniel Riesco*, Germán Montejano*, Narayan Debnath†
 *Dpto. de Informática - Universidad Nacional de San Luis – E. de Los Andes 950, San Luis, Argentina
 {araroq, liliana.arakaki}@yahoo.com, {driesco, gmonte}@unsl.edu.ar

†Computer Science Department - Winona State University
 Winona, MN 55987, USA
 ndebnath@winona.edu

Resumen

Object Management Group (OMG) ha producido un importante paso hacia el modelado del Proceso de Desarrollo de Software (PDS), el enfoque Model Driven Architecture (MDA). Sin embargo, no ha tenido el impacto esperado en la calidad de las instancias del modelo ni en los productos resultantes. La principal razón radica en el nivel de abstracción alejado de los dominios para los cuales deben generarse dichas instancias. En el caso de Diseño de Software (DS), no incluye abstracciones del proceso como tal y mucho menos de la actividad principal, la toma de decisiones. Presentamos aquí una línea de investigación que propone extender el modelo de transformaciones MDA con abstracciones inherentes a la actividad de síntesis y evaluación de alternativas para la toma de decisiones. Proponemos utilizar Patrones de Diseño de Software como el estándar de representación este tipo de conocimiento e incluir un metamodelo del mismo, para posibilitar su inyección en transformaciones MDA. Estas extensiones permitirán la generación de instancias del PDS, más cercanas a los dominios tecnológicos y de aplicación direccionados por los patrones, incrementando así el perfil ingenieril de la actividad.

Palabras clave: MDA, Arquitectura de Software, Model Driven Development, Model Driven Architecture, Proceso de Desarrollo de Software.

Contexto

El presente trabajo se desarrolla en el marco del proyecto ‘Ingeniería de software: aspectos de alta sensibilidad en el ejercicio de la profesión de ingeniero de software’, de la Fac. de Cs. Físico- Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL), bajo el número 22F222.

1. Introducción

La Ingeniería de Software ha producido avances realmente significativos a lo largo de su proceso evolutivo. Entre estos, la tecnología de componentes, la evolución arquitectural, la reutilización de

especificaciones de diseño y el modelado del PDS, donde destaca el enfoque MDA.

Los resultados benéficos se verifican a través de diferentes modelos aplicables al desarrollo de software. Sin embargo, existe aún un vacío importante en cuanto a la asimilación de estos resultados en modelos del PDS que impacten positivamente en el perfil ingenieril de la actividad. En el área de DS, la Ingeniería de Software provee un soporte amplio para la especificación de todos los artefactos resultantes de actividades del proceso, pero no ocurre lo mismo con el proceso propiamente dicho y mucho menos con su actividad característica y repetitiva: síntesis y evaluación de alternativas para la toma de decisiones. Entre las principales consecuencias que se derivan de este bajo perfil ingenieril, se destaca el pobre diseño arquitectural que resulta en arquitecturas complejas, con alto costo de mantenimiento y que se degradan significativamente durante su evolución.

El enfoque MDA proporciona un modelo del PDS que actúa como una meta-arquitectura que integra avances en las diferentes dimensiones mediante la contribución de metamodelos en el marco del estándar. El modelo incluye:

- ✓ Una arquitectura de metamodelado basada en MetaObject Facility (MOF) [1], un metalenguaje basado en una simplificación de las capacidades de modelado de clases de UML2 más un núcleo de capacidades para gestión de metamodelos.
- ✓ Una arquitectura del PDS que ordena el proceso en niveles de abstracción distribuidos entre modelos de negocio y de implementación
- ✓ Un modelo de transformación a nivel de metamodelos que permite evolucionar en el PDS a través de los niveles de abstracción

Este modelo permite contar con procesos de desarrollo con capacidad de soportar la actividad de desarrollo a partir del conocimiento contenido en sus abstracciones. Sin embargo, el nivel de abstracción

propuesto se encuentra aún demasiado alejado de los dominios tecnológicos y de aplicación para los cuales se pretende desarrollar las aplicaciones. Así, las instancias de PDS construidas, no incluyen instancias de abstracciones, incluidas en la meta-arquitectura, que garanticen un soporte efectivo a la actividad de desarrollo. Claramente, si las incluyese, permitiría construir instancias con una capacidad de soporte que elevaría significativamente el perfil ingenieril de la actividad [2]. En este sentido, en el año 2016 en [3], presentamos una línea de investigación que propuso definir la extensión de la meta-arquitectura aportada por el enfoque MDA. Su desarrollo durante 2016 derivó en una publicación con una propuesta de extensión de la meta-arquitectura [2] que incluyó:

- ✓ Una definición de DS que sintetiza el universo disponible e incorpora el enfoque de patrones adicionando estructura tanto a la definición del proceso como del artefacto resultante, la Arquitectura de Software.
- ✓ Un modelo extendido de transformaciones MDA que, propone la inyección del conocimiento representado por patrones en las ejecuciones de transformaciones.

1.1. Trabajos relacionados

Muchos trabajos de investigación enfocan desde el diseño arquitectural el problema e intentan aportar soluciones que mejoren el proceso por el cual se obtienen y mantienen las arquitecturas de software y, la mayoría lo hace sin considerar el PDS.

Dada la naturaleza del proceso de diseño, el razonamiento o la base lógica del curso de acciones asociado a la actividad de toma de decisiones se ha convertido en el hilo conductor de buena parte de estos trabajos. La mayoría de sus autores coinciden en diagnosticar el bajo perfil ingenieril que caracteriza actualmente al proceso de DS. Por ejemplo, en [4] Jansen y Bosch introducen el concepto ‘evaporación de conocimiento’ para referirse al hecho por el cual prácticamente todo el conocimiento e información acerca de las decisiones de diseño en las que está basada la arquitectura, está implícitamente embebido en la arquitectura misma, y carece de una representación de primera clase. Ellos afirman que, una vez tomada la decisión de diseño, no queda huella en la Arquitectura de Software que vincule al efecto resultante en la misma con la decisión de diseño que ocasionó dicho efecto.

Buena parte de estos trabajos se enfocan en la identificación y representación de este tipo de conocimiento y otros se enfocan además en su utilización para soportar al proceso de DS.

En la primera categoría se destacan los aportes de Jansen y Bosch quienes en [4], además de introducir el concepto de “evaporación de conocimiento”, definen Arquitectura de Software como una composición de decisiones arquitecturales de diseño, a las que definen como la descripción de adiciones, subtracciones y modificaciones arquitecturales. En este trabajo, coincidimos en el diagnóstico y en la definición de Arquitectura de Software. Inclusive en la vinculación de las decisiones en términos de acciones sobre la arquitectura. Sin embargo, lo hacemos utilizando abstracciones en términos de un mecanismo estándar de representación de este conocimiento y en el marco del enfoque MDA.

Otro trabajo destacado en esta categoría es el que presentan Ton That et al. en [5]. Allí los autores abordan el problema desde la misma perspectiva que Jansen y Bosch [4], pero se enfocan además en mantener el vínculo entre decisiones arquitecturales abstraídas por patrones y los efectos sobre las arquitecturas de software resultantes. En este caso, coincidimos con la utilización de patrones como forma estándar de representación de este tipo de conocimiento. También coincidimos con la necesidad de persistir el vínculo, sin embargo, en nuestro trabajo vamos más allá con esta persistencia. Proponemos tipificar el vínculo como una asociación clase-instancia y la representación de primera clase tanto del patrón como de la Arquitectura de Software. Esto hace posible garantizar la consistencia de las ocurrencias de la asociación y su tratamiento para operar sobre las mismas.

En la segunda categoría, se destaca el trabajo realizado por Zimmermann et al. quienes en [6] toman las definiciones de Jansen y Bosch [4] y proponen su representación mediante el uso de plantillas. Además, proponen utilizar este conocimiento para soportar al proceso de DS mediante un metamodelo común a la representación y el reforzamiento de decisiones.

A continuación, repasamos brevemente las propuestas de definición de DS y de extensión del modelo de transformaciones MDA presentadas en [2] y, a continuación, presentamos la línea de investigación que proponemos en el presente trabajo.

2. Definición de Diseño de Software

En esta sección presentamos la definición de DS resultante de trabajos desarrollados y presentados en [2]. Se apoya en las propuestas de Wand y Ralph [7] y de Frank Buschmann et al. [8], representativas del universo disponible y consistentes con la definición del término Diseño incluida en el glosario publicado en 2010 por el comité de estándares de IEEE [9]. Además, extiende estas definiciones con el enfoque de Patrones de Diseño de Software, a fin de considerar conocimiento potencialmente reutilizable por los mecanismos de toma de decisiones del proceso. Esta extensión, abstrae en la definición la esencia de la actividad de síntesis y evaluación de alternativas para la toma de decisiones (anticipa decisiones y acota el universo de evaluación de otras) en un esquema genérico que permite su concretización para resolver un problema de diseño.

Como se observa en la Fig. 1. *Modelo conceptual de diseño de software extendido*, el modelo original [11], especializado con conocimiento del dominio de DS, aportado por Buschmann et al. en [8], ha sido extendido adicionando la perspectiva de Patrones de Diseño de Software. El modelo idealmente supone que cada parte de la Arquitectura de Software de una aplicación, puede ser obtenida como una instancia de un patrón. Si bien no es ésta la realidad del estado actual, ni hay certeza de que lo vaya a ser en un estado futuro, sí hay importante evidencia de su efectividad en un buen número de casos. El modelo también distingue el fuerte grado de dependencia (ser instancia de) de la arquitectura respecto de Patrones de Diseño de Software en comparación con el resto de las dependencias.

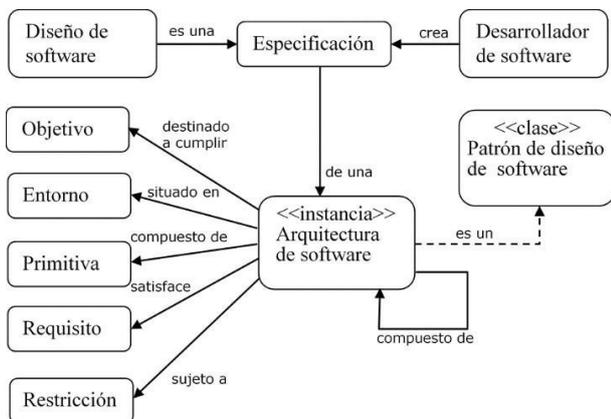


Fig. 1. Modelo conceptual de diseño de software extendido

Otro aspecto a considerar se refiere a la asociación que representa la capacidad de composición entre instancias de la clase Arquitectura de Software. Un aspecto notable es la coincidencia de esta capacidad con la capacidad equivalente entre patrones.

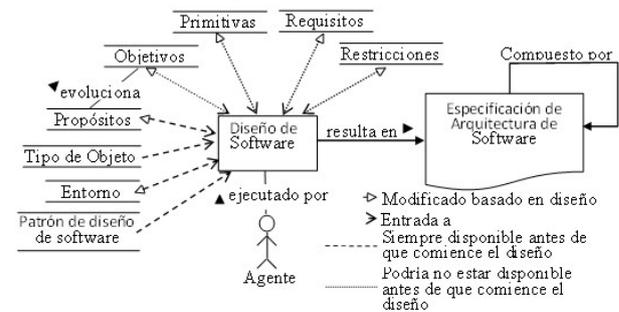


Fig. 2. Modelo conceptual del nivel contextual de diseño de software extendido con patrones

La Fig. 2. *Modelo conceptual del nivel contextual de diseño de software extendido con patrones*, muestra el nivel contextual del proceso de DS extendido con la perspectiva de Patrones de Diseño de Software. Esta extensión se refleja en la adición de una nueva entrada, Patrones de Diseño de Software, y en la estructuración del resultado mediante la asociación ‘compuesto por’ que permite la composición entre Arquitecturas de Software inducida por el mismo tipo de composición en Patrones de Diseño de Software. La influencia del patrón de diseño en la lógica del proceso de diseño está dada por la regla del patrón, que abstrae y simplifica la actividad de síntesis y evaluación de alternativas de la toma de decisiones.

3. El Modelo extendido

El enfoque MDA completa el modelo del PDS integrando artefactos y actividades de desarrollo mediante transformaciones. Éstas son definidas entre metamodelos para ser aplicadas a la transformación de uno o más modelos de entrada en otro de salida sin importar los niveles de abstracción definidos para la arquitectura del PDS. Una transformación puede ser utilizada para producir una representación a partir de otra, o para cruzar niveles capas arquitecturales.

La Fig. 3. *Ejemplo de Transformación Parametrizada* presenta un ejemplo de un patrón de transformación, tomado de la guía “MDA Guide Version 2.0” [10].

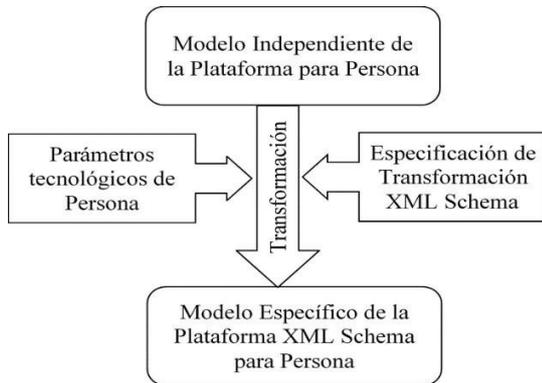


Fig. 3. Ejemplo de Transformación Parametrizada

3.1. Modelo extendido de transformación

La Fig. 3. *Ejemplo de Transformación Parametrizada* ilustra el concepto de transformación para un hipotético caso de transformación. El modelo extendido propuesto en [2] presenta un modelo de transformación cuyo origen y destino se encuentran en el mismo, o diferente, nivel del modelo arquitectural del PDS. Permite obtener otra representación del mismo modelo, con transformaciones dirigidas con reglas contenidas en la especificación de un Patrón de Diseño de Software.

La siguiente Fig. 4. *Modelo extendido de transformación* presenta el modelo presentado en [2]. Este modelo extiende el modelo presentado en la Fig. 3. *Ejemplo de Transformación Parametrizada*, con el conocimiento de la *Definición de Diseño de Software* discutida. Este conocimiento, en una instancia de la transformación, es inyectado como una representación de una instancia del metamodelo de un Patrón de Diseño de Software e interpretado por la transformación que lo utiliza como guía para construir la salida.

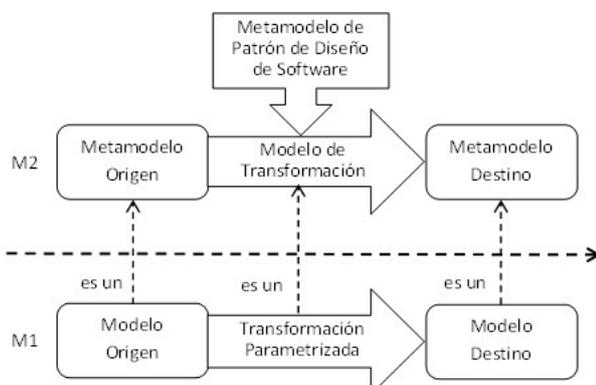


Fig. 4. Modelo extendido de transformación

El modelo asimila en la lógica de transformación la capacidad de incorporar decisiones anticipadas por el patrón y orientar la toma de otras decisiones en base a las restricciones, también aportadas por el patrón, sobre el universo potencial de evaluación de éstas.

4. Línea de Investigación y Desarrollo

Como describimos en la *Introducción*, el enfoque MDA, contribuido por OMG, proporciona un modelo del PDS que actúa como una meta-arquitectura, la cual permite integrar nuevo conocimiento mediante la contribución de metamodelos en el marco del estándar. En esa sección, también mencionamos que la dificultad para que este importante avance se materialice en la calidad de instancias del PDS generadas a partir del modelo y en los productos de software resultantes de la ejecución de las mismas, radica en el nivel de abstracción del modelo que no incluye abstracciones lo suficientemente cercanas a los dominios tecnológicos y de aplicación en que se desempeñan estas instancias. En consecuencia, y derivado de resultados producidos en trabajos previos de investigación [3] [2], presentamos aquí una línea de investigación que propone el desarrollo de un metamodelo de Patrones de Diseño de Software y de su asimilación en el modelo de transformación MDA extendido, propuesto en [2]. Para este objetivo, la línea de investigación propone:

- Definición y especificación del metamodelo para Patrones de Diseño de Software en términos del estándar MDA. Este trabajo involucra la especificación, como especializaciones de las construcciones aportadas por el enfoque MDA, de meta-definiciones para soportar la especificación de roles, comportamientos, asociaciones y reglas que definen a los Patrones de Diseño de Software.
- La extensión del modelo de transformaciones MDA, definido originalmente de un metamodelo origen en otro destino, como una transformación de un metamodelo origen más un metamodelo de Patrones de Diseño de Software en otro destino.

5. Resultados y Objetivos

Esta línea de investigación, se enmarca en el área de trabajo iniciada con las primeras líneas de investigación y trabajos desarrollados a partir del año 2012 [11] [12]. Desde entonces y hasta la actualidad, se han desarrollado diferentes trabajos que han

derivado en publicaciones de resultados [12] [2] y en trabajos de tesis en carreras de la UNSL.

Entre los trabajos en curso y planificados para el corriente año podemos destacar:

- 1) Durante el año 2016 se realizaron diferentes experiencias con:
 - a. Especificaciones de metamodelos
 - b. Especificaciones de Esquemas XML para representaciones XMI de instancias de metamodelos
 - c. Desarrollo de módulos reusables experimentales para transformaciones hipotéticas con capacidad para interpretar y procesar conocimiento contenido en representaciones XMI de instancias de modelos de patrones.
- 2) Actualmente se ha finalizado una tesis de la Maestría en Ingeniería de Software de la UNSL, para la cual se ha solicitado mesa, que presenta una propuesta integral de una meta-arquitectura que sintetiza resultados hasta aquí obtenidos y publicados en otros trabajos.
- 3) Para el año 2017 está planificado el inicio de trabajos de tesis en las carreras de Licenciatura en Ciencias de la Computación y de Ingeniería de Sistemas que se ocuparán del desarrollo de mecanismos de instanciación Arquitecturas de Software a partir de patrones basados en técnicas Orientadas a Objetos en un caso y en técnicas de XML Schemas en otro.
- 4) En el corriente año, están planeados los trabajos, y la publicación de sus resultados, propuestos aquí, que incorporen resultados de las experiencias descriptas en 1).

6. Formación de Recursos Humanos

En el área de la temática planteada, se desarrollan diferentes actividades que van desde la inclusión de trabajos prácticos, laboratorios y ejercicios de investigación, en materias de las carreras afines hasta el desarrollo de tesis y trabajos finales, algunos finalizados y otros en curso, en las mismas. También, en trabajos de extensión, consistentes en desarrollo de aplicaciones para el medio, se desarrollan experiencias manuales asistiendo al desarrollo con refinamiento de modelos con transformaciones dirigidas por patrones, cuyos resultados son retroalimentados a los trabajos de investigación.

7. Bibliografía

- [1] Object Management Group, OMG's MetaObject Facility, <http://www.omg.org/mof/>; 2016.
- [2] L. Roqué Fourcade et al., Software Design: Towards a Meta-architecture to Support Decision Making in the Definition of Software Architecture, *SEDE 2016*, Denver, USA, 2016.
- [3] L. Roqué Fourcade et al., Extendiendo la meta-arquitectura aportada por el enfoque MDA con conocimiento del dominio, *WICC 2016*, Entre Ríos, Argentina, 2016.
- [4] A. Jansen y J. Bosch, Software Architecture as a Set of Architectural Design Decisions, *WICSA 2005*, Pittsburgh, PA, USA, 2005.
- [5] Ton That et al., Preserving architectural decisions through architectural patterns, *Automated Software Engineering*, Singapore, 2016.
- [6] O. Zimmermann et al., Reusable Architectural Decision Models for Enterprise Application Development, *Soft. Architectures, Components and Applications*, Medford, Springer-Verlag Berlin Heidelberg, 2007, pp. 15-32.
- [7] P. Ralph y Y. Wand, A proposal for a formal definition of the design concept, *LNBIP 2009*.
- [8] F. Buschmann et al., Pattern-Oriented Software Architecture - A System of Patterns, Chichester, England, John Wiley & Sons, 2001.
- [9] IEEE Standards Association (IEEE-SA) Standards Board, ISO/IEC/IEEE 24765, Systems and software engineering — Vocabulary, IEEE Computer Society Press, 2010.
- [10] Object Management Group, MDA Guide revision 2.0, <http://www.omg.org/cgi-bin/doc?ormsc/14-06-01>. 2014.
- [11] L. Roqué Fourcade y L. Arakaki, Derivando el diseño a partir de especificaciones de requisitos basadas en Casos de Uso, *WICC 2012*, Posadas, Misiones, Argentina, 2012.
- [12] L. Roqué Fourcade y L. Arakaki, Derivando el Diseño a Partir de Especificaciones de Requisitos Basadas en Casos de Uso, *ASSE 2012*, La Plata, Argentina, 2012.

Fortalecimiento de la Calidad en Procesos de Software y Procesos de Gestión para la Mejora de las Sociedades del Conocimiento

Esponda Silvia ⁽¹⁾, Pasini Ariel ⁽¹⁾, Boracchia Marcos ⁽¹⁾,
Calabrese Julieta ⁽¹⁾, Muñoz Rocío ⁽¹⁾, Santiago Preisegger ⁽¹⁾, Estevez Elsa ⁽²⁾, Fillostrani Pablo ⁽²⁾,
Pesado Patricia ⁽¹⁾

⁽¹⁾Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 120 La Plata Buenos Aires
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

⁽²⁾Laboratorio de I&D en Ingeniería de Software y Sistemas de Información
Universidad Nacional del Sur (UNS) – Bahía Blanca
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

(sesponda, apasini, marcosb,jcalabrese, rmunoz,jspreisegger, ppesado) @lidi.info.unlp.edu.ar,
ecestevez@googlemail.com, prf@cs.uns.edu.ar

Resumen

El III-LIDI (Instituto de Investigación en Informática LIDI), en el marco de diferentes proyectos acreditados, posee un grupo dedicado a la investigación y desarrollo relacionado con la mejora de los procesos de gestión y el aseguramiento de la calidad en procesos de gestión y productos de software.

Los proyectos abarcan:

- **Mejora de los servicios de gobierno digital:** en unidades académicas de nivel universitario y en organismos públicos de gobierno
- **Mejora de procesos:** en el ámbito universitario a través de la certificación con IRAM-ISO 9001 y de los procesos de desarrollo del software mediante la combinación de herramientas de gestión de proyectos.
- **Calidad de producto:** Estudio de estándares internacionales y desarrollo de herramientas para asistir al proceso de evaluación.
- **Asesoramientos y consultorías.**

Palabras Claves

Ingeniería de Software – Calidad del proceso – Calidad del producto - Normas de Calidad – ISO

Contexto

El Instituto de Investigación en Informática LIDI mantiene desde el año 2004 una línea de investigación y desarrollo en calidad.

Esta línea de investigación se enmarca en el subproyecto– “*Mejora de Procesos en el desarrollo de Sistemas de Software y en Procesos de Gestión. Experiencias en PyMEs*” del proyecto “*Tecnología para Sistemas de Software Distribuidos. Calidad en Sistemas y Procesos. Escenarios educativos mediados por TICs*” (2014-2017), acreditado por el Ministerio de Educación de la Nación. Asimismo el III-LIDI participa en el proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de Buenos Aires (PIT-AP-BA) “*Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios*

bonaerenses” y de proyectos de la Facultad de Informática y de otros organismos.

El Instituto posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior y con empresas privadas del sector, interesadas en mejorar sus procesos de desarrollo aplicando mejoras.

Además, participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Introducción

El uso de diferentes tecnologías de información y comunicación (TICs) ha modificado el desarrollo de un gran número de actividades cotidianas, principalmente en las relaciones sociales, culturales y económicas. El gran volumen de información registrada constantemente por el uso de herramientas TICs, dan origen a la “sociedad de la información” donde todos pueden crear, acceder, utilizar y compartir información y el conocimiento, en donde se desarrolle y se mejore la calidad de vida”. Un paso más evolucionado de la información, es el análisis de la misma para obtener conocimiento, creando el concepto de sociedades del conocimiento como una conceptualización de las innovaciones de las TICs para transferir la información para el desarrollo de actividades de las sociedades modernas.

La calidad de los procesos, que acompañan el crecimiento de las TICs es fundamental para obtener información de calidad, que es la entrada a las denominadas sociedades de conocimiento.

Dentro de la línea de la investigación del proyecto, se destacan los siguientes ejes principales:

1- Mejora de los servicios de gobierno digital en unidades académicas de nivel universitario

En el marco de la mejora de la calidad de los servicios institucionales, se inició una evaluación de la calidad de los servicios que brindan las unidades académicas (UA) con autonomía para definir sus procesos académicos y de gestión. [3]

El gobierno universitario está compuesto por docentes, no docentes, alumnos y graduados. Todos en su conjunto representan a la comunidad universitaria que desarrollan sus actividades en el marco de las reglamentaciones que dispone dicho gobierno. Para llevar a cabo el cumplimiento de las reglamentaciones, la universidad pone a disposición de su comunidad un conjunto de servicios. Actualmente, varios de estos servicios son brindados a través del uso TICs.

Se define el concepto de Gobierno Electrónico Universitario (EGOV-U), cómo “el uso de las TICs como herramienta para mejorar los procesos y los servicios prestados por una universidad a los miembros de su comunidad”. [4]

Un Servicio Universitario se define como “el resultado de un proceso llevado a cabo por una universidad o, bajo un cierto control y regulación de ésta, por otra organización, destinado a satisfacer necesidades de los miembros de la comunidad universitaria”. Por ejemplo, una universidad presta servicios universitarios a los alumnos, al brindar información sobre el calendario académico; a los graduados, proponiéndoles cursos de posgrado; a los docentes, proveyéndoles concursos para sus cargos; y a los no docentes, registrando su asistencia. [4]

La prestación de los servicios se realiza a través de diferentes canales. Los canales

de entrega se clasifican en: tradicionales – por ejemplo, mostrador, teléfono, o fax; y electrónicos – por ejemplo, un sitio web, mail, dispositivos móviles, y redes sociales, entre otros. Dada la alternativa de canales, es importante destacar que no todos los servicios pueden ser entregados a través de todos los canales y que la selección del canal influye en el costo de entrega del servicio, el tiempo de respuesta para la entrega del servicio, así como en la satisfacción del receptor del servicio.

A fin de evaluar factores comunes, se seleccionó un conjunto básico de servicios que brindan las UAs, según sus usos, costumbres y reglamentaciones. Los servicios seleccionados y agrupados por claustro receptor de los mismos resultaron: 9 servicios destinados a alumnos, 9 a docentes, 5 a no docentes y 4 a graduados.

El nivel de madurez de los servicios públicos puede ser evaluado en base al grado de automatización y al soporte que dan las TICs al proceso que entrega el servicio. Naciones Unidas propone un modelo de madurez de cuatro niveles: 1) Emergente, 2) Mejorado, 3) Transaccional y 4) Integrado.

Adaptando el modelo propuesto por Naciones Unidas a los gobiernos universitarios, se define 1) **Emergente**: Los sitios de la universidad proveen información de políticas universitarias, gobernanza, reglamentaciones y los tipos de servicios universitarios que se proveen. 2) **Mejorado**: Los sitios de la universidad entregan servicios mejorados de una vía o de simples comunicaciones de dos vías en línea entre la universidad y los miembros de su comunidad. 3) **Transaccional**: Los sitios de la universidad involucran a los miembros de la comunidad universitaria en una comunicación de dos vías, se requiere alguna forma de autenticación electrónica de la identidad de la persona y

se procesan transacciones complejas. 4) **Integrado**: Un único punto de acceso al servicio. El servicio es complejo, intervienen varios actores de diferentes dependencias y pasa por diferentes estados, se realiza todo de forma on-line sin asistir a la UA en ninguna instancia.

El proceso de evaluación se compone de dos partes: 1) recopilación de la información de los servicios a evaluar 2) la evaluación por parte de los usuarios de los servicios.

2- Mejora de los servicios de gobierno digital en organismos públicos de gobierno

El III-LIDI participa en conjunto con la UNS en el proyecto PIT-AP-BA “*Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses*” que propone mejorar la eficacia de la implementación de las políticas sociales municipales mediante soluciones informáticas que permitan mejorar la entrega de servicios públicos de acción social basados en Internet, aplicaciones móviles, almacenamiento en la nube y otras tecnologías informáticas de punta.

Se propone caracterizar, documentar, capacitar y adoptar mejores prácticas para mejorar la calidad y accesibilidad de los servicios ofrecidos, y diseñar soluciones tecnológicas que permitan mejorar la eficiencia y eficacia de las interacciones gobierno-ciudadanos y de los procesos relacionados con la entrega de servicios públicos digitales de acción social.

3- Mejora en los procesos de gestión de la Facultad de Informática

El Área de Certificación de Calidad de la Facultad de Informática coordina con este proyecto, con el objeto de analizar, definir y establecer un plan a ser aplicado a

distintos procesos de la Gestión Universitaria.

Desde el año 2011, la Facultad de Informática ha iniciado el camino hacia la certificación de distintos procesos.

El curso de Nivelación a Distancia para el Pre Ingreso, logró en el año 2012 la certificación IRAM-ISO9001:2008 del Sistema de Gestión de Calidad (SGC) del "Diseño y realización del curso de Nivelación a Distancia para el Pre Ingreso a la Facultad de Informática" y en 2015 la Re-certificación

En el 2016 se ha obtenido la certificación IRAM-ISO9001:2008 del SGC del área de Concursos Docentes.

Durante el año 2017 se realizarán las auditorías de seguimiento de los procesos certificados.

4- Mejora de Procesos de gestión en el desarrollo de software

Existen diferentes herramientas de gestión de proyecto que facilitan el proceso de mejora en el desarrollo del software. Por ejemplo Trello, Jira, Redmine, etc. Por otro lado existen repositorios de documentos que proveen herramientas para mantener una Gestión de la Configuración ordenada. Por ejemplo GitHub. Cada una de estas herramientas provee un conjunto de prestaciones diferentes. Además las herramientas brindan un conjunto de servicios que permiten combinar las mismas para un mejor aprovechamiento de sus funcionalidades. [5]–[8]

Con el objetivo de asistir al equipo de desarrollo en el seguimiento del proyecto y almacenamiento de la documentación se analiza la posibilidad de brindar una herramienta que combine las funcionalidades que el equipo considere necesarias y las mismas se actualicen de forma automática en función de las modificaciones de los archivos del repositorio. Por ejemplo combinando las

tareas de un tablero de Trello con los archivos en GitHub

5- Calidad en productos.

La calidad de los productos de software se ha convertido en uno de los principales objetivos estratégicos de las organizaciones con la idea de lograr la competitividad que se espera en el mercado.

En este sentido, el Instituto ha estudiado la familia de las norma ISO/IEC 25000, que posibilita la certificación de los productos de software. En particular ISO/IEC 25040

La norma ISO 25040 define el proceso para llevar a cabo la evaluación de un producto de software. Dicha evaluación consta de una serie de pasos o tareas a seguir en las cuales se analizan diferentes aspectos y puntos de vista del producto.[1], [2]

A la hora de realizar una evaluación, el principal objetivo es obtener un informe detallado en el cual se disponga que posea la información necesaria para determinar el cumplimiento del objetivo de la evaluación.

Para ello se desarrolla una herramienta denominada SEP (Sistema de Evaluación de Producto). SEP propone una forma de poder obtener el informe mencionado anteriormente. La herramienta guía al usuario a lo largo de los diferentes ítems que la norma posee, controlando las evaluaciones pendientes, realizadas y por realizar facilitando la generación del informe final de la evaluación.

Líneas de investigación y desarrollo

- Análisis y estudio de normas y modelos orientados a la Calidad del Proceso, como por ejemplo IRAM - ISO/IEC 15504, IRAM - ISO/IEC 12207, ISO 9001 e ISO/IEC 90003. [9]–[12]

- Análisis y estudio de normas y modelos orientados a la Calidad del Producto, como por ejemplo IRAM - ISO/IEC 25000
- Evaluación de procesos en organismos públicos y privados según los requisitos de ISO 9001.
- Análisis, discusión y estudio de normas de calidad relacionadas con certificación de servicios gubernamentales.
- Análisis, discusión y estudio de mejoras de proceso en el desarrollo de software combinando herramientas de gestión de proyectos.
- Evaluación de madurez de los servicios de gobierno digital de una unidad académica y de organismos públicos gubernamentales.
- Capacitación y desarrollo de los documentos básicos de gestión de la calidad de productos.
- Se avanza en una herramienta para asistir en la evaluación del producto de software.
- Acciones de consultorías y asesoramiento en organismos públicos y privados
- Realización de cursos de actualización de la Norma IRAM-ISO 9001:2015 en el Instituto Argentino de Normalización y Certificación (IRAM).
- Se avanzó en la tesis de doctorado “Modelo de madurez de los servicios de gobierno electrónico en el ámbito universitario”.

Resultados obtenidos / esperados

- Se avanza en la evaluación de diferentes unidades académicas para obtener una visión general del estado de madurez de los servicios de gobierno digital universitario. En particular se realizó una encuesta a distintas UAs con el objetivo de recabar información acerca de los servicios de gobierno electrónico universitario prestados para determinar el nivel de madurez de los mismos.
- Se ha comenzado el relevamiento de los servicios públicos digitales de acción social en los Municipios de La Plata y Bahía Blanca.
- Mantenimiento de los procesos certificados IRAM-ISO 9001:2008 del Pre-Ingreso a Distancia de la Facultad de Informática y Concursos Docentes de la Facultad de Informática.
- Análisis de nuevos procesos de la Facultad, con posibilidad de ser certificados.
- Se avanza en una herramienta para la asistir a la gestión de proyecto, generando documentación automática en función de los avances del desarrollo, combinando herramientas como Trello y GitHub

Formación de recursos humanos

- Capacitación de los miembros del proyecto a través de diversos cursos del Instituto Argentino de Normalización y Certificación (IRAM).
- Se desarrollan tesis de doctorado y tesinas de grado en el área.
- Los integrantes de esta línea de Investigación participan en el dictado de asignaturas/cursos de grado/postgrado en la Facultad de Informática de la UNLP y en otras universidades del país. En particular, en la UNLP, se dicta la asignatura “Calidad de Sistemas de Software en Pequeñas y Medianas Empresas”.

Referencias

- [1] ISO, “ISO/IEC 25040:2011 □ Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process,” 2011.
- [2] IRAM and ISO, “IRAM-NM-ISO IEC9126-1 Information technology. Software engineering. Product quality. Part 1 - Quality model,” 2009.
- [3] A. Pasini and P. Pesado, “Quality Model for e-Government Processes at the University Level: a Literature Review,” *Proc. 9th Int. Conf. Theory Pract. Electron. Gov.*, pp. 436–439, 2016.

- [4] A. Pasini, E. Estévez, P. Pesado, and M. Boracchia, "Una metodología para evaluar la madurez de servicios universitarios," *Proc. Congr. XXII Congr. Argentino Ciencias la Comput. Congr. Argentino Ciencias la Comput.*, pp. 636–646, 2016.
- [5] T. Inc, "About Trello," <https://trello.com/about>, 2016.
- [6] M. Pinola, "Cómo utilizar Trello para organizar tu vida casi al completo," *Gizmodo*, 2015. [Online]. Available: <http://es.gizmodo.com/como-organizar-toda-tu-vida-utilizando-trello-1684529913>.
- [7] C. W. Tom Preston-Werner PJ Hyett, "GitHub," 2008. 2013.
- [8] Atlassian, "JIRA Software," *Atlassian Foundation*, 2016. [Online]. Available: <https://www.atlassian.com/software/jira>.
- [9] IRAM;ISO, "IRAM-ISO-IEC 15504 Information technology. Process assessment.," 2012.
- [10] IRAM; ISO, "IRAM-NM-ISO IEC 12207 Information technology. Systems and software engineering. Software life cycle processes.," 2008.
- [11] ISO, *IRAM - ISO 9001:2008 "Quality management systems -- Requirements."* 2008.
- [12] IRAM;ISO, "IRAM-ISO-IEC 90003 - Software engineering. Guidelines for the application of ISO 9001:2000 to computer software.," 2006.

Gestión Cuantitativa de Proyectos y Entrega Continúa en Entornos Ágiles

Dapozo, Gladys N.; Greiner, Cristina; Irrazabal, Emanuel; Medina, Yanina; Ferraro, María; Mascheroni, Agustín

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste
{gndapozo; cgreiner, eirrazabal, ferraro, yanina}@exa.unne.edu.ar

Resumen

Este proyecto de investigación aborda temas de gestión cuantitativa de proyectos y entrega continua en entornos ágiles buscando promover y/o generar métodos y herramientas que contribuyan a mejorar la calidad del proceso y del producto software. La automatización de los métodos de estimación permitirá a las empresas mejorar la precisión de la duración y costo de los proyectos de desarrollo bajo diferentes enfoques: tradicional y ágil, generando además información histórica que retroalimenta a los métodos de estimación. La gestión de incidentes, la sistematización de mecanismos de trazabilidad de los requerimientos y la incorporación de procedimientos para la entrega continua, contribuirá a mejorar la mantenibilidad del software, siendo este atributo uno de los más críticos en los ámbitos reales de producción de software. En todos los casos, se promueve el uso de técnicas de desarrollo ágil.

Palabras clave: Estimación software. Incidencias. Entrega Continua. Gestión de Requerimientos. Agilismo.

Contexto

Las diferentes líneas de I/D presentadas en este trabajo corresponden al proyecto PI-F10-2013 “Métodos y herramientas para la calidad del software”, acreditado por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE) para el periodo 2014-2017.

Introducción

Gestión cuantitativa

La gestión cuantitativa en el desarrollo de software proporciona una visión del grado de cumplimiento de metas, así como las causas que explican desviaciones significativas en procesos o productos [1]. La gestión de proyectos en base al conocimiento cuantitativo contribuye a la determinación de los aspectos de mayor relevancia, cuyo rendimiento afecta en forma significativa al logro de los objetivos del proyecto y la satisfacción de los clientes, obteniendo productos con un mayor nivel de calidad.

En este proyecto la gestión cuantitativa se enfoca en la estimación, la gestión de incidentes y en la trazabilidad de requerimientos.

Estimación

El desarrollo de software requiere de la estimación para controlar y administrar los recursos antes y durante el proyecto. Existen diferentes modelos de estimación que van desde las técnicas orientadas a los procesos hasta métodos paramétricos o algorítmicos, basados en datos históricos utilizados para definir y calibrar el modelo [2][3]. La técnica (o conjunto de técnicas) que se utiliza debe adaptarse a los datos disponibles y la naturaleza del problema en la estimación. En este sentido, es importante mantener diferentes técnicas de estimación de acuerdo con la metodología de desarrollo a utilizar.

La realidad de las empresas de la región NEA muestra que las áreas o empresas de software, en gran medida, desconocen los métodos paramétricos y

herramientas para la estimación de esfuerzo, duración y costo de los proyectos software [4], siendo lo usual la estimación por juicio de expertos.

En esta línea se estudian diversos métodos de estimación, principalmente aquellos orientados al desarrollo web. En particular, se consideró un método elaborado por el Centro de Estudios de Ingeniería de Software de la Universidad de la Frontera (CEIS-UFRO) que parte de una especificación de requerimientos basada en casos de uso y la productividad, llega a definir el costo de un producto de software transaccional web [5].

Gestión de incidentes

La medición de código fuente aporta a la gestión cuantitativa, por ejemplo, estableciendo relaciones en cuanto a la complejidad del software y la propensión a fallos, lo cual ofrece indicios sobre qué clases o porciones de código debe enfocarse la prueba de software.

Vinculado con el código fuente aparece el concepto de deuda técnica, la que se contrae cuando las actividades relacionadas con el desarrollo de software no se realizan con los niveles de calidad adecuados y por tanto puede disminuir la mantenibilidad del código u ocasionar mayores costos en el proyecto [6].

A medida que el proyecto crece en cantidad de requerimientos, tamaño del código o personas involucradas, el éxito resulta cada vez más difícil de alcanzar. Es importante que los equipos se mantengan dentro de presupuestos y tiempos planificados. Una forma efectiva para alcanzar las metas es que los equipos cuenten con herramientas que les permitan sistematizar aspectos tales como peticiones de cambios, incidencias y bugs. SWEBOK define como petición de cambio (Change Request, CR) a “una solicitud para ampliar o reducir el alcance del proyecto; modificar las políticas, procesos, planes o procedimientos; modificar los costos o presupuestos; o revisar los horarios. Una fuente de CR es

el inicio de medidas correctivas en respuesta a informes de problemas [7].

La International Software Testing SWEBOK define un bug como “un defecto en el código fuente. Un paso, proceso o definición de datos incorrectos en el programa. La codificación de un error humano en código fuente” [7].

En la literatura especializada surge que la asignación de incidentes es un aspecto crítico. En [8] se describe cómo la tarea de asignar un CR puede ser costosa en tiempo, y la reasignación doblemente costosa, señalando la falta de uso de una herramienta automatizada que asista en esta tarea, y la necesidad de considerar la información histórica para tomar decisiones efectivas, tales como la carga de trabajo actual y conocimiento de habilidades de sus desarrolladores. En [9] se describe la asignación de bug como un proceso social debido a la complejidad que implica determinar la persona más adecuada para resolver el problema, señalando la importancia de tener una base de conocimiento sobre la experticia de los desarrolladores y herramientas que abarquen información socio-técnica de la organización. En [10] se menciona el conocimiento de los desarrolladores y la repercusión en la carga de trabajo del equipo como desafíos claves en la asignación de un bug o incidentes.

Trazabilidad

Los procesos de evolución del software (adaptación a nuevos requerimientos), varían considerablemente dependiendo del tipo de software a mantener, los procesos de desarrollo utilizados en una organización y el personal implicado en el proceso. Por tanto, es imprescindible contar con una herramienta de gestión de requerimientos que acompañe este proceso, para minimizar el impacto de los cambios en el sistema y lograr que los mismos se realicen de manera planificada y controlada. Se reconoce que realizar el seguimiento a los requerimientos a lo largo del proceso de desarrollo de software no es tarea fácil. Todo artefacto

de software cambia en el tiempo por la evolución en las necesidades de los usuarios. Para minimizar el impacto causado por dicha evolución, la práctica de la trazabilidad ha sido estudiada e implementada con diferentes modelos y técnicas que permiten lograr mayor calidad en los productos de software [11]. El desafío es sin duda, acompañar la gestión de cambios con el suficiente nivel de trazabilidad que logre el equilibrio entre calidad de software y tiempo dedicado a la gestión.

Entrega continua

La Entrega Continua de Software, en inglés Continuous Delivery (CD), se define como un enfoque en el cual los equipos mantienen la producción de software en ciclos cortos de tiempo, asegurando que el producto pueda ser lanzado de manera fiable en cualquier momento [12]. La idea es poder lanzar a producción un producto software libre de defectos “*con solo apretar un botón*” [13].

Existe un concepto similar al CD, que es el de Despliegue Continuo, en inglés Continuous Deployment (DC). El DC es una actividad que consiste en lanzar cambios continuamente al ambiente de producción [14]. La principal diferencia se encuentra en la fiabilidad a la hora de lanzar una nueva versión del producto: el DC busca integrar código a producción una, dos, y muchas más veces en el mismo día, en cambio, el CD se centra en hacerlo con la certeza de que el producto que se está lanzando a producción tiene un alto grado de calidad y se encuentre libre de defectos.

En el intento de implementar el CD, muchas organizaciones terminaron alcanzando solamente el DC. Por ejemplo, algunas empresas como Facebook, Atlassian, IBM, Adobe, Microsoft y Flickr, han tenido éxito en la implementación de diferentes enfoques para realizar entregas en periodos cortos de tiempo. Sin embargo, diferentes autores señalan que el proceso de lanzar

rápidamente productos software a producción, aún presentan inconvenientes por resolver [15].

Uno de los principales problemas estaría en la calidad del producto software. Ésta puede disminuir, dado que, al realizarse los despliegues del sistema con mayor frecuencia, aparecen más defectos en el producto [13]. Por tanto, es esencial desarrollar un enfoque de priorización de los diferentes aspectos en la calidad del producto software, teniendo en cuenta la forma de trabajo actual de las empresas de desarrollo software.

Para adoptar este enfoque, en la literatura se utiliza el concepto de Tubería de Despliegue (DP - Deployment Pipeline). El DP es un estándar para automatizar el proceso de CD [16]. A pesar de que cada organización puede variar en la implementación de este estándar, el mismo se conforma de las siguientes actividades: Instalación, Compilación, Pruebas de Aceptación, Pruebas de Capacidad, Pruebas Manuales, Liberación a Producción.

Líneas de investigación y desarrollo

En la línea de Gestión Cuantitativa se propone:

Estimación

- Siguiendo con la línea de resultados obtenidos en [17] se trabajará en el desarrollo de una herramienta para automatizar la estimación de costos en proyectos web aplicando métodos de estimación ágiles. Con esta herramienta y la elaborada anteriormente, se podrán realizar estimaciones con diferentes técnicas orientadas a distintas metodologías de desarrollo, además de generar información histórica para mejorar la precisión.

Trazabilidad

Desarrollar una aplicación de gestión de proyectos en entornos de la Administración Pública para asegurar la

trazabilidad de los requerimientos a las actividades, tareas y el código fuente.

- *Gestión de incidentes*

Diseñar y construir una herramienta que permita realizar el seguimiento de las incidencias de los proyectos de software, aportando información de las características socio técnicas de los equipos involucrados. La herramienta permitirá gestionar un histórico de las incidencias, desde su reporte inicial hasta su corrección, siendo configurable el ciclo de vida de las mismas para que se adapte a las necesidades de la organización. A su vez, tendrá en cuenta la experiencia, carga de trabajo actual e involucramiento de los desarrolladores.

En la línea de *Entrega Continua en entornos ágiles*:

- Mejorar el algoritmo de pruebas para el desarrollo de compatibilidades web implementado en [18] y evitar escenarios de falsos positivos.
- Construir una técnica de Testing Continuo en base al framework para el desarrollo de pruebas REST [19].
- Desarrollar nuevos Estudios de Caso al implantar técnicas ágiles en entornos de la Administración Pública [20][21].

Resultados obtenidos

Los principales resultados de las actividades desarrolladas en estas líneas son:

- Se evaluaron cuatro metodologías específicas para la estimación de proyectos web y se realizó una aplicación que automatiza, simplifica y guía la aplicación de las mismas, que fue validada con datos experimentales. Los resultados indican que, a mayor especificidad en cuanto al tipo de proyecto y la adecuada configuración de parámetros que identifican el entorno de la aplicación, el ajuste de la estimación es mayor. Se espera que esta herramienta contribuya con la calidad del software permitiendo mejorar la

estimación de proyectos web, facilitando los cálculos, realizando un seguimiento de las estimaciones y aportando información histórica como feedback para estimaciones posteriores [22].

- Basado en los resultados obtenidos en [5], se realizó una aplicación web (CostEs) para estimar costos de desarrollo que implementa la técnica de estimación temprana y proporciona valores de tamaño, esfuerzo y costo [23][24]. La herramienta fue validada con los datos de un proyecto web finalizado, el portal del Gobierno de Corrientes, obteniendo resultados satisfactorios. Como punto fuerte, CostEs almacena toda la información referida al proyecto y al proceso de estimación, y de esta forma permite ir mejorando las estimaciones futuras.

En la línea de *Entrega Continua en entornos ágiles*:

- En [18] se describe el desarrollo de una herramienta para pruebas de compatibilidad entre navegadores web. La técnica propuesta es una iniciativa para automatizar el proceso de pruebas de compatibilidad mediante un algoritmo de comparación de imágenes. Para validarla, ha sido implementada en una empresa de desarrollo software de escala nacional que trabaja con desarrollo continuo de software. Los resultados demostraron que la herramienta permite acelerar el proceso de pruebas, mediante la automatización de pruebas de compatibilidad web. Por un lado, los tiempos de ejecución de pruebas disminuyeron un 82%, y esto también redujo el tiempo total del proceso de liberación de las versiones del sitio.
- En [19] se presentó un framework que permite el desarrollo de pruebas automatizadas para servicios REST. El tiempo de ejecución de un lote de pruebas que llevaba 3 horas, realizado manualmente con el equipo de 15 personas, pasó a ser de 15 minutos en

un servidor de integración continua 100% automatizado.

- En [20] y [21] se describe la implementación de la metodología de gestión de proyectos SCRUM en dos oficinas de sistema de administraciones públicas provinciales en Corrientes y Misiones. Como primer paso, se indicaron las características institucionales con el fin de establecer un diagnóstico del equipo de trabajo actual y el plan de implantación de SCRUM, así como la enumeración de las lecciones aprendidas al construir el plan. Como resultado de ello, la implementación tuvo una acogida general positiva por parte de todo el equipo de desarrollo y sus directivos, demostrando interés y predisposición para experimentar una nueva forma de trabajo en desarrollo ágil de software.

Finalmente, en el marco del proyecto como una línea de trabajo emergente, se trabajó durante el año 2016 en un proceso de explotación de información adaptado del modelo propuesto en [25], el cual combinó las etapas de la metodología CRISP-DM con actividades de la metodología COMPETISOFT. En [26], [27] y [28] se publicaron los resultados; abordando desde el punto de vista de la ingeniería de procesos una propuesta tecnológica para la Analítica Académica en la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste. Para ello se explotó la información de los alumnos de la carrera Licenciatura en Sistemas de Información disponibles en el sistema SIU guaraní, de amplio alcance nacional.

Formación de recursos humanos

En el Grupo de Investigación sobre Calidad de Software (GICS) están involucrados 5 docentes investigadores, 1 becario de investigación de pregrado y 1 tesista de doctorado. Tres alumnos de la carrera finalizaron la misma con proyectos vinculado a estos temas.

Referencias

- [1] Gou, L.; Wang, Q.; Yuan, J.; Yang, Y.; Li, M.; Jiang, N. "Quantitative defects management in iterative development with BiDefect". *Software Process Improvement and Practice*, 14(4), 227-241. 2009.
- [2] Andrés, J.; Fernandez-Lanvin, D.; Lorca, P. "Cost estimation in software engineering projects with web components development". 2015, vol.82, n.192
- [3] Mendes, E. "Using knowledge elicitation to improve Web effort estimation: Lessons from six industrial case studies," *Software Engineering (ICSE)*, 2012.
- [4] Dapozo, G. N.; Greiner, C. L.; Irrazábal, E.; Medina, Y; Ferraro, M. A.; Lencina, A. B. "Características del desarrollo de software en la ciudad de Corrientes". *CACIC* 2015.
- [5] Díaz Villegas, J. E.; Robiolo, G. "Método de estimación de costos de un producto de software Web". 43 JAIIO - ASSE 2014 - ISSN: 1850-2792
- [6] Cunningham, W. "The wycash portfolio management system," in *OOPSLA '92: Addendum to the proceedings on Object-oriented programming systems, languages, and applications (Addendum)*, 1992.
- [7] Bourque, P.; Fairley, R. E. *SWEBOK*, Tercera ed., IEEE, 2014, pp. 82-87.
- [8] P. Anselmo da Mota y Y. Cerqueira. "Towards Understanding Software Change Request". *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering*, 2013.
- [9] P. Guo y T. Zimmermann. "Not my bug! and other reasons for software bug report reassignments". *CSCW '11 Proceedings of the ACM 2011 Conference on*, 2011.
- [10] G. Bortis y A. van der Hoek. Teambugs: a collaborative bug. *CHASE '11 Proceedings of the 4th International*, pp. 69-71, 2011.
- [11] M. Tabares, f. Arango, R. Anaya. Una Revisión de Modelos y Semánticas Para la Trazabilidad de Requisitos. *Revista EIA*, ISSN 1794-1237 Número 6, p. 33-42. Diciembre 2006. Escuela de Ingeniería de Antioquia, Medellín. Colombia.
- [12] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in *IEEE Software* 03/2015. V. 32(2).

- [13] J. Humble and D. Farley. "Continuous delivery: reliable software releases through build, test, and deployment automation", 1st ed. Boston, US: Pearson Education, 2010.
- [14] H. H. Olsson, H. Alahyari, and J. Bosch. "Climbing the 'Stairway to Heaven'-A Multiple-Case Study Exploring Barriers in the Transition from Agile Development towards Continuous Deployment of Software" in 38th EUROMICRO Conference, 2012, pp. 392-399.
- [15] B. Fitzgerald and K. Stol, "Continuous Software Engineering and Beyond: Trends and Challenges," in 1st International Workshop on Rapid Continuous Software Engineering, 2014.
- [16] O. Prusak. "Continuous Testing: The Missing Link in the Continuous Delivery Process". Blaze Meter. 2015.
<https://blazemeter.com/blog/continuous-testing-missing-link-continuous-delivery-process>
- [17] G.N. Dapozo, Y. Medina, A. Lencina. "La práctica de la estimación en empresas y áreas de Sistemas". Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação. v.1, n. 4. 2015.
- [18] M. A. Mascheroni, M. K. Cogliolo, E. Irrazabal. "Automatización de pruebas de compatibilidad web en un entorno de desarrollo continuo de software". 45 JAIIO - ASSE 2016 - ISSN: 2451-7593 – Pág. 51-63.
- [19] M. A. Mascheroni, E. Irrazabal. Framework para la creación y ejecución de pruebas automatizadas sobre servicios REST. XXII Congreso Argentino de Ciencias de la Computación. CACIC 2016.
- [20] R. López Lovera, R. Morante, E. Irrazabal. Plan de implementación de la metodología SCRUM y primeros resultados en la Dirección de Desarrollo de Sistemas de la Municipalidad de Posadas. 45 JAIIO - SIE 2016 - ISSN: 2451-7534 Pág. 30-39.
<http://45jaiio.sadio.org.ar/sites/default/files/SIE-05.PDF>
- [21] C. Pinto Luft, E. Irrazabal. Hacia la implantación de metodología SCRUM en la Inspección General de Personas Jurídicas de Corrientes: primeras lecciones aprendidas. Anales del IV Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación. Universidad Nacional del Nordeste. ISBN 978-987-3619-15-1. 2016.
- [22] O. G. Pedrozo Petrazzini, C. Greiner. Herramienta para apoyar la estimación en el desarrollo de aplicaciones web. 45 JAIIO - EST 2016 - ISSN: 2451-7615 – Pág. 116-1130
<http://45jaiio.sadio.org.ar/sites/default/files/EST-1873.pdf>
- [23] Lencina, A.; Dapozo, G. N.; Medina, Y. Estimación de costo en el desarrollo de software. XXII Reunión De Comunicaciones Científicas y Tecnológicas 2016. Campus Corrientes. 4 y 15 de junio de 2016.
- [24] Lencina, A; Medina, Y.; Dapozo, G. N. Aplicación para estimar costos en proyectos de software. 45 JAIIO - ASSE 2016 - ISSN: 2451-7593 – Pág. 181-192.
<http://45jaiio.sadio.org.ar/sites/default/files/asse-12.pdf>
- [25] J. Vanrell. Un Modelo de Procesos para Proyectos de Explotación de Información. Tesis de Maestría en Ingeniería en Sistemas de Información. Buenos Aires. 2012.
- [26] Lopez, M.E.A., Dapozo, G.N., Greiner, C.L. Modelo de proceso para proyectos de explotación de información en el ámbito académico. XXII Reunión De Comunicaciones Científicas Y Tecnológicas 2016. Campus Corrientes. 2016.
- [27] Lopez, M.E.A., Dapozo, G.N., Greiner, C.L. Propuesta tecnológica como apoyo a procesos de Analítica Académica en FaCENA-UNNE. Anales del IV Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación. Universidad Nacional del Nordeste. ISBN 978-987-3619-15-1. 2016.
- [28] Lopez, M.E.A., Dapozo, G.N., Irrazabal, E.A., Greiner, C.L. Proceso de Explotación de Información para Analítica Académica en FaCENA-UNNE. Revista Latinoamericana de Ingeniería de Software. ISSN 2314-2642. Vol 4, No 6 (2016).

Gestión de Proyecto de Software: Un Método Basado en Gamificación para Mejorar la Calidad del Producto y Desempeño de Equipos de Desarrollo

María Soledad Bianciotti, Carlos Salgado, Alberto Sánchez, Mario Peralta
 Departamento de Informática Facultad de Ciencias Físico-Matemáticas y
 Naturales Universidad Nacional de San Luis
 Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
 e-mail: solebianciotti@gmail.com, {csalgado, alfanego, mperalta}@unsl.edu.ar

Resumen

El desarrollo o la implantación de software surgen para dar respuesta a alguna necesidad de la organización. Cuando la satisfacción del requerimiento implica el desarrollo o implementación de un software de mediana o gran complejidad se necesita, habitualmente, de la formación de un equipo de trabajo integrado por varias personas, cada una de las cuales intervendrá en una o varias fases del proyecto. Mantener ese grupo de personas comprometidas y motivadas para trabajar conjuntamente es una de las tareas que el líder de proyecto no debe perder de vista. Así como tampoco debe descuidar otros aspectos de la gestión de proyectos. En este trabajo se propone un método para mejorar la calidad del producto y el desempeño de los equipos de desarrollo de software, basado en técnicas de *gamificación* [2] [3] y en la aplicación y utilización de las mejores características de las normas ISO 9000-3 e ISO 10006 [5] e ISO 9001 [7] y su adecuación al mercado nacional.

El método propuesto ayudará a los distintos actores de un equipo de desarrollo a facilitar sus tareas diarias y armonizar las interrelaciones personales.

Palabras clave: Gamificación, Gestión de proyectos de software, Equipo de desarrollo.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Introducción

La combinación de algunas normas de calidad y técnicas de gamificación pueden ser herramientas útiles a la hora de mejorar la gestión de proyectos.

El concepto gamificación surgió en el año 2002 y tuvo difusión a mediados del año 2010 [2] Desde entonces y debido a su carácter transversal, ha sido aplicada a diversos ámbitos, como por ejemplo formación de recursos humanos, educación, marketing, innovación de productos y de procesos.

Herranz y Palacios en [1] plantean que... "La Gamificación, ..., se configura como un agente de cambio dentro de cualquier disciplina donde su aplicación pueda dirigir el comportamiento a los objetivos deseados. La Ingeniería del Software no

es ajena a esta circunstancia y puede ser objeto de mejora gracias a la Gamificación. Una Gamificación correctamente aplicada sobre un ámbito concreto de la Ingeniería del Software permitirá una mejora de los resultados, estimulará la resolución de problemas por parte de los usuarios y fomentará la colaboración entre todos los actores. Además, potenciará el desarrollo de nuevas ideas siempre desde un alineamiento de objetivos entre negocio y usuarios”...

El desarrollo de software es una actividad en la cual juegan un papel muy importante los recursos humanos. El desempeño de un equipo de desarrollo está ligado, no sólo a las competencias sino también a las motivaciones de cada uno de los integrantes del equipo. En [1] los autores manifiestan que...*”La motivación ha sido señalada como un factor crítico de éxito para los proyectos de software. Contar con aspectos motivadores que permitan mejorar el rendimiento de los equipos de desarrollo de software se presenta como uno de los principales retos de la gestión de la Ingeniería del Software en todo el mundo. La Gamificación permite definir mecanismos que canalizan la motivación de las personas al desarrollo de tareas con mayor productividad y rendimiento, fomentando las innovaciones a la vez que se potencia la colaboración y la participación”...* Acorde a lo expuesto, la gestión de proyectos y la mejora de procesos de software, entre otros, son ámbitos de la Ingeniería de Software en los cuales pueden ser aplicadas técnicas de gamificación.

Torres en [4] manifiesta que la gestión de proyectos es un ámbito en la cual la gamificación tiene cabida. Menciona que...*”el seguimiento del cumplimiento de objetivos, re-planificación de calendario, asignación de tareas,... se convierten en una autentica carga de trabajo que **tiende a evitarse** tanto por el project manager como por los miembros del equipo. En ese aspecto actúa la Gamification como solución”.*

Terán Panchi en [2] señala, como ventajas de la incorporación de la gamificación, el trabajo en equipo y la autonomía de los integrantes del grupo, la competencia entre los miembros, que deriva en mayor productividad y la valoración del desempeño de las personas involucradas.

El diseño de un sistema gamificado para gestionar proyectos genera, como señala Torres en [4], los beneficios de trabajar en base a objetivos claros, contar con retroalimentación constante, conseguir el trabajo colaborativo y el compromiso, incentivar a los miembros del equipo de desarrollo para que informen sus avances.

La norma ISO 10006 es una norma internacional que brinda orientación sobre la aplicación de la gestión de la calidad en los proyectos. Puede aplicarse a diferentes tipos de proyectos, en diferentes ámbitos y en forma independiente del producto o proceso involucrado [5] Diversas secciones de la mencionada norma dan sustento a lo propuesto en este trabajo. En la sección 5.2.2 de [5] se establece la necesidad de documentar los objetivos del proyecto detallándose lo que se desea lograr (considerando costo, tiempo y calidad del producto) y lo que se debe medir. En las secciones 5.2.3 de [5] se determina que debe seleccionarse un líder para el proyecto quien tendrá, entre otras responsabilidades, la de crear y mantener un ambiente interno en el cual el personal pueda llegar a involucrarse totalmente en el logro de los objetivos. En [6] se define al director del proyecto como la persona asignada por la organización ejecutante para alcanzar los objetivos del proyecto y se destaca el liderazgo como una de las capacidades con las que debe contar.

En relación a las personas que conformen el equipo dirigido por el líder seleccionado, la sección 5.2.4. de [5] indica que deberá proporcionárseles las herramientas, técnicas y métodos apropiados para permitirles efectuar el seguimiento y control de los procesos. En la sección 5.2.8. de [5] se establece que debe registrarse la información relativa al

progreso y desempeño del proyecto y que deben llevarse a cabo evaluaciones del desempeño y del avance. El análisis de la información obtenida será útil para tomar decisiones eficaces con respecto al proyecto y para revisar el plan de gestión del proyecto.

Aspectos importantes en el desarrollo de este trabajo son la motivación y el trabajo en equipo. La ISO 10006 en su sección 6.2.4. aborda estos temas señalando que el desempeño eficaz del equipo requiere que sus miembros, individualmente, sean competentes, estén motivados y estén dispuestos a cooperar con los demás. [5] Además plantea que la mejora del desempeño del equipo en su conjunto y de cada uno de sus integrantes se ve beneficiada cuando se brinda información que les permita reconocer la importancia de sus actividades en el logro de los objetivos del proyecto y los objetivos de calidad. En la sección 6.2.4. de [5] también se determina que la existencia de un trabajo en equipo eficaz, debe reconocerse y, cuando proceda, recompensarse.

La norma ISO 9001 especifica los requisitos para un sistema de gestión de la calidad. Estos requisitos son genéricos y aplicables a todas las organizaciones sin importar su tipo, tamaño o producto elaborado [7]. Las secciones de la mencionada norma que dan fundamento a lo planteado en este trabajo se relacionan con la gestión de recursos humanos, la realización del producto y la medición, el análisis y la mejora. En la sección 6.2.2. de [7] y 7.3 de [9], se indica la importancia de que el personal sea consciente de la pertinencia y relevancia de sus actividades y de cómo contribuyen al logro de los objetivos de calidad. En la sección 7.1 de [7] y 8.1 de [9], se indica la necesidad de planificar y desarrollar los procesos necesarios para la realización del producto. En la sección 7.3.1. [7] y 8.3.2 de [9], se hace referencia a la planificación y control del diseño y desarrollo del producto. En la sección 8.2.4 de [7] y 8.6 de [9], se establece la necesidad de realizar un seguimiento y medición de las

características del producto para verificar que se cumplen los requisitos. Esto se debe realizar en las etapas apropiadas del proceso de elaboración del producto que se planificaron. La norma ISO 9000-3 [8] [10] facilita la interpretación de la ISO 9001 en relación al desarrollo, implementación y mantenimiento de software.

Las secciones de las normas mencionadas dan sustento al desarrollo del método propuesto en este trabajo. Tanto en [5] como en [7] se expresa la importancia del trabajo en equipo, la buena comunicación, la motivación, el compromiso, el cumplimiento de objetivos, la medición de resultados, como características necesarias en el desarrollo de proyectos. En este trabajo se unifica lo que las normas de calidad exigen, en cuanto a la gestión de proyectos y a la calidad del producto, con técnicas de gamificación que faciliten la gestión y el desarrollo. La inclusión de tecnologías de la información y la comunicación se integran a este método como un recurso para hacer más ágiles las tareas de medición de resultados, obtención de información relevante para la mejora de los procesos de gestión y de desarrollo.

Líneas de Investigación, Desarrollo e Innovación

Considerando lo detallado anteriormente en esta línea de investigación se sugiere el desarrollo de un método para mejorar la calidad del producto y el desempeño de los equipos de desarrollo de software, basado en la incorporación de técnicas gamificación [2][3] y en la aplicación de las normas ISO 9000-3 [8] [10] e ISO 10006 [5] e ISO 9001 [7].

Se entiende por gamificación a la incorporación de mecánicas y dinámicas de juegos en ambientes o aplicaciones no lúdicas con la finalidad de lograr concentración, fidelización, compromiso, y otros valores comunes a los juegos. Se definen las dinámicas de juego como las necesidades y motivaciones que incentivan a que las personas realicen una acción y las

mecánicas como las estrategias que se utilizan para satisfacer esas necesidades (dinámicas). Las mecánicas de juego incluyen puntos, insignias de logros o niveles, bonos, misiones, retos, tablas de clasificación, entre otras. En cuanto a las dinámicas se utilizan recompensas, logros, competencia, etc.

El método que se propone incluye seleccionar e implementar aspectos de las mencionadas normas que puedan aplicarse para mejorar la gestión de proyectos de software y la calidad del producto. Seleccionar e implementar la plataforma o aplicación que gestione los proyectos de software. Seleccionar las técnicas de gamificación que mejor se adapten para mejorar la motivación del equipo de desarrollo y facilitar la gestión del proyecto. Definidas las técnicas de gamificación será necesario el desarrollo de una aplicación que sea capaz de aplicarlas a los resultados de la gestión de los proyectos. Esta aplicación deberá mostrar de manera dinámica los avances de los proyectos, los logros de los equipos de trabajo y de sus integrantes. La información obtenida servirá como retroalimentación para cada uno de los desarrolladores, para los líderes de cada proyecto, para los responsables del área de desarrollo, entre otros.

Resultados y Objetivos

Se ha estudiado y analizado las distintas características de cada una de las normas consideradas. Este estudio se enfocó en lograr una mejora en la gestión de proyectos de software y la calidad del producto.

Además en paralelo se ha llevado a cabo la selección y estudio de las técnicas de gamificación que mejor se adapten para mejorar la motivación del equipo de desarrollo y facilitar la gestión del proyecto.

La tarea que se está llevando a cabo consiste en el estudio de algunas aplicaciones que aplican o utilizan técnicas de gamificación. En base a ello, se está realizando un prototipo que se adapte a las condiciones de trabajo actuales. Todo esto se estudiará

sobre una empresa del medio en la ciudad de Córdoba.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación se está desarrollando una tesis de Maestría en Calidad del Software (Plan Ord. 017/09-CD) de la Ing. María Soledad Bianciotti, como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación.

Referencias

- [1] Eduardo Herranz y Ricardo Colomo Palacios - Artículo ResearchGate La Gamificación como agente de cambio en la ingeniería de Software.- Enero 2012
https://www.researchgate.net/profile/Eduardo_Herranz/publication/263737887_La_Gamificacin_como_agente_de_cambio_en_la_Ingeniera_del_Software/links/02e7e53ce2cd9dc860000000.pdf
Documento RPM-9.2-Articulo3.pdf
- [2] Ing. Christian Hugo Terán Panchi, Investigación Inclusión de una técnica de gamificación en la estructura operacional de un help desk aplicado a la categorización de tickets, para mejorar el nivel de servicio, compromiso y rendimiento del equipo de trabajo - Universidad de las Américas - Chile - 2015.
<http://dspace.udla.edu.ec/handle/33000/3380>. Documento UDLA-EC-TMGSTI-2015-11(S).pdf.
- [3] David Viteri - Investigación Desarrollo de un modelo para la gestión del cambio usando herramientas de gamificación con continuidad en el negocio. - Universidad Católica del Ecuador - Quito - 2015.
<http://repositorio.puce.edu.ec/handle/2000/9304> Documento DESARROLLO DE UN MODELO PARA LA GESTIÓN DEL CAMBIO USANDO HERRAMIENTAS DE GAMIFICACIÓN CON CONTIN.pdf.
- [4] Carlos Torres - <http://ctorresvalhondo.com/blog/>
- [5] INTE-ISO 10006 – Sistemas de gestión de la calidad - Directrices para la gestión de la calidad en los proyectos -

- Primera edición 2003-07-16
- [6] Gustavo Gabriel Maigua y Emmanuel Fernando López – Buenas prácticas en la dirección y gestión de proyectos informáticos - Editorial de la Universidad Tecnológica Nacional – edUTecNe - ISBN: 978-987-1896-01-1 -2012.
- [7] ISO 9001 – Sistemas de gestión de la calidad - Requisitos - Cuarta edición 2008-11-15
- [8] ISO 9003
-
<https://www.iso.org/standard/66240.html>
- [9] International Organization for Standardization - Correlation matrices between ISO 9001:2008 and ISO 9001:2015 – www.iso.org
- [10] IRAM-ISO/IEC 9003:2006.

Hacia un Modelo de Evaluación de Calidad para Datos Basado en la Norma ISO/IEC 25012

Carrizo Claudio*, Saldarini Javier*, Salgado Carlos+, Sanchez Alberto+, Peralta Mario+

*Grupo de I+D Calidad de Software - Facultad Regional San Francisco
Universidad Tecnológica Nacional

Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147
{ cjcarrizo77, saldarinijavier }@gmail.com

+Departamento de Informática Facultad de Ciencias Físico-Matemáticas y
Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
e-mail: {csalgado, mperalta, alfanego}@unsl.edu.ar

RESUMEN

Los datos representan el activo corporativo más importante, ya que a través del procesamiento de los mismos, permiten generar información de gran valor para las empresas en lo que respecta a las actividades de gestión, operación y toma de decisiones.

Uno de los problemas principales al que se enfrentan las empresas en general está relacionado con la falta de calidad de los datos que gestionan. Esto tiene un impacto directo en el desempeño de la empresa debido a que, al no contar con datos fiables, se cometen errores al tomar decisiones, se pierde competitividad y se daña la imagen corporativa. Por este motivo, se considera un factor clave poder garantizar la calidad de los datos.

El propósito de este trabajo consiste en elaborar un modelo basado en la Norma ISO/IEC 25012, el cual estará compuesto por características y sub-características de calidad de datos, las cuales deberán ser evaluadas a través de la aplicación de métricas e indicadores para un caso en concreto en el dominio de las empresas.

Con el modelo de datos propuesto se pretende brindar un aporte a la investigación científica y a las empresas

que deseen evaluar la calidad de sus datos en post de mejorar la gestión de los mismos.

Palabras clave: Modelo de Calidad – ISO 25012 – Métricas – Indicadores – Calidad de Datos.

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

Los datos se han convertido actualmente en uno de los activos más

valiosos para las empresas debido a que estos permiten generar información de gran valor para la gestión, operación y toma de decisiones. La calidad de la información es un factor clave para cualquier organización debido a que el acierto de las decisiones depende de la calidad de los datos en los que se basan dichas decisiones [1].

La falta de calidad de datos es uno de los principales problemas a los que se enfrentan los responsables de sistemas de información y las empresas en general, pues representa claramente uno de los problemas "ocultos" más graves y persistentes en cualquier organización. [2]. Una pobre calidad de datos tiene un gran impacto en el desempeño de la empresa, genera importantes sobrecostos, daña la imagen corporativa y trae consecuencias como las que se detallan a continuación:

- Errores en la toma de decisiones.
- Pérdida de competitividad.
- Impacto en la retención de clientes.
- Ineficiencia operacional
- Costos no planificados

De acuerdo a lo expuesto anteriormente, y entendiendo que la gestión de datos constituye un recurso estratégico para las empresas, se hace imprescindible garantizar la calidad de los datos. Disponer de datos con alto nivel de calidad no sólo es importante para las empresas en lo que se refiere a una adecuada gestión y toma de decisiones eficaces, sino también para lograr fortalecer la relación con el cliente de manera de poder fidelizarlo y lograr la satisfacción del mismo.

La calidad de los datos es un concepto multidimensional [4] y frecuentemente es

definida como "datos apropiados para el uso" [3, 5]. Esto quiere decir que el usuario es quién determina si un conjunto de datos, usados en una determinada tarea y en un contexto específico, pueden ser usados para el objetivo previsto. Según la norma ISO/IEC 25012 [6], la calidad de datos es el "grado en que las características de los datos satisfacen necesidades implícitas y establecidas cuando son usados en condiciones específicas". Lo expresado anteriormente, determina la relevancia que tiene la participación del usuario a la hora de definir si un conjunto de datos es de calidad. En consecuencia, se puede concluir que el término "calidad de datos" es un concepto relativo ya que, al variar las expectativas de los usuarios respecto a los mismos, varían también las características que éstos deben tener para ser considerados adecuados. Esto significa que, la calidad de los datos está asociada a un conjunto de características que permiten, por un lado, establecer un lenguaje común y, por otro lado, focalizar los problemas de calidad de los datos y las oportunidades de mejora.

Recientemente la Organización de Estandarización Internacional (ISO) en conjunto con la Comisión Electrotécnica Internacional (IEC), han presentado la familia de normas ISO/IEC 25000 (SQuaRE - System and Software Quality Requirements and Evaluation) [7], la cual proporciona una guía para establecer los requisitos y evaluación de calidad de productos de software. Esta familia está compuesta por 5 divisiones, una de ellas es la ISO/IEC 2501n que especifica los requisitos para un modelo de calidad. A su vez, dentro de esta división se encuentra la norma ISO/IEC 25012 que proporciona un modelo de calidad de datos en donde se establecen las características de calidad de datos que se deben tener en cuenta a la hora de evaluar

las propiedades de un producto de datos determinado. La Calidad del Producto de Datos se puede entender como el grado en que los datos satisfacen los requisitos definidos por la organización a la que pertenece el producto. Son precisamente estos requisitos los que se encuentran reflejados en el modelo de calidad de datos mediante 15 características, las cuales están clasificadas en 2 grandes categorías: Calidad de Datos Inherente: y Calidad de Datos Dependientes del Sistema [8].

El objetivo de este trabajo consiste en elaborar un modelo basado en la Norma ISO/IEC 25012 que estará compuesto por características y sub-características de calidad de datos, las cuales deberán ser evaluadas a través de la aplicación de métricas e indicadores en el dominio de las empresas a través de un caso en concreto. Para lograr esto será necesario:

- Identificar y seleccionar las características y sub-características de calidad de datos basadas en el Estándar ISO/IEC 25012 para la elaboración del modelo.
- Definir y/o adaptar métricas e indicadores para cada una de las características y sub-características del modelo de calidad de datos elaborado basándose en el estándar ISO/IEC 25024 [9].
- Evaluar el modelo a través de la aplicación de las métricas e indicadores en un caso en concreto en el dominio de las empresas.

El modelo de calidad de datos resultante será un instrumento de suma utilidad para las empresas ya que permitirá evaluar la calidad de los datos que manejan y

determinar las mejoras necesarias para asegurar que los datos almacenados tengan la calidad deseada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Una de las áreas prioritarias establecidas por la especialidad Sistemas de Información en la política de desarrollo de Ciencia y Tecnología de la Facultad Regional San Francisco es la de "Calidad de Software". Dentro de este contexto, la línea de investigación en la cual se está trabajando actualmente está relacionada con la temática "Calidad de datos".

Los principales ejes de trabajo de esta línea de investigación se detallan a continuación:

- Elaboración de un modelo compuesto por características y sub-características de calidad de datos para el dominio de las empresas.
- Elaboración de métricas e indicadores para cada una de las características y sub-características del modelo.
- Evaluación del modelo en un caso en concreto en el dominio de las empresas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Dentro de los resultados esperados para esta línea de investigación, se han logrado alcanzar algunos al momento de esta publicación. La propuesta consiste en poder obtener un modelo que permita evaluar a través de características y sub-características la calidad de los datos que son gestionados por las empresas. Para poder lograrla se han fijado los siguientes objetivos:

- Se han estudiado, identificado y seleccionado las características y sub-características de calidad de datos basadas en el Estándar ISO/IEC 25012.
 - Se ha elaborado un modelo de calidad de datos general con las mejores características y sub-características seleccionadas. Y complementada con características obtenidas de la experiencia y en la de expertos del medio.
 - Se están analizando distintas métricas de la bibliografía para poder utilizarlas, redefinirlas y/o adaptarlas al dominio que se estudiará. Para ello también se hace uso del estándar ISO/IEC 25024.
 - Evaluar el modelo a través de la aplicación de las métricas e indicadores definidos en un caso de estudio en concreto.
- Difundir y divulgar los resultados obtenidos

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos métodos de evaluación de calidad con la participación del Grupo de Investigación "Calidad de Software" perteneciente a la UTN Facultad Regional San Francisco, Córdoba.

Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) del Ing. Claudio Carrizo, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Como así también se están llevando a cabo algunas tesinas

de grado para la Licenciatura en Ciencias de la Computación y trabajos finales en el marco de la Ingeniería en Informática.

5. BIBLIOGRAFÍA

- [1] Calidad de Datos - ISO/IEC 25012. Recuperado el 07/09/16 <http://www.aqclab.es/index.php/calidad-datos-iso-25012>
- [2] Introducción a la Calidad de Datos. Recuperado el 15/09/16. <http://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/368784/introduccion-a-la-calidad-de-datos-definicion-control-y-beneficios>
- [3] D. Ballou and R. Wang. "Modeling information manufacturing systems to determine information product quality". *Management Science*. Vol. 44, pp. 462-484. 1998.
- [4] B.D. Klein. "User perceptions of data quality: Internet and traditional text sources". *Journal of Computer Information Systems*. Vol. 41, pp. 9-18. 2001.
- [5] D. Strong, Y. Lee and R. Wang. "Data Quality in Context". *Communications of the ACM*. Vol. 40, Issue 5, pp. 103-110. May, 1997.
- [6] ISO/IEC-25012. "ISO/IEC 25012: Software Engineering - Software Quality Requirements and Evaluation (SQuaRE) - Data Quality Model". 2008

- [7] La familia de normas ISO/IEC 25000. Recuperado el 05/10/2016. <http://iso25000.com/index.php/normas-iso-25000>
- [8] ISO/IEC 25012. Recuperado el 15/10/2016.
- [9] <http://iso25000.com/index.php/normas-iso-25000/iso-25012>
- [10] Norma ISO/IEC 25024:2015. Recuperado el 09/11/2016. <https://www.iso.org/standard/35749.html>

Herramienta de Calendarización para Proyectos Desarrollados Utilizando XP

Diego Alberto Godoy^a, Edgardo A. Belloni^b, Eduardo O. Sosa^c, Wilson F. Andres^d

Centro de Investigación en Tecnologías de la Información y Comunicaciones (C.I.T.I.C.)
Departamento de Ingeniería y Ciencias de la Producción-Universidad Gastón Dachary
Av. López y Planes 6519- Posadas, Misiones, Argentina. Teléfono: +54-376-4438677

^adiegodoy@citic.ugd.edu.ar, ^bebelloni@ugd.edu.ar, ^ceduardo.sosa@citic.edu.ar,
^dwilson.andres@citic.edu.ar

Resumen

En este trabajo se presenta un proyecto de investigación denominado “Simulación como Herramienta para la Mejora de los Procesos de Software Desarrollados con Metodologías Ágiles Utilizando Dinámica de Sistemas”, cuyo objetivo es construir modelos de simulación utilizando la disciplina de dinámica de sistemas y técnicas alternativas, que permitan dar soporte a la gestión de los procesos de desarrollo de software ágiles llevados a cabo utilizando diversas metodologías y prácticas ágiles. Particularmente se presentan los avances realizados hasta ahora en relación a construcción de una herramienta de calendarización para proyectos que utilizan la metodología Extreme Programming (XP).

Palabras claves: Administración de Proyectos de Desarrollo de Software; Extreme Programming (XP).

Contexto

El trabajo presentado en este artículo tiene como contexto marco el proyecto de investigación denominado “Simulación como Herramienta para la Mejora de los Procesos de Software Desarrollados con Metodologías Ágiles Utilizando Dinámica de Sistemas”, registrado actualmente en la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD) † y radicado en el Centro de Investigación en

Tecnologías de la Información y Comunicaciones de dicha universidad.

El mismo fue ratificado e incorporado como proyecto aprobado en el llamado a presentación interna de la UGD de proyectos de investigación N°6 mediante la Resolución Rectoral 24/A/15.

Entre las líneas con mayores resultados dentro del proyecto referido, se encuentran las de: (i) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software que utilizan Programación Extrema”, (ii) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software Bajo Scrum” y (iii) “Modelo de Simulación Dinámico de Gestión de Proyectos de Desarrollo de Software desarrollados con Crystal”. Este artículo se enfoca en la presentación de un línea para estudiar la calendarización de proyectos llevados a cabo Utilizando XP [1] y diseñar un prototipo de herramienta de calendarización de liberación e iteraciones en proyectos con XP.

Introducción

La calendarización o gestión de calendario (lo que en inglés se denomina scheduling) es un concepto que está presente en todo tipo de proyecto donde se manifiesta la necesidad de realizar la asignación temporal de recursos a actividades para lograr algún objetivo deseado [2]. Teniendo en cuenta esta premisa, se puede deducir que los proyectos de software no son una excepción. Una vez que se ha seleccionado

† Mediante Res.Rectoral UGD N° 04/1/12.-

un modelo de proceso apropiado, se han identificado las tareas de ingeniería del software que es preciso realizar, se ha estimado la cantidad de trabajo y el número de personas, se ha conocido la fecha límite, incluso se ha considerado los riesgos; se debe unir los puntos, creando una red de tareas de ingeniería del software que le permitirán tener el trabajo listo a tiempo. Básicamente, esto es la calendarización de proyectos de software [3]. Representa una de las tareas más difíciles para los gestores de proyectos. Usualmente las estimaciones previas son una base incierta para la calendarización de un nuevo proyecto [4]. Por lo tanto, gestionar eficientemente el calendario es un desafío crucial, que enfrenta el gestor y su equipo, para cumplir con la fecha límite, establecida en el comienzo de un proyecto nuevo.

Cabe mencionar, además, que la calendarización constituye un riesgo en el desarrollo de software. De hecho, el riesgo de calendario es considerado uno de los riesgos más comunes en proyectos de software [5]. De acuerdo a esto, es posible afirmar que una óptima planificación del calendario del proyecto, es esencial para mitigar o eliminar el riesgo de calendarización.

Considerando las cuestiones arriba mencionadas, se torna importante contar con herramientas que faciliten al equipo, llevar adelante la complejidad que significa administrar adecuadamente el calendario.

Si bien ya existen ciertos métodos maduros y factibles para la gestión de calendario de proyecto, como ser el Método del Camino Crítico (CPM), la técnica de evaluación y revisión de programa (PERT), el gráfico de Gantt, etc.; éstas son técnicas tradicionales, aplicables a todos los tipos de proyectos, y no están pensadas de acuerdo a las características específicas de los proyectos de software [6]. No obstante, dichas técnicas generalizadas de calendarización de proyecto, se aplican (un poco modificadas) a los proyectos de software. Para este fin se puede contar con varias herramientas de software, como por ejemplo AMS

Realtime, Microsoft Project, Viewpoint, etc. [3].

Las herramientas de software que se han nombrado arriba están basadas en los métodos y las técnicas de planificación antes mencionadas, y se utilizan particularmente para llevar adelante la calendarización de proyectos de desarrollo de software que emplean metodologías tradicionales, que por lo general son proyectos de gran envergadura, emprendido por grandes equipos.

Como resultado de la evolución de las metodologías de desarrollo de software, han emergido las denominadas Metodologías Ágiles, las cuáles sugieren un cambio revolucionario con respecto al enfoque de desarrollo convencional, a través de una serie de principios [7], cuyas características presentan un modo de planificación que no es adecuado abordar con las herramientas de calendarización tradicionales.

Según lo afirma Szöke [8], en los entornos ágiles se carece de un sólido soporte metodológico de calendario; y aunque están disponibles ciertas herramientas para planificación de proyecto (como Rally [9] y XPlanner [10]), una encuesta de Herramientas Ágiles [11] indica que tales herramientas presentan una escasa aceptación. La razón de esto puede ser explicada a raíz de que dichas herramientas poseen un débil soporte embebido de funciones tradicionalmente importantes de calendarización de proyecto, como por ejemplo la asignación de recursos [8].

Por lo tanto, tomando en consideración el panorama de calendarización de proyectos de software planteado anteriormente, sobre todo lo relacionado con las metodologías ágiles, se llevará adelante un proyecto de investigación, cuyo objetivo sea el de diseñar un prototipo de herramienta para la calendarización de proyectos de software ágiles desarrollados con la metodología particular Programación Extrema (XP) [1], que brinde ventajas con respecto a las características comunes que presentan las principales herramientas existentes para dicha metodología. A modo de cumplir

con este propósito, se combinarán las ideas de dos métodos de calendarización ágil. Uno es el método para calendarización de iteración ágil, presentado por Szöke [8], el cual promete mejorar significativamente el balanceo de carga de los recursos, producir la mayor calidad y el menor riesgo de calendario posible, y proveer la toma de decisiones más informadas y estables, debido a una producción de calendario optimizada. Y el otro es el método estadístico para planificación de una liberación (release) ágil, propuesto por Logue y McDaid [12], el cual permite manejar la incertidumbre en el valor del negocio de cada historia de usuario, el tamaño de la historia y los recursos disponibles, brindando la posibilidad de gestionar la planificación de qué funcionalidades incluir en las próximas liberaciones.

Línea de Investigación

Para esta línea se han planteado los siguientes objetivos:

Como objetivo general se propone diseñar un prototipo de calendarización de liberación y sus iteraciones en proyectos con XP.

Como objetivos específicos se realizarán los siguientes:

- Estudiar las generalidades de la calendarización en los entornos ágiles y en XP.
- Identificar características de las principales herramientas de calendarización ágil.
- Analizar los métodos de calendarización en los cuáles se basará el prototipo. Modelar el prototipo a través del modelado orientado a objetos usando notación UML [13].
- Desarrollar el prototipo de software. Realizar pruebas de la herramienta con datos históricos de proyectos reales.
- Evaluar la herramienta por medio de simulación.

Se pretende que la herramienta de calendarización esté basada en web, por lo que será accesible desde cualquier dispositivo que soporte navegación web.

La particularidad que presenta el software radica en el hecho de que se concentra exclusivamente en las funciones de calendarización de proyecto. El sistema requerirá el ingreso de datos (cantidad de miembros del equipo, recursos, liberaciones, historias de usuario, iteraciones estimadas, requerimientos, estimación de errores) que poseen los miembros del equipo de proyecto, y utilizando métodos concretos de calendarización, como los de [8] y [12], será capaz de obtener un plan de liberación óptimo y de producir un calendario de iteraciones optimizado, de más calidad y menor riesgo, brindando así, apoyo al equipo en la toma de decisiones más informadas y estables.

Al estar basado en web permitirá que los desarrolladores individuales puedan ingresar los datos de sus estimaciones en cualquier momento, incluso luego de haber realizado reestimaciones, y el software se encargará de generar el calendario resultado, el cual estará a la vista de todo el equipo.

Debido a que la herramienta se limitará a tratar específicamente con las funciones de calendarización, estará libre de la complejidad inherente a las herramientas de administración de todo el ciclo de vida del proyecto, y por consiguiente la interactividad e interfaces de usuario serán más simples e intuitivas, alineándose de esta manera a los principios ágiles.

Resultados

En este trabajo se presentaron avances de esta línea en que se han estudiado diversas herramientas para la calendarización de proyectos llevados a cabo Utilizando XP [1] con el objetivo de diseñar un prototipo de herramienta de calendarización de liberación e iteraciones en proyectos con XP.

En esta fase del trabajo se ha culminado la primera etapa, que incluye la adquisición de conocimientos sobre calendarización ágil y XP. También involucra el análisis detallado de los

métodos de calendarización, en los cuales se basará el prototipo.

En la segunda etapa se contempla el diseño del modelo siguiendo el modelado orientado a objetos, que se divide en captura de requisitos, diseño conceptual y modelado de comportamiento, utilizando el Lenguaje Unificado de Modelado UML [13]. Y la implementación de acuerdo a la arquitectura Modelo-Vista- Controlador (MVC).

Como trabajo futuro se espera avanzar, con la construcción de la tercera y última etapa abarca las pruebas del prototipo, y corridas de simulación a fin de comparar los resultados.

Como se ha mencionado anteriormente, las pruebas se realizarán con datos históricos de proyectos reales. Esos datos constituyen las Historias de Usuario para cada iteración y las tareas correspondientes a cada historia, las estimaciones de sus respectivos tamaños, la velocidad del proyecto, la cantidad de iteraciones y la duración de estas, la cantidad de desarrolladores y el factor de carga de los mismos, entre otros. A partir de esto se probará el desempeño de la herramienta para producir un calendario óptimo, es decir reduciendo el tiempo de culminación del plan al punto en el que aún se mantenga una alta probabilidad de completarlo, de igual manera al Modelo Dinámico de Simulación presentado en [14] [15].

Adicionalmente, se prevé adaptar el prototipo para comparación con los simuladores de modelos realizados para Scrum [16] [17] [18] [19].

Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, un Doctor en Ciencias Informáticas y Magister en Redes de Datos; un Doctor en Tecnologías de la Información y Comunicaciones Magister y Especialista en Ingeniería de Software; un Maestrando de Ingeniería de la Web, y seis

estudiantes en período de realización de trabajos finales de grado en el contexto de las carreras de Licenciatura en Sistemas de Información y de Ingeniería en Informática de la UGD. Actualmente, el número de tesinas de grado aprobadas en el contexto de este proyecto, es de cinco, y otras tres en proceso de desarrollo. El número de tesis de maestría terminadas relacionadas con este proyecto es de una.

Bibliografía

- [1] Kent Beck, *Una Explicación de la Programación Extrema. Aceptar el Cambio*. España: Addison Wesley, 2002.
- [2] Henri Casanova. (2013, Octubre) nii.ac. [Online]. <http://www.nii.ac.jp/userimg/lectures/20131010/Lecture1.pdf>
- [3] Roger S. Pressman, "Ingeniería del Software. Un enfoque práctico.," in *Ingeniería del Software. Un enfoque práctico.*, Mc-Graw Hill, Ed.: Mc-Graw Hill Interamericana, ch. 24, pp. 724-740.
- [4] Ian Sommerville, "Ingeniería del Software," in *Ingeniería del Software*, Pearson Addison Wesley, Ed. Madrid, España: Pearson Education Limited, 2005, ch. 5, p. 91.
- [5] Keshnee Padayachee, "An Interpretive Study of Software Risk Management Perspectives," in *SAICSIT*, 2002, pp. 118-127.
- [6] Jun-guang Zhang, "Schedule Management Method Study of Middle and Small Software Projects," in *IEEE Conference Publications*, 2011, pp. 1495-1498.
- [7] Ward Cunningham. (2001) Manifiesto por el Desarrollo Ágil de Software. [Online]. <http://www.agilemanifesto.org/is o/es/>

- [8] Ákos Szöke, "Decision Support for Iteration Scheduling in Agile Environments," in *Product- Focused Software Process Improvements*, Frank Bomarius, Ed. Berlín, Alemania: Springer Berlin Heidelberg, 2009, pp. 156-170.
- [9] Rally homepage. [Online]. <http://www.rallydev.com>
- [10] Xplanner homepage. [Online]. <http://xplanner.codehaus.org>
- [11] M. Dubakov and P. Stevens, "Agile Tools: The good, the bad, the ugly," *Agile Journal*, 2008.
- [12] Kevin Logue and Kevin McDavid, "Handling Uncertainty in Agile Requirement Prioritization and Scheduling Using Statistical Simulation," in *IEEE Conference Agile*, 2008, pp. 73-82.
- [13] Unified Modeling Language version 2.1. [Online]. <http://www.uml.org>
- [14] Godoy Diego Alberto y Kasiak Tamara., "Modelo dinámico de simulación para la gestión de proyectos de software desarrollados con XP," in *Actas XVIII Congreso Argentino de Ciencias de la Computación*, 2012, p. 10.
- [15] Tamara Kasiak y Godoy Diego Alberto, "Simulación de Proyectos de Software desarrollados con XP," *XIV Workshop de Investigadores en Ciencias de la Computación.*, 2012.
- [16] Diego Alberto Godoy, Edgardo A. Belloni, Henry Kotynski, Hector H Dos Santos, and Eduardo Omar Sosa, "Simulando Proyectos de Desarrollo de Software Administrados con Scrum," in *XVI Workshop de Investigadores en Ciencias de la Computación RedUNCI*, Ushuaia, 2014.
- [17] Diego Alberto Godoy, Eduardo Omar Sosa, Edgardo Belloni, and Henry Kotynski, "Simulación Dinámica de Gestión de Tareas en Proyectos Desarrollados Con Scrum," in *II Congreso Nacional de ingeniería informática/ingeniería de sistemas (CoNaIISI)*, Universidad Nacional de San Luis, San Luis, 2014, 2014.
- [18] Diego Alberto Godoy, "Diseño de un Simulador Dinámico de Proyectos de Desarrollo de Software que utilizan Metodología Scrum," UNLP, La Plata, Tesis de Maestría 2014.
- [19] Diego Alberto Godoy, Cristian Henry Kotyński, Edgardo Anibal Belloni, and Eduardo Omar Sosa, "Un Modelo de Simulación de Proyectos scrum con dinámica de sistemas.," in *WWW/INTERNET 2015 e COMPUTAÇÃO APLICADA 2015 FLORIANÓPOLIS*, Florianópolis, 2015, Organizada por IADIS International Association for Development of the Information Society.

Impacto del Factor Peopleware en la Implantación de Sistemas Informáticos

Marisa Panizzi^{1,2}, Oscar Bravo¹, Luis Catanzariti¹, Andrea Fleischman¹, Andrea Alegretti¹, Gabriela Velazquez¹, Gustavo Apaza¹, Nadia Suarez¹, Alejandro Hossian^{2,3}, Ramón García Martínez^{2,4}

¹ Departamento de Ingeniería en Sistemas de Información. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.

Medrano 951 (C1179AAQ), C.A.B.A, Argentina.

² Grupo de Investigación en Sistemas de Información. Departamento de Desarrollo Productivo Tecnológico.

Universidad Nacional de Lanús.

29 de Septiembre 3901, (B1826GLC), Lanús, Buenos Aires, Argentina.

³ Grupo de Investigación en Aplicaciones de Sistemas Inteligentes en Ingeniería. Universidad Tecnológica Nacional.

Facultad Regional Neuquén.

Av. Pedro Rotter S/N Barrio Uno, (8318), Neuquén, Argentina

⁴ Comisión de Investigación Científicas - CIC.

Calle 526 e/10 y 11. La Plata. Buenos Aires. Argentina

marisapanizzi@outlook.com, oscarbravo2066@gmail.com, lcatanzariti@gmail.com, licfleischman.af@gmail.com, andyalegretti@gmail.com, gav.sistemas@gmail.com, gustavo_10ah@hotmail.com, alejandrohossian@yahoo.com.ar, rgm@yahoo.com

Resumen

El presente trabajo de investigación pretende realizar aportaciones al proceso de implantación de sistemas informáticos con un foco en como el peopleware que interviene en el proceso, lo afecta. Se pretende identificar los elementos que componen un modelo de proceso de software; con énfasis en el proceso de implantación. Los enfoques de los modelos actuales no custodian el factor peopleware como un factor de éxito de los proyectos de software, en consecuencia no miran de manera integral las características técnicas, los aspectos humanos y de organización del proceso de software. En este proyecto se formula una investigación documental sobre procesos de implantación de sistemas informáticos embebidos en los modelos de procesos de Ingeniería de Software de los estándares actuales, entre los considerados están IEEE 1074, Métrica versión 3, MoProSoft, Proceso Unificado de Rational, Programación Extrema, Scrum, Método de Desarrollo de Sistemas Dinámicos, Proceso Unificado Ágil.

El proyecto de investigación buscará identificar debilidades en la dimensión peopleware de cada uno de estos estándares y formular aportaciones paliativas en la fase de implantación del modelo de proceso considerado. Se realizará pruebas de concepto en los casos de estudio y casos de validación identificados que corroboren el modelo de proceso propuesto.

Palabras clave: Peopleware, sistemas informáticos, implantación de sistemas, elementos del proceso de implantación.

Contexto

La investigación que se reporta en este artículo es financiada parcialmente por el Proyecto UTN4347 de la Secretaria de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional.

La línea de investigación que se desarrolla en este proyecto tiene el asesoramiento científico del Director del Grupo de Investigación de Sistemas

de Información de la Universidad Nacional de Lanús.

Introducción

La evolución de la ingeniería de software en cuanto a los diferentes estándares o metodologías para llevar a cabo la construcción del producto software ha permitido obtener un abanico de modelos de procesos, los cuales se aplican en función de cada tipo de proyecto de desarrollo de software. Sin embargo, se presentan varios problemas relacionados con las descripciones de procesos en los modelos de procesos de software o en los estándares [Vasconcelos y Werner, 1997], por lo tanto las mismas debilidades se las visualiza en el proceso de implantación de un sistema informático. Dentro de estos problemas, podemos mencionar que no cubren todos los subprocesos a considerar en la etapa de implantación, no manejan todos los elementos básicos (la organización del trabajo, proyectos, procesos, actividades, productos, personas, capacidades, funciones y herramientas de aplicación) y el enfoque utilizado no es lo suficientemente estructurado para servir como una guía. Para ser útil, una descripción del proceso de implantación debe establecer y organizar todos los detalles relacionados con el proceso real o propuesta con claridad.

En los proyectos de software, generalmente se resta importancia al proceso de implantación o entrega del sistema informático por tratarse de uno de los eslabones finales de la cadena productiva del producto software. Esta desvalorización del proceso se refleja en que el profesional al que se le asigna el rol de implantador o responsable de la entrega del producto software al cliente, no posee las competencias socio-técnicas necesarias para participar en dicho proceso.

Tradicionalmente, las representaciones de modelos de procesos de software se han centrado en tres aspectos elementales del proceso: la actividad, el artefacto y el agente (humanos y computarizados) [Mc Chesney, 1995]. Sin embargo, otras características que han sido

probadas empíricamente que tienen una gran influencia en el proceso de producción: los roles humanos, la organización del trabajo entre los seres humanos y las competencias de los recursos humanos [Boehm, 2000; Adelson & Soloway, 2007; Curtis et al., 2001; Sherdil & Madhavji, 1996]. Los roles están parcialmente tratados por los modelos de procesos de software existentes [Finkelstein et al., 1994].

La organización del trabajo es considerado como independiente de las características aplicadas a modelar el proceso de software [Min y Bae, 1997] o se ignora [Engels & Groenewegen, 1994], ya que la organización del trabajo es una preocupación de la organización y, por lo tanto, forma el entorno del proceso del software y no tiene que ser modelada de forma explícita. En consecuencia, los enfoques de los modelos actuales no lo hacen de manera integral y no modelan de forma conjunta las características técnicas y de peopleware.

Por lo tanto, en la actualidad los problemas existentes con los modelos de procesos de software son: i) No existe una definición de un modelo que abarca la representación conjunta de los procesos, productos, personas y organización, ii) No está formalizado en el proceso de software los aspectos como la organización del trabajo, las personas y sus interacciones y iii) No hay ningún proceso definido que incluya tanto la parte técnica y humana del proceso en el modelo [Acuña S., 2005].

Esta problemática presentada a nivel del proceso completo de construcción de software incluye el problema que se intenta abordar en este trabajo de investigación, que es resolver el proceso de implantación de los sistemas informáticos con una orientación integral considerando los aspectos técnicos y adicionando los aspectos de peopleware.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto tiene como objetivo evaluar el impacto del factor humano en los procesos de desarrollo de software haciendo foco en el

proceso de implantación de sistemas informáticos en las organizaciones. Se estudiarán los aspectos humanos, la definición de roles, las interacciones entre los diferentes roles, las capacidades de las personas que desempeñan esos roles, entre otros y como esta dimensión humana articula con el proceso de implantación de sistemas informáticos.

Resultados y Objetivos

Los procesos de implantación o instalación de sistemas informáticos cuentan de manera parcial con las descripciones del mismo para poder llevarlo a cabo, no manejan todos los elementos básicos (la organización del trabajo, proyectos, procesos, actividades, productos, personas, capacidades, funciones y herramientas de aplicación) y el enfoque utilizado no es lo suficientemente estructurado. Además de las debilidades mencionadas en dicho proceso se ha encontrado un área de vacancia transversal en todos los modelos o estándares existentes, la consideración de las capacidades del peopleware.

Pregunta Problema:

¿Se puede lograr cubrir la vacancia de un Modelo de Proceso para la implantación de un sistema informático que integre las actividades y sus tareas, las técnicas, las herramientas, los artefactos y los roles incorporando las capacidades de peopleware?

El objetivo de este proyecto es lograr un modelo de proceso de implantación de sistemas informáticos que integre las actividades, tareas, herramientas, técnicas y artefactos adicionando el factor peopleware.

Los objetivos específicos son:

1. Desarrollar una taxonomía de los procesos de implantación de sistemas informáticos basada en el conjunto de metodologías y estándares existentes.
2. Desarrollar un modelo de proceso de implantación de sistemas informáticos que identifique fases, tareas, técnicas, herramientas, artefactos y roles que logre independencia del resto de los procesos de construcción de software seleccionados para el proyecto.

3. Desarrollar un modelo transversal al modelo de proceso que solamente considere al peopleware que participara en el modelo de proceso de implantación de un sistema informático. En el mismo se consideraran los roles necesarios para el desarrollo de las actividades del proceso y las capacidades con las que deben contar dichos roles para la realización de dichas actividades.
4. Desarrollar un modelo de proceso integrado de implantación de sistemas informáticos que articule con las buenas prácticas que proponen los estándares de certificación de procesos de software.
5. Realizar pruebas de concepto en los casos de estudio y casos de validación identificados que corroboren el modelo de proceso propuesto.

Se ha logrado una primera investigación exploratoria de los siguientes modelos de proceso: IEEE-1074 [IEEE, 1997], Métrica v3 [Pae, 2001], Scrum [Palacio, 2015], Proceso Unificado de Rational [Péaire C. et al., 2007], MoProSoft [Secretaria de Economía México, 2005], Programación Extrema [Beck K., 2004], Método de Desarrollo de Sistemas Dinámicos [Agile Business Consortium, 2016] Proceso Unificado Ágil [Ambler S., 2016].

Formación de Recursos Humanos

El equipo de investigación se encuentra conformado por un Director, tres docentes que se inician en la formación como investigadores, una graduada, dos estudiantes, dos alumnos de posgrado. Se estiman dos tesis de Maestría y una tesis de Doctorado.

Referencias

de Vasconcelos F. and Werner C., "Software development processbreuse based on patterns". Proceedings of the Ninth International Conference on Software Engineering and Knowledge Engineering (June 1997) 97- 104.

McChesney, I. "Toward a classification scheme for software process modelling approaches".

- Information and Software Technology 37, 7 (1995) 363-374.
- Boehm B., Abts C., Brown A., Chulani S., Clark B., Horowitz E., Madachy R., Reifer D. and Steece B., Software Cost Estimation with COCOMO II. Prentice Hall (2000).
- Adelson B. and Soloway E., "The role of domain experience in software design". IEEE Transactions on Software Engineering 11, 11 (November 1985) 1351- 1360
- Curtis B., Hefley W. and Miller S., "People Capability Maturity Model@(P-CMMB). Version 2.0". Carnegie Mellon University, Software Engineering Institute. Maturity Model CMU/SEI-2001-MM-001. (2001).
- Sherdil K. and Madhavji, N. "Human-oriented improvement in the software process". Lecture Notes in Computer Science, Software Process Technology: Proceedings of the Fifth European Workshop 1149 (Springer-Verlag, 1996
- Finkelstein, J. Kramer and B. Nuseibeh (Eds.), Software Process Modelling and Technology. Chap. 4. Research Studies Press. (1994)
- Min S. and Bae D., "MAM nets: A Petri-net based approach to software process modeling, analysis and management". Proceedings of the Ninth International Conference on Software Engineering and Knowledge Engineering 78-86. (June 1997).
- Engels G. and Groenewegen L., "SOCCA: Specifications of coordinated and cooperative activities". In: A. Finkelstein, J. Kramer and B. Nuseibeh (Eds.), Software Process Modelling and Technology. Chap. 4. (Research Studies Press, 1994) 71-102.
- Acuña T. Juristo N., Moreno A., Mon A. A software Process Model Handbook for incorporating people's capabilities. United States of America: Springer Science+Business Media, Inc.(2005).
- IEEE 1074, 1997. IEEE Standard for Developing Software Life Cycle Processes. IEEE Std 1074 (Revision of IEEE Std 1074-1995; Replaces IEEE Std 1074.1-1995) (1997).
- PAe, Métrica versión.3. Portal de Administración Electrónica. Gobierno de España. (2001).
- Palacio J. Scrum Manager.. Gestión de Proyectos SCRUM Manager (Scrum Manager I y II). <http://www.scrummanager.net>. Página vigente al 22/02/2017 (2015).
- Pénaire C., Edwards M, Fernandes A., Mancin E. y Carroll K. IBM. Rational Software.. The IBM Rational Unified Process for Systems (2007).
- Secretaría de Economía México. MoProSoft. Modelo de procesos para la industria del software. Versión 1.3 (2005).
- Beck, K. Extreme Programming Explained: Embrace Change, 2da. Edición. Addison-Wesley (2004).
- Agile Business Consortium. DSDM. <https://www.agilebusiness.org>. Página vigente al 16/02/2017 (2016).
- Ambler S. The Agile Unified Process (AUP). <http://www.ambysoft.com/unifiedprocess/agileUP.html>. Página vigente al 15/02/2017 (2016).

Ingeniería de Requisitos para Proyectos CRM

Gladys Kaplan^{1,2}, Jorge Doorn^{2,3}, Guillermo Hindi¹, Gabriel Blanco¹, Gabriel Pousada⁴, Andrea Vera¹, Claudia Litvak¹, Nora Gigante¹, Mirian Taboada¹

¹Departamento de Ingeniería e Innovación Tecnológica, Universidad Nacional de La Matanza (UNLaM)

²Departamento de Ciencias Básicas, Universidad Nacional de Luján (UNLu)

³Escuela de Informática, Universidad Nacional del Oeste (UNO)

⁴Departamento de Ciencias Económicas, Universidad Nacional de La Matanza (UNLaM)

(gkaplan@unlam.edu.ar, jdoorn@exa.unicen.edu.ar, gblanco@unlam.edu.ar)

Resumen

Los sistemas CRM (Customer Relationship Management) tienen por objetivo mejorar la relación de las organizaciones con sus mejores clientes, asegurando su satisfacción y por lo tanto su fidelización a largo plazo añadiendo valor a la organización y a sus clientes. Varios de los fracasos en las implementaciones de sistemas CRM tienen su origen en la falta de discriminación entre la *Estrategia CRM* con el *Software CRM*. Es común que se utilice el acrónimo CRM como sinónimo de estrategia y de software. Este uso ambiguo del término trae aparejado algunos problemas de implementación, ya que a la hora de seleccionar un Software CRM se debe tener en cuenta la estrategia particular de cada organización con el objetivo de parametrizar la aplicación y definir su uso. En la presente investigación se utiliza el término *Proyecto CRM* para incluir ambos conceptos ya que se considera que no es

conveniente pensar en un Software CRM sin una estrategia previamente definida, aunque de manera informal. La elección del Software CRM tiene dos pilares inamovibles: por un lado se sustenta en las necesidades analíticas definidas en la estrategia comercial y por el otro en la cultura organizacional. También se debe tener en cuenta que toda la información que surja del Proyecto CRM debe ser integrada al sistema de información de toda la organización y ser accesible para todos aquellos que estén autorizados a hacerlo. Es la identificación de las necesidades analíticas y en la integración donde la ingeniería de requisitos (IR) [Neill 03] aporta las respuestas adecuadas para seleccionar y parametrizar un Software CRM. En la presente investigación se utiliza un proceso particular para dicha IR, esta estrategia es el Proceso de Requisitos basado en Escenarios [Leite 04].

Palabras clave: ingeniería de requisitos, proceso de requisitos basado en

escenarios, estrategia CRM, software CRM.

Contexto

Este proyecto de investigación es la continuación de otros donde se trabajó sobre el proceso de requisitos basado en escenarios [Leite 04] que se utiliza para realizar todo el análisis. Dicho proceso fue desarrollado desde el año 1995 a 2000 en proyectos realizados en la UB, luego desde 2001 a 2004 en UTN-FRBA y finalmente desde 2005 a la fecha en UNLaM. También existen trabajos en otras Universidades como por ejemplo la UNLP, UCA, UNICEN, UNO, PUC-Rio, UNPA entre otras. En estos proyectos se ha estudiado el proceso de requisitos basado en escenarios definiendo una estrategia, los procesos de construcción de los modelos que lo conforman, las actividades de verificación y validación, la de especificación, priorización y rastreabilidad de los requisitos.

Introducción

Como se mencionó en el resumen, es frecuente que se utilice CRM (Customer Relationship Management) como sinónimo de estrategia de negocio y de software. En la primera acepción de CRM se hace referencia a la Estrategia CRM [Garrido 11] definida por la organización para optimizar la relación con los mejores clientes y fortalecer la fidelización a largo plazo [Galbreath 88]. [Swift 01] [Chen 03] [Law 03]. La segunda acepción hace referencia a un sistema informático que

permite registrar y analizar dicha relación [Hashimura 11] [Bose 03].

El conocimiento de los clientes aporta valiosa información a las áreas de venta y de marketing, permite conocer el ratio de pérdida de clientes, las causas por las que abandonan a la empresa, su nivel de satisfacción y fidelización, los motivos por los que se sienten más o menos satisfechos o los servicios y productos que más consumen y utilizan. Con esa información, la planificación de las campañas comerciales se ajusta mejor a la realidad, el target está más definido y, por tanto, hay más posibilidades de éxito

El CRM requiere una filosofía de negocio centrada en el cliente y una cultura de empresa que apoye decididamente los procesos de marketing, ventas y servicios. Las aplicaciones CRM permiten implantar la gestión de la relación con los clientes cuando la empresa tiene el liderazgo, la estrategia y la cultura acertada. Otros definen la estrategia CRM de un modo más amplio: no sólo se trata de adoptar una estrategia de negocio centrada en el cliente, sino también de implementarla de tal modo que cambie la forma de trabajar de empleados y clientes para que, con el apoyo de nuevas tecnologías, se creen nuevas interacciones con nuevos procesos de trabajo.

Es necesario comprender que sin la estrategia es muy difícil o casi imposible realizar una selección adecuada de un Software CRM. En contrapartida en la utilización de un software CRM es la posibilidad de aprender a partir de los datos registrados y luego analizados del

CRM. Sin embargo para esta investigación se considera necesario tener una estrategia comercial y analizar cómo funciona con el cliente y ratificar o rectificar la misma de ser necesario.

Bajo estas premisas se aborda la IR identificando las necesidades de análisis que requiere cada organización en particular. La IR podrá definir los aspectos relevantes que debe satisfacer un Software CRM, las herramientas analíticas con las que se debe contar y las posibilidades de implementación que fluyan con la cultura de la organización.

La estrategia Leite et al. [Leite 04], consiste en la creación de un glosario denominando léxico extendido del lenguaje (LEL) [Hadad 09] y de un conjunto de escenarios [Leite 00] actuales que representan situaciones observables para la comprensión del universo de discurso. Luego se construye un conjunto de escenarios futuros [Doorn 02] que representan situaciones con el nuevo sistema del software. Finalmente la especificación de los requisitos del sistema de software basándose en el conocimiento adquirido y registrado en las etapas previas.

Líneas de Investigación, Desarrollo e Innovación

El proyecto se planificó en diferentes etapas. En la Primera etapa el objetivo fue analizar el mercado del Software CRM y el contexto en el que dichos sistemas de software fueron pensados para operar con éxito en una organización. Se analizaron los principales sistemas software

existentes para determinar el problema real por el cual fallan las implementaciones. Finalmente se va a aplicar el proceso de requisitos basado en escenarios para identificar las necesidades analíticas de la organización y determinar los servicios requeridos para el software. Se espera que el mismo proceso de requisitos sea una manera de colaborar con el usuario en la revisión de su estrategia comercial.

Resultados y Objetivos

Como resultado de la primera etapa se concluye que todos los CRM son iguales y aplican indistintamente a cualquier organización. Un análisis profundo del mercado actual determinó que los sistemas de software CRM están pensados para diferente tamaño de empresas, diferentes tipos de comercialización, usos particulares, etc.

En la segunda etapa se descubrió que la primera dificultad se presenta cuando la organización no tiene una estrategia CRM identificada con claridad, por lo tanto es casi imposible utilizar una herramienta para algo que no existe. Esto tiene como consecuencia la insatisfacción de los usuarios-clientes y la presencia de información insuficiente. Por otro lado, se determinó que la selección falla en muchos casos por no analizar en primer lugar las necesidades organizacionales que se satisfacen con el CRM analítico [Hashimura 11], tomando en cuenta las necesidades de información y la cultura de la organización.

Como puede observarse, la primera y

segunda etapa determinaron los resultados despejando dónde debía actuar la IR. Queda por delante la tercera y más significativa etapa del proyecto. Para el logro de este objetivo se seleccionarán algunas empresas vinculadas al Polo Tecnológico de UNLaM que tengan la necesidad de implementar un CRM. Se realizará la IR con el proceso de requisitos mencionado [Leite04] y se determinarán los requisitos para el software CRM desde la necesidad analítica de la organización. También se determinará de qué manera el proceso de requisitos aporta un mecanismo de revisión para la estrategia CRM. Por último se espera determinar una guía para identificar el software CRM que mejor se adecua a las necesidades particulares de cada organización.

Formación de Recursos Humanos

Se planifica la finalización de las tesis doctorales de Gabriel E. Blanco, d Claudia Litvak y de Gladys N. Kaplan.

Se espera que Guillermo Hindi, Gabriel Pousada y Miriam Taboada puedan categorizar como investigadores.

Se espera incorporar alumnos becarios en el segundo cuatrimestre de 2017.

Referencias

- [Bose 03] Bose, R., "Customer relationship management: key components for IT success". *Industrial Management & Data Systems*. Vol. 102 pp: 89-97. 2003.
- [Chen 03] Chen J., Popovich K., "Understanding Customer

Relationship Management", *Business Process Management Journal*, Vol. 9 pp. 672-688, 2003

- [Doorn 02] Doorn J., Hadad G., Kaplan G. (2002) *Comprendiendo el Universo de Discurso Futuro*, WER'02 - Workshop on Requirements Engineering, Valencia, Spain.
- [Galbreath 88] Jeremy Galbreath Tom Rogers, *Customer relationship leadership: a leadership and motivation model for the twenty first century business*", MCB UP Ltd, SSN: 0954-478X, 1988
- [Garrido 11] Aurora Garrido Moreno y Antonio Padilla Meléndez, "Estrategias CRM en empresas de servicios: recomendaciones directivas para su implementación", 2011
- [Hadad 09] Hadad GDS, Doorn JH, Kaplan GN (2009) *Explicitar Requisitos del Software usando Escenarios*. 12th Workshop on Requirements Engineering (WER'09), ISBN: 978-956-319-941-3, Chile, pp.63-74.
- [Hashimura 11] Hideki Erigh Hashimura, "Fundamentos para establecer una estrategia de CRM", 2011
- [Law 03]Law M, Lau T & Wong YH., "From customer relationship management to customer-managed relationship: unravelling the paradox with a co-creative perspective". *Marketing Intelligence & Planning*, 21(1): 51-60. 2003
- [Leite 00] Leite J.C.S.P., Hadad G.D.S., Doorn J.H., Kaplan G.N., "A Scenario Construction Process", *Requirements Engineering Journal*, Vol.5, N° 1, 2000, pp. 38-61.
- [Leite 04] Leite J.C.S.P., Doorn J.H., Kaplan G.N., Hadad G.D.S., Ridaio M.N., "Defining System Context using Scenarios", en el libro "Perspectives on Software Requirements", Kluwer Academic Publishers, EEUU, ISBN:

- 1-4020-7625-8, capítulo 8, pp.169-199, 2004.
- [Neill 03] Neill C.J., Laplante P.A., "Requirements Engineering: The State of the Practice", IEEE Software, Noviembre/Diciembre 2003, 20(6):40-45.
- [Swift 01] Ronald S. Swift, Accelerating Customer Relationships: Using CRM and Relationship Technologies, Prentice Hall Professional, 2001

Ingeniería de Software Dirigida por Modelos Aplicada a Sistemas Robóticos Usando los Estándares de la OMG

Claudia Pons, Gabriela Pérez, Roxana Giandini, Carlos Neil, Marcelo de Vincenzi

Comisión de Investigaciones Científicas de la Provincia de Buenos Aries (CIC)
Universidad Abierta Interamericana (UAI)
Facultad de Informática de la UNLP

Buenos Aires, Argentina

Resumen

El Desarrollo de software Dirigido por Modelos (MDD, Model Driven software Development) aparece como una alternativa viable para aplicar técnicas de ingeniería de software en el desarrollo de sistemas robóticos. Su uso logra un nivel de abstracción superior, permitiendo utilizar los estándares propuestos para robótica, y así obtener ventajas como generalidad, reutilización, claridad, expresividad. Estas son todas cualidades inherentes a un proceso de creación de software eficiente y eficaz. El objetivo general de esta investigación es contribuir al mejoramiento de los procesos de desarrollo de software de los sistemas robóticos, a través del análisis del paradigma de desarrollo MDD aplicando los estándares definidos por la OMG.

Palabras clave: sistemas robóticos, ingeniería de software, desarrollo dirigido por modelos, estándares de la OMG.

1. CONTEXTO

Los sistemas robóticos (RSS Robotic Software Systems) desempeñan un papel cada vez más importante en nuestra vida cotidiana. La

necesidad de sistemas robóticos en todos los entornos (industriales, educativos) aumenta y sus requisitos se vuelven más exigentes. Están hechos de diferentes componentes y sensores, lo que resulta en una arquitectura muy compleja y altamente variable. Actualmente, la mayoría de los sistemas de software de robótica se basan todavía en software propietario y están estrechamente ligados al hardware específico, las plataformas de procesamiento o la infraestructura de comunicación. En consecuencia, estos robots solo pueden ser armados, configurados y programados por expertos. Los enfoques tradicionales utilizados en el proceso de desarrollo de este tipo de sistemas están basados principalmente en codificar las aplicaciones sin ningún tipo de técnica de modelado. Y, aunque estas aplicaciones se utilizan en diferentes sistemas robóticos, se pueden identificar algunos problemas. Entre ellos, vale la pena mencionar que no hay documentación clara sobre las decisiones de diseño que se toman durante la fase de codificación, por lo que se dificultan tanto la evolución como el mantenimiento de estos sistemas. Además, al utilizar lenguajes de programación específicos perdemos la posibilidad de generalizar conceptos que pueden ser extraídos, reutilizados y aplicados

en otros sistemas, lo que nos permitiría evitarnos recodificar todo desde cero cuando se lo necesite.

Por otro lado, a medida que los sistemas robóticos crecen para ser cada vez más complejos, la necesidad de aplicar los principios de ingeniería de software para su proceso de desarrollo se convierte en un reto obligatorio en estos días.

Desde esta perspectiva, se acepta el hecho de establecer nuevos enfoques para satisfacer las necesidades del proceso de desarrollo de sistemas robóticos tan complejos como los de hoy. El desarrollo basado en componentes (component-based development o CBD), la arquitectura orientada a servicios (Service-oriented architecture o SOA) [5], así como la Ingeniería de software dirigida por modelos (MDE), y el modelado específico de dominio (DSM) son algunas de las tecnologías más prometedoras en el dominio de los sistemas robóticos.

Actualmente se promueve la integración de los componentes de los sistemas robóticos a través de la adopción de estándares de la OMG [1, 11]. Se han lanzado cuatro especificaciones: para la interacción en los sistemas robóticos (Robotic Interaction Service – ROIs) [3], para la localización de los sistemas robóticos (Robotic Localization Service - RLS), para el modelado de los componentes (Robotic Technology Component - RTC) [2] y para el despliegue dinámico y su configuración (Dynamic Deployment and Configuration for Robotic Technology Component -DDC4RTC).

Es en este contexto en el que el Desarrollo Dirigido por Modelos (MDD, Model Driven Development)[7, 8, 9] aparece como una alternativa viable para aplicar técnicas de ingeniería de software en el desarrollo de este tipo de sistemas. Su uso logra un nivel de abstracción superior, permitiendo utilizar los estándares propuestos para robótica, y así obtener ventajas como generalidad,

reutilización, claridad, expresividad. Estas son todas cualidades inherentes a un proceso de creación de software eficiente y eficaz.

2. OBJETIVOS, HIPÓTESIS Y RESULTADOS ESPERADOS

El objetivo general de esta investigación es contribuir al mejoramiento de los procesos de desarrollo de software de los sistemas robóticos, a través del análisis del paradigma de desarrollo MDD aplicando los estándares definidos por la OMG para el modelado de estos sistemas y complementado con componentes (CBD) y la arquitectura orientada a servicios (SOA).

3. LINEAS DE INVESTIGACIÓN

En este proyecto convergen dos líneas de investigación: Robótica y MDE.

La problemática descrita arriba ha llevado a la concepción de un conjunto de actividades con el objetivo de acercar a los investigadores de estas aéreas: Ingeniería de Software dirigida por Modelos por un lado, y Robótica por el otro. Ejemplos de estas actividades son el Workshop on Domain- Specific Languages and Models for Robotic Systems (DSLRob) lanzado en 2009, y el Workshop on Model-Driven Robot Software Engineering(MORSE) iniciado en 2013, ambos con el objetivo de incentivar la interacción de estas aéreas. Como resultado, en los últimos años varias plataformas han sido desarrolladas para proveer formas simples e intuitivas de construir las aplicaciones de software robótico. Esto incluye tanto prototipos académicos como así también productos comerciales. Paralelamente el organismo OMG que se ocupa de coordinar los estándares de software internacionalmente ha creado el

grupo Robotic Domain task Force, para ocuparse de estos temas.

En el ámbito industrial/comercial las plataformas para desarrollo robótico más conocidas son: Lego Mindstorms Evolution 3, Choregraphe, Robotino View 2 y Microsoft Robotics Developer Studio 4 (MRDS4).

En el ámbito académico hemos encontrado muchas propuestas teóricas que aplican el paradigma CBD para el desarrollo de sistemas robóticos. Otras propuestas han aplicado SOA para construir estos sistemas. También el paradigma MDD/MDA/DSL ha sido aplicado de manera teórica a la construcción de robots (ver nuestro survey en [21]).

Todos estos trabajos proveen un punto de partida robusto para el desarrollo de nuestro proyecto.

4. PROPUESTA

Este trabajo tiene como objetivo general estudiar los sistemas robóticos y cómo aplicar técnicas de ingeniería de software para desarrollarlos. Este objetivo está siendo abordado a través de los siguientes sub-objetivos:

- _ Estudiar las propuestas de la OMG para el modelado de sistemas robóticos, y analizar como pueden ser aplicadas en el desarrollo de los sistemas robóticos.
- _ Analizar las interrelaciones entre los distintos estándares.
- _ Analizar cómo pueden ser integrados los paradigmas SOA y CBD con lo propuesto por la OMG
- _ Estudiar como los modelos realizados en dichos estándares pueden ser transformados para crear otros modelos, y eventualmente, crear código.

5. METODOLOGÍA DE INVESTIGACIÓN Y DESARROLLO PROPUESTA

A partir del estudio de los sistemas robóticos actuales, se observa que son artesanales, y que su desarrollo esta basado principalmente en el código. A pesar de que existen propuestas para el modelado de los sistemas robóticos, aun no se aplican y tampoco cubren la posibilidad de la integración de dichas propuestas.

Nuestra hipótesis de trabajo consiste en la posibilidad de integrar los estándares con otras técnicas de ingeniería de software para poder aplicar MDD en el desarrollo de los sistemas robóticos.

Respecto a los lenguajes de modelado y transformación nos basaremos en los estándares de la OMG, en particular MOF [6], UML, OCL y QVT[10].

Finalmente, todos nuestros resultados teóricos serán plasmados en herramientas de desarrollo de software de código abierto, preferentemente sobre la plataforma Eclipse [15].

6. ESQUEMA DE PLAN DE TRABAJO C/ACTIVIDADES

El proyecto viene desarrollándose a través de las siguientes actividades:

- _ Estudio de los lenguajes estándares para modelado de sistemas robóticos
- _ Estudio del estándar como base para el modelado de software, en particular MOF (Meta Object Facilities), RTC (Robotic Technology Component), ROIs (Robotic Interaction Service), DDC4RTC (Dynamic Deployment and Configuration for Robotic Technology Component), RLS (Robotic Localization Service).
- _ Estudio del estándar para modelado de servicios SOA.

- _ Analizar las relaciones entre dichos estándares y como se complementan para ser utilizados en el proceso de desarrollo.
- _ Analizar mecanismos de transformación entre modelos
- _ Plasmar lo analizado en herramientas de desarrollo de software de código abierto
- _ Evaluación de la propuesta mediante la aplicación de la misma en proyectos reales.

7. FORMACIÓN DE RECURSOS HUMANOS Y RESULTADOS

El equipo está integrado por cinco investigadores senior, provenientes de 2 universidades (UNLP y UAI), quienes combinan sus conocimientos en robótica e ingeniería de software. El resto del equipo está integrado por estudiantes de pregrado y postgrado. En el contexto de este proyecto se están desarrollando 5 tesis de licenciatura en Informática, 5 tesis de Maestría en TI y 5 tesis doctorales (1 financiada por CONICET y 1 por el programa DoctorAR). Se trabaja además en colaboración con la Universidad de Viena a través de un convenio bi-lateral financiado por el Mincyt [16]. Los resultados preliminares del proyecto se han publicado en [17], [18], [19], [20], [21] y [22].

8. REFERENCIAS

1. OMG Robotics-DTF - Object Management Group - <http://robotics.omg.org/> (consultado en 2015)
2. Documento de especificación RTC 1.0 - Object Management Group - <http://www.omg.org/spec/RTC/1.0/> (consultado en 2015)
3. Documento de especificación Robotic Interaction Service (RoIS) – OMG - <http://www.omg.org/spec/RoIS/> (consultado en 2015)
4. Robotic Localization Service (RLS) – OMG - <http://www.omg.org/spec/RLS/>
5. Documento de especificación SOA – OMG - <http://www.omg.org/technology/readingroom/SOA.htm> (consultado en 2015).
6. Documento de especificación OMG's MetaObject Facility (MOF) Home Page - <http://www.omg.org/mof/>
7. Stahl, M Voelter. Model Driven Software Development. John Wiley, ISBN 0470025700.
8. Kleppe, Anneke G. and Warmer Jos, and Bast, Wim. MDA Explained: The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
9. Object Management Group, MDA Guide, v1.0.1, omg/03-06-01(2003).
10. MOF 2.0 Query/View/Transformations - OMG Adopted Specification. March 2005. <http://www.omg.org>.
11. Object Management Group (OMG) <http://www.omg.org>
12. Ledeczki, A., Bakay, A., Maroti, M., Volgyesi, P., Nordstrom, G., Sprinkle, J., Karsai, G. Composing Domain-Specific Design Environments. IEEE Computer 34 (2001)
13. Kleppe, Anneke. MCC: A Model Transformation Environment. A. Rensink and J. Warmer (Eds.): ECMDA-FA 2006, LNCS 4066, pp. 173 – 187, Spain, June 2006.
14. Atlas Model Weaver Project Web Page. <http://www.eclipse.org/gmt/amw/>, 2005.
15. Proyecto Eclipse - <https://eclipse.org/> (consultado en 2015).
16. Project Title: “Adoxx Meta-Model Compiler. Modelling Methods for Robotic Systems”. Programa de

- Cooperación Científico-Tecnológica entre el Ministerio de Ciencia, Tecnología e Innovación Productiva de la República Argentina (MINCyT) y el Ministerio de Ciencia e Investigación de la República de Austria (BMWF).
17. C. Pons, G. Pérez, R. Giandini, G. Baum . A Model-Driven Approach to Constructing Robotic Systems. *Journal of Computer Science & Technology*. Vol. 14 - No. 1 – April 2014 - ISSN 1666-6038.
 18. Jerónimo Irazábal and Claudia Pons. Metamodel independence in Domain Specific Modeling Languages. "Communications in Computer and Information Science CCIS Series" vol.411, "Software and Data Technology" pp.140-154, ISBN 978-3-642-45403-5. Springer-Verlag Berlin Heidelberg (2013).
 19. Omar Martinez Grassi, Claudia Pons, Gabriel Baum. Variable-Based Analysis for Traceability in QVT-R Model Transformations. *CibSE - SET 2015 (CibSE -Software Engineering Track)*. XVIII Conferencia Iberoamericana en "Software Engineering", CibSE 2015. Perú. April 2015.
 20. Gabriela Pérez, Jerónimo Irazábal, Claudia Pons y Roxana Giandini. Applying MDE tools to defining domain specific languages for model management. *SADIO Electronic Journal of Informatics and Operations Research*. ISSN 1514- 6774 vol. 12, no. 1 (Sept2013).
 21. Claudia Pons, Roxana Giandini, Gabriela Arévalo . A systematic review of applying modern software engineering techniques to developing robotic systems.. Vol. 32 No. 1 de 2012. *Revista Ingeniería e Investigación*. Tri-annual ISSN: 0120- 5609. Indexada en el ISI.
 22. Claudia Pons, Gabriela Pérez, Roxana Giandini and Gabriel Baum: "Applying MDA and OMG Robotic Specification for Developing Robotic Systems".Published in *Lecture Notes in Computer Science*. Springer. Volume 9959 2016. *System Analysis and Modeling. Technology-Specific Aspects of Models*. Saint-Malo, France, October 3-4, 2016. Editors:Jens Grabowski, Steffen Herbold ISBN: 978-3-319-46612-5 (Print) 978-3-319-46613-2 (Online)

Ingeniería de Software para Sistemas Distribuidos

Patricia Pesado, Rodolfo Bertone, Pablo Thomas, Luciano Marrero, Ariel Pasini, Lisandro Delia, Nicolás Galdamez, Eduardo Ibañez, Cesar Estrebou, Alejandra Ripodas, Verónica Aguirre, Rocío Muñoz, Germán Cáseres, Santiago Medina, Matías Dell’Oso, Juan Manuel Paniego, Martín Pi Puig, Sebastián Rodríguez Eguren

Instituto de Investigación en Informática LIDI (III-LIDI)
50 y 120 La Plata Buenos Aires Facultad de Informática – Universidad Nacional de La Plata

Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

{ppesado, pbertone, pthomas, lmarrero, apasini, ldelia, ngaldamez, eibanez, cesarest, aripodas, vaguirre, rmunoz, gcaseres, smedina, mdelloso, jmpaniego, [mpipuiug_seguren](mailto:mpipuiug_seguren@lidi.info.unlp.edu.ar)}@lidi.info.unlp.edu.ar

Resumen

Esta línea de investigación y desarrollo del III-LIDI está centrada en los sistemas de software distribuidos analizando metodologías, técnicas y herramientas de Ingeniería de Software orientadas a la resolución de diferentes clases de sistemas, en particular sistemas de E-government.

Esta línea ha significado la transferencia de conocimiento a diferentes sectores, tanto públicos como privados o mixtos.

Palabras claves: Sistemas Distribuidos – Ingeniería de Software - Metodologías de Desarrollo – E- Government

Contexto

La línea de Investigación que se presenta en este trabajo está enmarcada en el Proyecto F016-Tecnologías para Sistemas de Software Distribuidos. Calidad en Sistemas y Procesos. Escenarios educativos mediados por TICs (2014-2017), del Instituto de Investigación en Informática LIDI de la Facultad de Informática, acreditado por el Ministerio de Educación. En particular en el subproyecto “Métodos y Procesos para la gestión de Sistemas de Software Distribuidos. Aplicaciones”.

Asimismo el III-LIDI participa en el proyecto PIO (CONICET-UNLP) “Construcción de un Sistema Integrado de Gestión del Riesgo

Hídrico en la Región del Gran La Plata” y proyectos de la Facultad de Informática y posee diversos acuerdos de cooperación con varias Universidades de Argentina y del exterior y con organismos públicos y empresas privadas.

Además, participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Introducción

Hoy en día los sistemas se desarrollan en diferentes ámbitos, una operación puede iniciarse desde un dispositivo y finalizarse desde otro. Por ejemplo, se puede iniciar un mail desde un celular y finalizarlo desde una computadora. O iniciar la reproducción de una película desde un Smart-TV y continuar viéndola, en otro momento, desde una Tablet. También avanza la concepción de los sistemas ubicuos, incorporando la posición geográfica a los dispositivos, permitiendo procesar la información en función de la ubicación del dispositivo. Por ejemplo buscar restaurantes cercanos desde mi ubicación.

Todas estas actividades pueden realizarse porque existen equipos capaces de procesar información y un medio por el cual estos dispositivos pueden conectarse. Estos sistemas se encuentran contenidos en la definición que proponen Tanenbaum y Van Steen [28]. para un sistema distribuido:

“Una colección de computadoras independientes que aparecen al usuario como un solo sistema coherente”.

La ingeniería de los sistemas distribuidos además de los conceptos de los sistemas tradicionales debe tener en cuenta los conceptos de:

Transparencia: Proporcionar al usuario y a las aplicaciones una visión de los recursos del sistema como gestionados por una sola máquina virtual. La distribución física de los recursos es transparente Pueden describirse diferentes aspectos de la transparencia: identificación, ubicación, replicación, paralelismo entre otras.

Escalabilidad: capacidad del sistema para crecer sin aumentar su complejidad ni disminuir su rendimiento.

Fiabilidad y Tolerancia a fallos: su capacidad para realizar correctamente y en todo momento las funciones para las que se ha diseñado.

Consistencia: La necesidad de mantener un estado global consistente en un sistema con varios componentes, cada uno de los cuales posee su propio estado local.

Una de las temáticas que importa particularmente en esta línea de investigación es E-Government.

El gobierno electrónico consiste en el uso de las tecnologías de la información y el conocimiento en los procesos internos de gobierno en la entrega de los productos y servicios del Estado tanto a los ciudadanos como a la industria. Debe centrarse en la inclusión de los ciudadanos de una manera participativa a través de las tecnologías de la información y la comunicación [15] [16] [17]. Al concepto clásico de E-Government que se ha enfocado en poner los servicios tradicionales del Estado al alcance del ciudadano (Consultas en línea, Gestión de trámites, Expedientes digitales, Voto Electrónico, Consultas populares, etc) [18][19][20][26] se agrega la concepción de E-Citizen, es decir un ciudadano capacitado para interactuar con el Estado, empleando Tecnología. [21][22][23][24][25][27].

Desde el año 2003 el Instituto trabaja en aplicaciones en esta área, entre las cuales se destacan prototipos de hardware y software de distintos tipos de votaciones (urnas electrónicas, ambientes de votación, comunicaciones, entre otras)[11][12][13][14].

Un sistema de votación electrónica utiliza las TICs (Tecnologías de la Información y Comunicación) para llevar a cabo el proceso de emisión de votos, haciendo el recuento de votos simple y rápido.

Existen numerosas implementaciones que apoyan el proceso de votación, algunos de ellos destinado sólo a ciertas fases del proceso, mientras que otros a resolver todas las actividades que se llevan a cabo durante una elección. Estos últimos se denominan sistemas de voto electrónico.

Los requisitos básicos de un sistema de votación electrónica son:

- Sólo los que están en el registro electoral debe ser capaz de votar;
- Ningún votante debe ser capaz de tomar el lugar de otro, o votar más de una vez;
- El voto debe ser secreto, la persona debe emitir su voto y éste no debe ser identificable;
- Los votos en blanco deben ser posibles.

Una vez que el proceso electoral está habilitado por el consejo de autoridades de la Junta Electoral, el proceso de votación incluye tres etapas principales:

1. identificación del votante,
2. registro del voto,
3. una vez finalizado el acto electoral, realizar es escrutinio.

El software desarrollado para el voto electrónico debe ser auditable, lo que significa que las características de código abierto en un sistema operativo de software libre con la firma digital serían útiles para la fiabilidad.

Se han analizado y caracterizado modelos y tecnologías posibles, realizando varias experiencias prácticas. En particular se han analizado tres variantes: Voto electrónico

presencial (VEP), Voto electrónico semipresencial (VESP) y Voto remoto (VER)

VEP: Una de las soluciones más integrales para la registración del sufragio, es la urna electrónica, un dispositivo que permite al elector visualizar las opciones de voto a través de una pantalla táctil, verificar su elección con el voto impreso que se muestra en una ventana y confirmar el mismo para actualizar los contadores de votación y deslizar automáticamente el voto en una urna alojada en el “cajero/kiosco”. Estos sistemas son llamados Sistemas de Registro Electrónico Directo (DRE). El VEP se organiza en puestos de votación distribuidos, en los que existe un padrón de votantes en la mesa de autoridades donde se presenta físicamente el elector, conectada a la urna electrónica. Finalizada la elección se realizará el recuento automático de los votos de la urna.

VESP: Combina las características del voto presencial en cuanto a los puestos de votación donde se identifican los electores y efectúan su opción de voto; con la transmisión remota del voto realizado a una urna distante físicamente, donde se guarda el respaldo papel y se registran los resultados y sobre la que se realizará el recuento de los votos recibidos una vez finalizada la elección.

VER: Consiste en una aplicación WEB, que puede ser utilizada por el elector a través de un navegador web desde un dispositivo conectado a internet en cualquier lugar del mundo. El voto del elector es transmitido por una conexión segura a un servidor donde se almacenan los resultados. Una vez finalizada la elección, las autoridades electorales realizan el recuento a través de la aplicación web.

Líneas de Investigación y Desarrollo

Se detallan a continuación las principales líneas de investigación y desarrollo que representan una continuidad respecto de años anteriores:

- Metodologías de especificación, validación y desarrollo de SSD.
- Metodologías ágiles de desarrollo utilizando frameworks propios y disponibles de uso libre, con diferentes tecnologías.
- Evaluación de técnicas para mejorar el proceso de Ingeniería de Requerimientos.
- Lenguajes y ambientes para procesamiento distribuido.
- Reingeniería de sistemas complejos que migran por downsizing a esquemas cliente-servidor distribuidos.
- Sistemas basados en tecnologías Cloud. Enfoque a las aplicaciones de E-Citizen.
- Sistemas distribuidos para el tratamiento de información no estructurada (Big- Data). Aplicaciones.
- Voto electrónico presencial y Voto por Internet. Arquitecturas adaptadas a la legislación vigente. Inclusión de electores.
- Sistemas de identificación segura en tiempo real. Identificación biométrica.
- Seguridad y tolerancia a fallas en sistemas de E-Government.
- Certificación de software y hardware para sistemas de E-Gov. Normas.
- Nuevas tecnologías de E-Gov, en particular M-Gov (Mobile Government)
- Conceptos de procesamiento distribuido. Arquitectura, comunicaciones y software. Middleware.

Resultados esperados/obtenidos

Algunas transferencias realizadas y en curso:

- **Voto Electrónico Presencial.** Desde 2003, se utilizan adaptaciones de la urna electrónica en las elecciones de alumnos de distintas universidades (los mismos pueden votar a “Claustro de estudiantes y Centro de estudiantes” o “Sólo Centro” de acuerdo a su condición), incluyendo alumnos no videntes. También en elecciones de graduados, de profesores, de no docentes, en congresos para votación de las mejores presentaciones o en eventos para la elección de las mejores propuestas. En 2017 se utilizará una nueva arquitectura de Urna Electrónica con utilización de microcontroladores del tipo Raspberry Pi 2.
- **Voto Electrónico Semipresencial.** Desde 2007 se utiliza voto semi-presencial en elecciones estudiantiles de la UNLP en sedes regionales.
- **Voto Electrónico Remoto:** Desde 2007 se utiliza voto remoto en elecciones de distintas organizaciones públicas y privadas (unidades de investigación, consejos y cajas profesionales), para diferentes conformaciones de autoridades. En particular durante el año 2016 se utilizó el sistema en la Caja de Farmacéuticos de La Plata.
- **Sistema de Alerta temprana.** El III-LIDI desarrolló un Sistema WEB de Alerta Temprana para la prevención de inundaciones, en el marco del proyecto PIO CONICET-UNLP que muestra los datos recibidos de un conjunto de pluviómetros. El objetivo general del proyecto es contar con un sistema de alerta temprano que permitirá monitorear no sólo las precipitaciones sino el crecimiento de cada cuenca hidrológica de la región. El proyecto busca comunicar estaciones de censado pluviométrico con una base de datos central y que la información producida sea visible en la web desarrollada a tal efecto.
- **Sistema de Gestión Integral para el CONICET La Plata.** Este sistema, que abarca las áreas de Recursos Humanos; Comercio Exterior, Compras y Patrimonio; Administración Contable; Tesorería. Además dispone de funcionalidad utilizable en forma distribuida por todos los Centros de Investigación dependientes del CCT La Plata.
- **Plataforma de Educación a Distancia.** El sistema WebUNLP fue desarrollado al comienzo del milenio por el III-LIDI. Desde ese momento fue utilizado por numerosas cátedras de la Facultad de Informática, la Universidad Nacional de la Plata, y otras universidades del país. El proyecto contó con mantenimiento preventivo y correctivo, pero el transcurso del tiempo demostró que era necesario un mantenimiento perfecto que incorporara una serie de nuevas funcionalidades al entorno. Se ha desarrollado una nueva plataforma denominada Ideas, como evolución de WebUNLP.
- **Sistema de Monitoreo de Plagas y Enfermedades Frutihortícolas.** En colaboración con el INTA y SENASA, se está desarrollando un sistema para realizar el monitoreo de plagas y enfermedades en plantaciones de frutas y hortalizas.

Metas/Resultados esperados (en continuidad con el proyecto que se viene realizando) :

- Desarrollar soluciones a problemas concretos de software de sistemas distribuidos, poniendo énfasis en el desarrollo de metodologías y herramientas específicas para clases de aplicaciones.
- Analizar metodologías aplicables a Sistemas Distribuidos utilizando frameworks de desarrollo, específicos para diferentes entornos.
- Ampliar herramientas que soporten la evaluación de calidad, eficiencia y relación de esfuerzo/costo. Esta meta

relacionada con el subproyecto de gestión de la calidad del Instituto.

- Continuar evolucionando los prototipos de voto electrónico presencial y remoto utilizables en diferentes modelos de elecciones/consultas.
- Transferir al sector productivo nacional.
- Formar recursos humanos de grado y postgrado en la temática.
- Conformar vínculos y acuerdos de cooperación con otras universidades del país y del exterior.
- Publicar los resultados y participar en eventos científicos/tecnológicos para compartirlos.

Formación de Recursos Humanos

Los integrantes de esta línea de investigación dirigen Tesinas de Licenciatura y Tesis de Postgrado en la Facultad de Informática, y Becarios del Instituto en temas relacionados con el proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP y en otras universidades del país y del exterior.

Referencias

- [1] G. Coulouris. Distributed Systems – Concepts and Design. Addison-Wesley. 1994.
- [2] R. Pressman. Ingeniería de Software: Un Enfoque Práctico. McGraw-Hill. 2002
- [3] Pleeger. Ingeniería de Software: Teoría y Práctica. Prentice-Hall. 2002
- [4] Stephen Kan. Metrics and Models in Software Quality Engineering (2nd Edition). Addison Wesley. 2003
- [5] Offutt J., “Quality Attributes of Web Software Applications”. IEEE Software: Special Issue on Software Engineering of Internet Software 19 (2):25-32, Marzo / Abril 2002.
- [6] Wu, Y. y Offutt, J. “Modeling and testing web-based Applications”.
<https://citeseer.ist.psu.edu/551504.html>: 1-12, Julio 2004
- [7] Piattini, M; Oktaba, H; Pino, F; Orozco, M; Alquicira, C. COMPETISOFT. Mejora de Procesos

Software para Pequeñas y Medianas Empresas y Proyectos. Editorial RaMa. ISBN; 978-84-7897-901-1. 2008

[8] Ingeniería de Software, Ian Sommerville, Pearson, Addison Wesley, 2006

[9] Software Engineering Institute. Disponible en <http://www.sei.cmu.edu/risk/>

[10] A spiral model of software development and enhancement. B. Boehm, IEEE Computer, 21(5), 61- 72. 1988

[11] Feierherd G., De Giusti A., Pesado P., Depetris B. “Una aproximación a los requerimientos del software de voto electrónico de Argentina”. CACIC 2004.

[12] Pesado P., Feierherd G., Pasini A. “Especificación de Requerimientos para Sistemas de Voto Electrónico”. CACIC 2005.

[13] Pesado P., Pasini A., Ibáñez E., Galdámez N., Chichizola F., Rodríguez I., Estrebou C., De Giusti A. “E-Government- El voto electrónico sobre Internet”. CACIC 2008.

[14] Carri J., Pasini A., Pesado P., De Giusti A. “Reconocimiento biométrico en aplicaciones de E-Government. Análisis de confiabilidad / tiempo de respuesta.” CACIC 2007.

[15] Center for democracy and technology E-Government Handbook. 2002.
<http://www.cdt.org/egov/handbook/>

[16] Jones A., Williams L. “Public Services and ICT - FINAL REPORT. How can ICT help improve quality, choice and efficiency in public services?”. London: The Work Foundation. 2005.

[17] London: National Audit Office. “Better Public Services through e-government”. Report HC 704-I Session 2001-2002.

[18] Washington DC: Office of Management and Budget. “E-Government Strategy: Simplified Delivery of Services to Citizens”. OMB. 2002.

[19] “Citizen Centric Government: Global Best Practice in Delivering Agile Public Services to Citizens and Businesses”. London: Gov3 Ltd. 2006.

[20] Brussels: European Commission. “The Role of eGovernment for Europe's Future”. Report COM(2003) 567 Final. Europe. 2003.

[21] Wang L., Bretschneider S., Gant J. “Evaluating Web-based e-government services with a citizen-centric approach”. Proceedings of the 38th Hawaii International Conference on System Sciences. 2005.

[22] Transforming Public Services: The Next Phase of Reform. Edinburgh: Scottish Executive”. Scotland 2006.

[23] Clift S. “e-Government and Democracy: Representation and Citizen Engagement in the Information Age”. 2004.

[24] O'Donnell S., McQuillan H., Malina A. “eInclusion: expanding the Information Society in Ireland. Dublin: Government of Ireland. Information Society Commission”. 2003.

[25] Juma C., Yee-Cheong L. “Reinventing global health: the role of science, technology and innovation”. Lancet 2005.

[26] Laguna A., Ferri Tormo R., Hernandez V., Peñarrubia J. “gCitizen: uso de tecnologías Grid para la interoperabilidad entre Administraciones Públicas”. IX Jornadas sobre Tecnología de la Información para la Modernización de las Administraciones Públicas. Sevilla. 2006.

[27] Brunner J. “Educación: escenarios de futuro. Nuevas Tecnologías y sociedad de la información”. PREAL, Santiago de Chile. 2000.

[28] Andrew S. Tanenbaum, Maarten van Steen “Distributed Systems: Principles and Paradigms”, Second Edition, 2016. Pearson – Prentice Hall. ISBN 0-13-239227-5

Integración de Arquitectura de Software en el Ciclo de Vida de las Metodologías Ágiles. Una Perspectiva Basada en Requisitos

Mg. Mirta E. Navarro¹, Mg. Marcelo P. Moreno², Lic. Juan Aranda³, Lic. Lorena Parra⁴,
Lic. Jose R. Rueda⁵, Juan Cruz Pantano⁶

Departamento de Informática - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan

¹mirtaenavarro@yahoo.com.ar ²mpmoren@gmail.com ³juanaranda@live.com ⁴lorenaparra152@yahoo.com.ar
⁵josericardorueda@hotmail.com ⁶juancruz871@hotmail.com

Resumen

Las metodologías Ágiles se centran en el trabajo en equipo, la adaptabilidad y colaboración dentro del grupo de software y también entre los miembros del grupo y los usuarios finales. El uso de las Metodologías Ágiles (MA), ha marcado una tendencia [1] en su adopción al desarrollo de proyectos de software dado las necesidades cambiantes y la espera de beneficios en el menor tiempo posible. En general, pero también desde la perspectiva de los requisitos, esto hace que las MA, típicamente eviten un trabajo inicial sustancial, suponiendo que los requisitos siempre cambian y continúan cambiando a lo largo del ciclo de vida del proyecto.

La Arquitectura de Software (AR), en tanto, es una manifestación de decisiones de etapas muy tempranas del diseño sobre un sistema [2]. Estas decisiones tempranas llevan un peso importante con respecto al desarrollo del resto de un sistema, ya que condicionan otras decisiones que siguen, y en el caso de cambios, implican ramificaciones posteriores. Esto supone una captura de requisitos que no tengan cambios sustanciales en las etapas intermedias y finales del desarrollo de un proyecto.

Este tratamiento con enfoques diferentes en las primeras etapas (y también en otros aspectos), ha sido uno de los factores que ha causado la sensación de que las MA y la AS van en direcciones diferentes y no pueden coexistir juntas [3].

Sin embargo, en los últimos cinco años esta tendencia esta cambiando [4,5], hasta el punto

que ha surgido el concepto “*Arquitectura Ágil*” (AA). En la AA, se enfatiza fuertemente en el concepto de los “*Requisitos Significantes para la Arquitectura*” (RSA) o “*Architecturally Significant Requirement*” ASR, por sus siglas en inglés. En ese sentido, esta línea de investigación tiene como objetivo indagar y validar el alcance de hasta dónde pueden confluir la AS con las MA, poniendo foco en la captura de requisitos y en los procesos de identificación de los RSA.

Palabras clave: *Software Architecture, Agile methodologies, Information Systems Architecturally Significant Requirements (ASR)*

Contexto

El presente trabajo se encuadra dentro del área de I/D de la IS y de los SI, y forma parte de una etapa del proyecto de investigación: “*Integración de Metodologías Ágiles y Arquitecturas de software en el Desarrollo de Sistemas de Información*”, presentado en WICC 2016 [6] ejecutado a partir de Enero de 2016, con una duración de dos años y que tiene como unidades ejecutoras al Departamento e Instituto de Informática de la FCEFYN de la UNSJ. El proyecto se encuentra acreditado y financiado por la Secretaria de Ciencia y Técnica UNSJ.

El grupo de investigación tiene una trayectoria de 16 años en diferentes proyectos vinculados a Metodologías de Desarrollo y Tecnologías, con numerosas publicaciones en

diferentes ámbitos, y con la formación de recursos humanos en el área de interés.

Introducción

El lugar que ocupa la AS en el ciclo de vida de desarrollo de software es difícil de definir. Pero, puesto que las arquitecturas describen el espacio de solución de un sistema, tradicionalmente se la ha considerado como una parte temprana de la fase de diseño. Si bien la AS puede ser vista desde diferentes focos contextuales y de diferentes niveles de abstracción, el interés de este trabajo está centrado en los aspectos relacionados en los requisitos. Así, desde esa mirada, la AS, debe garantizar que la arquitectura sea realizable mediante una implementación adecuada que satisfaga los requisitos del sistema. En las MA, es justamente allí, en la consideración de los requisitos, donde los aspectos arquitectónicos se suelen pasar por alto. En ese sentido, normalmente, se identifican y captan los requisitos, tanto funcionales como no funcionales, y luego se intenta definir una arquitectura de software que los cumpla.

Sin embargo, esta es una tarea compleja en particular con la identificación de los requisitos no funcionales y con los requisitos mal elicitados (que no consideran aspectos AS) y puede derivar en una realización arquitectónica inadecuada. Además, se debe considerar que algunos requisitos tienen un efecto mucho más profundo en la arquitectura que otros. En este escenario, es donde cobra fuerza un concepto reciente que va de la mano con la Arquitectura Ágil, los RSA, que son requisitos que tendrán un efecto profundo en la arquitectura, y que como postula [4], si llegasen a faltar, la Arquitectura final resultante será muy diferente.

Los Requisitos de Importancia Arquitectónica

Los requisitos de software se dividen principalmente en requisitos funcionales y

requisitos no funcionales. Los requisitos funcionales corresponden a las características deseadas de un sistema; en tanto que no funcionales especifican las propiedades requeridas de un sistema. Los requisitos no funcionales pueden estar caracterizados por diferentes factores tales como pueden ser los atributos de calidad, de seguridad, de disponibilidad, restricciones del sistema, metas, características de usabilidad, etc.

Un Requisito de Importancia Arquitectónica es un requisito que tendrá un importante efecto en la arquitectura, y que como se dijo, si están ausentes, la arquitectura resultante será totalmente diferente. Tomando la definición de [7] los RSA *“son aquellos requisitos que tienen un impacto medible en una arquitectura de sistemas de software”*. Por lo tanto, no sería posible diseñar una arquitectura adecuada, si no se elicitaban adecuadamente los RSA. En general, en la identificación y especificación de los RSA intervienen las principales partes interesadas del sistema, como los usuarios finales, los desarrolladores, los administradores y los mantenedores.

El problema es que los RSA generalmente toman la forma de requisitos no funcionales, pero en algunos casos también toman la forma de requisitos funcionales. Además, los RSA suelen ser subjetivos, relativos y también interactúan entre ellos. Son subjetivos, porque pueden ser vistos, interpretados y analizados de manera diferente por diferentes personas y en diferentes contextos; son relativos, porque la importancia de cada RSA se determina a veces a partir de su relación con otros RSA en un contexto dado; y se considera que interactúan entre ellos en el sentido de que al intentar alcanzar un RSA particular, puede a su vez afectar (en forma favorable o desfavorable) a otros RSA. Los RSA son más difíciles de comprender que los requisitos funcionales [7], por lo que en general pasan desapercibidos o no obtienen suficiente atención por adelantado, factor que es más acentuado cuando se utilizan MA, donde generalmente se expresan (muchas veces en forma contradictoria) de manera

informal durante el análisis de requisitos. También son difíciles de validar cuando el proyecto está finalizado. Por ello, la identificación de los RSA, es una tarea difícil, que lleva mucho trabajo y es poco clara para los que no son expertos en arquitecturas.

Para abordar esta situación, algunos autores [4,5 y 7] han propuesto una serie de métodos y estrategias de identificación y captura de RSA a partir de documentos de requisitos tradicionales de una MA, los que serán utilizados por este grupo para analizarlos y evaluarlos.

Estrategias de captura de RSA.

Las formas de obtener RSA a partir de los documentos de requisitos, son varias. Chen [7] propone un marco para caracterizar los RSA, sobre la base de un estudio empírico que realizó con expertos. Sus hallazgos pueden mejorar el entendimiento de los requisitos y las interacciones con la arquitectura. En general, las estrategias que propone son:

- Categorizar las decisiones de diseño de arquitectura que los arquitectos de software tienen que tomar. Estas decisiones toman la forma de tablas resumidas. A partir de allí proporcionan una lista de requisitos para buscar cual de ellos podría afectar a ese tipo de decisión. Si un requisito afecta a alguna decisión de diseño, es por definición un RSA.
- Entrevistas a interesados.
- Comprensión de objetivos de negocios.

En tanto que [4 y 5] proponen utilizar enfoques tales como escenarios para caracterizar diferentes atributos, en particular los de calidad, y también plantean utilizar aproximaciones basadas en frameworks de evidencias.

Finalmente debe quedar claro que las consideraciones sobre requisitos RSA, no alcanzan a especificar la totalidad sobre las arquitecturas, y que un trabajo adicional deberá ser realizado para la adecuada especificación de las AS.

Tópico de investigación, Desarrollo e Innovación

Como se ha manifestado en [6], los posibles beneficios de una integración de Arquitectura de Software en las Metodologías Ágiles no es un tema lo suficientemente explorado. Sin embargo [4 y 5], han propuesto últimamente algunas publicaciones relacionadas en lo que se denomina Arquitectura Ágil, donde exploran con detenimiento los RSA.

Esta línea de investigación toma como punto de partida los hallazgos de [4,5 y 7] que son la base para discutir y llevar a cabo más investigaciones sobre RSA. Pensamos que esos métodos y enfoques pueden ser complementarios entre sí, y a la vez podrían ser combinados con los hallazgos que han sido ampliamente investigados por los autores de este artículo, en el área de las metodologías ágiles.

Finalmente los aportes que pudieran surgir, serán utilizados como guías para elicitar, analizar implementar y evaluar situaciones reales en proyectos de tesis y tesinas que han surgido a partir de inquietudes emergentes en proyectos de investigación anteriores que este grupo llevó a cabo [8 y 9].

Resultados y Objetivos

Los resultados que se esperan obtener al finalizar la investigación son:

- Analizar las características particulares de la relación de los requisitos tradicionales con los RSA.
- Evaluar las propuestas de [4,5 y 7], utilizando documentos de requisitos tradicionales de las MA proponiendo, eventualmente, alternativas que surjan de la línea de investigación, y aplicar y evaluar esos hallazgos en un contexto real en tesis o tesinas.
- Continuar con la línea trazada en el proyecto [10], cuyo objetivo principal es establecer la incorporación de arquitecturas

al proceso de desarrollo dirigido por metodologías ágiles, con el propósito de que esa integración favorezca el diseño de Sistemas de Información.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por cinco docentes-investigadores, y seis alumnos adscriptos. En el periodo 2015-2016 se ha asesorado: una tesis de grado finalizada y cuatro tesis de grado en proceso, una de las cuales se está desarrollando en el nuevo tópico de investigación que se presenta en este trabajo.

Además, se espera realizar direcciones, a los alumnos adscriptos, en diversos trabajos orientados al desarrollo de sistemas con metodologías ICONIX-SCRUM, adoptando el enfoque de integración con arquitecturas desde la perspectiva de los RSA.

Con los resultados de la presente investigación, se harán actividades de divulgación en publicaciones y presentaciones en eventos nacionales e internacionales y también en cursos de postgrado y actualización.

Referencias

- [1] Canós J., Letelier P. “Metodologías Ágiles en el Desarrollo de Software”. JISBD 2003. España.
- [2] Urquiza Yllescas, J.F., et al, “Las Metodologías Ágiles y las Arquitecturas de Software”. Coloquio Nacional de Investigación en Ing. De Sofá. 2010, León, México.
- [3] Nord RL, Tomayko JE. “Software architecture-centric methods and agile development.” IEEE Software 2006; 23(2):47–53.
- [4] Babar MA, Brown AW, Mistrik I, Agile Software Architecture Aligning Agile Processes and Software Architectures. 2013 ISBN 978-0-12-407772-0.
- [5] Bass L, Clements P, Kazman R. Software architecture in practice 3rd ed. Addison-Wesley ISBN 978-0-321-81573-6.

- [6] Navarro, M; Moreno, M; Aranda, J; Parra, L; Rueda, J. “Integración de Metodologías Ágiles y Arquitecturas de software en el Desarrollo de Sistemas de Información.” XVIII Workshop de Investigadores en Ciencias de la Computación. 2016, Concordia, Entre Ríos.
- [7] Chen L, Babar MA, Nuseibeh B. Characterizing architecturally significant requirements. IEEE Software 2013; 30:38–45.
- [8] “Aplicabilidad de Metodologías y Tecnologías en el desarrollo de Sistemas de Información” Cod: 21/E 979. FCEFN. Navarro, Mirta, et al. 2014-2015
- [9] “Convergencia de Tecnologías Informáticas y Metodologías para la Implementación de Sistemas de Información” Cod: 21/E/871 FCEFN. Navarro, Mirta, et al. 2012-2013.
- [10] “Integración de Metodologías Ágiles y Arquitecturas de software en el Desarrollo de Sistemas de Información” Cod: 21/E 1027. FCEFN. Navarro, Mirta, et al. 2016-2017

Interfaz Gráfica de Usuario: el Usuario como Protagonista del Diseño

M. Claudia Albornoz, Mario Berón, Germán Montejano

*Departamento de Informática/Universidad Nacional de San Luis-U.N.S.L./San Luis/Argentina
Ejército de los Andes 950, Tel: +54 (0266) 4520300; int 2102
{albornoz,mberon,gmonte}@unsl.edu.ar*

Resumen

La Interfaz Gráfica de Usuario (GUI por su nombre en inglés, Graphical User Interface) es parte fundamental de cualquier aplicación; al comenzar a trabajar con una computadora el usuario comienza a interactuar con la Interfaz, ya sea la del sistema operativo, la de un software en particular o la de cualquier sitio web. Es donde comienza la interacción hombre-computadora. El diseño de la GUI no se lo debe considerar como una tarea secundaria y sin importancia; por el contrario el equipo de desarrollo debe contar con integrantes especializados en el tema.

En ocasiones, a partir de la GUI se puede determinar si una aplicación será utilizada o no para resolver los problemas para los cuales fue diseñada. En el actual mundo informatizado se debe ofrecer al usuario una Interfaz que lo ayude a concretar las tareas de manera rápida, sencilla y satisfactoria. Es la Interfaz la responsable de ofrecer una interacción fluida y agradable.

El objetivo del presente trabajo es mostrar la importancia del diseño de la GUI desde el punto de vista del usuario; explicar qué es la Interacción Persona-Computadora, fundamentar la importancia del Diseño Centrado en el Usuario y la Ingeniería de la Usabilidad y cómo lograr una interacción satisfactoria entre el usuario y la computadora, contemplando los estados emocionales del usuario.

Palabras Claves:

Interfaz Gráfica de Usuario (GUI), Interacción Persona-Computadora (IPC), Diseño Centrado en el Usuario (DCU), Ingeniería de la Usabilidad, Computación Afectiva

Contexto

Este trabajo de investigación se encuentra enmarcado dentro de una de las líneas de investigación del Proyecto de Incentivos cuyo título es: “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad”, código P-031516, de la Facultad de Ciencias Físico, Matemáticas y Naturales, de la Universidad Nacional de San Luis. Éste proyecto es la continuación de diferentes proyectos de investigación, los cuales han logrado importantes vínculos con diversas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. Introducción

La Interfaz es fundamental en toda aplicación y/o sitio web. Es la responsable de transmitir o hacerle saber al usuario lo que es capaz de hacer el producto. En la actualidad gran parte de la información se encuentra digitalizada, por esto los usuarios (en su mayoría inexpertos) se enfrentan a interactuar con productos de software permanentemente. Para aprovechar los beneficios de la tecnología, debe existir una buena interacción entre el usuario y la computadora. Pensando en los usuarios se debe diseñar la GUI [1].

La GUI se la puede definir como la parte de una computadora y su software que el usuario puede ver, oír, tocar, hablar, o de otra manera entender o dirigir [2]. La Interfaz de usuario tiene dos componentes: entrada y salida. La entrada es cómo el usuario comunica sus necesidades o deseos a la computadora. Para esto utiliza los componentes de entrada: teclado, ratón, el trackball, el dedo (para pantallas táctiles) y la voz (para instrucciones habladas). La

salida es cómo la computadora transmite los resultados al usuario. Hoy en día, el mecanismo más común de salida de la computadora es la pantalla, también se utiliza la voz y el sonido.

Cualquier usuario, sin conocimiento previo, puede operar un sistema si la Interfaz está bien diseñada y construida. Una Interfaz mal diseñada obstaculiza la ‘usabilidad’, es decir aquella característica ‘*que hace que la aplicación sea fácil de utilizar y fácil de aprender*’ [2].

Al hablar de ‘usabilidad’ se puede considerar que la Interfaz presenta las siguientes características (Fig. 1):

- *Satisfacción*: que el usuario trabaje en un entorno en el que se sienta cómodo.
- ☐ *Efectiva*: hace lo que se espera que haga.
- ☐ *Eficiente*: realiza las tareas en el tiempo necesario y sin errores.



Fig 1: Usabilidad

El objetivo es lograr una Interfaz ‘usable’ para mejorar la Interacción Persona Computadora, concepto que se tratará en la siguiente sección.

1. INTERFAZ GRÁFICA DE USUARIO

Esta línea de investigación abarca el estudio de los siguientes temas:

1.1 I.P.C.: Interacción Persona-Computadora

Las personas interactúan con las interfaces constantemente: al usar el celular, en el cajero automático, la computadora, la cámara fotográfica digital, el GPS del auto, etc. Además de la interacción física que realiza el usuario (al teclear, al mover el mouse, al tocar la pantalla digital, etc.) se debe tener en cuenta el nivel cognitivo necesario para que el usuario comprenda el protocolo de interacción que presenta la interfaz [3].

Para el usuario sólo existe la Interfaz, no le importa cómo se concretan las tareas, cómo están programadas o implementadas. Los usuarios, a través de Internet, pueden probar, obtener sistemas y herramientas de software; ellos se han convertido en clientes exigentes y críticos: esperan un alto grado de elaboración en las Interfaces Gráficas de Usuario (GUI), luego le dan importancia al funcionamiento del sistema y al código [9].

Gracias a la invasión tecnológica en la vida de los usuarios, en los últimos tiempos surgió una nueva disciplina que se encarga de estudiar cómo lograr una fluida interacción usuario-máquina. Es lo que se conoce como Interacción Persona-Computadora (IPC o IPO por Interacción Persona-Ordenador), también conocida como Human Computer Interaction (HCI).

Se puede considerar que la IPC es el punto de encuentro de las ciencias humanas (como la psicología, la pedagogía, sociología y otras) con la tecnología o las ciencias exactas como la informática. Se focaliza en estudiar la interacción entre usuarios y sistemas informáticos; cuyo objetivo es proporcionar bases teóricas, metodológicas y prácticas para el diseño y evaluación de aplicaciones interactivas. Esta tarea se realiza para que puedan ser usadas de forma eficaz, segura, eficiente y satisfactoria [4].

Como definición, se puede decir que: la interacción persona-computadora es el intercambio observable de información, datos y acciones entre un humano y la computadora, y viceversa [5]. La IPC surge para educar a los primeros trabajadores que se vieron obligados a usar la computadora, ya sea en organismos gubernamentales o grandes empresas. Con el tiempo se fue perfeccionando con el fin de ayudar a todos los usuarios, incluyendo personas con capacidades diferentes [6]. Al referirse a proporcionar bases teóricas, metodológicas y prácticas para el diseño y evaluación de aplicaciones interactivas, se habla de Ingeniería de la Usabilidad y el Diseño Centrado en el Usuario. Ambos abarcan un

conjunto de procesos y metodologías que aseguran el cumplimiento de los niveles de usabilidad requeridos en la aplicación. Conceptos que se tratarán a continuación.

1.2 Ingeniería de la Usabilidad

La Ingeniería de la Usabilidad es multidisciplinar; se nutre de la informática, de la psicología, de la lingüística, de la sociología, de la antropología y del diseño industrial. Este término se utiliza desde la década del 80 para designar a una nueva disciplina, que se ocupa de proporcionar “métodos sistemáticos y herramientas para la compleja tarea de diseñar interfaces de usuario que sean fácilmente comprensibles, rápidamente aprendible y fiablemente operables” [7]. Para el usuario la Interfaz es ‘la aplicación’ o ‘el sistema’ en sí, porque es lo que ve y con lo que interactúa. Si la interacción no es fluida, se la considera inútil, no efectiva, las funcionalidades y la utilidad son limitadas. Cuando esto ocurre el usuario se confunde y se frustra.



Fig 2: Metodología conceptual y esquemática de la Ingeniería de la Usabilidad.

El objetivo de la Ingeniería de la Usabilidad es minimizar la sobrecarga cognitiva y perceptiva de los usuarios. Utiliza un método de diseño iterativo con prototipado rápido (necesariamente deberá contar con herramientas de ayuda), cuyo ciclo es: "análisis, diseño, implementación, evaluación" (Fig. 2), que se repite varias veces con el fin de ir mejorando progresivamente la aplicación. La etapa de evaluación del prototipo, la cual se realiza con usuarios reales a cada repetición del ciclo, es de suma importancia para obtener resultados dignos de una ingeniería. En ésta etapa el usuario es el protagonista.

1.3 Diseño Centrado en el Usuario

El Diseño Centrado en el Usuario (DCU) es el término que se utiliza para describir el diseño en el que el usuario influye en el resultado final. Se considera que es una filosofía y un proceso. Una filosofía porque sitúa al usuario en el centro con la intención de desarrollar un producto adecuado a sus requerimientos, necesidades y un proceso de diseño porque se centra en los factores cognitivos de las personas y como éstos intervienen en sus interacciones con los productos [8].

El concepto Diseño Centrado en el Usuario nació tras la investigación de Donald A. Norman en la Universidad de California San Diego (UCSD); autor del libro *User centered system design: new perspectives on human-computer interaction*, y posteriormente se consolidó con el libro de Norman *The design of everyday things*. Obra referente en el ámbito del DCU y las disciplinas afines como la interacción persona-computadora, el diseño y la experiencia de usuario.

El usuario final se ve involucrado en cada etapa o fase del proceso de desarrollo, con esto se garantiza que el producto se ajuste a sus necesidades. Ésta forma de diseño le otorga al usuario el rol central del desarrollo. En ocasiones, el usuario es un miembro más del equipo de diseño. El objetivo es obtener productos fáciles de usar, efectivos y eficientes.

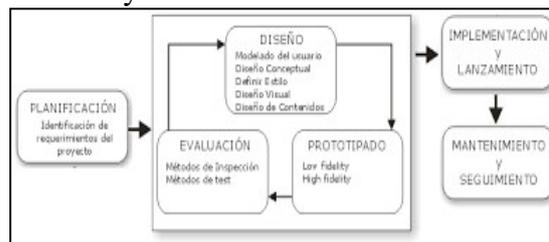


Fig. 3: Etapas del DCU

Esta metodología consta de diferentes etapas (Fig. 3) [10]. Si bien estas etapas corresponden a todo el proceso de desarrollo, es obvio que el diseño y desarrollo de la Interfaz se ven involucrados en ellas. Como es evidente, el usuario se ve involucrado en cada etapa o fase.

1.4 Computación Afectiva: las Emociones

En los últimos tiempos han surgido nuevas investigaciones para lograr no solo una interacción fluida, entre el usuario y la computadora, sino también en que sea satisfactoria y porque no, placentera.

Es por esto que los usuarios no se los considera sólo como seres con habilidades cognitivas y físicas, también se deben tener en cuenta los estados emocionales (miedos, esperanzas, valores, etc.). Hoy el objetivo es lograr una excelente interacción entre el usuario y la computadora; así surge una nueva disciplina: la Computación Afectiva.

Ésta disciplina es una rama de la Inteligencia Artificial, desarrolla métodos computacionales orientados a reconocer emociones humanas y generar emociones sintéticas. Surge ante la necesidad de optimizar la interacción entre personas y computadoras, pero también se inscribe en la investigación de los procesos inteligentes [11]. La afectividad es fundamental en el comportamiento y la comunicación de las personas. La interacción humana siempre incluye emociones, estados de ánimo, afectos los cuales se transmiten de manera explícita (verbalmente) o implícita (no verbal) a través de gestos, expresiones, actitudes [12]. Esta información que se transmite con actitudes, expresiones y gestos es de gran valor y producen un gran efecto en la comunicación, aún en la comunicación usuario-computadora.

El hecho de que las computadoras puedan comprender nuestras emociones y a su vez que puedan “expresar” (o simular) emociones propias, sería un paso importante para establecer un cambio cualitativo en la interactividad.

2 Actividades Llevadas a Cabo en la Línea de Investigación

Hasta el momento, en la línea de investigación descrita previamente, se están llevando a cabo las siguientes tareas:

- Revisión Sistemática de la literatura referente al tema de diseño y construcción de Interfaces Gráficas de Usuario; con el fin de evidenciar la importancia de su diseño y cómo lograr una óptima y satisfactoria interacción con los usuarios. Haciendo una mención especial respecto al rol del usuario en el diseño de la Interfaz.
- Estudio de Métodos de Evaluación Multicriterios. Esta tarea tuvo como objetivo seleccionar un método de evaluación para evaluar interfaces la calidad de las interfaces gráficas de usuario. Hasta el momento se ha encontrado que el método LSP (Logic Scoring of Preference) se adapta muy bien para la evaluación de las GUI.
- Elaboración de Criterios de Evaluación. Esta tarea se lleva a cabo con el objetivo de establecer cuáles son las características que las interfaces gráficas de usuario deben tener para que sean consideradas adecuadas para su utilización por parte del usuario. Estos criterios serán la entrada al método de evaluación multicriterio mencionado en el ítem precedente.
- Estudio de Reglas, Técnicas y Principios de Diseño de GUI. Esta tarea tiene como finalidad proveer la información necesaria para que los equipos de desarrollo de GUI puedan desarrollar GUIs adecuadas para el usuario.
- Análisis de la Influencia Emocional, el objetivo es determinar cómo influye el estado emocional de los usuarios en la interacción con las computadoras y, a su vez, cómo el diseño de la Interfaz influye en el estado emocional del usuario.

3 Conclusiones y Trabajos Futuros

Luego de la lectura de diferentes trabajos respecto de cómo mejorar la interacción usuarios-computadoras se puede observar que el diseño de una Interfaz

Gráfica de Usuario no es una tarea secundaria. Además, no sólo involucra un grupo de expertos en programación; sino que es fundamental que el equipo integre a los usuarios. Diversos estudios demuestran que cada vez son más las investigaciones que se realizan con el objetivo de lograr una óptima interacción. No sólo dejando en claro qué es lo que necesita el usuario (requerimientos), sino también lo que desean y sus estados emocionales.

Como resultado, ésta investigación pretende determinar cómo lograr una GUI 'usable', comprensible, satisfactoria; con el análisis de los diferentes tópicos que se involucran en la Interacción Persona-Computadora. .

Futuros trabajos: i) Enriquecer el estudio sistemático de la literatura; ii) Implementar un método de evaluación multicriterio para evaluar GUIs, iii) Definir criterios de evaluación de interfaces gráficas de usuario, iv) Elaborar un manual con las Reglas, Técnicas y Principios de Diseño de GUI y v) Profundizar el estudio sobre la influencia emocional.

4 Formación de Recursos Humanos

En el contexto de esta línea de investigación se están llevando a cabo diferentes tesis de grado, trabajos finales integradores de ingeniería y trabajos de especialización. En lo que respecta a tesis de grado una de ellas finalizó exitosamente y con repercusión internacional. En dicho trabajo se implementó el método de evaluación Logic Scoring of Preference (LSP). Por otra parte, se está realizando un trabajo final integrador de ingeniería en donde se implementa el método de evaluación AHP (Analytic Hierarchy Process). Ambos trabajos se van a utilizar para evaluar GUIs y también para analizar las ventajas y desventajas de cada método para evaluar el tipo de sistemas estudiados en esta línea de investigación.

En Ingeniería de Software, a nivel de Especialización, se puede decir que se

está desarrollando un trabajo final en donde se pretende integrar todos los conceptos abordados en la línea de investigación. También se encuentran en desarrollo tesis de grado y postgrado que aborden las temáticas de la línea de investigación.

Referencias

- [1] Chamba, Carrión, and Rober Alonso. *Estándar de usabilidad para la interfaz gráfica de usuario en los proyectos de desarrollo de software*. BS thesis. 2014.
- [2] Galitz, Wilbert O. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. Wiley.com, 2007.
- [3] Tonj Granollers i Saltiveri. *MPIu+ a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares*. Universitat de Lleida, 2007. ISBN: 978-84-690-6378-1
- [4] Mascheroni, M. A., Greiner, C. L., Petris, R. H., Dapozo, G. N., and Estayno, M. G. *Calidad de software e ingeniería de usabilidad*. 2012. In XIV Workshop de Investigadores en Ciencias de la Computación.
- [5] Díaz, F. J., Harari, I., and Amadeo, A. P. *Guía de recomendaciones para diseño de software centrado en el usuario*. 2013. Editorial de la Universidad Nacional de La Plata (EDULP).
- [6] Olson, G. M. and Olson, J. S. Human-computer interaction: Psychological aspects of the human use of computing. 2003. Annual review of psychology, 54(1):491-516.
- [7] Lorés, Jesús, and Toni Granollers i Saltiveri. *La Ingeniería de la Usabilidad y de la Accesibilidad aplicada al diseño y desarrollo de sitios web*. Universitat de Lleida, 2004.
- [8] Domingo, Muriel Garreta, and Enric Mor Pera. *Diseño centrado en el usuario*, 2010.
- [9] Marín, José M. Belmonte. *Ingeniería de la Usabilidad. Aplicada al desarrollo de un portal web administrado dinámicamente*. Trabajo Final de Carrera. Universitat de Lleida. 2003.
- [10] Yusef Hassan & Francisco J. Martín Fernández & Ghzala Iazza. *Diseño Web Centrado en el Usuario: Usabilidad y Arquitectura de la Información*. "Hipertext.net", núm. 2, 2004. <<http://www.hipertext.net>>
- [11] Causa, Emiliano, and Andrea Sosa. *La computación afectiva y el arte interactivo*. Área Transdepartamental de Artes Multimediales 52 (2007).
- [12] González, J. A., Gardezabal, L., and Vitoria, N. G. *Mediación emocional aplicada en sistemas de comunicación aumentativa y alternativa*. Inteligencia Artificial: revista iberoamericana de inteligencia artificial, 6(16):65-70. 2002.

La Gestión del Conocimiento en Pequeñas y Medianas Fábricas de Software en el Área Metropolitana de Buenos Aires

Straccia, Luciano; Maulini, Adriana; Pytel, Pablo; Masci, Marcelo; Vegega, Cinthia; Pollo-Cattaneo, Ma. Florencia

Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS)
 Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
 Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires Tel +54 11 4867-7511

lstraccia@frba.utn.edu.ar, adri.maulini@gmail.com, ppytel@gmail.com,
 marcelomasci@gmail.com, cinthiavg@yahoo.com.ar, flo.pollo@gmail.com

Resumen

El conocimiento constituye un recurso estratégico y su gestión otorga ventajas competitivas. Existen diversos modelos que han sido creados para guiar la Gestión del Conocimiento en una organización, pero es necesario trabajar específicamente sobre modelos para las fábricas de software. En este contexto, el objetivo del proyecto es definir e implementar un modelo de gestión del conocimiento para las pequeñas y medianas fábricas de software en el Área Metropolitana de Buenos Aires.

Palabras clave: Gestión del Conocimiento, Software Factory, Ingeniería de Software, PyME.

Contexto

La Universidad Tecnológica Nacional (UTN) ha definido diversos Programas de Investigación + Desarrollo + Innovación (I+D+i), entre los cuales se encuentra el Programa Tecnología de las Organizaciones, que tiene por finalidad “articular las distintas temáticas

relacionadas con la gestión del conocimiento, la innovación y los sistemas de gestión de la calidad e integrados aplicables a las organizaciones”, incluyendo la gestión del conocimiento en las organizaciones como área prioritaria [1].

En el marco de las actividades del Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS) de la UTN-FRBA, con dependencia del Departamento de Ingeniería en Sistemas de Información, se comienza una nueva línea de trabajo en el campo de la Gestión del Conocimiento en pequeñas y medianas empresas de software mediante la definición e implementación de un modelo de gestión del conocimiento. De esta manera, se articula dentro de los objetivos de GEMIS en el campo de la Informática, la generación de nuevos conocimientos en el área de la Ingeniería de Software.

Introducción

Conocimiento y gestión del conocimiento

El conocimiento “se constituye en el recurso estratégico más importante, y la

habilidad para generarlo, adquirirlo, codificarlo, transferirlo, aplicarlo y reutilizarlo, se ha convertido en la competencia sustancial para la obtención de una ventaja competitiva sostenible” [2]. Diversos autores como Nonaka y Takeuchi [3], Davenport [4], Davenport y Prusak [5] y Wiig [6] han presentado diversas definiciones asociadas al término conocimiento. Pérez y Urbáez [7] afirman que existen dos campos epistemológicos sobre la naturaleza del mismo: una perspectiva objetivista y otra basada en que el conocimiento es esencialmente personal e inmerso en las prácticas individuales y organizacionales. Los mismos autores, sustentándose en Seaton y Bresó [8], indican que también existen dos enfoques o clasificaciones principales para las distintas definiciones: un enfoque organizacional y un enfoque económico.

Según Díaz y Millán [9] las organizaciones se preocupan cada vez más por mantener el talento inmerso en sus empresas, por ende la Gestión del Conocimiento (GC) y el Capital Intelectual (CI) se convierten en ventajas competitivas para las organizaciones modernas. Rueda Martínez [10] indica que no existe un consenso definitivo respecto del término GC y que la mayoría de las definiciones académicas se pueden clasificar en tres grupos: la GC entendida como explotación de un recurso a disposición de la organización, aquella con una perspectiva humana-productiva y la conceptualizada en base a una descripción del proceso de producción y aplicación del conocimiento.

Existen diversos modelos que han sido creados para guiar el proceso de GC en una organización. Según Pérez y Urbáez [7], estos modelos son el resultado de propuestas teóricas y de investigaciones de experiencias en empresas, que “buscan aproximarse a una explicación de cómo

debe organizarse, gestionarse y fluir el conocimiento en la organización”.

Las fábricas de software

Con base en conceptos y caracterizaciones referentes a las fábricas de software (*Software Factory*, SF) de autores como Nomura y otros [11], Santos y Soares [12], Fernstrom y Narfelt [13] y Greenfield y Short [14], se puede inferir que una SF es un modelo de referencia, plantilla o forma de trabajo, que se puede aplicar en empresas que desarrollan productos de software, basado en equipos de trabajo multidisciplinarios, conformados por profesionales especializados, con roles y responsabilidades claras, que siguen procesos bien definidos, para producir una familia de productos de software.

Un área que necesita ciertamente de la GC, sobre todo en la era tecnológica, son las entidades dedicadas al desarrollo de software, incluyendo a las SF. Según Aurum y otros [15], los desarrolladores de software procesan conocimiento de alto valor, el cual es dinámico y evoluciona con la tecnología y la cultura organizacional y consideran que la GC debe ser diferente para la ingeniería de software debido a la alta especialización de los profesionales, la alta rotación del personal, y la evolución de las tecnologías. Para Rus y Lindvall [16], cuando se trata de procesos ingenieriles, existen dos nuevos tipos de conocimiento: el embebido en los productos (artefactos), porque son el resultado de actividades creativas y de alto intelecto; y el meta- conocimiento, que es el conocimiento acerca de los productos y procesos. Además consideran que la GC debe incluir la incorporación de conocimiento externo a la compañía debido a la constante evolución de las tecnologías utilizadas durante los procesos de desarrollo de software y que el

conocimiento en la ingeniería de software es muchas veces implícito y debido a su complejidad es difícil convertirlo en explícito. Siguiendo en la línea de los desafíos de la GC en la Ingeniería de Software, Dingsoyr y Smite indican que cuando existen proyectos globales, con grupos divididos en locaciones geográficas diferentes, “las distancias temporales y geográficas afectan en primera instancia la habilidad para compartir y acceder al conocimiento, mientras que las distancias socio-culturales introducen desafíos para unificar las formas en que el conocimiento se comparte y mantiene” [17].

En lo que respecta a la localización geográfica de la producción de software, alrededor del 80% se localiza en el Área Metropolitana de Buenos Aires y el 20% en el resto del país [18], por lo cual se ha definido como alcance del proyecto a dicha región metropolitana.

Modelos de gestión del conocimiento

Pérez y Urbáez [7], indican que los investigadores en las últimas décadas, a través de sus propuestas teóricas y de sus investigaciones de experiencias en empresas, han intentado aproximarse a explicar cómo debe organizarse, gestionarse y fluir el conocimiento en la organización. Gómez [19] sugiere que los modelos de GC existentes pueden ser clasificados en tres categorías: almacenamiento, acceso y transferencia de conocimiento; sociocultural; tecnológicos.

Rus y Lindvall [16] describen los principales problemas que puede enfrentar una organización de este tipo en cuanto a GC y sus posibles soluciones, sin embargo no plantea ningún modelo. En cuanto a modelos de gestión del conocimiento, Díaz y Millán [9] analizan algunos modelos académicos existentes

de GC y de capital intelectual, y narran experiencias al aplicarlos en diversas universidades de Latinoamérica y España. Por su parte Pérez y Urbáez [7], analizan y caracterizan diversos modelos diferentes de GC y realizan un análisis comparativo entre ellos.

Por otra parte, en lo relacionado a la GC en el desarrollo de software, Pons y otros [20] proponen un modelo de GC para mejorar el desarrollo de equipos de proyectos informáticos. Sin embargo este modelo no toma en cuenta las características de una SF y determinados problemas relacionados a la GC en desarrollo de software, tales como la integración de procesos de tercerización de servicios o desarrollo, que implica particularidades en el proceso de transferencia de conocimientos que es aquel “mediante el cual el conocimiento propiedad de una persona, grupo o entidad emisora es comunicado a otra persona, grupo o entidad receptora que tiene una base común o conocimiento similar al del emisor” [21]. Varios estudios han demostrado que la transferencia efectiva del conocimiento técnico puede tener un significativo impacto en la efectividad de la tercerización [22].

Líneas de Investigación, Desarrollo e Innovación

Dentro del ámbito de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN-FRBA) se ha conformado en el año 2009 el Grupo GEMIS, integrado por un equipo de docentes, alumnos y graduados con interés en la sistematización de conocimientos y su promoción sobre el campo de la Ingeniería en Sistemas de Información y la Ingeniería de Software incluyendo sus aplicaciones y abordajes

metodológicos en todo tipo de escenarios. Estas actividades se han venido llevando a cabo dentro del marco de cuatro PID asociados al equipo.

Los autores del presente trabajo han desarrollado a lo largo de los últimos años en el ámbito académico y también en el ámbito profesional de la Ingeniería en Sistemas de Información, diversas tareas asociadas a la temática de Gestión del Conocimiento. Además algunos de los autores forman parte del cuerpo docente del Seminario de Modelos de Organizaciones y Sistemas de Información (MOSI) de la carrera de posgrado Maestría en Ingeniería en Sistemas de Información de la UTN- FRBA entre cuyos contenidos se encuentra la temática asociada a la nueva línea de investigación y dirigen trabajos de especialización y acompañan el desarrollo del plan de tesis de los estudiantes de posgrado involucrados.

Teniendo en cuenta los antecedentes de trabajos realizados para analizar las problemáticas presentes en la GC en la industria del software, se ha considerado necesario proponer una nueva línea de trabajo que se estará desarrollando en el marco de un nuevo PID que tenga como objetivo excluyente la definición e implementación de un modelo de gestión del conocimiento para las pequeñas y medianas fábricas de software en el área Metropolitana de Buenos Aires.

Resultados y Objetivos

En el marco de la Especialización y Maestría en Sistemas de Información, con la dirección de autores del presente trabajo, tres estudiantes se encuentran desarrollando sus trabajos finales de especialidad (que tendrán su continuidad como tesis de maestría) denominados "Modelos de gestión del conocimiento y su aplicación en software factories",

"Método de evaluación de transferencia de conocimiento en los procesos de Outsourcing para las PyMEs de la región de CABA y Gran Buenos Aires" y "Características de un sistema inteligente para la gestión del conocimiento en el área de Tecnología Informática".

En los proyectos de investigación y desarrollo trabajados por el momento en GEMIS el conocimiento ha sido objeto de estudio en trabajos tales como [23] y [24] asociados a la Ingeniería del Conocimiento y en otros casos se ha trabajado sobre la sistematización del conocimiento, especialmente mediante metodologías asociadas al desarrollo de software [25]. Si bien estos trabajos no son propios de temáticas de GC, entendemos que pueden resultar de utilidad como base para la línea de trabajo presentada.

La finalidad de este proyecto es darle continuidad a los objetivos de GEMIS mediante la definición e implementación de un modelo de gestión del conocimiento para las pequeñas y medianas fábricas de software en el Área Metropolitana de Buenos Aires. Se espera: a) identificar modelos de gestión del conocimiento vigentes en el mundo; b) identificar los requisitos necesarios para las organizaciones para la implementación de los modelos; c) caracterizar a las pequeñas y medianas empresas de la industria del software; d) identificar problemas para la implementación de los modelos de gestión del conocimiento en las pequeñas y medianas empresas de la industria del software; e) construir un modelo de gestión del conocimiento adaptado a las pequeñas y medianas empresas de la industria del software; y f) implementar el modelo de gestión del conocimiento en diversas organizaciones y analizar sus resultados.

Formación de Recursos Humanos

El equipo se encuentra conformado por investigadores formados, tesistas de maestría, graduados de grado y alumnos de la carrera de Ingeniería en Sistemas de Información. Esta nueva línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para que asciendan dentro de la carrera de investigadores, además de fomentar la aplicación de las temáticas en su propia actividad profesional; y plantea la integración de alumnos avanzados de grado y posgrado con posibilidades de articular Proyectos Finales, Trabajos Finales Integrador de Especialidad y Tesis de Maestría de las carreras del área.

Referencias

- [1] SeCTyP (2017). Programas I+D+I. UTN. Secretaría de Ciencia, Tecnología y Posgrado.
- [2] Gelaf, Graciela (2010). Abordajes creativos en situaciones de crisis organizacionales. Contaduría General de la Nación. Tucumán, Argentina.
- [3] Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford university press
- [4] Davenport, T. (1998). Knowledge Management and the broader firm: Strategy, advantage, and performance. En J. Liebowitz, Knowledge Management Handbook.
- [5] Davenport, T. y Prusak, L. (1998). Working Knowledge: How organizations manage what they know. Cambridge, MA: Harvard Business School Press.
- [6] Wiig, K. (1993). Knowledge Management Foundations: Thinking about thinking – How people and organizations create, represent, and use knowledge. Arlington, TX: Schema.
- [7] Pérez, V. A., & Urbáez, M. F. (2016). Modelos teóricos de gestión del conocimiento: descriptores, conceptualizaciones y enfoques. Entreciencias: diálogos en la Sociedad del Conocimiento, 4(10), 201-227.
- [8] Seaton, Carlos y Bresó, Salvador (2001). El desarrollo de un sistema de gestión del conocimiento para los institutos tecnológicos. Revista Espacios Digital, Vol. 22 (3), Venezuela.
- [9] Díaz, M. T. R., & Millán, J. J. G. (2013). Gestión del Conocimiento y Capital Intelectual, a través de modelos universitarios. Económicas CUC.
- [10] Rueda Martínez, M. I. (2014). La gestión del conocimiento y la ciencia de la información: relaciones disciplinares y profesionales. Tesis Doctoral, Universidad Carlos III, Getafe
- [11] Nomura, L., Spínola, M. M., Tonini, A. C., & Hikage, O. K. (2007). A model for defining software factory processes. In 19th International Conference on Production Research.
- [12] Santos, S. C. D., & Soares, F. S. (2013). Authentic assessment in software engineering education based on PBL principles: a case study in the telecom market. In Proceedings of the 2013 International Conference on Software Engineering. IEEE Press.
- [13] Fernstrom, C., Narfelt, K. H., & Ohlsson, L. (1992). Software factory principles, architecture, and experiments. IEEE Software, 9(2), 36-44.
- [14] Greenfield, J., & Short, K. (2003). Software factories: assembling applications with patterns, models, frameworks and tools. ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications (pp. 16-27). ACM.
- [15] Aurum, A., Jeffery, R., Wohlin, C., & Handzic, M. (Eds.). (2013). Managing software engineering knowledge. Springer Science & Business Media.
- [16] Rus, I. y Lindvall, M. (2002). Knowledge management in software engineering. IEEE software.
- [17] Dingsoyr, T., & Smite, D. (2014). Managing knowledge in global software development projects. IT Professional.
- [18] Motta, J., Morero, H. y Borrastero, C. (2016). La política industrial en el sector de software de Argentina durante los años 2000. Univ de San Andrés. Bs.As.
- [19] Gómez, D. R. (2006). Modelos para la creación y gestión del conocimiento: una aproximación teórica. Educar 37 (25 - 39).
- [20] Pons, N. L., Pérez, Y. P., Stiven, E. R., & Quintero, L. P. (2014). Diseño de un modelo de gestión del conocimiento para mejorar el desarrollo de equipos de proyectos informáticos. Revista española de documentación científica, 37(2), 5.
- [21] Orejuela, A.R. (2005). Un modelo integral para evaluar el impacto de la transferencia de conocimiento interorganizacional en el desempeño de la firma. En Revista Estudios Gerenciales. Nro 95. Abril-Junio 2005. Universidad ICESI: Murcia, España.
- [22] Zeinab, N. (2016) Effects of Human Factor on the Success of IT Outsourcing.
- [23] Pytel, P., Ramón, H., Tomasello, M., Britos, P., Pollo Cattaneo, M. F., Vegega, C., Rodríguez, D., García-Martínez, R. (2011). Propuesta de Aplicación de Técnicas de Representación de Conocimiento en el Análisis de Requisitos Software. Actas 1er Seminario Argentina-Brasil de Tecnologías.
- [24] Pytel, P., Uhalde, C., Ramón, H. Castello, H., Tomasello, M., Pollo-Cattaneo, M., Britos, P., García-Martínez, R. (2011). Ingeniería de Requisitos Basada en Técnicas de Ingeniería del Conocimiento. Workshop de Investigadores en Ciencias de la Computación.
- [25] Straccia, L.; Pytel, P.; Pollo-Cattaneo, M.F (2016). Metodología para el desarrollo de software en proyectos de I+D en el nivel universitario basada en Scrum. XXII Congreso Argentino de Ciencias de la Computación.

¿Las Redes Sociales Cumplen con los Criterios de Accesibilidad?

Castro Valeria, Ortiz Claudia, Chapetto Viviana,
Balleto Carmen, Rossi Bibiana

Departamento de Ciencias Básicas, Universidad Nacional de Luján (UNLu)

castro.vales@gmail.com, cortiz@unlu.edu.ar, vchapetto@yahoo.com.ar,
carmenballeto@gmail.com, brossi@unlu.edu.ar

Resumen

En este trabajo se presenta un análisis de accesibilidad de los sitios web de las redes sociales Facebook, YouTube, Twitter, Google+ e Instagram. Los datos se recolectaron entre agosto del 2015 y marzo del 2016, teniendo en cuenta la normativa respecto a accesibilidad web vigente en Argentina. Se utilizaron herramientas automáticas de evaluación tales como Wave, Examinator y A-Checker, y se complementó su evaluación de forma manual como lo sugiere la norma.

Palabras clave: accesibilidad web, accesibilidad, redes sociales accesibles.

Contexto

La presente investigación es parte de las actividades del proyecto Accesibilidad de los sitios Web en Argentina del Departamento de Ciencias Básicas de la Universidad Nacional de Luján, presentado en WICC 2016 [Rossi, y otros, 2016]. El objetivo del proyecto es analizar el grado de accesibilidad de portales de distintas áreas, de acuerdo con lo establecido en la Ley 26.653.

Este trabajo se realizó con la participación de una alumna de la carrera de la Licenciatura en Sistemas de Información, bajo la dirección del equipo de investigación y se concretó en un Trabajo Final que ha sido defendido en Diciembre de 2016.

Introducción

En Argentina se sanciona, en noviembre de 2010, la “Ley de Accesibilidad de la Información en las Páginas Web” [Ley 26.653, 2010]. Y en el año 2014 se aprueban la “Norma de Accesibilidad Web 2.0” y los “Niveles mínimos de conformidad”, [ONTI, 2014]. Las normas se basan en las WCAG 2.0, [W3C, 2008] que definen tres niveles de accesibilidad, A, AA y AAA. La normativa argentina asignó 4 puntos a cada uno de los 25 criterios de conformidad del nivel A. De acuerdo con esta normativa, se considera que un sitio es accesible si cumple como mínimo 50 puntos del nivel A en una primera etapa. Asimismo la normativa prevé una segunda etapa en la que el requisito mínimo será el cumplimiento de 80 puntos del nivel A. [ONTI, 2014].

Es importante destacar que en los estándares internacionales se considera que un sitio es accesible si cumple el nivel AA (que incluye el cumplimiento del nivel A).

En este trabajo se analizó si las redes sociales Facebook, YouTube, Twitter, Google+ e Instagram (redes sociales más utilizadas hacia el año 2015) [Mander, 2015], respetan las normas de accesibilidad exigidas en Argentina.

Se consideró las redes Sociales como tema de interés para el análisis por ser de uso masivo y por esa misma razón cobra importancia el cumplimiento de los criterios de accesibilidad para que todas las personas tengan igualdad de oportunidades en el uso del servicio.

La evaluación se realizó sobre la página de inicio o portada de cada uno de los sitios, habiendo iniciado previamente sesión de usuario.

Líneas de Investigación, Desarrollo e Innovación

En el proyecto de investigación, el análisis del cumplimiento de los sitios web se organiza por áreas temáticas, por ejemplo: instituciones financieras, de salud, universidades, entre otras. El área de interés seleccionada para el presente trabajo ha sido Redes Sociales.

Para llevar a cabo esta investigación, los pasos que se siguieron fueron: seleccionar el área de interés a evaluar, definir las muestras y evaluar las aplicaciones informáticas a utilizar.

Las herramientas de evaluación empleadas fueron: Wave, Examiner y A-Checker, ya que permiten acceder a la evaluación de los sitios con la sesión de usuario iniciada. [A-Checker, (s.f.)], [Examiner, (s.f.)], [WAVE, (s.f.)]

Para la definición de los requisitos y criterios de cumplimiento se tomó como base las Normas de Accesibilidad Web 2.0, recomendadas por la ONTI.

El enfoque de la investigación es de tipo cuantitativo y el tipo de investigación es descriptiva. [Hernández Sampieri y otros, 2006].

El procedimiento seguido para el estudio de cada sitio fue: [Castro, 2016]

1. Realizar el análisis de cada sitio con las tres herramientas automáticas.
2. Registrar los resultados arrojados por cada herramienta.
3. Analizar el cumplimiento de los criterios de conformidad de nivel 'A' manualmente:
 - 3.1. Analizar la cantidad de fallas y criterios incumplidos, y determinar el grado de adecuación a la normativa, así como los puntos de falla más comunes.
 - 3.2. Analizar las alertas informadas por las diferentes herramientas, ya que pueden ofrecer "falsos positivos". Se evaluó para cada criterio, si se cumple o no, accediendo manualmente al sitio web de la red social y determinando si efectivamente se adapta a la normativa.
4. Establecer, en base a los resultados arrojados por las herramientas y el análisis manual, si el criterio se cumple

(SI), no se cumple (NO), o no aplica al sitio en estudio (NA).

- Asignar a cada criterio cumplido o que no aplica, un valor de 4 puntos (puntaje designado por la Norma 2.0).
- Sumar los puntos obtenidos, para comprobar el puntaje alcanzado en cada caso y verificar si cumplen con los 50 puntos mínimos de nivel 'A', exigidos por la normativa argentina.

Resultados y Objetivos

Para el análisis de los datos, se compararon los resultados de cada herramienta por sitio, pudiendo detectar la cantidad de errores que presenta cada uno (Tabla 1).

Red Social	WAVE		A-CHECKER			EXAMINATOR
	Errores	Alertas	Probl. Conoc.	Probl. Pot. (Muy mal + Mal)	Regular	
Facebook	54	32	33	1094	221	28
YouTube	405	60	1	970	528	2
Twitter	159	19	254	519	466	1
Google +	164	79	1	589	40	4
Instagram	70	22	61	820	109	80

Tabla 1. Errores y alertas detectados automáticamente

La red social con mayor cantidad de problemas encontrados automáticamente fue Youtube (Gráfico 1). A-Checker identificó un solo problema, pero detectó una gran cantidad de advertencias (Gráfico 2).

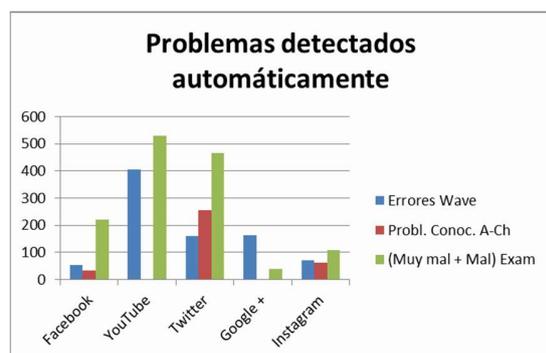


Gráfico 1. Problemas detectados automáticamente



Gráfico 2. Advertencias a verificar manualmente

Criterios	F	T	Y	G	I
1.1.1-contenido no textual	N	N	N	N	N
1.2.1-solo audio y video	N	N	N	N	N
1.2.2-subtítulos	N	N	N	N	N
1.2.3-adiod descripción	NA	N	N	N	N
1.3.1-información y relaciones	N	N	N	N	N
1.3.2-secuencia significativa	N	X	X	N	X
1.3.3-características sensoriales	NA	N	X	X	N
1.4.1-uso del color	X	NA	NA	X	NA
1.4.2-control audio	NA	NA	X	NA	NA
2.1.1-teclado	N	X	X	X	X
2.1.2-sin trampa teclado		X	NA	X	X
2.2.1-límite tiempo ajustable	N	N	NA	NA	NA
2.2.2-pausar, detener, ocultar	NA	NA	NA	NA	NA
2.3.1-tres destellos	X	X	N	X	X
2.4.1-saltar bloques		N	N	N	NA
2.4.2-página titulada	N	N	X	X	X
2.4.3-orden de foco	N	X	X	X	X
2.4.4-propósito vínculo	N	N	N	N	N
3.1.1-idioma		X	X	X	N
3.2.1-con foco		X	X	N	X
3.2.2-con entrada datos	N	NA	N	N	N
3.3.1-identificación errores	X	X	X	X	X
3.3.2-instrucciones o etiquetas	N	N	N	N	N
4.1.1-interpretación	X	N	N	N	N
4.1.2-nombre, rol, valor	N	N	N	N	N
Cumplidos	8	8	9	9	8
No cumplidos	13	13	12	13	12
No aplica	4	4	4	3	5

Tabla 2. Resultados generales por criterio

F(Facebook), T(Twitter), Y(YouTube), G(Google+), I(Instagram)

En la Tabla 2, se muestran los resultados de la evaluación manual. Se observan los criterios cumplidos (X), los No cumplidos (N), y aquellos que no aplican (NA) por corresponder a características no presentes en los sitios analizados.

Las herramientas automáticas mostraron que Google+ e Instagram no presentaban gran cantidad de errores, en comparación con las demás redes sociales, y que Twitter era la red que menos advertencias tenía para revisar manualmente. Sin embargo, una vez concluida la evaluación manual, se pudo observar que las cinco redes sociales cumplen e incumplen similar cantidad de criterios: 8 o 9 cumplidos y 12 o 13 incumplidos.

En la Tabla 3, se muestran los puntos obtenidos por sitio, calculados con la siguiente fórmula: $(Total SI + Total NA) * 4$. Si no se consideraran los criterios que NA, podría suceder que el sitio obtuviera un puntaje muy bajo, pero no necesariamente por incumplimiento.

Sitio	SI	NO	NA	Total
YouTube	9	12	4	52
Instagram	8	12	5	52
Google+	9	13	3	48
Facebook	8	13	34	48
Twitter	8	13	4	48

Tabla 3. Puntaje final de las redes sociales.

Se obtuvo como resultado que YouTube e Instagram cumplen con el mínimo exigido, para el nivel “A” de conformidad actualmente vigente en Argentina, ya que superan los 50 puntos. El resto de los sitios está próximo a alcanzar dicho puntaje. En todos los casos están lejos de alcanzar 80

puntos, que es el mínimo a considerar por la normativa argentina en una segunda etapa. Y si tenemos en cuenta el total de criterios de nivel A, o sea 100 puntos, están aún más lejos.

Si se observa detenidamente la Tabla 2, los cinco primeros criterios y los tres últimos, no son cumplidos por las cinco redes sociales, salvo el caso de Facebook en la cual no aplica el criterio 1.2.3 (audio descripción) y sí cumple el 4.1.1 (interpretación). Tampoco cumplen el criterio 2.4.4 (propósito de un vínculo). El incumplimiento de estos criterios afecta a las personas ciegas o con problemas visuales, ya que están relacionados con incluir alternativas textuales, descripciones, subtítulos, etiquetas o instrucciones cuando el contenido requiera entrada de datos. Además, los componentes del sitio deben estar bien estructurados, para que sean correctamente interpretados por los diferentes dispositivos de navegación y de asistencia. Al no cumplir los criterios mencionados, son muchas las personas a quienes se les dificulta participar en las “comunidades sociales”.

Algo para destacar es que la mayoría de las redes estudiadas cumplen los criterios relacionados con su uso a través del teclado (o dispositivos similares), lo cual es altamente conveniente tanto para personas con problemas de visión como de motricidad.

Formación de Recursos Humanos

Integran el proyecto un profesor asociado, cinco docentes auxiliares, y alumnos. La

formación de recursos humanos se concreta a través de:

- la propuesta de temas y tutores para el desarrollo del Trabajo Final de Carrera de los alumnos de la Licenciatura en Sistemas de Información de la UNLu
- el desarrollo de temas de investigación para la tesis de maestría de los docentes y auxiliares que participan del proyecto.

Durante 2016, dos alumnos han comenzado a desarrollar su trabajo final de carrera, uno de los cuales ha concluido exitosamente su trabajo en diciembre de 2016, fecha en la que defendió su tesis.

Uno de los docentes auxiliares se encuentra definiendo su propuesta de trabajo final de maestría.

Los docentes auxiliares están adquiriendo experiencia en la dirección de trabajos finales de carrera y en el proceso de investigación y publicación de resultados.

Referencias

[A-Checker, (s.f.)] Web Accessibility Checker. (s.f.). Recuperado el 28/2/2017, de <http://achecker.ca/checker/index.php>

[Castro, 2016] Accesibilidad de los sitios web de las Redes Sociales en Argentina, Trabajo Final Lic. en Sistemas de Información. Universidad Nacional de Luján, 12/2016

[Examinator, (s.f.)] Evaluación de la Accesibilidad Web. (s.f.). Recuperado el 11/8/2015, de <http://examinator.ws/>

[Hernández Sampieri y otros, 2006] Metodología de la Investigación. Mc Graw Hill. Quinta Edición. ISBN: 978-607-15-0291-9

[Ley 26.653, 2010] Ley N° 26.653. Accesibilidad de la Información en las Páginas Web. Centro de Documentación e Información, Ministerio de Economía y Finanzas Públicas. Argentina, 3/11/2010.

[Mander, 2015] Instagram still lags Twitter as the fifth biggest social network. Globaleindex. Consulta 5/12/2015, de <http://www.globalwebindex.net/blog/instagram-still-lags-twitter-as-the-fifth-biggest-social-network>

[ONTI, 2014] Disposición N° 2/2014. Requisitos de accesibilidad para contenidos en la Web. Oficina Nacional de Tecnologías de la Información. Buenos Aires, 12/8/2014. Consulta 15/1/2017 de <http://servicios.infoleg.gob.ar/infolegInternet/anexos/230000-234999/233667/norma.htm>

[Rossi, Ortiz, Chapetto, 2016] XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina). ISBN: 978-950-698-377-2. Consulta 18/2/2017 de <http://sedici.unlp.edu.ar/handle/10915/52990>

[WAVE, (s.f.)] Web Accessibility Evaluation Tool. (s.f.). Consulta 11/8/2015, de <http://wave.webaim.org/>

[W3C, 2008] Pautas de Accesibilidad de Contenido Web 2.0. 11 de Diciembre de 2008. Recuperado el 10/7/2016 de <http://www.codexemplar.org/traduccion/pautas-accesibilidad-contenido-web-2.0.htm#text-equiv>

Líneas de Productos de Software: Hacia la Creación de Técnicas y Métodos para Mejorar el Reuso

Agustina Buccella, Juan Luzuriaga, Alejandra Cechich, Matias Pol'la, Maximiliano Arias, Marcos Cruz Rodolfo Martinez, Rafaela Mazalu
 GIISCO Research Group
 Departamento de Ciencias de la Computación
 Universidad Nacional del Comahue Neuquen,
 Argentina agustina.buccella@fi.uncoma.edu.ar

Resumen

El estudio de nuevas técnicas y métodos para fomentar el desarrollo de software basado en reuso ha tomado gran importancia en los últimos años dentro del área de la ingeniería de software, dando lugar a numerosos paradigmas, como el desarrollo orientado a servicios, desarrollo basado en componentes, desarrollo de Líneas de Productos de Software (LPS), entre otros. A diferencia de un desarrollo que realiza algún tipo de reutilización, todos estos paradigmas se basan en planear el reuso, de modo que los desarrollos futuros puedan hacer un uso efectivo de diferentes artefactos de software previamente desarrollados. En particular, la Ingeniería de Líneas de Productos de Software (ILPS) aplica un reuso orientado a dominios como mecanismo para lograr una estructura común de servicios (llamada similitudes), en la cual se basan todos los productos a ser desarrollados. A su vez, agrega una estructura variable (llamada variabilidad) que permite brindar mayor flexibilidad logrando crear productos con bases similares pero adaptados a las necesidades específicas de los diferentes usuarios u organizaciones pertenecientes a esos dominios.

En nuestro proyecto, nos basamos en la

creación y aplicación de nuevas técnicas y recursos para lograr una mejora dentro de un proceso integral de desarrollo de LPSs. En particular nos enfocamos en la definición de técnicas que diseñan y aplican artefactos de software dentro de una jerarquía de dominios que comparten generalidades y particularidades en forma de servicios comunes y variables. A su vez, nos proponemos sentar las bases para el desarrollo de un ambiente inteligente que permita crear LPSs dentro de la misma jerarquía mediante la selección y composición de componentes reusables.

Palabras Clave: Líneas de Productos de Software - Artefactos de Software - Dominios - Reuso

Contexto

La línea presentada se inserta en el contexto de los siguientes proyectos y acuerdos de cooperación:

- *Programa: Desarrollo de Software Basado en Reuso - Parte II (04/F009)*. En proceso de acreditación, para el periodo 2017-2020. Directora: Dra. Alejandra Cechich.
 - *SubProyecto: Reuso Orientado a Dominios - Parte II*. Incluido dentro del Programa. Directora: Dra. Agustina Buccella, Co-director: Mg. Juan Manuel Luzuriaga.
- *Acuerdo de Cooperación* entre el

Laboratorio de Investigación en Ecología Bentónica y el Laboratorio de Parasitología e Histopatología de Moluscos del Instituto de Biología Marina y Pesquera Almirante Storni.

- *Proyecto de Extensión Universitaria* (RESOL-2016-2555-E-APN-SECPU#ME) entre integrantes docentes del proyecto de Reuso Orientado a Dominios - Parte II e integrantes del Museo Ciencias Naturales - Secretaría de Extensión Universitaria. Director: Juan Porfiri.

Introducción

En un desarrollo basado en Líneas de Productos de Software (LPS) [6, 11, 14] se aplican diferentes mecanismos de acuerdo a dos fases bien diferenciadas: la *ingeniería de dominio* y la *ingeniería de aplicación*. En la primera fase, el análisis de dominio esta dirigido a identificar, capturar y organizar toda la información fuente recopilada de los sistemas existentes, los expertos del dominio, libros de texto, etc. Como resultado, se obtiene una plataforma de software con un conjunto de componentes reusables y configurables que proveen un entorno en común. A su vez, dicha plataforma posee un conjunto de variabilidades que permiten justamente adaptarla a las necesidades particulares de las aplicaciones a ser creadas [12, 17]. En la segunda fase, la *ingeniería de aplicación*, se instancia dicha plataforma, mediante la configuración de la variabilidad definida. Finalmente, la ultima actividad de esta fase devuelve un producto de software en particular. Como se puede observar a partir de estas dos fases, el éxito del desarrollo de una LPS depende de la identificación, uso y administración de los artefactos reusables. Por lo tanto, la aplicación de técnicas específicas para la sistematización del reuso

es crucial aquí. Sin embargo, lograr un reuso efectivo no es una tarea sencilla requiriendo esfuerzos especiales, los cuales muchas veces son particulares a determinados dominios. De esta manera, los dominios o contextos en la cual la LPS es desarrollada resulta fundamental a la hora de realizar componentes orientados al reuso.

En general existen dominios con diferentes granularidades, es decir, que involucran mayor o menor nivel de detalle en la cantidad de servicios incluidos. Algunos dominios de granularidad alta se consideran generales ya que abarcan un conjunto muy amplio de aspectos comunes a tener en cuenta. Sin embargo, crear una LPS sobre estos dominios es una tarea muy compleja ya que habría que contemplar demasiadas variabilidades o particularidades, lo que atentaría contra las ventajas establecidas en estos desarrollos. De esta manera, surge la necesidad de dividirlos en una serie de subdominios o ramas mas específicas, cada uno de ellos centrado en sus propias particularidades [7].

En particular en este proyecto estamos interesados en dos subdominios del dominio geográfico, el de *ecología marina* y el de *paleontología*. Particularidades de esos dominios son:

- *Dominio Ecología Marina*: aquí se estudian las relaciones de todos los organismos que viven en el hábitat de la vida marina y la interacción de estos con su ambiente, considerando los factores bióticos y abióticos del entorno. Los primeros involucran aquellos factores inertes (sin vida) tales como la temperatura, productos químicos, la salinidad, la luz, etc. Los factores bióticos incluyen tanto la flora como la fauna (organismos vivos) que sobreviven en un ecosistema así como las

relaciones entre los diversos organismos. Los estudios en el ámbito de la ecología marina abarcan el examen de los microorganismos unicelulares, el entorno donde las especies sobreviven, el impacto de la actividad humana y los efectos globales de la contaminación. De esta manera, el conjunto de actividades que se deben realizar abarca todos los estudios y análisis necesarios para arribar a conclusiones acerca de la vida y la conservación de los organismos. Para realizar estos estudios, en general se recopila la información necesaria a través de censos o campanas sobre las especies que viven en el mar. Dentro del dominio de ecología marina hemos realizado varios esfuerzos para la estandarización de los servicios provistos, cuyos trabajos pueden verse en [8, 10].

- *Dominio de Paleontología:* este dominio estudia todas las formas de vida pretéritas de nuestro planeta como por ejemplo animales invertebrados (esponjas, corales, braquiópodos, moluscos, artrópodos, equinodermos, etc.), animales vertebrados (peces, anfibios, reptiles, aves y mamíferos), plantas, hongos, bacterias, etc. Los estudios de los individuos fósiles pueden ser múltiples, desde análisis anatómicos, filogenéticos, tafonómicos, estratigráficos, paleoambientales, paleobiogeográficos, etc. Por otro lado, también se hacen dibujos de los individuos o se toman numerosas imágenes fotográficas para acompañar las publicaciones científicas o didácticas. Todos

estos estudios permiten obtener una variada información de las formas de vida que habitaron nuestro planeta hace miles o millones de años atrás. Los organismos colectados producto de salidas de campo y excavaciones son depositados en las colecciones de los museos correspondientes. Allí pasan a ser catalogados y a formar parte del inventario de la institución que los tiene bajo su resguardo. Aquí, existen varios esfuerzos de estandarización, entre los que podemos citar a la norma ISO 21127:2014¹, LIDO² y CIDOC³ los cuales definen y clasifican servicios y protocolos requeridos para gestión de colecciones.

Al mismo tiempo, hay que considerar que este enfoque de LPS basado en niveles de dominio debe poseer una forma sencilla de construcción en base al reuso de componentes de dominio previamente desarrollados. De esta manera, es necesario contar con un conjunto de reglas y patrones de uso definidos que permita el reuso dentro del dominio geográfico considerando la jerarquía de dominios preestablecida. Propuestas dirigidas hacia este tema definen entornos inteligentes de ensamblaje de componentes que soportan la creación de nuevas plataformas LPSs en base a reglas de composición. Algunas herramientas dirigidas a soportar estos entornos son DOPLER⁴ o S.P.L.O.T⁵. Sin embargo, estos entornos deben ser adaptados a las realidades del dominio geográfico y del reuso dentro del mismo.

¹ Information and documentation – A reference ontology for the interchange of cultural heritage information

² LIDO - Lightweight Information Describing Objects Version 1.0 - <http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/>

³ CIDOC Conceptual Reference Model Version 6.0 - <http://www.cidoc-crm.org>

⁴ <http://ase.jku.at/dopler/>

⁵ <http://www.splot-research.org/extensions/fundp/fundp.html>

Líneas de investigación y desarrollo

En los últimos años se han realizado varios avances en el modelado de LPS para el dominio geográfico general y en el subdominio de ecología marina en particular. Dentro de estos avances hemos definido la base para la división o jerarquía de dominios en subdominios más manejables [7,8] que mejoraron la complejidad inherente del dominio geográfico general. Para esto, se ha trabajado dentro del subdominio de ecología marina, en el cual hemos desarrollado una taxonomía de servicios [8] basada en los estándares geográficos definidos por el OGC⁶ y la ISO⁷. Al mismo tiempo hemos definido una metodología de desarrollo [7,9], la cual se basa en la construcción de artefactos de software para cada una de las actividades, comenzando con la taxonomía previamente mencionada. Son justamente estos artefactos los que guían a las actividades de manera de mejorar el desarrollo de sistemas en el dominio geográfico y en sus subdominios incluidos. En la Figura 1 se muestran las actividades junto con los artefactos de software generados y utilizados en cada uno [9].

Dichas actividades son parte de la ingeniería de dominio de un desarrollo de LPS. En la primera actividad se asiste a los ingenieros de software en la búsqueda de los servicios de la taxonomía más adecuados para cumplir con los requerimientos del dominio (formulados por los usuarios expertos). Trabajos realizados en la misma han sido publicados en [2, 4, 5]. Luego, en la segunda actividad [13], se diseñan las funcionalidades, en artefactos denominados hojas de datos funcionales. Las mismas son creadas utilizando la lista de servicios

sugeridos (provenientes de la actividad anterior) y la arquitectura de referencia (previamente diseñada por los ingenieros de software y desarrolladores). A su vez, también se deben validar estas hojas de datos (con sus modelos de variabilidad) en cuanto a inconsistencia o incompatibilidades en el diseño [16]. La tercera actividad, derivación *de componentes*, crea una estructura inicial de componentes basados en las funcionalidades que fueron traducidas a archivos XML [1, 15]. Dicha traducción es semi-automática involucrando tecnologías existentes y permitiendo llegar a una implementación en código como parte de la última actividad [3].

Actualmente, nos encontramos trabajando en dos tareas específicas. La primera, la cual está incluida dentro de la primera actividad de la Figura 1, intenta, mediante técnicas de lenguaje natural y recuperación de información, proporcionar los medios para poder elegir y evaluar variantes dentro de la construcción de LPSs, intentando reducir el esfuerzo que conlleva incorporar nuevas estrategias. A su vez, estamos avanzando dentro de la segunda actividad, en aspectos de validación y verificación de las funcionalidades y de las variabilidades incluidas. Dichos avances se enfocan en el análisis automático de modelos de variabilidad.

Resultados y Objetivos

El objetivo principal de la línea de investigación es *Desarrollar técnicas y herramientas que mejoren el desarrollo de software basado en reuso a nivel de dominios, favoreciendo el desarrollo de un*

⁶ <http://www.opengeospatial.org/>

⁷ <http://www.iso211.org/>

ambiente inteligente que permita crear LPSs según taxonomías de servicios.

En base a los trabajos previos realizados, se propone continuar la formalización de los servicios específicos dentro de un nuevo subdominio del dominio geográfico (el subdominio paleontológico), de manera de definir la forma en que los componentes son diseñados e implementados para poder ser reusados en otros subdominios. El ambiente inteligente deberá permitir crear nuevas LPS en base a la combinación de componentes reusables y específicos dentro de cada subdominio.

Formación de Recursos Humanos

El proyecto reúne aproximadamente a 13 investigadores, entre los que se cuentan docentes y alumnos de UNComa, y colaboradores. A su vez, el proyecto cuenta actualmente con un doctor y un magíster. Varios de los docentes-investigadores de GIISCo-UNComa han terminado o se encuentran próximos a terminar carreras de postgrado. Uno de ellos se encuentra finalizando el último año de la Beca CONICET Tipo II y otros poseen una beca CONICET Tipo I. A su vez, este año se incorporan 3 alumnos como becarios EVC-CIN.

Referencias

- [1] M. Arias, A. Buccella, and A. Cechich. Towards semi-automatic component derivation from an spl variability model. In *CONAIIISI 2015: 3er Congreso Nacional de Ingeniería Informatica/Sistemas de Informacion*, BsAs, Argentina, 2015.
- [2] M. Arias, A. Buccella, and A. Cechich. Búsqueda de funcionalidades basada en expansión de consultas para lps. In *Proceedings of the CACIC'16: XXII Congreso Argentino de Ciencias de la Computacion*, San Luis, Argentina, 2016.
- [3] M. Arias, A. Buccella, and A. Cechich. Smooth transition from abstract to concrete spl components: a client-server implementation for the geographic domain. In *Proceedings of the IEEE ARGENCON'16: El Congreso Bienal de la Seccion Argentina de IEEE*, Buenos Aires, Argentina, 2016. GRSS: IEEE Geoscience and Remote Sen-sing Society.
- [4] M. Arias, A. de Renzis, A. Buccella, A. Cechich, and A. Flores. Búsqueda de servicios para asistir en el desarrollo de una línea de productos de software. In *16 Simposio Argentino de Ingeniería de Software (ASSE'15) en el marco de las 44 Jornadas Argentinas de Informatica (JAIIO)*, Rosario, Argentina, 2015.
- [5] M. Arias, A. Renzis, A. Buccella, A. Flores, and Cechich. Minería basada en clasificación de componentes reusables sobre líneas de productos de software. *IEEE Latin America Transactions*, 14(2), 2016.

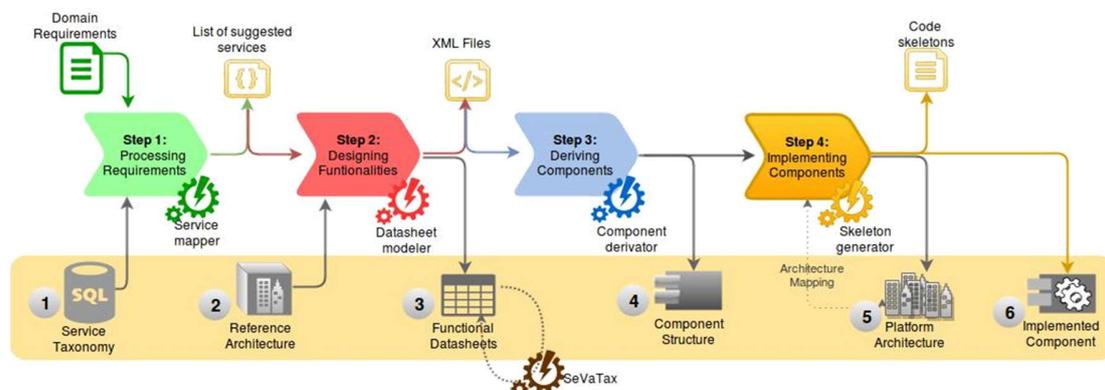


Figura 1. Actividades y artefactos de software involucrados en la ingeniería de dominio de una LPS

- [6] J. Bosch. *Design and use of software architectures: adopting and evolving a product-line approach*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 2000.
- [7] A. Buccella, A. Cechich, M. Arias, M. Pol'la, Doldan, and E. Morsan. Towards systematic software reuse of gis: Insights from a case study. *Computers & Geosciences*, 54(0):9- 20, 2013.
- [8] A. Buccella, A. Cechich, M. Pol'la, M. Arias, Doldan, and E. Morsan. Marine ecology service reuse through taxonomy-oriented SPL development. *Computers & Geosciences*, 73(0):108 – 121, 2014.
- [9] A. Buccella, M. Pol'la, M. Arias, and A. Cechich. Taxonomy-based annotations for software product line development. In *17 Simposio Argentino de Ingeniería de Software (ASSE'16) en el marco de las 45 Jornadas Argentinas de Informática (JAIIO)*, Buenos Aires, Argentina, 2016.
- [10] A. Buccella, M. Pol'la, A. Cechich, and A. Arias. A variability representation approach based on domain service taxonomies and their dependencies. In *International Conference of the Chilean Society of Computer Science (SCCC'14)*, Talca, Chile, 2014. IEEE Computer Society Press Boston.
- [11] P. C. Clements and L. Northrop. *Software Product Lines : Practices and Patterns*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.
- [12] M. Galster, D. Weyns, D. Tofan, B. Michalik, and P. Avgeriou. Variability in software systems - a systematic literature review. *IEEE Transactions on Software Engineering*, 40(3):282–306, 2014.
- [13] M. Mancuso, A. Buccella, A. Cechich, M. Arias, and M. Pol'la. Datasheet modeler: Una herramienta de soporte para el desarrollo de funcionalidades en líneas de productos de software. In *XXI Congreso Argentino en Ciencias de la Computacion*, Junin, Argentina, 2015.
- [14] K. Pohl, G. Bockle, and F. J. van der Linden. *Software Product Line Engineering: Foundations, Principles and Techniques*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
- [15] M. Pol'la, M. Arias, A. Buccella, and A. Cechich. Un sistema de anotaciones para la especificacion de componentes de una línea de productos de software. *Revista Tecnología y Ciencia de la Universidad Tecnológica Nacional*, 6(1):116–122, 2015.
- [16] M. Pol'la, A. Buccella, A. Arias, and A. Cechich. Sevatax: Service taxonomy selection & validation process for spl development. In *XXXIV International Conference of the Chilean Society of Computer Science (SCCC'15)*, Santiago, Chile, 2000. IEEE
- [17] M. Sinnema and S. Deelstra. Classifying variability modeling techniques. *Information and Software Technology*, 49:717–739, 2007.

Modelado de Negocios Orientado a Aspectos con AOP4ST

Fernando Pincirolí, Laura Zeligueta

Instituto de Investigaciones
Facultad de Informática y Diseño
Universidad Champagnat
Mendoza, Argentina, +54 (261) 424-8443
{pincirolifernando, zeliguetalaura}@uch.edu.ar

Resumen

A lo largo de la historia de la informática hemos visto, y vemos sucederse, nuevas propuestas tecnológicas, cambios de paradigmas, etc. Y, al mismo tiempo, también hemos sido testigos tanto del éxito de muchas de ellas como también de las que pasaron al olvido. Un ejemplo de esto han sido las bases de datos orientadas a objetos, las que más allá de la enorme popularidad del paradigma y de lo promisorios que sonaban los beneficios que ofrecían, no tuvieron la presencia en la industria que era de esperarse [1].

El paradigma de la orientación a aspectos también se encuentra en una fase en la que sus promesas son mayores que sus realidades. Por esta razón es que en algún momento nos planteamos el desafío de encontrar una manera de que sus beneficios pudieran verse materializados más rápidamente en la industria. Así es que diseñamos AOP4ST, acrónimo de Aspect-Oriented Process for a Smooth Transition, que es un proceso marco de desarrollo de software orientado a aspectos que ofrece la posibilidad de aprovechar los beneficios de la orientación a aspectos en las diferentes fases del ciclo de vida del desarrollo de software mediante la aplicación real y completa del paradigma, pero a través de herramientas y técnicas

hoy ampliamente difundidas y aceptadas en la industria.

En un primer proyecto de investigación impulsamos AOP4ST, y en este nuevo proyecto nos dedicamos a profundizar el primero de los modelos que lo componen, el modelo de negocio. En la fase de modelado de negocio todavía no existe una cantidad de propuestas lo suficientemente numerosas y con gran consenso [2].

Así, el objetivo de nuestro proyecto consiste en establecer un conjunto de reglas que permitan el empleo del paradigma orientado a aspectos en la fase de modelado de negocio, mediante el empleo de BPMN y de herramientas disponibles en el mercado, en forma totalmente estándar, realizando un aporte real a la disciplina de modelado de procesos de negocio.

Palabras clave: orientación a aspectos, procesos de desarrollo de software, early aspects, modelado de negocio, BPMN, AO4BPMN, separación de incumbencias, composición de incumbencias, resolución de conflictos, AOP4ST.

Contexto

El Instituto de Investigaciones de la Facultad de Informática y Diseño “Prof. Carlos Olivera”, de la Universidad

Champagnat, tiene una línea de investigación en el área de Ingeniería de Software. En ella se han desarrollado diversos proyectos de investigación en “*Desarrollo de Software Orientado a Aspectos*” (AOSD). El primero de ellos, “*Procesos de desarrollo de software de calidad basados en aspectos*”, se llevó a cabo con la UTN Facultad Regional Mendoza, continuó en nuestra Universidad con el proyecto “*Definición de criterios para la detección temprana de aspectos en el modelado de negocios y el desarrollo de los requisitos*” (presentado en WICC’15) y que continuó con este nuevo proyecto, “*Modelado de procesos de negocio orientados a aspectos con BPMN*”, iniciado a comienzos de 2016 y que tiene como alcance la primera fase del ciclo de vida del desarrollo de software, al plantearse como objetivo el definir un proceso de modelado orientado a aspectos; sus primeros avances fueron presentados en WICC’16. También se realizaron diversas publicaciones en congresos nacionales e internacionales. Este proyecto cuenta con la financiación de la Universidad Champagnat y recibe aportes de la empresa Aconcagua Software Factory S.A. de la provincia de Mendoza.

Introducción

Ya se han discutido largamente los beneficios del paradigma orientado a aspectos en las diferentes fases del ciclo de vida del desarrollo de software, pero no existe una propuesta completa que cubra todas las fases del ciclo de vida de manera uniforme, tal como lo demuestra el estudio de Aws et al. [4]. Por este motivo es que hemos elaborado una propuesta llamada AOP4ST, con la que pretendemos cubrir ese espacio.

AOP4ST cubre el ciclo de vida del desarrollo de software desde el modelo de negocio en adelante. En este proyecto de

investigación hemos querido hacer especial hincapié en ese modelo, intentando ofrecer una alternativa de modelado que, más allá de que sea el punto de partida para el resto del ciclo de vida, sea completa en sí misma, además de otras virtudes, y que los beneficios del empleo de la orientación a aspectos puedan verse reflejados en cada uno de los modelos desde el comienzo.

Así, esperamos poder administrar las incumbencias que están *desparramadas y enredadas* en cada uno de los niveles de abstracción a lo largo del ciclo de vida completo.

En esta etapa de modelado de negocio procuramos lograr modelos –y, por consiguiente, sistemas al final del ciclo de vida– más modulares, mantenibles, reusables, extensibles, comprensibles, etc.

En el modelo de negocio pretendemos alcanzar esos mismos objetivos por medio de la separación de las cuestiones que son específicas del dominio del problema de aquellas otras que no lo son, razón por la que se las conoce como *incumbencias transversales*, por estar presentes no solamente en la mayoría de los procesos de negocio a los que atraviesan en forma transversal, sino también a través de diferentes dominios de problema.

Así, tras detectar las *incumbencias transversales* en el modelo de negocio es posible encapsularlas para separar el conocimiento específico del dominio del problema de aquellos otros elementos que no pertenecen a él.

Esto que acabamos de mencionar se conoce como *enfoque asimétrico*, ya que se hace una diferencia entre los intereses específicos del dominio del problema de aquellos que no lo son, pero existe también un *enfoque simétrico*, que consiste en la separación de las incumbencias que pertenecen al propio dominio de problema. Las incumbencias del primer enfoque se

encapsulan como *aspectos impares* y las del segundo como *aspectos pares*.

Si bien se puede observar una aparente inclinación hacia el enfoque asimétrico, las prácticas habituales de modelado de procesos de negocio también van preparando el terreno para la detección y el encapsulamiento de los aspectos pares.

Además de la detección y del encapsulamiento de las incumbencias transversales, en nuestra propuesta tenemos en cuenta su posterior composición en base a reglas, la resolución de los posibles conflictos y la mejora global de nuestra propuesta en el concierto de las propuestas existentes.

Líneas de Investigación, Desarrollo e Innovación

Los proyectos de investigación que se vienen sucediendo en nuestra Facultad presentan cuatro ejes de investigación, donde los tres primeros se ven integrados con el cuarto:

1. *Modelado de procesos de negocio orientados a aspectos.*
2. *Separación y composición de incumbencias con resolución de conflictos.*
3. *Especificación de requisitos y gestión de incumbencias con casos de uso.*
4. *Procesos de desarrollo de software orientados a aspectos.*

Resultados y Objetivos

Tras un año de proyecto, hemos avanzado en nuestra técnica de modelado de procesos orientado a aspectos siguiendo el criterio de evaluación de este tipo de metodologías de modelado propuesta por Jalali [5]. Inicialmente habíamos hecho una evaluación de nuestra propuesta utilizando este método, pero luego la mejoramos a partir de los requisitos no

cubiertos que exige la técnica de Jalali. Así, logramos ubicar a la fase de modelado orientado a aspectos de AOP4ST al tope de las técnicas evaluadas por ese autor (Tabla 1), trabajo que está en proceso de publicación [10].

Propuesta	Calificación
AOP4ST	12,46
Jalali et al. [2]	10,87
Cappelli et al. [6]	7,47
AO4BPMN [7]	6,14
Jabeen et al. [8]	6,14
Patiniotakis et al. [9]	5,47

Tabla 1. Calificaciones de las propuestas de modelado de negocio orientado a aspectos.

Aun así, todavía quedan oportunidades de mejora, ya que AOP4ST y la propuesta de Jalali satisfacen requerimientos diferentes, por lo que tenemos un buen margen para seguir enriqueciéndonos con esta última.

Entre los principales requisitos que debe cumplir los enfoques de modelado orientados a aspectos que siguieron la evaluación mencionada, se encuentra principalmente su *capacidad de modelado y notación*. En AOP4ST hemos partido de la notación *AO4BPMN* presentada por Charfi et al. [7], hoy ampliamente aceptada, y que si bien consideramos muy adecuada, tuvimos que realizarle algunos ajustes para que se mantenga dentro de la notación estándar *BPMN 2.0* y para poder indicar la inserción de los advices en múltiples join points en forma simultánea al estilo *AspectJ*.

Para la *detección de incumbencias* en esta etapa de modelado de procesos de negocio ya habíamos publicado nuestras ideas en la revista “Perspectivas em Ciências Tecnológicas - Revista do Curso de Ciência da Computação” [11].

También contemplamos los *enfoques asimétrico y simétrico* desde esta etapa del

ciclo de vida [12] [13]. Tras la separación de incumbencias y su refactorización, hemos desarrollado dos modelos, de join points y de incumbencias, ideas que nos permitan alcanzar dos objetivos: analizar los modelos completamente integrados y detectar y resolver los posibles conflictos que podrían aparecer.

Con respecto al análisis de los modelos integrados, hemos elaborado un conjunto de *reglas de composición*. Para ello estamos recorriendo la totalidad de la notación de BPMN 2.0, incluyendo la combinación de elementos, la vista dinámica de los procesos, los procesos con instancias múltiples y los patrones de modelado de procesos de negocio, sumado a la revisión de una amplia casuística que nos permita emitir nuestras opiniones adecuadamente. Estas reglas fueron aceptadas en nuestro trabajo presentado en el SCCC CLEI 2016 y que se encuentra disponible en IEEE Xplore [14].

En cuanto a la *resolución de conflictos*, nos fue publicado un trabajo el año pasado en Elsevier Science, en Procedia Computer Science series, que se encuentra disponible on-line [15], y que presentamos en Madrid. Aunque fuera de este proyecto, aplicamos estos mismos criterios para los sistemas interactivos, cuyos resultados fueron presentados y aceptados en CACIC 2016 [16], artículo que fue seleccionado entre los mejores trabajos para una publicación especial.

Para el análisis de las relaciones entre aspectos hemos avanzado con dos modelos específicos. El primero de ellos es el *modelo de join points*, que permite analizar los puntos de los procesos en los que se insertan los advices, la relación entre advices que se componen y la posibilidad de alertar sobre los puntos de los procesos en los que se deben aplicar soluciones ante conflictos que fueran aplicadas en otros puntos de los procesos. El segundo modelo es el *modelo de incumbencias*, que

presenta la relación entre incumbencias al nivel de los procesos, también para favorecer el análisis y toma de decisiones para la mejora de las incumbencias detectadas hasta ese momento.

También hemos presentado AOP4ST por completo, incluyendo esta fase de modelado de procesos, en el IEEE 11 CCC realizado en año pasado en Colombia [17] y estamos desarrollando un estudio de mapeo sistemático para ser publicado en una revista internacional, cuyo protocolo está disponible en ArXiv [18].

Los resultados de este proyecto ya han sido aplicados en proyectos de Aconcagua Software Factory S.A., una de las principales fábricas de software de Mendoza y que se aplicarán en un proyecto de BPM para la Administración Tributaria de Mendoza.

Formación de Recursos Humanos

El proyecto apunta a la formación del equipo de profesores del área de Ingeniería de Software y de los alumnos y egresados de la Universidad Champagnat.

AOP4ST es el tema central de la tesis del doctorando Fernando Pinciroli, en el Doctorado en Ciencias Informáticas de la Universidad Nacional de San Juan, bajo la dirección del Dr. Raymundo Forradellas, de la Universidad Nacional de Cuyo, y la codirección del Dr. José Luis Barros Justo, de la Universidad de Vigo, España.

También se está elaborando la tesis de Gustavo Albino, de la Maestría en Ingeniería de Software de la Universidad Nacional de San Luis, y dos tesinas de grado de alumnos de nuestra Universidad.

Referencias

- [1] N. Leavitt, "Whatever Happened to Object-Oriented Databases?," Computer, California, vol. 33, pp. 16–19, 2000.

- [2] A. Jalali, P. Wohed, C. Ouyang, and P. Johannesson, "Dynamic Weaving in Aspect Oriented Business Process Management," OTM 2013 Conf., vol. 8185, pp. 2–20, 2013.
- [3] B. Johansson, B. Andersson, and N. Holmberg, *Perspectives in Business Informatics Research*. Springer, 2014.
- [4] M. Aws, S. Zarina, and M. A. Noorazeen, "Systematic Review on Aspect-Oriented UML Modeling: A Complete Aspectual UML Modeling Framework," *J. pf Appl. Sci.*, vol. 13, no. 1, 2013.
- [5] A. Jalali, "Assessing Aspect Oriented Approaches in Business Process Management," in *Perspectives in Business Informatics Research, 13th International Conference, BIR 2014, 2014*, pp. 231–245.
- [6] C. Cappelli, F. M. Santoro, J. C. S. do P. Leite, T. Batista, A. L. Medeiros, and C. S. C. Romeiro, "Reflections on the modularity of business process models: The case for introducing the aspect-oriented paradigm," *Bus. Process Manag. J.*, vol. 16, no. 4, pp. 662–687, 2010.
- [7] A. Charfi, H. Müller, and M. Mezini, "Aspect-oriented business process modeling with AO4BPMN," *Lect. Notes Comput. Sci.*, vol. 6138 LNCS, pp. 48–61, 2010.
- [8] A. Jabeen, S. Tariq, Q. U. A. Farooq, and Z. I. Malik, "A lightweight aspect modelling approach for BPMN," *Proc. 14th IEEE Int. Multitopic Conf. 2011, INMIC 2011*, pp. 255–260, 2011.
- [9] I. Patiniotakis, N. Papageorgiou, Y. Verginadis, D. Apostolou, and G. Mentzas, "An aspect oriented approach for implementing situational driven adaptation of BPMN2.0 workflows," *Lect. Notes Bus. Inf. Process.*, vol. 132 LNBIP, pp. 414–425, 2013.
- [10] F. Pinciroli, J. L. Barros Justo, and R. Forradellas, "Aspect-Oriented Business Process Modeling Approaches: An assessment of AOP4ST," (en publicación).
- [11] F. Pinciroli, "Considerações acerca da mineração de aspectos," *Perspect. em Ciências Tecnológicas*, vol. 5, no. 5, pp. 83–101, 2016.
- [12] J. Bálík and V. Vranić, "Symmetric aspect-orientation: some practical consequences," *Proc. 2012 Work. Next.*, pp. 7–11, 2012.
- [13] D. C. Collell, "Aspect-Oriented Modeling of Business Processes," no. September, 2012.
- [14] F. Pinciroli, "Aspect-oriented business process composition rules in AOP4ST," in *35th International Conference of the Chilean Computer Science Society (SCCC 2016)*, 2016, pp. 1–6.
- [15] F. Pinciroli, "Improving software applications quality by considering the contribution relationship among quality attributes," *Procedia Comput. Sci.*, vol. 83, pp. 970–975, 2016.
- [16] F. Pinciroli, "An HCI quality attributes taxonomy for an impact analysis to interactive systems design and improvement," in *XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*, 2016.
- [17] F. Pinciroli, "Aspect-Oriented Process for a Smooth Transition," in *Ph.D. Symposium of the IEEE 11 CCC*, 2016.
- [18] F. Pinciroli, L. Zeligueta, M. I. Lund, and J. L. Barros Justo, "Systematic Mapping Protocol - Coverage of Aspect-oriented Methodologies for the Early Phases of the Software Development Life Cycle," 2017.

Modelo de Gestión de la Calidad Orientado a Empresas de Software y Servicios Informáticos (SSI) de la República Argentina

Marina C. Gette, Alberto Sánchez, Carlos Salgado, Mario Peralta.
Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina
marinagette@gmail.com, {alfanego, csalgado, [mperalta](mailto:mperalta@unsl.edu.ar)}@unsl.edu.ar

Resumen

Los modelos de calidad son referencias que las organizaciones utilizan para mejorar su gestión. Los modelos, a diferencia de las normas, no contienen requisitos que deben cumplir los sistemas de gestión de la calidad sino directrices para la mejora. Existen modelos de calidad orientados a la calidad total y la excelencia, modelos orientados a la mejora, modelos propios de determinados sectores e incluso modelos de calidad que desarrollan las propias organizaciones [1].

Por ello, se propone un modelo de Gestión de la Calidad orientado a empresas del Sector Software y Servicios Informáticos, basado en los 7 principios de la calidad de las organizaciones [2]: Enfoque al cliente, Liderazgo, Enfoque a procesos, Compromiso de las personas, Mejora, Toma de decisiones basada en la evidencia y Gestión de las relaciones.

Para el desarrollo del modelo propuesto se tuvo en cuenta la estructura de la Norma Internacional ISO 9004 [3] e información actualizada del Sector Software y Servicios Informáticos brindada por

Cámara de Empresas de Software y Servicios Informáticos de la República Argentina. El objetivo del modelo es brindar a las empresas de desarrollo de software de la región una herramienta que les permita posicionarse en un nivel altamente competitivo en el mercado actual, mediante la producción de software de calidad.

Palabras Claves: Modelo de Gestión de Calidad, Calidad, mejora continua.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

Introducción

La Calidad es grado en el que un conjunto de características inherentes a un objeto (producto, servicio, proceso, persona, organización, sistema o recurso) cumple con los requisitos. [4].

La Calidad establece que es el grado predecible de uniformidad y fiabilidad a un bajo costo y que se ajuste a las necesidades del mercado. La calidad no es otra cosa más que "una serie de cuestionamiento hacia una mejora continua". [4].

La calidad es entregar a los clientes y a los compañeros de trabajos, productos y servicios sin defectos y hacerlo a tiempo. [4].

La calidad es el conjunto de características que satisfacen las necesidades de los clientes. Además según Juran, la calidad consiste en no tener deficiencias. La calidad es "la adecuación para el uso satisfaciendo las necesidades del cliente". [4]

El concepto de Calidad ha evolucionado con el transcurso del tiempo y se extiende a las empresas en su conjunto, pasando a tener prioridad máxima cuando el cliente tiene mayores posibilidades de elección, motivo por el cual la exigencia sobre los productos y servicios que compra es cada vez mayor.

En la actualidad, para enfrentar estos desafíos de productividad y competitividad, las organizaciones del sector software y servicios informáticos analizan la utilización de modelos de gestión como herramienta para cubrir las necesidades y expectativas de todas las partes interesadas, que estimulen la mejora sistemática y continua como así

también la innovación a largo plazo en la organización.

Para gestionar mejor la organización en su conjunto surgió la familia de normas ISO 9000. Esta familia de normas recoge los requisitos que debe cumplir una organización para tener un buen sistema de gestión [5] pero también existen los modelos de gestión de la calidad, éstos no tienen carácter de norma. Su objetivo es materializar la filosofía de la calidad total de modo que ayuden a las organizaciones a gestionarse en el día a día, de tal modo que toda la organización se oriente hacia el logro de la excelencia. [5]

La adopción de un modelo de gestión de la calidad como puede ser el Modelo para la gestión de excelencia de empresas [6] lleva implícito un compromiso de mejora continua lo cual permite lograr mejores niveles de efectividad y por ende productividad.

Líneas de investigación

En los últimos años se evidenció un significativo incremento en la cantidad de empresas del sector de Software y Servicios Informáticos (SSI) en la República Argentina, según datos oficiales, actualizados por el Observatorio de Empleo y Dinámica Empresarial del Ministerio de Trabajo, el número de empresas creció un 132% en 10 años consecutivos [7].

Debido a este fenómeno, las empresas SSI se ven en la necesidad de diferenciarse de la competencia para captar nuevos clientes como así también mantener el crecimiento al ritmo de los avances tecnológicos y los requerimientos del mercado.

Muchas empresas del sector no se

encuentran aplicando un modelo de gestión integrado que esté orientado a los clientes, el liderazgo, los procesos, los resultados y la mejora continua lo cual las aleja de la excelencia y por ende son poco productivas. Si bien un importante porcentaje de empresas de software, alrededor del 60% según los últimos resultados publicados [7], poseen algún tipo de certificación no cuentan con un modelo a seguir para su éxito a largo plazo.

Se propone en esta línea de investigación diseñar un Modelo de Gestión de la Calidad para el sector Software y Servicios Informáticos basado en el enfoque de Gestión por Procesos orientado a los clientes, el liderazgo directivo, el personal, los resultados y la mejora continua.

La utilización de un Modelo de referencia para la gestión de la Calidad en las empresas SSI tendrá como mayor beneficio la mejora continua en sus procesos lo cual impactará directamente en la satisfacción de las partes interesadas, mayormente en los clientes.

Tres son los propósitos generales que componen el Modelo de gestión de la calidad para las empresas SSI:

1. Contribuir en la mejora continua de los procesos de las empresas SSI como elemento clave para su desempeño considerando los resultados.
2. Brindar un aporte a las empresas SSI de la República Argentina que se alineen con el modelo para obtener un reconocimiento de Calidad.
3. Servir de referencia para un posterior proceso de evaluación como herramienta de mejora continua.

Resultados obtenidos y objetivos

En base a todo lo expresado anteriormente, se definió un modelo formado por 6 componentes interrelacionados los cuales pueden ser considerados por las empresas SSI de la República Argentina para poder utilizarlo como referencia y así mantener el éxito a largo plazo.

Como componente número uno se encuentra la Gestión para el éxito sostenido para empresas SSI, aquí se analizan dos aspectos fundamentales: el primero es el entorno de la organización, el entorno es una combinación de factores y de condiciones internas y externas que pueden afectar al logro de los objetivos de una organización y a su comportamiento hacia las partes interesadas.

Las empresas SSI deben realizar seguimientos constantes de su entorno que, primeramente, los deben tener correctamente identificados.

Como segundo aspecto se encuentra: el Conocimiento de las necesidades y expectativas de las partes interesadas, esto significa que se evalúa cómo la empresa SSI examina los requerimientos, preferencias y expectativas de las partes interesadas ya sea tanto de su ambiente interno como del ambiente externo. Y cómo utiliza esa información para entender y anticiparse a sus necesidades. Para que las empresas SSI aumenten su capacidad de respuesta a las necesidades y expectativas de las partes involucradas deben realizar revisiones identificando los posibles cambios de las mismas y asignar prioridades para poder crear planes de acción para su desarrollo y mantenimiento.

Como componente número dos se encuentra la Estrategia y Política: Las empresas SSI deben desarrollar procedimientos para establecer, adoptar y mantener políticas eficaces y estrategias para lograr los objetivos establecidos por la Alta Dirección, en el ámbito de sus valores, visión y misión.

Como componente número tres encontramos la Gestión de los Recursos: Las empresas SSI deben identificar los recursos internos y externos necesarios para lograr sus objetivos a corto y largo plazo, explorando continuamente su propio ambiente interno y externo. Los recursos que utilizan las empresas SSI para desarrollar sus actividades son:

- Recursos económicos y financieros.
- Personas en la organización.
- Proveedores y aliados.
- Infraestructura.
- Ambiente de trabajo.
- Conocimiento, información y tecnología.
- Recursos naturales.

Gestión de los procesos es considerado el cuarto componente: Todas las actividades que se llevan a cabo en una empresa SSI pueden y deben considerarse como procesos, son específicos de cada una y varían según el tamaño y nivel de madurez. Para desempeñarse de manera eficaz y eficiente, una empresa SSI, tiene que identificar y realizar planificaciones de sus procesos, revisarlos de forma regular y realizar los ajustes pertinentes a fin de lograr sus objetivos.

El Seguimiento, medición, análisis y revisión forma parte del quinto componente del modelo: Las empresas SSI realizan seguimientos continuos del entorno en el cual se encuentran inmersos,

miden los valores obtenidos, los analizan y vuelven a revisar de manera regular para tomar decisiones eficaces y mejorar así el desempeño a largo plazo.

Como último elemento del modelo tenemos la Mejora, innovación y aprendizaje: Las empresas SSI deben establecer procedimientos de mejora continua en diferentes elementos como ser sus productos, procesos, interfaces, la estructura organizacional, el sistema de gestión, en aspectos humanos y culturales, en infraestructura, ambiente de trabajo, la tecnología y las relaciones con las partes interesadas. Para poder realizar innovaciones en sus productos y/o servicios utilizan como base el aprendizaje continuo de sus casos de éxito y fracasos.

Próximos Trabajos

Se identifican los principales trabajos futuros y/o derivados del presente trabajo de investigación:

Desarrollo de las bases y condiciones del modelo a la calidad para empresas SSI para postularse a reconocimientos de Calidad para el sector.

Método de evaluación para aplicar el modelo del presente trabajo de investigación.

Desarrollo de herramientas de autoevaluación que sirvan de referencia para determinar el nivel de madurez actual de la empresa SSI y constituyan una base para la mejora continua.

Formación de Recursos Humanos

Bajo esta línea, en el grupo de investigación se está desarrollando una tesis de Maestría en Calidad del Software (Plan Ord. 017/09-CD) de la Ing Marina mmmm C Gette, como así también se están

llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación.

Referencias

- [1] Asociación Española para la Calidad
<http://www.aec.es/web/guest/centro-conocimiento/modelos-de-calidad>
- [2] ISO, “ISO Sistema de Gestión de la Calidad – Fundamentos y Vocabularios 9000:2015”, 2015.
- [3] ISO, “Gestión para el éxito sostenido de una organización – Enfoque de Gestión de la Calidad”, 9004:2009.
- [4] Piattini, Mario; García, Félix; García, Ignacio; Pino, Francisco. Calidad de Sistemas de Información: Alfaomega RA-MA, 2011.
- [5] Sangüesa, Mateo, Ilzarbe, Teoría y Práctica de la Calidad: Thomson Editores Spain, 2006.
- [6] Modelo para una Gestión de Excelencia Empresas, Sector privado, Publicado por Fundación Premio Nacional a la Calidad, edición 2013.
- [7] Cámara de la Industria Argentina de Software,
<http://www.cessi.org.ar/opssi>

Modelo de Sistema de Gestión de Calidad para un Instituto de Investigación Universitario

Jorge A. Silvera, Daniel Arias Figueroa, Gustavo Gil, Loraine Gimson, Valeria González

Centro de Investigación y Desarrollo en Informática Aplicada (C.I.D.I.A.)

Facultad de Ciencias Exactas – Universidad Nacional de Salta – Salta - Argentina

jsilvera@unsa.edu.ar, gdgil@unsa.edu.ar, daaf@cidia.unsa.edu.ar,

loraine@cidia.unsa.edu.ar, vgonzalez@cidia.unsa.edu.ar

RESUMEN

El presente trabajo propone un modelo de software resultado de la aplicación de la metodología de desarrollo de software Webml para la gestión de la calidad del El Centro de Investigación y Desarrollo en Informática Aplicada (C.I.D.I.A.), dependiente de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta (U.N.S.a).

Palabras claves:

Normas de Calidad ISO 9.001, Sistema de Gestión de Calidad (SGC), Institutos de Investigación, Herramientas informáticas de apoyo al SGC. Metodología WebML (Web Modeling Language).

CONTEXTO

El Centro de Investigación y Desarrollo en Informática Aplicada C.I.D.I.A. funciona en la Facultad de Ciencias Exactas de la Universidad Nacional de Salta desde junio del año 2.002, fue aprobado bajo RES. C.D. CS. Ex. N° 167/02 – Expediente N° 8545/01. El mismo nace como fruto de las experiencias anteriores como la del Laboratorio de Investigación y Desarrollo en Tecnologías Informáticas L.I.D.T.I. que funcionaba desde el año 1.997 y que dependía del Departamento de Matemática de la misma Facultad.

Entre los objetivos generales del C.I.D.I.A. podemos mencionar:

- Nuclear y apoyar a un importante grupo de docentes, profesionales e investigadores en áreas directamente vinculadas con las Nuevas Tecnologías Informáticas, facilitando las tareas de investigación y desarrollo.
- Desarrollar investigaciones científicas individuales o interdisciplinarias, principalmente en los campos de las Redes de Comunicaciones de Datos, Internet, Informática aplicada en Educación y la Ingeniería de Software.
- Promover la capacitación en nuevas tecnologías informáticas a todos los niveles.

Dentro de las actividades que desarrolla podemos destacar los cursos de especialización, actualización profesional y postgrado, a los fines de incrementar el nivel de conocimiento y facilitar así la formación continua, como así también las acciones y transferencias al medio.

El presente trabajo se lleva a cabo dentro del marco del proyecto de investigación Nro: 2278/0 “Estudio de la familia de normas iso 9.000 y su aplicación a centros educativos”, aprobado en el año 2.014 por el consejo de investigación de la Universidad Nacional de Salta.

1. INTRODUCCION

Se planteó desde dos puntos de vistas diferentes, por un lado se estudian y definen directrices que guían a los centros e institutos

universitarios de investigación, en cuanto a la implementación de un sistema de gestión de calidad (SGC) eficaz que cumpla los requisitos de la norma ISO 9001:2008. Por otro lado se plantea el análisis, diseño e implementación de un software de apoyo al SGC para el Centro de Investigación y Desarrollo en Informática Aplicada (C.I.D.I.A.), utilizando como metodología Web Modeling Language (WebML) y la herramienta CASE WebRatio. Se espera que la presente investigación sirva de referencia para cualquier implementación de gestión de la calidad en el ámbito de la investigación formal universitaria.

A modo de breve reseña histórica, en el año 1.987 se obtuvo la primera referencia europea a la calidad en investigación y desarrollo (I+D), que analiza la aplicabilidad de la norma ISO 9.001:2008 a las actividades de I+D [Alonso, P. M. 2005].

La implementación exitosa de un sistema de calidad (SGC) aporta un gran número de beneficios a las organizaciones en general, y en este caso concreto a las organizaciones vinculadas a la investigación universitarias, que apuestan por esta estrategia, logrando no solo reducir sus costos de manera razonable, sino que además ayudan a lograr la tan preciada satisfacción de sus usuarios, lo cual nos brinda una gran motivación para los integrantes de la misma.

Promover la calidad en investigación es tratar de mejorar de forma continua sus prácticas de forma que permitan:

- Garantizar los resultados y productos obtenidos.
- Asegurar la trazabilidad de los procesos que se ejecutan en la organización.

Cabe aclarar que no trata de la calidad de la investigación en sí misma, la cual es evaluada por pares investigadores en los campos de cada área científica, sino de los procesos que

se llevan a cabo en el instituto de investigación, en este caso el C.I.D.I.A.

Desde el momento de su creación, los requisitos de la norma internacional ISO 9.001:2008 son genéricos y aplicables a todas las organizaciones sin importar su tipo, tamaño o producto o servicio suministrado. En particular el SGC en un instituto de investigación universitario también debe tener en cuenta a los servicios de administración y gestión que deben proveer apoyo a los grupos de investigación, facilitando que los mismos lleven a cabo sus actividades.

La gestión de la calidad en la investigación debe ser sobre todo flexible y adaptada a las necesidades específicas de los investigadores, pudiendo conceptualizarse tres fases asociadas a la investigación:

- Definición de los objetivos: En esta fase hay que lograr la identificación de los criterios de satisfacción de las diferentes partes interesadas.
- Realización de la investigación: Las cuestiones a tener en cuenta en esta fase tienen que ver con los diferentes procesos a ejecutar, sus interacciones y gestión eficiente.
- Valoración y puesta en valor de los resultados: Esta fase es esencial pues contribuye al reconocimiento del organismo de investigación y en ella se concretan los esfuerzos y la razón de ser de los investigadores.

La representación de las tres fases principales de la investigación junto con los principios de gestión de calidad aplicables se muestra en la siguiente figura:

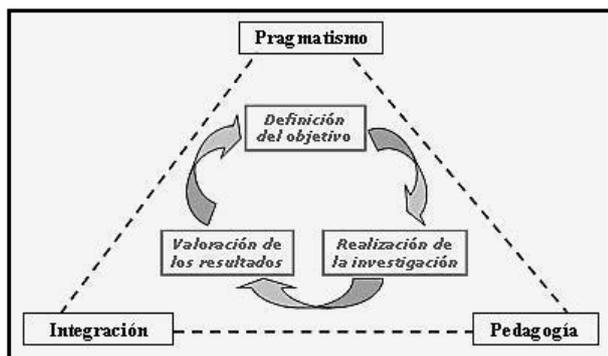


Figura 1- Fases principales de la investigación.

Podemos definir una metodología para implementar el SGC en el área de la investigación, utilizando el esquema anterior de manera sistemática, dando lugar a un ciclo de mejora continua, tal y como se representa en la siguiente figura:

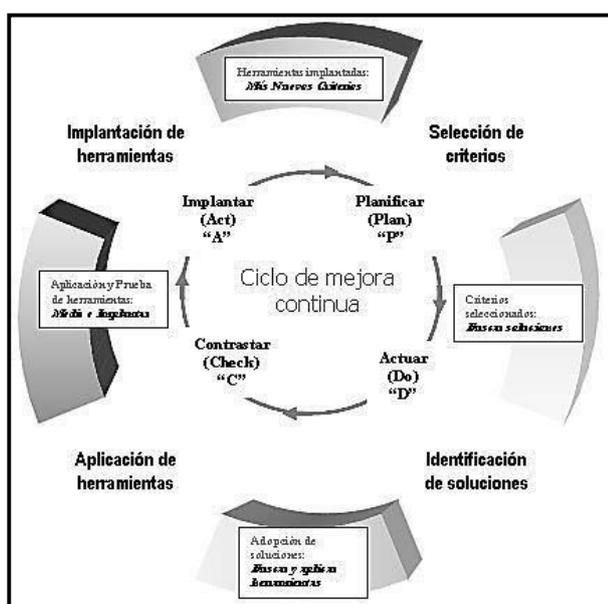


Figura 2- Ciclo de mejora continua.

Como hemos podido observar, en base a la investigación bibliográfica, un sistema de gestión de la calidad se basa en la gestión de muchos documentos, de manera estricta y minuciosa. Por esta razón es fundamental el apoyo de la tecnología informática. No alcanza con un procesador de textos y un espacio de almacenamiento compartido en un servidor de archivos.

Todo documento relacionado con el sistema de gestión de la calidad debe ser desarrollado

a través de un proceso perfectamente documentado. Además, una vez aprobado, debe estar disponible para todos quienes participan en el SGC, y por otra parte no disponible para quienes no están involucrados en el. A su vez las sucesivas revisiones deben quedar claramente identificadas así como los cambios realizados. También, si los documentos obsoletos se mantienen en el SGC para poder ser consultados deben quedar claramente identificados como obsoletos para impedir que sean utilizados como actuales, por error. Por ejemplo, debe impedirse su modificación incluso a personal autorizado para generar documentos.

Esto es claramente el manejo de una base documental asociada a un proceso de decisión y de elaboración conjunta de los que típicamente se realizan con herramientas de trabajo en grupos.

En función a lo expresado, tomando como referencia las especificaciones y directrices investigadas por el C.I.D.I.A. para la aplicación de la familia de norma ISO 9.001:2008 se realizó el análisis y diseño de un SGC, utilizando la metodología WebML, logrando la implementación de un prototipo funcional de una herramienta integrada y genérica para acompañar una implementación de calidad bajo los requisitos de la norma ISO 9001:2008 en un centro educativo. Posteriormente el prototipo se convirtió en producto final.

Para la especificación de requerimientos de calidad se utilizó el estándar ISO/IEC 25030:2007 que es parte del standard internacional SQuARE, de esta manera los requisitos de usuarios son expresados como requisitos funcionales.

El enfoque de la metodología WebML combina componentes tradicionales bien conocidos por los desarrolladores, como el diseño conceptual de datos usando el modelo Entidad Relación y la especificación de los casos de usos usando UML, con nuevos conceptos y métodos para el diseño de

hipertextos, que son fundamentales para el desarrollo web. No obstante, el valor del enfoque de la propuesta no está en los componentes individuales, sino en la definición de un marco sistemático para que las actividades de desarrollo de aplicaciones web pueda ser organizado de acuerdo a los principios fundamentales de la ingeniería de software, de modo que todas las tareas encuentren un soporte adecuado a partir de los conceptos, notaciones y técnicas propuestas por la metodologías. La característica distintiva de este marco de desarrollo es el énfasis en el modelo conceptual, el cual ha sido probado con éxito en muchos campos de la ingeniería del software, y en diseño de base de datos, donde el modelo Entidad Relación ofrece una notación de alto nivel e intuitiva para la comunicación de los requisitos de información entre los diseñadores y no técnicos, y es la base para la creación de esquemas de bases de datos de alta calidad.

En esencia, WebML consiste en gráficos simples y conceptuales para expresar un hipertexto como un conjunto de páginas y operaciones, WebML representa una página como una estructura compuesta por unidades de contenido y links.

WebML propone que un sitio web conste de tres grandes partes conceptuales, la estructura, el hipertexto y la presentación.

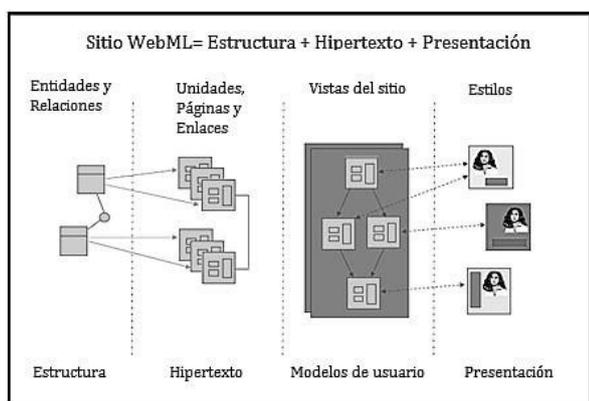


Figura 3- Conceptos principales de WebML.

La metodología está compuesta por la creación de los Modelos de Datos, Hipertexto, Presentación y Personalización. En este

trabajo se desarrollaron todos los modelos de la metodología y para la implementación se utilizó la herramienta WebRatio con licencia de uso gratuito para el modelado de procesos de negocios (BPM).

En síntesis, los resultados obtenidos a la fecha son los siguientes:

- El Análisis y diseño de un Sistema de Gestión de Calidad para el C.I.D.I.A. junto a la especificación de requerimientos, aplicados a los procesos específicos relacionados a la investigación.
- Un prototipo funcional de un Sistema de Gestión de Calidad para institutos de investigación

3. LINEAS DE INVESTIGACION y DESARROLLO

Los principales ejes temáticos que se están investigando son los siguientes:

- ✓ Tecnología Informática aplicada en Educación.
- ✓ Gestión de Calidad aplicada a Institutos de Investigación Universitaria.
- ✓ Herramientas informáticas para la implementación de un SGC ISO 9.001.
- ✓ Aplicación de la metodología WebML (Web Modeling Language) para el diseño de una plataforma integral que apoye a la implementación de SGC ISO 9.001.

4. FORMACION DE RECURSOS HUMANOS

La estructura del equipo de investigación es de 5 (cinco) miembros incluidos el Director y Co-director.

Uno de sus miembros obtuvo la Especialidad en Ingeniería de Software de la Universidad Nacional de La Plata, con el trabajo “Sistema de Gestión de Calidad bajo Normas ISO”.

Otro de sus miembros alcanzo el título de Licenciado en Análisis de Sistemas, otorgado por de la Universidad Nacional de Salta, con la tesis “Análisis y Diseño de un Sistema de Gestión de Calidad basado en la Norma ISO 9.001”.

Además, otro de sus miembros se encuentra realizando el trabajo final del Magister en Administración de Negocios de la Universidad Católica de Salta denominada “Calidad ISO en Centros de Capacitación y Formación”.

Continuamos con la dirección de tesis de grado de alumnos de la carrera Licenciatura en Análisis de Sistemas de la Universidad Nacional de Salta.

4. RESULTADOS OBTENIDOS/ESPERADOS

En el presente trabajo se ha abordado la gestión de los sistemas de calidad desde una perspectiva de los centros de investigación y capacitación.

Consideramos que, además de todas las herramientas disponibles, la aplicación de la familia de normas ISO 9001:2008 junto con el apoyo de un modelo adecuado de software, con la especificación de requerimientos adecuado, se constituye en una estrategia importante para alcanzar la satisfacción de los alumnos e investigadores.

Actualmente se está gestionando la posibilidad de alcanzar la certificación por algunos de los entidades certificadoras, tales como el IRAM y BUREAU-VERITAS.

Por último se está estudiando los impactos de la nueva norma ISO 9.001:2.015, incorporando nuevos requisitos a cumplir que deben ser estudiados y evaluados, buscando el impacto en los SGC implementados anteriormente.

5. BIBLIOGRAFIA

- Arias Figueroa, Daniel y otros (2.011). Normas ISO y su Aplicación en Centros Educativos. XIII Workshop de Investigadores en Ciencias de la Computación. ISBN 978-950-673-892-1.
- Ishikawa, Kaoru (1.997). ¿Qué es el control de la Calidad?: La modalidad Japonesa. Bogotá: Grupo Editorial Norma.
- Fernández, A. (2.006.). Implantación de un sistema de gestión de la calidad Norma ISO 9.001:2.001. Centro de la Calidad de Asturias/ Instituto de Fomento Regional.
- Norma ISO 9.001 (2.008) Elaborada por el Comité Técnico ISO/TC176 de ISO.
- Guía de Interpretación de la IRAM-ISO 9.001:2008 para la educación. IRAM 30.000:2.001.
- Garbarini, R., Cigliuti, P., Burstyn, A., Pollo-Cattaneo. (2.013). Implementación de un Sistema de Gestión de Calidad y Servicios en Laboratorio Universitario de Ingeniería en Sistemas de Información Implementación de un Sistema de Gestión de Calidad y Servicios en Laboratorio Universitario de Ingeniería en Sistemas de Información. VIII Congreso de Tecnología en Educación y Educación en Tecnología. Quilmes. ISBN 978-987-1676-04-0.
- Orthustegui, F. y otros. (2.011). Análisis del Impacto de la Implementación de un Sistema de Gestión de la Calidad Interno para los Laboratorios de Enseñanza e Investigación de la UNLaM. XXXVI Jornadas IRAM Universidades – XXIII - Foro UNILAB. UNSL.
- Alonso, P. M. (2.005). Revista Madri+d, Número 32. Calidad En Investigación 1ª Parte. De qué trata la Gestión de Calidad en Investigación.
- Alonso, P. M. (2.005). Calidad En Investigación 2ª Parte. Aproximación metodológica a la mejora de las actividades de investigación. Revista Madri+d, Número 33.
- Web Modeling Language, sitio oficial de la Metodología <http://www.webml.org>.

- WebRatio, sitio oficial de la herramienta <http://www.webratio.com>.
- Centro de Investigación y Desarrollo en Informática Aplicada, sitio oficial <http://cidia.unsa.edu.ar>.
- Norma ISO/IEC 25010 (2.007). Software Engineering–Software Product Quality Requirements and Evaluation (SQuaRE) – Quality model and guide elaborada por el Comité Técnico ISO/IEC JTC 1/SC 7 de ISO.

Prácticas de la Ingeniería de Requerimientos en las Organizaciones de Desarrollo

Alejandro Oliveros

UNTREF – Departamento de Ciencia y Tecnología

Valentín Gómez 4752, Caseros, Buenos Aires, Argentina

aoliveros@untref.edu.ar

Resumen

Se presenta una línea de investigación en el campo de la Ingeniería de Requerimientos que se ha enfocado en identificar las prácticas que efectivamente utilizan en el desarrollo de software. Se han desarrollado distintos *survey* en organizaciones que desarrollan software para el mercado o para una organización específica (de la que forme parte o no), así como para el mercado argentino o el exterior. En los mismos se investigaron las prácticas de requerimientos que se utilizan en los proyectos de desarrollo de software. Identificar estas prácticas provee información que permite ajustar planes de transferencia de tecnología, actividades académicas y de investigación a las posibilidades y necesidades de las organizaciones de desarrollo de software. A lo largo del desarrollo de las investigaciones se han formado recursos humanos. Los próximos pasos se orientan a ampliar las muestras y el alcance de los cuestionarios, incorporar el estudio de casos y *grounded theory*.

Palabras clave: Ingeniería de requerimientos, Prácticas, técnicas

Contexto

La Facultad de Ingeniería y Ciencias Exactas de la Universidad Argentina de la Empresa (UADE) tiene definida la Ingeniería de Requerimientos como una de las líneas de investigación del Instituto de Tecnología y relacionada con la carrera de Ingeniería Informática. Desde 2010 viene desarrollándose esta línea de investigación. En este contexto se desarrolló la investigación

en las prácticas reales de Ingeniería de Requerimientos (IR de ahora en más) orientadas a conocer el desempeño en el campo de la IR de las organizaciones que se desempeñan en la Argentina.

Introducción

La IR se ha consolidado paulatinamente como disciplina de la ingeniería de Software, ha crecido en su reconocimiento como tal y ha producido un relevante cuerpo de conocimiento, prácticas y experiencia. Prueba de ello son el crecimiento y desarrollo de conferencias internacionales específicas (una de ellas regional), la inserción de workshops dedicado a IR en conferencias generales de Ingeniería de Software, la consolidación de publicaciones dedicadas exclusivamente a IR y el creciente espacio que ocupa en las publicaciones de Ingeniería de Software en general. La IR ha acompañado las solicitudes a las que la ha sometido el desarrollo de software.

Las aplicaciones de software se han transformado en partes esenciales de las organizaciones empresarias, gubernamentales o sin fines de lucro. Muchas de estas organizaciones están en el negocio del software, aunque no son plenamente conscientes. A esta expansión del papel del software se asocia el desarrollo tecnológico propio de la disciplina. En un trabajo reciente [1] de amplio reconocimiento por la comunidad académica de IR, se han mencionado las siguientes tendencias y necesidades como las más importantes del estado actual de la IR:

- la mayor escala que están alcanzado los sistemas de software,

- integración más estrecha entre el software y su entorno,
- mayor autonomía del software para adaptarse a su entorno y
- creciente globalización del desarrollo de software

Este nuevo ambiente en el que se encuentra el software impacta en los stakeholders y afectan los procesos y prácticas de la IR. También es impactado el campo de la investigación en IR, para el que se han formulado las siguientes recomendaciones:

- los investigadores deben trabajar con los profesionales,
- los investigadores no deben descuidar la evaluación y la investigación empírica,
- las organizaciones de desarrollo deben proporcionar datos sobre los proyectos de fuertes de la industria.

En enfoque metodológico propuesto en esta línea de investigación se encuentra encuadrado en los métodos experimentales de la Ingeniería de Software [2] o más exactamente en la Ingeniería de Software Basado en Evidencia (EBSE, por sus iniciales en inglés).

El proceso de transferencia de enfoques relacionados con la IR desde la comunidad académica a las empresas, y en especial en las medianas y pequeñas, dista de alcanzar el potencial disponible en la IR [3].

La transferencia de las prácticas a las organizaciones de desarrollo (para el mercado o para un cliente específico) es un proceso complejo en el que el “efecto demostración” desempeña un papel importante. La ausencia de experiencia propia en las prácticas de la ingeniería de software es la principal restricción a su aceptación por las organizaciones de desarrollo [4]. Conocer el estado de las prácticas de requerimientos permite identificar buenas prácticas difundidas en la industria, lo que contribuye a su instalación en otras organizaciones mediante el uso del “efecto demostración” [1]. Por otra parte, desde el punto de vista de

la investigación, este conocimiento ayuda a conocer los problemas reales de la industria [1].

El foco del proyecto se encuentra en conocer el estado de las prácticas de elicitación de requerimientos en las organizaciones de desarrollo de software, ya sea que trabajen para el mercado en general o para un cliente en particular (interno o externo) en especial en el contexto de las aplicaciones Web.

El que sigue se presentan investigaciones similares que se han realizado para investigar las prácticas de las organizaciones de desarrollo de software mediante *surveys*.

En 1993 comenzó una tradición de investigaciones de las prácticas de IR en las organizaciones que desarrollan software para el mercado o para consumo propio. En [2] entrevistaron a 87 desarrolladores y gerentes mediante 35 entrevistas. Los entrevistados se relacionaban con 23 proyectos de investigación desarrollados en 10 organizaciones. El foco de la investigación estaba en las prácticas relacionadas con la especificación de requerimientos. Se utilizaron entrevistas en profundidad como metodología de obtención de los datos.

En Nueva Zelanda se desarrollaron varias investigaciones para establecer las prácticas de la IR. En 2000 se reportó una investigación [3] de las prácticas de IR en la industria del software. La investigación se enfocaba en las fases posteriores al proceso de elicitación de requerimientos. Se realizaron entrevistas a 24 empresas por vía telefónica. En 2000 se informó de una investigación en empresas de desarrollo mediante entrevistas personales a 12 empresas [4]. También se investigó la evolución en el tiempo de las prácticas de IR en Nueva Zelanda. En 2011 se publicó una investigación con el propósito, entre otros, de establecer la evolución de las prácticas de IR entre 2000 y 2005 [5]. Se compararon el resultado de entrevistas a 14 empresas con los de tres estudios que consideraron 24, 5 y 12 respuestas.

En [6] se investigaron 27 firmas de la industria de software con base en Malasia mediante un survey. El objetivo fue investigar las prácticas de Ingeniería de Software en especial las de IR.

En 2002 se realizó una investigación en Estados Unidos de las prácticas de elicitación, especificación y validación de requerimientos [7], [8]. Se recopilaron 194 respuestas a un cuestionario Web.

En nuestra región se ha desarrollado un interesante nivel de actividad en la investigación de las prácticas de IR. En la Argentina en 2002 se informó de una investigación [9] en la que 38 estudiantes de posgrado de dos localizaciones contestaron un cuestionario remitido por correo electrónico. Esta investigación se enfocó en las técnicas de elicitación y las fuentes utilizadas en el proceso de elicitación. En 2015 se reiteró la investigación, esta vez en una sola localización con 13 estudiantes de posgrado. En [10] se presentó el estudio de la evolución.

En Brasil en 2013 se informó una investigación [11] de las actividades de IR de las empresas de desarrollo instaladas en Porto Digital (similar a nuestros polos informáticos) de Pernambuco. Se entrevistaron 22 empresas del polo, aunque la unidad de análisis fue el desarrollador. Nuevamente en Pernambuco se desarrolló un survey [12] en un universo de 230 empresas de software. Se obtuvieron 48 respuestas a un cuestionario administrado en línea. El foco de la investigación fue el uso y experiencia de las empresas en el área de Gestión de Requerimientos.

En la comunidad de IR también se han desarrollado investigaciones orientadas a aspectos específicos de la IR. Se trata de investigaciones basadas en *survey* y que se orientan a terrenos específicos de la RE como la reutilización [13], rastreabilidad [14] o priorización de requerimientos [15].

La investigación de las prácticas de IR en las organizaciones de desarrollo tiene una larga trayectoria. Sin embargo no se han encontrado muchos trabajos sobre el tema. En

particular las investigaciones no han puesto el foco en el proceso de elicitación de requerimientos ni en la evolución de las prácticas a los largo del tiempo. Son escasos los estudios que atienden a estos aspectos.

Líneas de investigación y desarrollo

El objetivo general de la línea de investigación es el estudio de las prácticas de IR en las organizaciones de desarrollo de software de nuestro país, en especial en el desarrollo de aplicaciones Web. Desde el punto de vista de las actividades de la IR el foco se pone la elicitación de requerimientos.

El énfasis en las organizaciones locales de desarrollo se origina en que este tipo de prácticas se encuentran fuertemente asociadas con aspectos culturales de los usuarios y desarrolladores. Aspectos que deben identificarse para poder entender el modo de uso de instancias específicas de técnicas de tipo universal.

Los objetivos más detallados del proyecto son:

- Establecer el estado del arte de las prácticas de Ingeniería de Requerimientos vigentes en las organizaciones de desarrollo de software.
- Desarrollar casos de estudio para analizar en profundidad las prácticas de Ingeniería de Requerimientos de organizaciones de desarrollo de software
- Formar recursos humanos en los enfoque metodológicos de la Ingeniería de Software Basada en Evidencia (EBSE)

Resultados obtenidos/esperados

Los resultados obtenidos se han presentado en varias conferencia [16], [17], [18], [19] y [20]. Los resultados obtenidos en los distintos estudios aún no han sido agregados. Sin embargo puede asegurarse que predominan las prácticas más tradicionales, no se utilizan

las nuevas técnicas más adecuadas a las aplicaciones que se desarrollan actualmente (como las aplicaciones Web). El resultado que emerge es que tenuemente se han introducido prácticas más modernas y que al menos hay un mayor conocimiento de las mismas.

Parte de estos resultados se basaron en trabajo experimental desarrollado por estudiantes sobre la base de encuestas a organizaciones de desarrollo y que formaron parte de varios Proyectos Finales de Ingeniería. Así se produjo el material empírico para los análisis de las prácticas de Ingeniería de Requerimientos en la Argentina que se presentó en [18], [19] y [20].

En cuanto a resultados esperados hay varias alternativas. Se desarrollará una Revisión Sistemática de Literatura (SLR, por sus iniciales en inglés) para investigar la literatura sobre las prácticas reales de IR en las organizaciones. De esa forma se establecerá un estado del arte definitivo en cuanto a las prácticas utilizadas en el desarrollo de software, así como establecer una línea de evolución en el tiempo de las mismas. Este resultado, valioso en sí mismo, permitirá comparar la evolución de nuestro país con la de otras geografías.

La agregación de los datos obtenidos así hasta ahora en el proyecto se acompañará de un survey con una muestra mayor y con un cuestionario de mayor alcance que utilizados hasta ahora. Con esto se podrá tener una visión más completa del estado actual de las prácticas de IR utilizadas en el desarrollo.

Por último, contribuyendo al punto mencionado, se desarrollará un caso de estudio [21] de las prácticas de una organización de desarrollo de software de gran envergadura, utilizando el enfoque de grounded theory [22]. Con ello se obtendrá una visión cualitativa que permitirá completar los resultados obtenidos mediante los *survey* realizados.

Formación de recursos humanos

Los cuatro coautores de [17], [18], [19] y

[20], son egresados de la carrera de Ingeniería Informática que se iniciaron en la investigación científica. Asimismo cinco graduados comenzaron el proceso de formación como investigadores en la línea de investigación de ingeniería de requerimientos, aunque no exactamente en los aspectos mencionados en este documento.

Referencias

- [1] B. H. C. Cheng and J. M. Attle, "Research Directions in Requirements Engineering," in *Proceeding FOSE '07 2007 Future of Software Engineering*, Minneapolis, MN, USA, 2007, pp. 285–303.
- [2] Lubars, Mitch, C. Potts, and C. Richter, "A Review of the State of the Practice in Requirements Modeling," in *Proceedings of IEEE International Symposium on Requirements Engineering*, San Diego, CA, USA, 1993, pp. 2–14.
- [3] L. Groves, R. Nickson, G. Reeve, S. Reeves, and M. Utting, "A survey of software development practices in the New Zealand software industry," in *Proceedings. 2000 Australian Software Engineering Conference*, Canberra, ACT, 2000, pp. 189–201.
- [4] U. Nikula, J. Sajaniemi, and H. Kälviäinen, "A State-of-the-Practice Survey on Requirements Engineering in Small- and Medium-Sized Enterprises," Telecom Business Research Center. Lappeenranta University of Technology, Lappeenranta, Finland, Research Report 1, 2000.
- [5] A. Talbot and A. Connor, "Requirements Engineering Current Practice and Capability in Small and Medium Software Development Enterprises in New Zealand," in *Proceedings 2011 Ninth International Conference on Software Engineering Research, Management and Applications*, Baltimore, MD, 2011, pp. 17–25.
- [6] A. Tahir and R. Ahmad, "Requirement Engineering Practices – an Empirical Study," presented at the 2010 International Conference on Computational Intelligence and Software Engineering (CiSE), Wuhan, 2010, pp. 1–5.
- [7] C. J. Neill and P. A. Laplante, "Requirements Engineering: The State of the Practice," *IEEE Software*, vol. 20, no. 6, pp. 40–45, Dec. 2003.

- [8] P. A. Laplante, C. J. Neill, and C. Jacobs, "Software Requirements Practices: Some Real Data," in *Proceedings of the 27th Annual NASA Goddard/IEEE Software Engineering Workshop (SEW-27'02)*, Goddard, 2002.
- [9] L. Antonelli and A. Oliveros, "Fuentes utilizadas por desarrolladores de software en Argentina para elicitar requerimientos," in *Proceedings del 5to. Workshop de Ingenieria de Requerimientos (WER'02)*, Valencia, España, 2002, pp. 106–116.
- [10] A. Oliveros and L. Antonelli, "Fuentes y técnicas para elicitación de requerimientos. Un estudio de evolución de la práctica," in *Proceedings XVIII Ibero-American Conference on Software Engineering*, Lima, Perú, 2015, pp. 563–577.
- [11] D. Arruda, R. Soares, D. Vieira, R. Ferreira, T. Cabral, and M. Lencastre, "Engenharia de Requisitos: Um Survey realizado no Porto Digital, Recife/Brasil.," in *Proceedings of the XVII Ibero-American Conference on Software Engineering*, Pucón, Chile, 2014, pp. 477–490.
- [12] J. Melo, "Requirements Engineering: A survey performed in Recife-PE/Brazil companies," *Requirements Engineering Research Group. Engenharia de computacao. UPE-POLI*, 2014. [Online]. Available: <http://www.serg.ecomp.poli.br/req/index.php/pt-BR/publicacoes/22-survey2014-01a>. [Accessed: 12-Jul-2014].
- [13] Y. Chernak, "REQUIREMENTS REUSE: THE STATE OF THE PRACTICE," in *Proceedings 2012 IEEE International Conference on Software Science, Technology and Engineering*, Herzlia, 2012, pp. 46–53.
- [14] P. Rempel, P. Mäder, and T. Kuschke, "An empirical study on project-specific traceability strategies," in *Proceedings 2013 21st IEEE International Requirements Engineering Conference (RE)*, Rio de Janeiro, 2013, pp. 195–204.
- [15] R. B. Svensson *et al.*, "Prioritization of Quality Requirements: State of Practice in Eleven Companies," in *2011 19th IEEE International Requirements Engineering Conference (RE)*, Trento, Italy, 2011, pp. 69–78.
- [16] A. Oliveros, J. Zuñiga, R. Wehbe, S. Rojo, and J. Rousselot, "Requerimientos para Aplicaciones Web," presented at the Workshop de Investigadores en Ciencias de la Computación, Rosario, Argentina, 2011.
- [17] A. Oliveros, F. J. Danyans, and M. L. Mastropietro, "Prácticas de Ingeniería de Requerimientos en el desarrollo de aplicaciones Web," in *Proceedings of the XVII Ibero-American Conference on Software Engineering*, Pucón, Chile, 2014, pp. 491–505.
- [18] A. Oliveros, F. J. Danyans, and M. L. Mastropietro, "Stakeholders en los requerimientos de aplicaciones Web," in *Anales del 15° Simposio Argentino de Ingeniería de Software*, La Plata, Argentina, 2012, pp. 240–252.
- [19] A. Oliveros, F. J. Danyans, and M. L. Mastropietro, "Stakeholders of Web Applications Requirements," *EJS*, vol. 14, no. 1, Jun. 2015.
- [20] A. Oliveros, F. Napolillo, and F. López Infesta, "Requirements in Web applications development," in *IEEE CACIDI 2016 - IEEE Conference on Computer Sciences*, Buenos Aires, Argentina, 2016.
- [21] P. Runeson, M. Höst, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering*. Hoboken, New Jersey: John Wiley and Sons, 2012.
- [22] Glaser, Barney G. and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New Jersey, USA: ALdine Transaction, 2008.

Programa de I+D+I en Ingeniería de Explotación de Información

Informe de Avances Periodo 2011-2016 y Previsión Periodo 2017-2019

Ramón García-Martínez, Hernán Merlino, Darío Rodríguez, Sebastian Martins, Ezequiel Baldizzoni, Santiago Bianco

Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información
Grupo de Ingeniería de Explotación de Información y Grupo Investigación en Sistemas de Información

Departamento de Desarrollo Productivo y Tecnológico. Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús. Argentina. Tel +54 11 5533 5600 Ext. 5194
rgm1960@yahoo.com

Resumen

El Programa de I+D+I en Ingeniería de Explotación de Información tiene como objetivo general construir conocimiento aplicable a problemáticas del área sobre los procesos, las metodologías y las técnicas utilizadas para: ordenar, controlar y gestionar la tarea de encontrar patrones de conocimiento en grandes masas de información. El programa evoluciona en la línea de tiempo a partir de los problemas abiertos que se identifican en los proyectos y de los requerimientos que plantea el sistema productivo regional.

Palabras clave: Minería de datos, Explotación de Información, Ingeniería de Procesos, Ingeniería de Explotación de Información.

Contexto

El Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús, realizó un ordenamiento de las actividades de I+D+I, estructurándolas en Líneas de Investigación [DDPyT, 2016]. Cada Línea de Investigación se estructura en Programas de I+D+I. En particular, la Línea de Investigación en el Campo de Productividad en Informática se estructura en los siguientes Programas de I+D+I: Ingeniería de Sistemas de Información, Ingeniería de Sistemas Inteligentes, Ingeniería de Explotación de Información, Ingeniería de Espacios Virtuales de Trabajo, e Ingeniería de Desarrollo de Sistemas Especiales.

Introducción: Resultados Periodo 2011-2016

En el Periodo 2011-2016 se acreditaron y desarrollaron los siguientes Proyectos de I+D+I:

- Ingeniería de Proyectos de Explotación de Información para PyMEs (Código: 33A105, Periodo:2011-2012)
- Aseguramiento de la Calidad para Proyectos de Explotación de Información (Código 33B102, Periodo 2012-2013)
- Instrumentos para la Gestión de Proyectos de Explotación de Información (Código: 33A167, Periodo:2013-2014)
- Procesos de Explotación de Información Georeferenciada (Código: 33B133, Periodo 2015-2016)
- Modelos de Proceso para Ingeniería de Explotación de Información para PYMES: Abordaje Ágil y Abordaje Robusto (Código: 33A205, Periodo 2015-2016)

En el Proyecto UNLa 33A105 *Ingeniería de Proyectos de Explotación de Información para PyMEs* [GISI-UNLa, 2011] se desarrollaron los siguientes artefactos de Ingeniería de Proyectos de Explotación de Información: [a] una batería de técnicas de educación y formalismos de documentación de requerimientos; [b] la versión preliminar de un modelo de procesos y se comenzó a explorar las métricas asociadas; [c] un modelo de ciclo de vida; y [d] un mapa de actividades.

En el Proyecto UNLa 33B102 *Aseguramiento de la Calidad para Proyectos de Explotación de Información* [GISI-UNLa, 2012] se revisaron los paradigmas existentes sobre calidad en proyectos de software aplicables a proyectos de explotación de información, y se planteó su adaptación para lograr el Aseguramiento de la Calidad en este tipo de Proyectos.

En el Proyecto UNLa 33A167 *Instrumentos para la Gestión de Proyectos de Explotación de Información* [GISI-UNLa, 2013] se desarrollaron mediante la metodología de prototipado evolutivo, los siguientes instrumentos para el control y la gestión de Proyectos

de Explotación de Información para PyMEs: [a] test de viabilidad, [b] procedimiento de derivación del problema de negocio en problema de explotación de información, [c] formalismos para documentación de requisitos, [d] técnicas y métodos de aseguramiento de la calidad del proceso y del producto, y [e] método de estimación de recursos.

En el Proyecto UNLa 33B133 *Procesos de Explotación de Información Georeferenciada* [GISI- UNLa, 2014] se sentaron las bases para caracterizar los procesos de explotación de información georeferenciada asociados a detección de patrones y regularidades importantes en artefactos de información geográficos, y se identificaron las tecnologías de Sistemas Inteligentes que se pueden utilizar para los procesos caracterizados.

En el Proyecto UNLa 33A205 *Modelos de Proceso para Ingeniería de Explotación de Información para PYMES: Abordaje Ágil y Abordaje Robusto* [GISI-UNLa, 2015] se desarrolló un modelo de proceso que propone fases, tareas, técnicas de representación y procedimientos de ejecución de la tarea; y que soporta de forma completa e integrada las actividades de administración y desarrollo de proyectos en dos versiones, una orientada a microproyectos y otra orientada a proyectos regulares.

En el Periodo 2011-2016 en el marco de los Proyectos de I+D+I que integraban el Programa se desarrollaron y defendieron las siguientes tesis doctorales y de maestría:

- Procedimientos de Explotación de Información en Identificación de Datos Faltantes, con Ruido o Inconsistentes (Tesis Doctoral Universidad de Malaga, 2014)
- Viabilidad y Estimación de Proyectos de Explotación de Información (Tesis Doctoral Universidad Nacional de La Plata, 2014)
- Gestión de la Educación Superior en Contextos de Masividad Basada en Tecnologías Inteligentes de Transformación de Información (Tesis Doctoral Universidad Nacional de Córdoba, Defensa pendiente)
- Método de Estimación de Esfuerzo para Proyectos de Explotación de Información. Herramientas de Validación (Tesis de Magister Universidad Politécnica de Madrid/ ITBA, 2011)
- Un Modelo de Procesos para Proyectos de Explotación de Información (Tesis de Magister UTN-FRBA, 2012)

- Descubrimiento de Conocimiento Mediante la Integración de Algoritmos de Explotación de la Información (Tesis de Magister UTN-FRBA, 2012)
- Propuesta de Métricas para Proyectos de Explotación de Información (Tesis de Magister UTN-FRBA, 2014)
- Proceso de Conceptualización del Entendimiento del Negocio para Proyectos de Explotación de Información (Tesis de Magister UTN-FRBA, 2014)
- Proceso de Identificación de Comportamiento de Comunidades Educativas basado en Resultados Académicos (Tesis de Magister Universidad Nacional de La Plata, 2016)

En la Tesis *Procedimientos de Explotación de Información en Identificación de Datos Faltantes, con Ruido o Inconsistentes* [Kuna, 2014] se establece una taxonomía relacionada con los métodos, técnicas y algoritmos de detección de valores anómalos en bases de datos, analizando las ventajas y desventajas de cada una de ellos, y se proponen procedimientos de explotación de información que combinados entre sí permitan detectar los campos que tienen valores atípicos en bases de datos, combinando distintas técnicas de minería de datos, entre ellas específicamente las de clustering, algoritmos de inducción, redes bayesianas, principios de la teoría de la Información, entre otras, para lograr mejorar de esa manera la calidad de los datos.

En la Tesis *Viabilidad y Estimación de Proyectos de Explotación de Información* [Pytel, 2014] se proponen dos procesos ad-hoc que permitan en el marco de los Proyectos de Explotación de Información: [a] evaluar la viabilidad de a partir de la información disponible al comienzo del proyecto, y [b] estimar los recursos necesarios con particular énfasis en su utilización en PyMEs.

En la Tesis *Gestión de la Educación Superior en Contextos de Masividad Basada en Tecnologías Inteligentes de Transformación de Información* [Díaz, 2017] propone el uso de Tecnologías Inteligentes de Explotación de la Información y otras herramientas del campo de la Inteligencia Computacional para construir categorías y enriquecer teorías a la luz del Paradigma Interpretativo de investigación en las Ciencias Sociales. En particular, a partir de la experiencia piloto, se proporcionará un mejor conocimiento del estudiante de cursos masivos de esta Universidad en sus características socio- económicas y de rendimiento académico.

En la Tesis *Método de Estimación de Esfuerzo para Proyectos de Explotación de Información. Herramientas de Validación* [Pytel, 2011] se propone una herramienta de software que ayude a validar el método de estimación de esfuerzo para proyectos de Explotación de Información denominado Método Matemático Paramétrico de Estimación para Proyectos de Data Mining (DMCoMo). La herramienta permite generar un banco de pruebas con las características de proyectos de Explotación de Información relevantes para el método de estimación, y así luego analizar el comportamiento del método de estimación seleccionado.

En la Tesis *Un Modelo de Procesos para Proyectos de Explotación de Información* (Tesis de Magister UTN-FRBA, 2012) Se busca desarrollar un modelo de procesos para proyectos de explotación de información, atendiendo los siguientes objetivos específicos: [a] Establecer diferencias y relaciones entre los proyectos de explotación de información y los proyectos clásicos, [b] Determinar límites, alcance y componentes del modelo MoProSoft aplicables a Proyectos de Explotación de Información, [c] Identificar los procesos específicos de los proyectos de explotación de información a través de las distintas metodologías existentes que se aplican a dichos proyectos y que no estén cubiertos por los modelos de procesos de software considerados, [d] Seleccionar los procesos a incluir en el nuevo modelo, y [e] Desarrollar el nuevo modelo basándose en MoProSoft a partir de la documentación de los procesos identificados y utilizando el patrón de procesos especificado por MoProSoft.

En la Tesis *Descubrimiento de Conocimiento Mediante la Integración de Algoritmos de Explotación de la Información* [Lopez Nocera, 2012], sobre la premisa que en Minería de Datos se parte del supuesto que los dominios son equivalentes o directamente indistinguibles entre sí a los efectos experimentales; se postula que esto no es así, y que la caracterización de los dominios en uso durante la fase experimental al estudiar algoritmos de minería de datos tiene amplia injerencia en los resultados observados y finalmente obtenidos. La postulación es corroborada con resultados experimentales.

En la Tesis *Propuesta de Métricas para Proyectos de Explotación de Información* [Basso, 2014] se parte de la premisa que las métricas usuales de la Ingeniería de Software y la Ingeniería del Conocimiento no son adecuados para los proyectos de explotación de información, ya que las etapas de desarrollo y los

parámetros utilizados son de naturaleza y características diferentes. En ese contexto, se propone un conjunto de métricas aplicables al desarrollo de un proyecto de Explotación de Información para PyMEs, centrado en el Modelo MoProPEI.

En la Tesis *Proceso de Identificación de Comportamiento de Comunidades Educativas basado en Resultados Académicos* [Cigliuti, 2016] se propone el uso de procesos de explotación de información para identificar el comportamiento de comunidades educativas basado en resultados académicos.

Objetivos: Líneas de Trabajo, Desarrollo e Innovación Periodo 2017-2019

Para el periodo 2017-2019 las Líneas de Trabajo, Desarrollo e Innovación se encuentran articuladas a través de los siguientes Proyectos de I+D+I que, ha la fecha de esta comunicación, cuentan con acreditación institucional y han comenzado su desarrollo:

- Integración de Herramientas y Procesos de Ingeniería de Explotación de Información (Código: AH2016 33A 80020160400001LA, Periodo: 2017- 2018)
- Identificación de Patrones de Adicción a Sustancias Psicoactivas Basada en Ingeniería de Explotación de Información (Código: AH2016 33A 800201605000 02LA, Periodo: 2017-2018)
- Ingeniería de Proyectos de Explotación de Grandes Volúmenes de Datos (Código PIO UNLa-CONICET 22420160100032CO, Periodo: 2017-2018)

En el Proyecto *Integración de Herramientas y Procesos de Ingeniería de Explotación de Información*, se busca desarrollar un marco integrado de herramientas conceptuales y procesos de utilizables en producción para llevar adelante las actividades de Ingeniería de un Proyecto de Explotación de Información. El marco propuesto se orienta a contribuir a que los profesionales informáticos dispongan de guías que ordenen el trabajo y permitan establecer y ofrecer hitos para los entregables de proyectos en el área de ingeniería de explotación de información orientadas a inteligencia de negocio.

En el Proyecto *Identificación de Patrones de Adicción a Sustancias Psicoactivas Basada en Ingeniería de Explotación de Información*, se busca explorar la utilización de Ingeniería de Explotación de Información como abordaje alternativo a la identificación de patrones de consumo de sustancias

psicoactivas tomando como insumo la información disponible en encuestas del INDEC con énfasis en la identificación de [a] variables descriptoras concurrentes con el consumo de sustancias psicoactivas, y [b] patrones de consumo y de consumidores de sustancias psicoactivas describibles en términos de instancias de las variables concurrentes con el consumo de sustancias psicoactivas.

En el Proyecto *Ingeniería de Proyectos de Explotación de Grandes Volúmenes de Datos* se busca desarrollar el cuerpo de conocimiento necesario para las actividades tempranas de Ingeniería de Proyectos de Explotación en Grandes Volúmenes de Datos (PEGVD) con focalización en su transferencia al sector PyMEs de Informática. El proyecto busca proveer un marco integrado de trabajo para PEGVD que contempla generar la versión de producción de las siguientes herramientas: [a] un modelo de procesos de decisión estratégica, control y gestión y desarrollo para PEGVD, [b] una batería de técnicas de educación y formalismos de documentación de requerimientos para PEGVD, [c] un test de viabilidad para PEGVD, [d] un modelo de ciclo de vida para PEGVD, [e] un mapa de actividades para PEGVD, [f] un método de estimación de recursos para PEGVD, [g] un procedimiento de derivación del problema de negocio en problema de exploración en grandes volúmenes de datos, y [h] un conjunto de técnicas y métodos de aseguramiento de la calidad del proceso y del producto para PEGVD.

En el Periodo 2017-2019 en el marco de los Proyectos de I+D+I que integran el Programa se prevé desarrollar y finalizar las siguientes tesis doctorales y de maestría:

- Integración de Métodos en Ingeniería de Explotación de Información (Tesis Doctoral Universidad Nacional de La Plata, Finaliza 2017)
- Metodología para la Especificación de Requisitos en Proyectos de Explotación de Información (Tesis Doctoral Universidad Nacional de La Plata, Finaliza 2018)
- Integración de Métodos de Descubrimiento de Conocimiento Embebido en Fuentes de Información Desestructuradas (Tesis Doctoral Universidad Nacional de La Plata, Finaliza 2019)
- Formulación de Procesos para una Ingeniería de Explotación de Información Espacial (Tesis Doctoral Universidad Nacional de La Plata, Finaliza 2020)

- Propuesta de Proceso para Ingeniería de Explotación de Información Centrado en Métodos y Técnicas - Núcleo IMMAT (Tesis de Magister UTN-FRBA, Finaliza 2018)

En la Tesis *Integración de Métodos en Ingeniería de Explotación de Información* [Martins, 2014] se busca sistematizar el conocimiento existente sobre ingeniería de explotación de información y formular una propuesta unificada de modelo de proceso que integre todos los métodos, técnicas y herramientas de administración y desarrollo de proyectos en esta subdisciplina informática.

En la Tesis *Metodología para la Especificación de Requisitos en Proyectos de Explotación de Información* [Pollo-Cattaneo, 2010] se busca definir una metodología de educación de requisitos para proyectos de explotación de información en el marco del ciclo de vida utilizado en este tipo de proyectos. La tesis se enfocará a construir una metodología que permita la educación de requisitos para proyectos de explotación de información teniendo presente los distintos procesos que involucra el ciclo de vida usado en este tipo de proyectos. Se considera que la tesis implica realizar investigación básica en la definición del proceso de educación de requisitos en proyectos de explotación de información.

En la Tesis *Integración de Métodos de Descubrimiento de Conocimiento Embebido en Fuentes de Información Desestructuradas* se busca construir una familia de métodos de extracción de conocimiento tal que dado una estructura de información inicial como entrada, que contienen conocimiento embebido, genere un conjunto de piezas de conocimiento (reglas de producción, subgrafos de una red semántica, entre otros). Tiene los siguientes Objetivos Específicos: (i) Realizar una comparación entre los distintos métodos de extracción de conocimiento relevados en la literatura, indicando, para diversas estructuras de información, la calidad y los tiempos de ejecución asociados a cada uno; (ii) Identificar fortalezas y debilidades en los diversos métodos evaluados en OE1 con el fin de detectar las condiciones particulares bajo las cuales es conveniente utilizar un método por sobre otro y (iii) Desarrollar una familia de métodos integrados de extracción de conocimiento que exhiban un mejor comportamiento que los métodos individuales integrados.

En la Tesis *Formulación de Procesos para una Ingeniería de Explotación de Información Espacial*, dadas las particularidades de la Explotación de

Información Espacial en contraste con la Explotación de Información convencional, resulta de interés proponer un conjunto de Procesos de Explotación de Información Espacial relevantes para la inteligencia de Negocio. Estos Procesos de Explotación de Información Espacial deberán incluir las fuentes de Información necesarias, tratamiento de los datos y los algoritmos o familias de algoritmos de Explotación de Información a ser utilizados para el descubrimiento de Patrones en la Información Espacial.

En la Tesis *Propuesta de Proceso para Ingeniería de Explotación de Información Centrado en Métodos y Técnicas - Núcleo IMMAT*, se busca desarrollar la adaptación del núcleo SEMAT, que fue pensado para proyectos de Ingeniería de Software, a proyectos de Ingeniería de Explotación de Información con el fin de aportar a esta última, buenas practicas que colaboren en obtener los mejores resultados, de mejor calidad y optimizando los recursos.

Formación de Recursos Humanos

En este Programa participan 3 investigadores formados, y un investigador formado adscripto. A la fecha de esta comunicación se han radicado en el Programa: 7 tesis doctorales (4 defendidas), 7 tesis de magíster (6 defendidas), y 8 Trabajos Finales de Licenciatura (5 defendidos).

Referencias

- Baldizzoni, E. 2016. *Investigación en Progreso: Propuesta de Proceso para Ingeniería de Explotación de Información Centrado en Métodos y Técnicas - Núcleo IMMAT*. Revista Latinoamericana de Ingeniería de Software 4(1): 55-58. ISSN 2314-2642.
- Basso, D. 2014. *Propuesta de Métricas para Proyectos de Explotación de Información*. Tesis de Magíster en Ingeniería de Sistemas de Información. Facultad Regional Buenos Aires. Universidad Tecnológica Nacional.
- Cigliuti, P. 2016. *Proceso de Identificación de Comportamiento de Comunidades Educativas basado en Resultados Académicos*. Tesis de Magíster en Tecnología Informática Aplicada en Educación. Facultad de Informática. UNLP.
- DDPyT, 2016. *Líneas de Investigación - Programas I+D+I - Grupos de I&D y Producción Científica*. <http://ddpyt.unla.edu.ar/documentos/DDPyT-Lineas-IyD.pdf>
- Díaz, L. 2014. *Gestión de la Educación Superior en Contextos de Masividad Basada en Tecnologías Inteligentes de Transformación de Información*. Tesis Doctoral en Administración Pública en Desarrollo. Universidad Nacional de Córdoba. <http://sistemas.unla.edu.ar/sistemas/gisi/GISI-TD-2014-02.htm>
- GISI-UNLa, 2009. *Proyecto UNLa 33A081 Sistemas de Información para Inteligencia de Negocio*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33A081.htm>
- GISI-UNLa, 2011. *Proyecto UNLa 33A105 Ingeniería de Proyectos de Explotación de Información para PyMEs*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33A105.htm>
- GISI-UNLa, 2012. *Proyecto UNLa 33B102: Aseguramiento de la Calidad para Proyectos de Explotación de Información*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33B102.htm>
- GISI-UNLa, 2013. *Proyecto UNLa 33A167 Instrumentos para la Gestión de Proyectos de Explotación de Información*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33A167.htm>
- GISI-UNLa, 2014. *Proyecto UNLa 33B133 Procesos de Explotación de Información Georeferenciada*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33B133.htm>
- GISI-UNLa, 2015. *Proyecto UNLa 33A205 Modelos de Proceso para Ingeniería de Explotación de Información para PYMES: Abordaje Ágil y Abordaje Robusto*. <http://sistemas.unla.edu.ar/sistemas/gisi/UNLa-33A205.htm>
- Kuna, H. 2014. *Procedimientos de Explotación de Información para la Identificación de Datos Faltantes con Ruido e Inconsistentes*. Tesis Doctoral en Ingeniería de Sistemas y Computación. Universidad de Málaga.
- Lopez Nocera, M. 2012. *Descubrimiento de Conocimiento Mediante la Integración de Algoritmos de Explotación de la Información*. Tesis de Magíster en Ingeniería de Sistemas de Información. UTN-FRBA.
- Martins, S. 2014. *Integración de Métodos en Ingeniería de Explotación de Información*. Tesis Doctoral en Ciencias informáticas en Desarrollo. Universidad Nacional de La Plata. <http://sistemas.unla.edu.ar/sistemas/gisi/GISI-TD-2014-01.htm>
- Peralta, F. 2014. *Proceso de Conceptualización del Entendimiento del Negocio para Proyectos de Explotación de Información*. Tesis de Magíster en Ingeniería de Sistemas de Información. UTN-FRBA.
- Pollo-Cattaneo, F. 2010. *Metodología para la Especificación de Requisitos en Proyectos de Explotación de Información*. Tesis Doctoral en Ciencias informáticas en Desarrollo. Universidad Nacional de La Plata. <http://sistemas.unla.edu.ar/sistemas/gisi/GISI-TD-2010-02.htm>
- Pytel, P. 2011. *Método de Estimación de Esfuerzo para Proyectos de Explotación de Información. Herramienta Para Su Validación*. Tesis de Magíster en Ingeniería del Software. Convenio Universidad Politécnica de Madrid e Instituto Tecnológico Buenos Aires.
- Pytel, P. 2014. *Viabilidad y Estimación de Proyectos de Explotación de Información*. Tesis Doctoral en Ciencias informáticas. Universidad Nacional de La Plata.
- Vanrell, J. 2012. *Un Modelo de Procesos para Proyectos de Explotación de Información*. Tesis de Magíster en Ingeniería de Sistemas de Información. UTN-FRBA.

Nota: Se puede acceder a la Producción Científica del Programa a través de la página web del Grupo de Ingeniería en Explotación de Información:

<http://sistemas.unla.edu.ar/sistemas/gisi/GIEL.html>

Propuesta de Investigación: Diagnóstico e Impacto de las TI/SI en Pymes de la Región Centro

Carrizo, Blanca Rosa ⁽¹⁾, Abet, Jorge Eduardo ⁽²⁾, Colazo Carlos ⁽³⁾

⁽¹⁾⁽²⁾ *Facultad Regional Córdoba, Universidad Tecnológica Nacional
Maestro Marcelo López esq. Cruz Roja Argentina. C.P.: 5016*

bcarrizo@frc.utn.edu.ar / bcarrizo@yahoo.com.ar

⁽³⁾ *Facultad Regional Villa María, Universidad Tecnológica Nacional*

Resumen

La presente propuesta se fundamenta en un proyecto de investigación cuyo objetivo es estudiar, analizar, describir y diagnosticar recursos de TI/SI (Tecnologías/Sistemas de Información) en las Pymes cordobesas, diseñando indicadores que reflejen la situación actual y permitan describir modelos que optimicen su gestión competitiva.

En las pymes tiene importancia el buen funcionamiento de la administración, que implica desarrollar y documentar procesos, proveer una infraestructura tecnológica y de sistemas adecuada, como así también la formación y conducción de un buen equipo de colaboradores. Surge entonces, la necesidad de conocer su desempeño regional, el entorno que las rodea y precisar el grado de desarrollo tanto tecnológico como organizacional; aspectos que son esenciales ya que permiten predecir y evidenciar su crecimiento y competitividad.

En este contexto, y en un mundo dominado por la tecnología y la innovación, el uso de TI/SI es de vital importancia al momento de definir estrategias de negocio para poder ser más productivas y aumentar su grado de eficiencia.

Palabras Claves: Pymes / Tecnologías (TI)/Sistemas (SI) / TIC / Contexto Tecnológico

Contexto

El proyecto IFN 4346 forma parte del pool de proyectos del Grupo de Investigación del Dpto. Industrial en Control Avanzado de Procesos y Producción (GICCAP) reconocido como Grupo de la UTN mediante la Res. N°

816/2011 del Consejo Superior de la UTN con fecha del 07 de Julio de 2011 y el mismo forma parte de las mejoras asumidas en el proceso de acreditación de la carrera Ingeniería Industrial.

1. Introducción

El 99,7% de las empresas de nuestro país son micro, pequeñas y medianas. En conjunto generan 70% del empleo privado registrado de la Argentina. En los últimos años no la pasaron bien. Entre 2007 y 2015, la participación de las pymes en el empleo registrado cayó 3 puntos. De 2012 a 2015 su tasa de crecimiento fue negativa. En esos cuatro años perdimos 8500 pymes. Y no se crearon empleos.

Las estadísticas aturden: “sólo el 30% de estos emprendimientos sobreviven a la segunda generación, el 7% a la tercera y hay solo 32 con más de 200 años en el mundo” y no es que estas empresas fracasen por lo difícil que es resistir cambios económicos de por sí, sino por ser familiares. El emprendedor debe dejar de serlo para convertirse en empresario, y es ahí donde parecen empezar las fallas”. “Al emprendedor le cuesta delegar, y eso termina por asfixiar el crecimiento de la empresa y por resentir los vínculos”

Es un desafío importante para nosotros los profesionales para contribuir al crecimiento y desarrollo de estas empresas arraigadas a la figura patriarcal, teniendo en cuenta que una ventaja competitiva de las pymes es su capacidad de adaptabilidad gracias a su estructura pequeña; también son aspectos positivos la posibilidad de especializarse en nichos de mercado y la flexibilidad que les

permite detectar nuevos procesos, productos y mercados. Otra ventaja es la capacidad dinámica y la gran potencialidad de crecimiento.

"Estas empresas tienen la capacidad de cambiar rápidamente su estructura productiva en el caso de variar las necesidades de mercado, lo cual es mucho más difícil en una gran compañía".

En general, puede afirmarse que las Pymes presentan elevadas tasa de implementación en la incorporación y la difusión de las TIC más básicas como el acceso a Internet y la existencia de PC. Sin embargo, al desagregar estos indicadores aparecen diferencias que muestran cierto grado de heterogeneidad y que se manifiestan en incrementos de las brechas digitales entre empresas⁶⁵. En este sentido, resulta importante mencionar que las empresas de menor tamaño relativo revisten una dependencia respecto de la infraestructura y que las grandes por el contrario son las que traccionan dicha infraestructura al lugar donde se instalan.

Desde la perspectiva de la difusión de las TI/SI y las brechas digitales dadas por las diferencias en el acceso a herramientas de infraestructura básica se expresan tanto para el nivel regional como para el nivel productivo y develan una problemática a tener en cuenta a la hora de dirigir los esfuerzos en materia de políticas de innovación y desarrollo. (4)

Existen datos de ciertas Pymes donde se evidencia rasgos de heterogeneidad en el uso de las PC. El uso de la PC resulta más intensivo para las regiones del Centro y Pampeana. Además, en materia de acceso a Internet, el tamaño de las firmas por nivel de facturación resulta un elemento diferenciador ya que el nivel de acceso aumenta con el tamaño de las empresas. Cabe aclarar que el nivel de acceso a Internet resulta considerablemente elevado, superando el 90% en casi todas las regiones del país, excepto en el NEA donde alcanza al 85%. Además, la difusión de Internet presenta diferencias entre los sectores industriales, y los que mayor acceso presentan son los sectores más intensivos en ingeniería y automotriz donde el 98% y el 95% de las

firmas, respectivamente acceden a Internet, mientras que este porcentaje se reduce a casi el 92% para el resto de los sectores industriales. (5)

Existen Observatorios del comportamiento de las Pymes en distintas regiones de Nuestro País realizado en conjunto entre Universidad Cámaras Pymes y el propio gobierno. Estos observatorios no incluyen la región Córdoba y gran Córdoba.

Los Observatorios PyME Regionales constituyen un instrumento operativo para monitorear de manera permanente la demografía industrial y empresaria de las PyME de las distintas regiones de Argentina. Los OPR son impulsados por la Fundación Observatorio PyME y se constituyen a partir de pactos territoriales entre universidades, líderes empresarios y gobiernos provinciales y/o municipales. La Università di Bologna es responsable de la dirección científica del proyecto y aplica una metodología uniforme que permite crear un mecanismo continuo de análisis comparativo del papel de las PyME en cada región.

La información que generan los Observatorios es fundamental a la hora de realizar un diseño más ajustado de políticas regionales, evaluar el desempeño empresarial y conocer el papel que juegan en los mercados de factores, bienes y servicios en cada territorio.

Las características Generales de esta relevan datos la evolución de la producción de datos y actualización permanente de los mismos, comparan los datos a nivel temporal, regional e internacional y comparan la participación directa de los actores locales en la producción del dato y en la propuesta de política pública.

2. Líneas de Investigación, Desarrollo e Innovación

- Bases de Datos y Almacén de Datos.
- Minería de Datos.
- Diseño de Encuestas a medida
- Relevamiento y análisis de datos
- Software libre para el aprendizaje automático o Minería de Datos.
- Elaboración de metodología relacionada con la aplicación de Minería de Datos, aplicada al ámbito industrial.

- Implementación de técnicas de clasificación, usando software especializado.
- Métricas de calidad para el modelo de conocimiento obtenido.

3. Resultados Obtenidos/Esperados

El Objetivo Principal de esta propuesta es estudiar, analizar, describir y diagnosticar recursos de TI/SI en las Pymes cordobesas determinando indicadores que reflejen la situación actual y permitan el diseño de modelos que optimicen su gestión competitiva.

Objetivos Específicos:

- Trazar un mapa inicial descriptivo de la situación general actual de las Pymes frente a la administración de las TI/SI.
- Evaluar el grado de inserción de las TI/SI en la gestión de las Pymes bajo análisis.
- Diagnosticar posibles problemas que afectan la informatización de estos casos bajo análisis.
- Formular estrategias de solución que permitan optimizar la gestión de las TI/SI.
- Evaluar la factibilidad técnica, económica y operativa de implementar soluciones ya formuladas.
- Diseñar modelos metodológicos basados en técnicas que reúnan las mejores prácticas para cada perfil de Pyme.
- Transferir herramientas, metodologías y estrategias estudiadas, evaluadas y/o diseñadas tanto al Sector Pyme como al seno de las cátedras involucradas
- Difundir los resultados y conclusiones obtenidos al ámbito científico y académico mediante su presentación y publicación en Congresos.
- Formar recursos humanos mediante el sistema de becas de alumnos y graduados así como a través de la dirección de trabajos de Práctica Supervisada y Tesis de carreras de grado.
- Diseñar un instrumento operativo para monitorear las actividades de las Pymes en la región Córdoba, desde el seno de la Universidad.

La metodología concebida como un conjunto de pasos en un orden determinado que permiten el logro de un objetivo, se

fundamenta en algunas etapas de la metodología de Sistemas de Información y se sustenta en trabajos de indagación y reflexión crítica en el seno del equipo de trabajo así como con especialistas en el ámbito de Congresos y reuniones académicas; en la revisión del estado del arte en la temática y de otras disciplinas relacionadas (escenario industrial) así como en el estudio de casos y la práctica de investigaciones de campo en Pymes de distintos rubros de la región.

La investigación a realizar será inicialmente de tipo exploratorio, tomando luego carácter descriptivo y correlacional, con intervenciones en las Pymes testigo en un formato de investigación-acción. Se contará para ello con datos proveniente de fuentes primarias y secundarias que serán recolectadas y evaluadas con métodos cualitativos y cuantitativos.

En función de los objetivos propuestos, se plantean los siguientes momentos metodológicos:

Etapa de reconocimiento.

Aporta un primer contacto con la Pyme bajo estudio (previa selección de una muestra representativa de la población), su estructura, su cultura, sus necesidades y expectativas. Permite establecer los límites necesarios para identificar desde y hasta dónde trabajar, confeccionar los objetivos principales y planificar el relevamiento. Para acercarse a los SI, existen ciertas pautas establecidas y otras no definidas que giran en torno a políticas, estrategias, planes de ejecución, etc., que ayudan a formar una mirada general sobre la empresa, la problemática del negocio, comprender su dominio, sus tecnicismos y su vocabulario específico.

Para poder acceder a todo ello, se pueden concertar entrevistas, confeccionar cuestionarios y check list, realizar observaciones, etc., y solicitar toda la documentación del negocio en la que describan misión, visión, valores, etc., y donde pueda revelarse su estructura (organigramas, manuales, modelos de procesos o eventos).

Etapa de relevamiento.

Se vale de diversas técnicas de recolección (entrevistas, cuestionarios, encuestas, y todo

formalismo utilizado en la Organización) y documentación (cursogramas, tablas de decisión, modelos de procesos, datos, eventos de dominio, etc.) de la información necesaria (circuitos administrativos, requerimientos y requisitos, registros, etc.) para la construcción de un modelo de análisis (modelo de la realidad). En relación con las técnicas de recolección de información, la entrevista en todas sus variantes, es de las más importantes. Puede resultar provechoso, también, tener acceso a las características de la empresa por medio de organigramas, manuales institucionales, House Órganos glosarios de empresa, en los que se encuentran definidos los objetivos, misión, visión y políticas generales. Es decir, toda documentación existente en la empresa, puede servir como fuente. Con respecto a las técnicas de documentación, son muy útiles los resultados de entrevistas, reconocimientos, informes, etc. Cabe resaltar que, en esta etapa, las herramientas de modelado son usadas para reflejar la realidad y hacer un diagnóstico, por ende, tendrán una técnica particular asociada que, en algunos casos, podrá diferir de la técnica usada para la misma herramienta en una fase anterior.

Etapas de Diagnóstico.

Determina las causas que dan origen al problema y lo define, y establece alternativas. El diagnóstico se hace en base al modelo construido en el "Reconocimiento" y debe ser lo suficientemente objetivo y riguroso para poder hacer un juicio de valor correcto, eficaz y eficiente, y con el fin de encontrar los verdaderos problemas que afectan a la organización en cuestión. Poner en palabras el problema, definirlo en un modo apropiado, constituye el primer paso en la búsqueda de una solución.

Etapas de Análisis.

Mediante el uso de herramientas pertinentes a la selección y evaluación de hardware y software (matriz de homogeneización, método analítico, benchmarking, entre otros), se evaluarán parámetros homogéneos a identificar como: Sistema Operativo, Software de oficina, Software de Aplicación discriminado en Software de Uso General o Enlatado y Software desarrollado a medida;

así como las características técnicas del hardware que los soporta, entre los más representativos.

Estudio de Factibilidad.

Es el momento de evaluar las alternativas originadas en el punto anterior. En virtud del principio de equifinalidad, es posible arribar al mismo destino por varios caminos; el estudio de factibilidad, pues, elige el camino más apropiado según una serie de criterios establecidos en función de aspectos económico-financieros, técnico-operativos, políticos, legales, derivados de la cultura propia de la organización, etc. Tras el análisis de las alternativas se procede a analizar si la propuesta implementada hoy está basada en algunos parámetros como: compra o alquiler de Hardware; compra, adaptación o desarrollo de Software; capacitación, formalización de procesos no informatizados; tercerización de procedimientos (outsourcing); telecomunicaciones; entre otros.

Etapas de Diseño.

En función de los hallazgos obtenidos en etapas anteriores, se segmentarán las Pymes bajo estudio, en forma representativa (por ej.: alimenticias, metalmecánica y de servicios) y se elaborará un primer Informe a priori de los hallazgos encontrados en esta instancia. Se prevee cargar los datos y procesarlos en un soporte adecuado (Software de uso estadístico como SPSS) que permita comparar los principales parámetros representativos de cada perfil de Pyme, con la finalidad de hacer inferencias fiables de este estudio. Cabe aclarar que, si es necesario retroalimentar el análisis de más datos, se repite el proceso las veces que sea necesario. Dado que el rol del profesional es ajeno a la Organización, la objetividad y e imparcialidad prevalecerán en esta etapa.

Etapas de discusión de Informes:

- **Interna:** en el seno de las cátedras involucradas se analizarán y evaluarán los resultados obtenidos y las conclusiones a las cuales se arribó, constatando la veracidad de las fuentes relevadas y la fiabilidad de los datos registrados.

Las cátedras involucradas son: Administración de Recursos (asignatura

troncal de 4° Nivel de Ing. en Sistemas), Informática I (asignatura de 1° Nivel de Ing. Industrial) y Fundamentos de Informática (1° Nivel de Ing. Mecánica).

- **Externa:** en el ámbito científico y académico mediante su presentación y publicación en Congresos Nacionales e Internacionales.

Etapas de Transferencia:

- **Interna:** continuar con la formación de recursos humanos mediante el sistema de becas de alumnos y graduados; así como de la dirección de trabajos de Práctica Supervisada y Tesis de carreras de grado.
- **Externa:** proponer el diseño de un instrumento operativo para monitorear las actividades de las Pymes en la región Córdoba, desde el seno de la Universidad; así como ofrecer servicios de consultoría a Pymes del medio a nivel asesoramiento TI/SI.

Cabe aclarar que, una etapa puede superponerse con otra, en función del nivel de avance de cada una y que el proceso de retroalimentación o feedback

4. Formación de Recursos Humanos

La estructura del PID IFN 4346. y los respectivos niveles de categorización de cada integrante son:

Directora: Ing. Blanca Carrizo (“C” y “IV”).

Co-Directores:

Ing. Jorge E. Abet (“C” y “IV”).

Ing. Carlos Colazo (“C” y “III”).

Integrantes:

Ing. Adriana Olmedo (“E” y “IV”).

Ing. Beatriz Kunda (“E” y “V”).

Ing. Simieli, Paola

Ing. Olivo Aneiros, Federico Raúl

Becarios:

Flageat, Soledad

Huide, Gabriela Alejandra

Relevamiento en empresas del medio a través de Encuestas y Trabajos Prácticos de las cátedras de Informática y Fundamentos de Informática, de Ing. Industrial y Mecánica respectivamente.

Se compartirán análisis y experiencias profesionales desde la cátedra Administración de Recursos de Ing. Sistemas de Información

5. Bibliografía

- Colonnello, J. C. S. M. B. Innovación Empresaria En Argentina Difusión de TICs en las Pymes-Segunda parte-Juan Carlos Scarabino-Mariana Belén Colonnello. Consejo Asesor Académico, 79.
- Lorenzo, C. A., Elisondo, L., & Errandosoro, F. (2010). Uso de TIC en empresas PyMEs de la Cámara Empresaria de la ciudad de Tandil. *Técnica administrativa*, 9(41), 1.
- Peirano, F., & Suárez, D. (2004). Estrategias empresariales de uso y aprovechamiento de las TICs por parte de las PyMEs de Argentina en 2004. In Ponencia presentada en el 33 JAIIO, Simposio sobre la Sociedad de la Información.
- Scarabino, J. C., & Colonnello, M. B. (2009). Innovación empresaria en Argentina. Difusión de TICs en las PyMEs. *Invenio: Revista de investigación académica*, (22), 93-108.
- Trujillo, M. L. (2007). Planeación estratégica de tecnologías informáticas y sistemas de información. Universidad de Caldas.

Reuso Orientado a Servicios: Compatibilidad y Complejidad de Servicios

Andres Flores, Alejandra Cechich, Martin Garriga,
 Marcelo Moyano, Alan De Renzis, Diego Anabalon, Franco Corgatelli
 Grupo de Investigación en Ingeniería de Software del Comahue (GIISCo)
<http://giisco.uncoma.edu.ar>
 Departamento de Ingeniería de Sistemas – Facultad de Informática
 Universidad Nacional del Comahue
 Buenos Aires 1400, (8300) Neuquén
 Contacto: [andres.flores, alejandra.cechich]@fi.uncoma.edu.ar

Resumen

El reuso de artefactos software brinda oportunidades para proveedores y clientes, tanto para acelerar el proceso de desarrollo de software como para establecer oferta de productos reusables. El paradigma de Computación Orientada a Servicios (SOC), promueve el desarrollo de aplicaciones distribuidas en ambientes heterogéneos, que son construidas ensamblando o componiendo servicios reusables, que se publican a través de una red y se acceden mediante protocolos específicos. SOC ha sido ampliamente adoptado bajo su implementación con la tecnología de Servicios Web, que provee flexibilidad de ejecución remota que oculta las plataformas específicas de ejecución y permite descentralizar los procesos de negocios. SOC requiere la publicación de servicios en un registro (UDDI de acuerdo a Servicios Web), los cuales luego son identificados y evaluados para una aplicación en desarrollo. Sin embargo, aún este proceso necesita métodos exhaustivos y eficientes, tanto para identificación como para selección de servicios, en el cual se puede considerar la aplicación de técnicas de Pruebas de Software y el uso de dos conceptos actuales: Orquestación y Coreografía de servicios.

Palabras Clave: Ingeniería de Software basada en Reuso – Software Orientado a Servicios – Servicios Web – Calidad de Software – Verificación y Validación.

Contexto

La línea presentada se inserta en el contexto de los siguientes proyectos y acuerdos de cooperación:

- 04/F001: “Reuso Orientado a Servicios”, sub-proyecto del Programa “Desarrollo Orientado a Reuso”. Financiado por UNCo. (2013-2016).
- PICT-2012-0045: “Mecanismos y Herramientas para Grid Híbridos Orientados a Servicios”. Financiado por ANPCyT. (09/2013-03/2017).
 - Investigaciones conjuntas con ISISTAN-UNICEN, Tandil.
- Acuerdo de Cooperación con el Grupo Alarcos, Escuela Superior de Informática, Universidad de Castilla-La Mancha, España.

Introducción

Actualmente la industria de software observa cómo mediante el reuso de software se puede alcanzar un proceso de desarrollo de software acelerado y confiable al basarse en artefactos software que ya han sido probados en diferentes contextos de aplicación. Para ello se

adopta el concepto denominado “tercerización”, por medio del cual se acuerdan contratos para adquisición y provisión de artefactos software reusables y se establecen relaciones comerciales entre vendedores y clientes. Por lo tanto, desde el punto de vista de un cliente implica la posibilidad de acelerar el desarrollo de un producto software para reducir el lanzamiento al mercado, y desde el punto de vista de un proveedor implica la posibilidad de observar sus productos con una perspectiva nueva que los coloque dentro del mercado de artefactos reusables.

Un paradigma que promueve altamente el reuso de software se denomina Computación Orientada a Servicios (SOC), donde la funcionalidad a ser reusada adopta la forma de servicios, o unidades lógicas que presentan entornos heterogéneos de ejecución y pueden ser ensambladas para formar otras unidades lógicas de mayor nivel de abstracción que resuelvan (directamente o en parte) los procesos de negocios para un contexto de aplicación [SH05,PTDL07]. El paradigma SOC encontró una plataforma potencial de aprovechamiento mediante la Web, desde donde se desarrolló la tecnología de Servicios Web [NSS03, Wetal05], cuya base es el esquema estándar XML y el Lenguaje de Descripción de Servicios Web (WSDL), que facilitan ampliamente el desarrollo y mantenimiento de especificaciones formales de servicios. Así el paradigma SOC bajo la implementación con Servicios Web ha logrado su amplia adopción en la industria, principalmente bajo la flexibilidad de ejecución remota que permite a las compañías descentralizar aún más sus procesos de negocios y la ventaja de que las plataformas específicas de ejecución se encuentran ocultas, por lo cual no se requiere de inversiones adicionales en tecnología (incluyendo

costos y esfuerzo de aprendizaje) al adquirir funcionalidad de terceras partes. El beneficio que la tecnología de Servicios Web provee al paradigma SOC se ha denominado “relación sin responsabilidad”, donde una aplicación cliente no requiere asumir cómo se ha implementado el servicio con el que se comunica. Sin embargo, los proveedores de servicios tienen la responsabilidad de evaluar la calidad de los productos ofrecidos como servicios y los consumidores de servicios a su vez deben ser capaces de identificar tal calidad que influirá sobre las aplicaciones en desarrollo. Para ello, las estrategias de Pruebas de Software deben ajustarse a este contexto específico donde los servicios están acordados como cajas negras que sólo permite evaluar el comportamiento y cualidades observables externamente [BDN10, Z08].

El funcionamiento concreto del paradigma SOC se basa en la Arquitectura orientada a Servicios (SOA) [SH05] que se encuentra compuesta por tres actores principales: un proveedor, un consumidor y un registro de servicios; donde el proveedor desarrolla y publica servicios en el registro, para que luego el consumidor busque servicios y establezca una comunicación con el proveedor. Sin embargo, la búsqueda de servicios publicados en un registro UDDI (según la tecnología de servicios Web) [OASIS04], en general requiere invertir un esfuerzo considerable para distinguir servicios candidatos que satisfagan los requerimientos de la aplicación cliente [NSS03, Wetal05]. En particular, cuando varios candidatos ofrecen funcionalidades similares se requieren métodos eficientes de selección de servicios que discriminen tanto aspectos funcionales como no-funcionales, considerando además las interacciones válidas para un servicio candidato en función de los procesos de

negocio que implementará la aplicación cliente. En particular el ensamblaje de servicios considerando procesos de negocio e interoperabilidad de servicios plantea el uso de dos conceptos de reciente investigación: Orquestación y Coreografía de servicios [P03, Wetal05]. El primero relacionado a una aplicación particular que describe un proceso de negocios específico, y el segundo relacionado a las interacciones válidas que pueden ocurrir entre distintos servicios predestinados a intervenir en una colaboración. En este contexto se cuenta actualmente con diversos lenguajes de descripción y frameworks de ejecución, tales como BPEL4WS [OASIS07], y WSCDL (Web Services Choreography Description Language) [W3C05].

Líneas de Investigación, Desarrollo e Innovación

El perfil de esta línea puede definirse en base a las actividades de investigación y transferencia, a las que da soporte el grupo GIISCo. Los temas específicos consideran los desafíos diferentes asociados al crecimiento de la Tecnología de la Información y las Comunicaciones. Actualmente, abordamos los siguientes aspectos:

- Complejidad y legibilidad de servicios
- Compatibilidad y selección de servicios.
- Adaptación y Composición de servicios.
- Testing de servicios.
- Herramientas para evaluación, selección, composición y testing de servicios.
- Definición de aplicaciones en dominios específicos.

Resultados y Objetivos

En [FCGMR16] hemos enumerado una serie de contribuciones anteriores. Durante el año 2016, hemos profundizado la investigación en aspectos de *evaluación de compatibilidad y complejidad de servicios*, generando métodos y herramientas enfocados en las interfaces y comportamiento dinámico de los servicios. Este avance se ha efectuado en colaboración con investigadores de ISISTAN (UNICEN) [AGF16, AGFCZ16, ARBFC16, ARGFCZ16, GFMCZ16, GRAFC16, GRLFCZ16, MZMAF16, RGFCMZ16a-16b-16c, RGFCZ16].

Las líneas de investigación convergen en el tratamiento del desarrollo de software basado en el reuso de servicios desde la perspectiva de las aplicaciones orientadas a servicios. Una aplicación orientada a servicios implica una solución de negocio que consume servicios de uno o más proveedores y los integra en un proceso de negocio [SW04]. Además puede verse como una aplicación basada en componentes que integra dos tipos de componentes: internos localmente empujados en la aplicación, y externos estática o dinámicamente enlazados a algún servicio [CMZC14]. No solamente se enfocará en el reuso de servicios individuales, sino también en la composición de servicios como forma de tercerizar una funcionalidad. Se adoptará la visión de proceso de negocio para la definición de comportamiento, donde se aplicará testing de servicios para una evaluación dinámica. Se complementará el modelo de selección y composición de servicios mediante las últimas plataformas y avances tecnológicos incluyendo semántica y estandarización. Se prevee la aplicación de estos modelos y las herramientas de soporte a dominios específicos, con particular énfasis en

aquellos que requieran rigurosidad como aporte de validación efectiva. La visión de esta línea de investigación se resume en:

“Definir técnicas y herramientas para la mejora del desarrollo de software, en función del reuso de servicios web. La definición de modelos de identificación y selección de servicios, y la posibilidad de composición de servicios”.

Formación de Recursos Humanos

Este proyecto se compone de 11 investigadores, entre los que se cuentan docentes y estudiantes del Grupo GIISCO de UNComa y asesores externos. Algunos de los docentes–investigadores se encuentran realizando carreras de postgrado. Se cuenta actualmente con 3 doctores (1 investigador asistente CONICET), 1 postdoctorado (becario CONICET), 2 doctorandos (becarios CONICET) y 1 maestrando entre los miembros del proyecto. Dirección de Tesis de Grado durante 2016: 2 tesis.

Referencias

- [AGF16] Anabalón, D., Garriga, M., Flores A. (2016). *Modeling Web Service Selection for Composition as a Distributed Constraint Optimization Problem (DCOP)*. Simposio Argentino de Ingeniería de Software ASSE. Argentina.
- [ARGFCZ16] Anabalón, D., Garriga, M., Flores, A., Cechich, A., Zunino, A. (2016). *Pattern-based Integrability on Service Oriented Applications*. ACM SugarLoaf LA-PLOP 2016. [in press]
- [ARBFC16] Arias, M., De Renzis, A., Buccella, A., Flores, A., Cechich, A. (2016). *Classification-based Mining of Reusable Components on Software Product Lines*. IEEE Latin America Transactions, 14(02): 870–876. **(Indexed SCI, IF JCR2014: 0.326)**.
- [ARGFCZ16] Anabalón, D., De Renzis, A., Garriga, M., Flores, A., Cechich, A., Zunino, A. (2016). *Testing-supported Case-based Reasoning for Web Service Selection*. IEEE Argencon. Argentina.
- [BDN10] Baresi, L.; Di Nitto, E. (2010). *Test and Analysis of Web Services*. Springer.
- [CMZC14] Crasso, M., Mateos, C., Zunino, A., Campo, M. (2014). *EasySOC: Making Web Service Outsourcing Easier*. International Journal on Information Sciences, 259: 452–473.
- [GFMZ16] Garriga, M., Flores, A., Mateos, C., Cechich, A., Zunino, A. (2016). *RESTful Service Composition at a Glance: a Survey*. Journal of Network and Computer Applications, 60(C): 32–53. Elsevier. **(Indexed SCI, IF JCR2014: 2,229)**.
- [GRAFC16] Garriga, M. Rozas, K. Anabalón, D. Flores, A. Cechich, A. (2016). *RESTful Mobile Architecture for Social Security Services: A Case Study*. IEEE CLEI – XLII Conferencia Latinoamericana de Informática. Chile.
- [GRLFMZ16] Garriga, M., De Renzis, A., Lizarralde, I., Flores, A., Mateos, C., Cechich, A., Zunino, A. (2016). *A Structural-Semantic Web Service Selection Approach to Improve Retrievability of Web Services*. Information Systems Frontiers. [in press] **(Indexed SCI IF JCR2015: 1.450)**.
- [MZMAF16] Mateos, C., Zunino, A., Misra, S., Anabalón, D., Flores, A. (2016). *Keeping Web Service Interface Complexity Low Using an OO Metric-based Early Approach*. IEEE CLEI – XLII Conferencia Latinoamericana de Informática. Chile. [BEST PAPER]

- [NSS03] Nagappan, R.; Skoczylas, R.; Sriganesh, R. (2003). *Developing Java™ Web Services: Architecting and Developing Secure Web Services Using Java*. Wiley Publishing Inc.
- [OASIS04] OASIS Consortium (2004). *UDDI Version 3.0.2*. UDDI Spec Technical Committee Draft, October.
- [OASIS07] OASIS Standard (2007). *Web Services Business Process Execution Language Version 2.0*. <http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html>
- [P03] Peltz, C. (2003). *Web Services Orchestration and Choreography*. IEEE Computer, 36(10): 46–52.
- [PTDL07] Papazoglou, M.; Traverso, P.; Dustdar, S.; Leymann, F. (2007). *Service-Oriented Computing: State of the Art and Research Challenges*. IEEE Computer, 40(11): 38–45.
- [RGFCMZ16a] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., Zunino, A. (2016). *Assessing Readability of Web Service Interfaces*. IEEE CLEI – XLII Conferencia Latinoamericana de Informática. Chile. [BEST PAPER]
- [RGFCMZ16b] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., Zunino, A. (2016). *Assessing Web Services Interfaces with Lightweight Semantic Basis*. Computing and Informatics. [in press] (**Indexed SCI IF JCR2015: 0.504**).
- [RGFCMZ16c] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Mateos, C., Zunino, A. (2016). *A Domain Independent Readability Metric for Web Service Descriptions*. Computer Standards & Interfaces, 50: 124–141. Elsevier [early view] (**Indexed SCI, IF JCR2015: 1,268**).
- [RGFCZ16] De Renzis, A., Garriga, M., Flores, A., Cechich, A., Zunino, A. (2016). *Case-based Reasoning for Web Service Discovery and Selection*. ENTCS, Electronic Notes on Theoretical Computer Science, 321: 89–112. Special Issue of Best papers of XLI CLEI'15. Elsevier.
- [SH05] Singh M.; Huhns, M. (2005). *Service-oriented computing: Key concepts and principles*. IEEE Internet Computing, 9(1): 75–81.
- [SW04] Sprott, D.; Wilkes, L. (2004). *Understanding Service-Oriented Architecture*. The Architecture Journal. MSDN Library. Microsoft Corporation, <http://msdn.microsoft.com/en-us/library/aa480021.aspx>
- [W3C05] W3C Candidate Recommendation (2005) *Web Services Choreography Description Language Version 1.0*. <http://www.w3.org/TR/ws-cdl-10/>
- [Wetal05] Weerawarana, S.; Curbera, F.; Leymann, F.; Storey, T.; Ferguson, D. (2005). *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More*. Prentice Hall PTR.
- [Z08] Zhou, X. (2008). *Testing and Verifying Web Services. From the Researcher's Perspective*. VDM Verlag.

Selección de Atributos de Calidad de Datos en Sistemas de Gestión de Aprendizaje bajo la Familia de Normas ISO/IEC 25000

Estela Fritz¹; Germán Montejano^{1 2}; Pablo García¹

¹Departamento de Matemática
 Universidad Nacional de La Pampa
 Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
 Tel.: +54-2954-245220 – Int. 7125
 [fritzem, pablogarcia]@exactas.unlpam.edu.ar

²Departamento de Informática
 Universidad Nacional de San Luis
 Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
 Tel.: +54-2652-424027 – Int. 251
 gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

Resumen

Los Sistemas de Gestión de Aprendizaje (*LMS* por su nombre en inglés *Learning Management Systems*) se han tornado relevantes en la última década, principalmente en los niveles de educación terciario y universitario. Aunque también se ha extendido su utilización a los niveles de capacitación del ámbito empresarial

Un *LMS* no sólo constituye un repositorio de materiales curriculares de un curso. En la actualidad, su aplicación varía desde la funcionalidad como un sitio complementario de un curso curricular, hasta la posibilidad de implementar un curso no presencial con actividades sincrónicas y asincrónicas para los estudiantes.

Por otra parte, el docente o responsable de dicho curso, obtiene información sobre actividades realizadas por los alumnos, como encuestas, evaluaciones, tareas que pueden ser subidas al sitio del curso. El sistema le provee además información estadística relevante acerca de numerosos aspectos entre los cuales figuran los mencionados más arriba, organizada y presentada en formato portable a otras aplicaciones.

A su vez los alumnos pueden recibir sus calificaciones, correcciones a sus trabajos, respuestas a sus consultas, notificaciones del docente, entre otras.

De todo lo dicho anteriormente se puede inferir que un *LMS* almacena importante cantidad de datos de forma estructurada, los cuales deben atender a las normas para integridad y seguridad de los datos establecidas en los estándares ISO/IEC.

Palabras clave: Sistemas de Gestión de Aprendizaje, Calidad de los Datos, Normas ISO/IEC sobre calidad de productos de software.

Contexto

Por Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa se acredita el Proyecto de Investigación: "Aspectos de Seguridad en Proyectos de Software", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (UNLPam). El mismo es dirigido por el Doctor Germán Antonio Montejano y codirigido por el Magister Pablo Marcelo

García e incluye a la Licenciada Silvia Gabriela Bast y la Profesora Estela Marisa Fritz como investigadoras.

El presente trabajo se enmarca en el Proyecto de Investigación: "*Aspectos de Seguridad en Proyectos de Software*", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa y en el Proyecto de Investigación "Ingeniería de Software, Conceptos, Métodos y Herramientas en un Contexto de Ingeniería de Software en Evolución" de la Universidad Nacional de San Luis.

Introducción

La ausencia de un formalismo específico para medir la calidad del software y la falta de acuerdo sobre su significado han dado lugar a variados esfuerzos por lograr consenso en estos aspectos. Así, varios autores y organizaciones han elaborado definiciones al respecto.

La IEEE define calidad de software como "el grado en el cual un software posee una combinación de atributos deseados" [2]. Albin especifica que la calidad de un sistema de software es "una característica directamente relacionada con la habilidad del sistema para satisfacer requerimientos funcionales y no funcionales, tanto implícitos como explícitos" [3]. Una definición similar se presenta en [4], donde se establece que la calidad de un sistema de software es "la concordancia del sistema con los requisitos funcionales y de rendimiento explícitamente establecidos, con los estándares de desarrollo explícitamente documentados y con las características implícitas que se esperan de todo software desarrollado profesionalmente"

En todas las definiciones precedentes se destaca la importancia de evaluar la calidad del software en función del estudio de los atributos asociados a los requerimientos del sistema. Un atributo de calidad (AC) es una propiedad específica de un sistema de software que puede asumir un valor cualitativo o cuantitativo, el cual es medible u observable [3]. Debido a que entre los distintos AC existen relaciones e

interacciones, no es posible que un sistema complejo alcance un AC en forma aislada. [5]

1. Serie ISO/IEC 25000: SQuaRE (Software Product Quality Requirements and Evaluation)

La familia ISO/IEC 25000 es una familia de normas que tiene por objetivo la creación de un marco de trabajo común para evaluar la calidad del producto software. Es el resultado de la evolución de otras normas anteriores y se encuentra compuesta por cinco divisiones.

La serie SQuaRE define un marco teórico que contextualiza múltiples aspectos de la calidad del software.

Se busca incorporar indicadores específicos de modo que la seguridad se convierta en un ítem de relevancia entre los criterios para la evaluación de la calidad del software, principalmente en lo que se refiere a la integridad de los datos.

El presente trabajo propone la selección de algunos aspectos de los *LMS* para ser empleados como atributos para evaluar calidad. Se han tomado como base los atributos de calidad establecidos en los estándares de ISO/IEC 25010 [7], ISO/IEC 25012: 2008[8], y además ISO/IEC 25040 [9] que definen un modelo general de calidad de productos de software, un modelo para calidad de producto de datos almacenados de forma estructurada dentro de un sistema y definición del proceso para llevar a cabo la evaluación de calidad, respectivamente.

Se han tenido en cuenta sólo algunas normas de dicha familia, principalmente aquellas referidas específicamente a seguridad de los datos, y que se describen brevemente a continuación:

1.1 Norma ISO/IEC 25010

El modelo de calidad del producto definido por la ISO/IEC 25010 se encuentra compuesto por ocho características de calidad que se dividen a su vez en sub-características y pueden observarse en la siguiente figura:

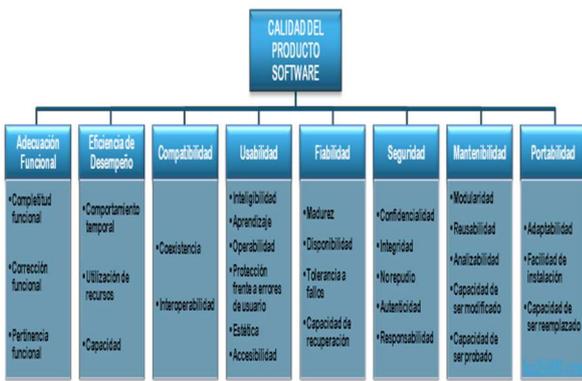


Fig. 1 Modelo de calidad del producto software

De todas estas características se pone énfasis en seguridad con las sub-características: *confidencialidad – integridad – no repudio – autenticidad – responsabilidad.*

1.2 Norma ISO/IEC 25012

El modelo de Calidad de Producto de Datos definido por el estándar ISO/IEC 25012 se encuentra compuesto por las 15 características que se muestran en la siguiente figura y se dividen en dos grandes categorías:

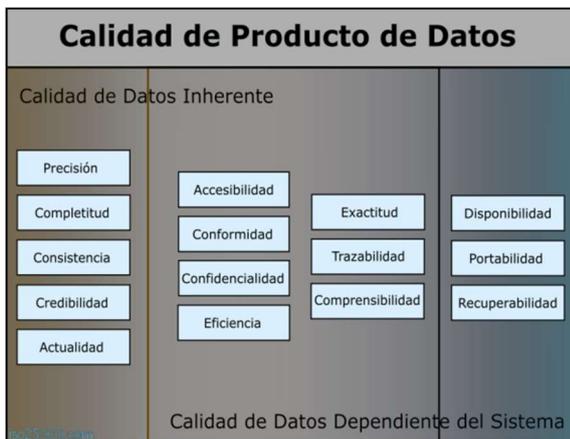


Fig. 2 Modelo de Calidad de Producto de Datos

Calidad de datos Inherente y Calidad de Datos Dependiente del Sistema, haciendo esta última categoría referencia al grado con el que la Calidad de Datos es alcanzada y preservada a través de un sistema informático cuando los datos son utilizados bajo condiciones específicas.

1.3 Norma ISO/IEC 25040

Define el proceso para llevar a cabo la evaluación del producto software. Dicho proceso de evaluación consta de un total de cinco actividades, como puede observarse en la siguiente figura:

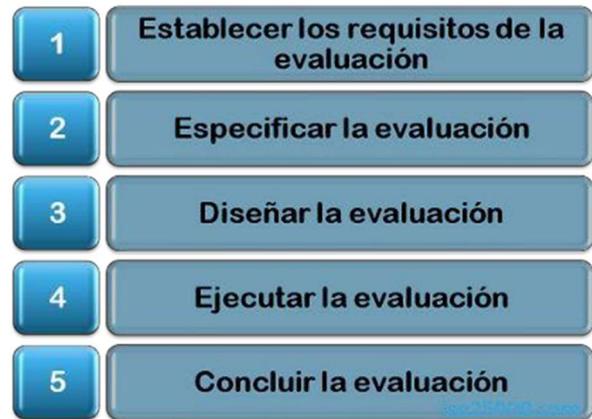


Fig. 3 Proceso para la Evaluación del producto Software

Líneas de Investigación, Desarrollo e Innovación

El grupo de trabajo investiga, básicamente sobre tres campos relacionados:

- Protección del anonimato de los votantes en sistemas de voto electrónico ([10]).([12])
- Integridad de los datos de un sistema de e-Voting ([11]).
- Integridad de las bases de datos almacenados de forma estructurada dentro de un sistema, más específicamente un sistema de gestión de aprendizaje.

El presente trabajo corresponde a la tercera línea de investigación

Resultados y Objetivos

Como puede observarse en la fig. 1., la norma ISO/IEC 25010 propone un esquema con ocho características, cada una con sub-características. Sin embargo, los atributos no se encuentran definidos en el estándar debido a que su aplicabilidad varía según el tipo de producto de software analizado.

Esos atributos específicos se desprenden de los aspectos establecidos en la norma. Estos aspectos conformarán el Árbol de criterios para aplicar el método LSP

propuesto en [6] para aplicar a la evaluación de la integridad de los datos. Luego es necesario fijar atributos de calidad específicos, que conformarán los criterios elementales y se derivan del árbol de criterios.

Así, aspectos del *LMS* relacionados con esos criterios elementales [6] o atributos específicos [5] serán insumo para el cálculo de la Preferencia Global [6]. Algunos de esos atributos se mencionan a continuación:

- Número de usuarios que han mantenido derechos y privilegios de acceso históricos, en cada categoría.
- Registro de cuentas huérfanas o inactivas, cuantificado.
- Registro de accesos no autorizados a los datos, que hubiesen sido identificados.
- Cuantificación de accesos no autorizados en un período (en un mes/ en un año).
- Registro estadístico de las vistas de un usuario de un curso a los distintos materiales.

La elección de LSP como método para realizar la evaluación, constituye la segunda actividad que establece la norma ISO/IEC 25040. El cálculo de la Preferencia Global con los propios operadores, se corresponde con la actividad 3 establecida en la norma.

A futuro, se prevé evaluar la calidad de Producto Datos en el *LMS* instalado en la Facultad de Ciencias Exactas y Naturales de la UNLPam.

Para ello se aplicará la norma ISO/IEC 25012, teniendo en cuenta los pasos para la Evaluación establecidos en ISO/IEC 25040 y el método LSP.

Formación de Recursos Humanos

En el marco del presente proyecto se presentan los siguientes puntos relacionados con la formación de recursos humanos

- La Lic. BAST, Silvia alcanzó el grado de Magíster en Ingeniería de Software, tras la defensa de su tesis titulada “Optimización de la Integridad de Datos en Sistemas de E-Voting”. La tesis fue dirigida por el Dr. Germán Montejano y el Mg. Pablo García, directores del proyecto de investigación y recibió la calificación de “sobresaliente”.

- La Lic. BAST, Silvia completó el cursado de la carrera de posgrado Doctorado en Ingeniería de Software de la Universidad Nacional de San Luis. Se presentará el Plan de Tesis Doctoral durante 2017.
- El Mg. GARCIA, Pablo completó el cursado de la carrera de posgrado Doctorado en Ingeniería de Software de la Universidad Nacional de San Luis. Se presentará el Plan de Tesis Doctoral durante 2017.
- El Mg Pablo M. GARCÍA obtuvo una beca de perfeccionamiento docente otorgada por el CAFB-BA (Centro Asociados para el Fortalecimiento de Posgrados Brasil Argentina). La beca incluyó una estadía desde el 18 de octubre hasta el 18 de noviembre de 2016 en la Universidad Federal de Minas Gerais, Brasil, con el fin de realizar cursos de perfeccionamiento docente, académico, científico y tecnológico en informática. Coordinador del Proyecto: Dr. Roberto Uzal. Resolución N° 505/16 de la decana de la FCEyN de la UNLPam.
- Estela Marisa Fritz: completó su etapa de capacitación en un tema en el que no era experta. Durante 2017 realizará todos los aportes relacionados con la temática de generadores aleatorios, insumo necesario para los nuevos avances en el proyecto. Dado que se espera, a futuro, completar el modelo de voto electrónico e implementarlo, es necesario producir avances en el tema mencionado.

Referencias

- [1] Uzal R., van de Graaf J., Montejano G., Riesco D., García P.: “Inicio de la Línea de Investigación “Ingeniería de Software y Defensa Cibernética”. Memorias del XV Workshop de Investigadores en Ciencias de la Computación 2013 (WICC 2013).Ps.769 - 773. ISBN: 9789872817961. 2013.
- [2] IEEE Std. 1061-1998 IEEE Standard for a Software Quality Metric Methodology-Description. “Institute of Electrical Engineering”, 1998

[3] **S. Albin**, The art of software architecture: design methods and techniques. John Wiley & Sons, 2003.

[4] R. Pressman, Software Engineering: A Practitioner's Approach, 7th ed. McGraw-Hill, 2010

[5] **Blas M. Julia, Gonnet S., Leone, H.** “Una Taxonomía de Atributos de Calidad para la Evaluación de Arquitecturas de Software por medio de Simulación” Conference: 2° CoNaIISI Congreso Nacional de Ingeniería en Informática / Sistemas de Información, ISSN: 2346-9927, 13 y 14 de noviembre 2014 San Luis-Argentina.

[6] **Dujmovic J.J.** A Method for Evaluation and Selection Hardware and Software Systems. *The 22nd Int'l Conference for the Resource Management and Performance Evaluation of Enterprise CS. CMG 96 Proceedings, 1:368-378, 1996*

[7] **International Standard ISO/IEC 25010** Systems and software engineering Systems and software Quality Requirements and Evaluation (SQuaRE) System and software quality models. 2011

[8] **International Standard. ISO/IEC 25012** Software Engineering Software Quality Requirements and Evaluation (SQuaRE) Data Quality Model. 2008

[9] **International Standard ISO/IEC 25040** Systems and software engineering Systems and software Quality Requirements and Evaluation (SQuaRE) Evaluation process. 2011

[10] **García P. M.** “Una optimización para el Protocolo Non Interactive Dining Cryptographers”. ISBN-13: 978-3-639-85270-7. ISBN-10: 3639852702. EAN: 9783639852707. Editorial Académica Española (<https://www.eae-publishing.com/>) 2017

[11]. **Bast S., Montejano G., García P., Fritz E.:** “Evaluación de la integridad de datos en Sistemas de e-Voting”. XVII Workshop de Investigadores en Ciencias de la Computación (WICC 2015). 16 y 17 de abril de 2015. Facultad de Ciencias Exactas. Universidad Nacional de Salta. ISBN:978-987-633-134-0.

[12] **García P., Montejano G., Bast S., Fritz E.:** “loss of votes in NIDC Applying Storage in Parallel Channels”. Congreso Argentino de Ciencias de la Computación, CACIC 2016. San Luis, 3 al 7 de octubre de 2016. Universidad Nacional de San Luis (UNSL). Se obtiene

distinción como MEJOR EXPOSITOR del Workshop de Seguridad Informática. El trabajo ingresa en categoría “Selected Paper”

Selección de Metodologías Ágiles e Integración de Arquitecturas de Software en el Desarrollo de Sistemas de Información

Mg. Mirta E. Navarro¹, Mg. Marcelo P. Moreno², Lic. Juan Aranda³, Lic. Lorena Parra⁴,
Lic. José R. Rueda⁵, Juan Cruz Pantano⁶

Departamento de Informática - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan

¹mirtaenavarro@yahoo.com.ar ²mpmoren@gmail.com ³juanaranda@live.com ⁴lorenaparra152@yahoo.com.ar
⁵joseruardorueda@hotmail.com ⁶juancruz871@hotmail.com

Resumen

La Ingeniería de Software (IS) se ha convertido imprescindible en el ámbito organizacional, su desenvolvimiento y gestión depende en gran medida de los Sistemas de información (SI) y de las Tecnologías de la información (TI).

En los últimos años, se ha impuesto el uso de las Metodologías Ágiles (MA), marcando una tendencia en su adopción al desarrollo de proyectos de software. La causa principal es que en ambientes donde las necesidades de las organizaciones y la tecnología cambian rápidamente, las metodologías tradicionales predictivas han demostrado ser poco eficientes para atender los requerimientos de clientes y usuarios, limitando la competitividad y a la obtención de mayores beneficios en la producción de bienes o en la prestación de servicios, en el menor tiempo posible, y es en ese escenario donde las MA han ganado bastante popularidad siendo una muy buena solución para proyectos a corto plazo, en especial, aquellos proyectos en donde los requisitos están cambiando constantemente.

La Arquitectura de Software (AS) comprende elementos de software, las propiedades externamente visibles de aquellos elementos y su interrelación para satisfacer la funcionalidad y requerimientos deseados. Si bien en los últimos años, ha comenzado a cobrar una mayor importancia dentro de los estudios

propiciados en la IS, el desarrollo de la AS es una práctica poco común para algunos desarrolladores, en especial si se utilizan metodologías de desarrollo ágiles, en las que esas actividades no se consideran relevantes.

En este trabajo se presenta un avance de la exploración, comparación y selección bajo distintos criterios de dos metodologías ágiles, con el propósito de incluir actividades de diseño de arquitecturas de software, que permita generar un modelo genérico de integración, que pueda ser aplicado a otras metodologías bajo el dominio agilista, sin dejar de lado las consideraciones de calidad, riesgo y costos asociados a la integración pretendida.

Palabras clave: *Software Architecture, Agile methodologies, Information Systems*

Contexto

El presente trabajo se encuadra dentro del área de I/D de la IS y de los SI, aborda el análisis bajo distintos criterios y selección de metodologías ágiles, uno de los dos ejes fundamentales del proyecto de investigación: “Integración de Metodologías Ágiles y Arquitecturas de software en el Desarrollo de Sistemas de Información”, presentado en WICC 2016 [2]. El proyecto se encuentra acreditado y financiado por la Secretaria de Ciencia y Técnica Universidad Nacional de San

Juan (UNSJ). Iniciado en Enero de 2016, con una duración de dos años, que tiene como unidades ejecutoras al Departamento e Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNSJ. El grupo de investigación tiene una trayectoria de varios años en diferentes proyectos vinculados a Metodologías de Desarrollo y Tecnologías, con numerosas publicaciones en diferentes ámbitos, y con la formación de recursos humanos en el área de interés.

Introducción

La Ingeniería de Software se ha convertido en una disciplina imprescindible debido a que la mayoría de los proyectos de Sistemas de o atributo de calidad, esto ha motivado la integración de actividades, técnicas y patrones de arquitectura en procesos de desarrollo ágiles.

Esta idea de integración de metodologías ágiles y arquitectura de software constituye el objetivo principal del Proyecto de investigación referenciado. Para su consecución se plantearon actividades de análisis, comparación y selección de metodologías ágiles, exploración de arquitecturas y propuesta de un modelo de integración.

Metodologías Ágiles

Son metodologías adaptativas, que permite llevar a cabo, proyectos de desarrollo de software, adaptándolo a los cambios como una oportunidad para mejorar el sistema e incrementar la satisfacción del cliente, considerando la gestión de cambios como un aspecto inherente al propio proceso de desarrollo software y, permitiendo de este modo, una mejor adaptación al entorno, maximizando la inversión y reduciendo los costos, ya sea para variar parte de su funcionalidad, añadir otra nueva, o por ejemplo, adaptar el sistema a un nuevo dominio de aplicación[1]. Las metodologías ágiles se caracterizan por el desarrollo iterativo e incremental, las entregas frecuentes, la

Información que se impulsan dependen en gran medida de la aplicabilidad de Metodologías de la elección de la AS y de las actuales características dinámicas y variables de las organizaciones que están reconsiderando las bases sobre las que se sustenta el desarrollo software.

Para el desarrollo de software existen numerosas propuestas metodológicas que inciden en distintas dimensiones del proceso, desde las metodologías tradicionales predictivas hasta las actuales metodologías giles que se han impuesto en el desarrollo de proyectos de software, marcando una tendencia su adopción.

Por otra parte las arquitecturas en el desarrollo de software están cobrando cada vez mayor importancia con

simplicidad de la implementación; priorizan los requisitos, comunicación continua con el cliente y la cooperación entre los desarrolladores.

Estas metodologías constituyen una solución con una elevada simplificación, que a pesar de ello no renuncia a las prácticas esenciales para asegurar la calidad del producto. Por lo tanto, para ejecutar un desarrollo exitoso con metodologías ágiles, es importante comprender los requisitos del cliente, intentando extraer al máximo los deseos del usuario para entregar un producto lo más cercano a la realidad. Al ser flexibles y colaborativas, se ajustan fácilmente a cambios de requisitos, de hecho, el cambio en los requerimientos es una característica esperada y deseada, al igual que las entregas constantes al cliente y la retroalimentación por parte de él. Tanto el producto como el proceso son mejorados frecuentemente y la capacidad de entrega continua y en plazos breves permite al cliente verificar in situ el desarrollo del proyecto y familiarizarse con la funcionalidad del producto progresivamente y comprobando si el sistema cumple con los requerimientos, mejorando de esta forma su satisfacción. [2].

Además, el desarrollo en ciclos de corta duración favorece que los riesgos y dificultades se repartan a lo largo del desarrollo del producto, principalmente al comienzo del mismo que permite analizar las dificultades.

Teniendo en cuentas, características, fases, colaboración con el cliente, iteraciones desarrollo incremental, elementos, ciclo de vida, roles y otros factores preponderantes en la comparación realizada de ocho metodologías ágiles surge la elección de ICONIX y SCRUM, para efectuar la integración de una arquitectura de software flexible y adaptable a dichas metodologías, que permita construir un sistema o producto de calidad. Las mencionadas metodologías se ajustan fácilmente a cambios de requerimientos, de hecho, el cambio en los requerimientos es una característica esperada y deseada, al igual que las entregas constantes al cliente y la retroalimentación por parte de él. Tanto el producto como el proceso son mejorados frecuentemente y la capacidad de entrega continua y en plazos breves permite al cliente verificar in situ el desarrollo del proyecto y familiarizarse con la funcionalidad del producto progresivamente y comprobando si el sistema cumple con los requerimientos, mejorando de esta forma su satisfacción. [2]. [3].

Arquitectura de Software.

La Arquitectura de Software reúne todos los requerimientos técnicos y operacionales, se compone de elementos de software, las relaciones entre ellos, las propiedades de ambos (elementos y relaciones) y de patrones arquitectónicos que guían a esta organización. Estas incumbencias van más allá de los algoritmos y estructuras de datos de los programas. [4] La arquitectura es principalmente importante para satisfacer los requisitos no funcionales, que están relacionados a los atributos de calidad como el rendimiento, seguridad y escalabilidad.

Los componentes fundamentales, en la AS, son los patrones arquitectónicos, siendo un conjunto de principios que proporcionan un marco para el diseño, proveen la forma a una aplicación mejoran la partición, ayudan a definir las características básicas y de comportamiento de una aplicación. Por ello, es necesario

investigar, analizar conocer claramente las características, fortalezas y debilidades de cada patrón de arquitectura y seleccionar aquel que se adapte a la solución de un problema determinado.

Dado la importancia que existe sobre el papel de la AS en las MA, los autores Edwin Rafael Mago y Germán Harvey Alférez [11] presentan una interesante propuesta para gestionar la arquitectura de software en SCRUM, tal propuesta consiste adicionar un sprint inicial llamado "*Sprint 0*" al inicio del ciclo de desarrollo para analizar y diseñar la generalidad del sistema, que satisfaga los requisitos y sea entendible por los miembros del equipo desde sus diferentes puntos de vista durante el desarrollo. Un punto clave, es reutilizar artefactos de software creados a partir de la arquitectura para ser más ágiles en el desarrollo de productos específicos. Construye en el Sprint 0 la arquitectura de forma iterativa mediante un análisis preliminar de los conductores arquitectónicos (requisitos funcionales, de calidad y del negocio), y de un estudio de factibilidad del proyecto. El proceso de análisis de requisitos será entonces influenciado por las preguntas generadas durante el diseño arquitectónico. El resultado del Sprint 0 es un documento inicial que explica la arquitectura mediante un proceso de descomposición basado en los atributos de calidad del software.

Pero sería importante desarrollar un modelo genérico de estructura de software que pueda ser integrada a cualquier metodología ágil desde las fases iniciales de desarrollo y que incluya flexibilidad ante los cambios y adapten los criterios agilistas. Esto conlleva a decidir el paradigma más efectivo para integrar la arquitectura a los métodos ágiles, en el desarrollo de SI, respetando los criterios agilistas y que se distingan claramente los beneficios de aplicarla en todas las etapas de desarrollo.

Los autores del presente trabajo están elaborando un modelo que desde las MA se puede implantar una arquitectura que sea ágil también, basada en herramientas, patrones y

componentes reutilizables que tenga en cuenta los requisitos funcionales, no funcionales y atributos de calidad.

Líneas de investigación y Desarrollo

Es de profundo interés para este grupo de investigación trabajar con la metodología ágil ICONIX Y SCRUM para lograr la integración de una Arquitectura de software que sea soportadas por ellas y no disminuyan sus capacidad, principios y criterios agilistas.

Los posibles beneficios de una integración de AS en las MA no es un tema lo suficientemente explorado en ambientes académicos, habiéndose encontrado escasa cantidad de publicaciones relacionadas.

Por otro lado, el interés de indagar en posibles estrategias de integración y posterior evaluación de riesgos/costos, surgieron de inquietudes emergentes en proyectos de investigación anteriores que este grupo llevó a cabo en [3 y 9], siempre siguiendo la línea de las Metodologías de desarrollo, en especial de las ágiles.

ICONIX: es una metodología ágil con un claro proceso simplificado, iterativo e incremental que unifica un conjunto de métodos de orientación a objetos con el propósito de abarcar todo el ciclo de vida de un proyecto, está adaptada a los patrones de UML, su proceso dirigido por casos de uso. Posee cuatro fases y cuatro hitos de fácil comprensión y aplicación. Cada paso que se realiza está definido por un requisito, se define la trazabilidad como la capacidad de seguir una relación entre los diferentes artefactos de software producido.

SCRUM: Sigue el desarrollo de las MA, parte de la visión, del concepto general del producto, y sobre ella el equipo produce de forma continua incrementos en la dirección apuntada por la visión; y en el orden de prioridad que necesita el negocio del cliente.

Comprende cinco fases y se construye el producto de forma incremental a través de

Información, tendiendo a lograr un equilibrio en la integración y reducir la exposición al riesgo.

iteraciones breves Estas iteraciones (en Scrum llamadas sprints) se repiten de forma continua hasta que el cliente da por cerrado el producto. Estas iteraciones son la base del desarrollo ágil, y Scrum gestiona su evolución en reuniones breves diarias donde todo el equipo revisa el trabajo realizado el día anterior y el previsto para el siguiente. [3, 9].

En cuanto al análisis de las variables deseables que deben poseer las metodologías de software, las dos presentan virtudes destacables, Scrum es una que las favoritas por poseer certificaciones. Con relación a la evolución de un sistema después de mucho tiempo, Iconix es la que posee documentación suficiente para cumplir con esta condición.

Las dos metodologías cumplen casi sin diferencias entre ellas, el hecho de que deben cubrir el ciclo entero de desarrollo de software, integrar las distintas fases del ciclo de desarrollo, poder emplearse en un entorno variables de proyectos Software

Resultados y Objetivos

Los resultados que se esperan obtener al finalizar la investigación son:

- Proponer un modelo de Arquitectura de Software sin contraponerse con las postulaciones del manifiesto de las metodologías Ágiles.
- Aplicar la integración de Metodología Ágil y la Arquitectura de Software propuesta al desarrollo de diferentes SI de pequeñas organizaciones.
- Creación de un marco de referencia que permita analizar riesgos y estimar costos de la implementación de SI guiada por la integración propuesta.

El objetivo principal es establecer la incorporación de la AS al proceso de desarrollo dirigido por MA, con el propósito de que esa integración, respete los criterios, valores y principios del dominio de aplicación agilista, favorezca el diseño de Sistemas de

Formación de Recursos Humanos

El equipo de trabajo está compuesto por cinco docentes-investigadores, y seis alumnos adscriptos. En el periodo 2016- 2017 se han asesorado: una tesis de maestría aprobada y finalizada, una tesis de maestría en proceso, una tesis de grado finalizadas y cuatro tesis de grado en proceso, tres de las cuales se están desarrollando en el nuevo tópico de investigación que se presenta en este trabajo.

Además, se espera realizar direcciones, a los alumnos adscriptos, en dos trabajos orientados al desarrollo de SI con metodologías ICONIX-SCRUM, adoptando el enfoque de integración con Arquitectura de Software.

Con los resultados de la presente investigación, se harán actividades de divulgación en publicaciones y presentaciones en eventos nacionales e internacionales y también en cursos de postgrado y actualización

Referencias

- [1] Canos J., Letelier P. “Metodologías Ágiles en el Desarrollo de Software”. JISBD 2003. España
- [2] Navarro, Moreno, Aranda, Parra, Rueda “Integración de Metodologías Ágiles y Arquitecturas de Software en el desarrollo de Sistemas de Información” XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016) Entre Ríos, Argentina) ISBN: 978-950-698-377-2.
- [3] Navarro, Moreno, Aranda, Parra, Rueda. Proyecto de Investigación “Aplicabilidad de Metodologías y Tecnologías Informáticas en el Desarrollo de Sistemas de Información” Código: 21/E 979 –FCEFN- UNSJ – CICTCA- SIGEVA 2014-2015.
- [4] Urquiza Yllescas, J.F., et al, “Las Metodologías Ágiles y las Arquitecturas de Software”. Coloquio Nacional de Investigación en Ing. De Sofá. 2010, León, México.
- [5] Manifiesto for Agile Software Development Kent Beck, Mike Beedle, et al. www.agilemanifesto.org/ Accedido el 22 de Febrero de 2016.
- [6] Breivold,H.P., Sundmark, D., Wallin, P. and Larsson, S., “What Does Research Say about Agile and Architecture?”, en Proceedings of the Fifth International Conference on Software Engineering Advances (ICSEA), USA, (2010).
- [7] ISO/IEC/IEEE42010. IEEE Std 1471:2000, “Recommended Practice for Architectural Description of Software intensive Systems”.
- [8] Martin Fowler. “Patterns of Enterprise Application Architecture” Addison- Wesley. 2003 1st Edition. ISBN-13: 007-6092019909.
- [9] Navarro- Rueda- Moreno – Aranda- Parra- “Convergencia de Tecnologías Informáticas y Metodologías para la Implementación de Sistemas de Información” Cod: 21/E/871 FCEFN.. 2012-2013. – CICTCA- SIGEVA.
- [10] Scrum y XP desde las trincheras - proyectalis.com Una historia de guerra Ágil SCRUM Y XP DESDE LAS TRINCHERAS Cómo hacemos Scrum Henrik Kniberg Prólogos de Jeff Sutherland y Mike Cohn ISBN: 978-1-4303-2264-1
- [11] Edwin Rafael Mago - Germán Harvey Alférez “El Papel de la Arquitectura de software en Scrum “ Publicado en SG #30 <https://sg.com.mx/revista/30/el-papel-la-arquitectura-software-scrum#.WNiAUtLhB1s>.

Tecnología Mobile Aplicada a las Instituciones Educativas

Miguel Alfredo Bustos, Norma Beatriz Perez y Mario M. Berón

Departamento de Informática / Facultad de Ciencias Físico Matemáticas y Naturales / Universidad Nacional de San Luis (UNSL)
Ejercito de los Andes 950, D5700HHW San Luis, +54-0266 4520300 - Int. 2102
{mabustos, nbperez, mberon}@unsl.edu.ar

RESUMEN

El mercado *mobile* presenta un crecimiento exponencial en los más diversos sectores. Éste mercado ofrece un amplio abanico de dispositivos con características de gran trascendencia que se adaptan a las necesidades de cada usuario. Por ello, los gigantes tecnológicos como Google, Samsung, Apple, entre otros, invierten en este mercado gran parte de sus recursos económicos para el desarrollo, cada vez más revolucionario, de hardware y software. Esto permite satisfacer las necesidades, incluyendo las más exigentes, por parte de los consumidores de estas tecnologías digitales. La aceptación de la tecnología *mobile* en la sociedad se ve reflejada en los sectores de industria, negocio, gobierno y educación. La tecnología *mobile* ha impulsado el paradigma *m-learning*. Este paradigma se emplea para un aprendizaje focalizado, colaborativo entre sus pares, etc.

En este trabajo se describe la línea de investigación que aborda el estudio de las tecnologías *mobile* aplicadas a la educación superior. Dicho estudio se fundamenta en el análisis de la innovación, desarrollo, generación, evaluación y seguridad de las aplicaciones. Por ello, se lleva adelante evaluaciones y análisis de aplicaciones empleando diferentes métodos, técnicas y herramientas automatizadas.

Palabras clave: Tecnología digital, Mobile, Métodos de Evaluación.

CONTEXTO

La presente línea de investigación se enmarca en el Proyecto de Investigación: Ingeniería de Software: *Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad* – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos.

1. INTRODUCCIÓN

En la actualidad, las universidades, emplean recursos didácticos de enseñanza haciendo uso de tecnologías digitales como por ejemplo pizarras digitales, smartphones, etc. Esto se debe a la amplia aceptación por parte de los usuarios adeptos a estos recursos tecnológicos; al fácil e inmediato acceso a información precisa; a los bajos costos de estas tecnologías; al soporte de almacenamiento a través de la nube [1]; al acceso a Internet a través de la banda ancha; al Internet de las Cosas (IoT) [2], entre otros factores de importancia. Por ello, las universidades adaptan recursos tecnológicos para la creación e incorporación de redes de conocimiento que promulguen la utilidad de contenidos virtuales así como el uso de sistemas de información.

Los recursos, así como las aplicaciones *mobile* deben ser evaluados para posibilitar una adecuada implementación brindando un porcentaje de satisfacción apropiado según los requerimientos de los consumidores. Es por ello, que el uso de técnicas, métodos y herramientas permiten llevar a cabo una implementación eficiente, fiable y automatizada de las aplicaciones *mobile*. La línea de investigación aborda los ejes fundamentales que se describen de manera sucinta a continuación.

11 TECNOLOGÍAS DE SOFTWARE Y HARDWARE APLICADAS A LA EDUCACIÓN

El crecimiento del mercado *mobile* permite automatizar las actividades cotidianas (entretenimiento, comunicación, operaciones bancarias, etc.) que los usuarios realizan a través de innumerables aplicaciones que se encuentran a disposición de ellos. En este mercado surgen, diariamente, gran cantidad de aplicaciones con el objetivo de satisfacer las necesidades de la sociedad.

Las tecnologías digitales (smartphones, smartTVs, etc.), incorporan características relevantes como: procesador, memoria, cámara de alta definición, Sistemas Operativos (SOs) [3], [4]. Los SOs, ver Tabla 1, que utilizan estas tecnologías permiten interactuar con una heterogeneidad de aplicaciones según los requerimientos de los usuarios.

SO	DESCRIPCIÓN
ANDROID	Desarrollado por Google Basado en Linux Última versión Android 7.0
IOS	Desarrollado por Apple Basado en Linux Última versión, iOS 9
WINDOWS PHONE	Desarrollado por Microsoft Basado en Windows Última versión Windows 10
FIREFOX OS	Desarrollado por Mozilla Corporation Basado en HTML5 con Linux Última versión, Firefox OS 2.5
UBUNTU TOUCH	Desarrollado por Canonical Ltd Basado en Linux Última versión Meizu PRO 5 Ubuntu

Tabla 1: *Mobile* - Sistemas Operativos

Las aplicaciones que soportan estos SOs se clasifican en aplicaciones de mensajería, multimedia, entretenimiento, educación, entre otras categorías. Las aplicaciones educativas permiten a los usuarios ejercitar sus conocimientos interactuando con diferentes temáticas de manera colaborativa resolviendo tareas en común con sus pares, resolver ejercicios *on-line*, tomar notas, agendar fechas importantes, entre otras innumerables actividades.

Expertos en el desarrollo de tecnologías *mobile* diseñan una amplia variedad y cantidad de aplicaciones que son utilizadas por la diversidad de usuarios. Por ello, los programadores llevan adelante el desarrollo e implementación de aplicaciones utilizando entornos de desarrollo integrado (IDE) [5]. Los IDEs son seleccionados por los programadores según criterios de eficiencia, simplicidad, adaptabilidad, etc. Sin lugar a duda, existe gran variedad de IDEs [6] que permiten adaptarse a las diferentes preferencias de programadores. Los IDEs, ver Tabla 2, se diferencian según el tipo de SO en que se desea desarrollar las aplicaciones *mobile*; este factor ha perdido relevancia debido al surgimiento de diferentes plataformas de desarrollo que permiten adaptarse a la diversidad de SOs de manera simultánea.

ANDROID	Android Studio [7] – Eclipse [8] AIDE - DroidEdit
IOS	Xcode [9] - Xamarin
WINDOWS PHONE	Microsoft Visual Studio [10] Windows Phone App Studio

Tabla 2: IDEs – Aplicaciones *mobile*

12 MÉTODOS DE EVALUACIÓN

La calidad de un producto de software es indispensable para el éxito en el mercado. Un software de calidad incorpora características bien definidas de tal manera de garantizar la eficiencia respecto a los requerimientos de los usuarios. Evaluar un producto de software implica probar no solo el producto final sino también los procesos de diseño y desarrollo.

Actualmente, existen numerosas herramientas automatizadas de evaluación multicriterio [11]. Los autores de esta línea de investigación estudian los métodos de evaluación multicriterios que permiten evaluar con un alto nivel de confiabilidad las aplicaciones *mobile* aplicadas a la educación superior. Las herramientas FOP [12] como Logic Scoring of Preference (LSP) [13], permiten seleccionar la mejor alternativa entre diferentes sistemas, detectar vulnerabilidades, bugs, implementación adecuada y eficiente, así como comparar sistemas generando un ranking, etc.

LSP se adapta a diferentes contextos de evaluación lo cual permite ser utilizado en aplicaciones *mobile*. Para aplicar LSP se debe definir: i) Árbol de Criterios (AC); ii) Funciones de Criterios Elementales y; iii) Estructura de Agregación.

FOP es una técnica que posibilita gestionar la variabilidad en una familia de sistemas de software. FOP permite definir tres tipos de variabilidad: i) Métricas; ii) Contexto y; iii) Satisfacción:

El análisis de los métodos mencionados anteriormente permitió detectar que LSP tiene propiedades destacables que lo hacen el candidato ideal para la evaluación de aplicaciones *mobile*. Particularmente LSP permite volcar en el proceso de evaluación la experiencia del experto en el dominio de aplicación.

13 SEGURIDAD EN PLATAFORMAS MOBILE

La seguridad es indispensable en cualquier sistema informático, en particular, en el campo de la tecnología *mobile* [14], [15]. Las plataformas *mobile* incorporan sistemas de seguridad que permite proteger a sus usuarios de aplicaciones maliciosas cuyo fin es quebrantar la privacidad como así también realizar tareas no deseadas.

La plataforma Windows *Mobile* en sus inicios de desarrollo no disponía de un sistema de seguridad para sus aplicaciones.

Es por ello, que los usuarios debían ser precavidos al descargar e instalar una aplicación. Por otro lado, las aplicaciones de iOS requieren ser verificadas por Apple previo a su instalación. Es importante destacar que Apple permite descargar/installar aplicaciones que se encuentren en su tienda oficial. No obstante, Android tiene un sistema de seguridad no centralizado que protege a sus usuario. La seguridad en Android se fundamenta en tres pilares: 1) Android basado en Linux: aprovecha las bondades de la seguridad que dispone el SO Linux. Esto permite impedir que aplicaciones maliciosas afecten al hardware; 2) Firma Digital: toda aplicación basada en Android debe ser firmada con un certificado digital que identifique al autor. Por otro lado, al modificar una aplicación que se encuentra disponible en la tienda, el autor debe firmar nuevamente y 3) Permisos: Son innumerables las aplicaciones que necesitan interactuar con información del dispositivo (como por ejemplo: Contactos, multimedia, etc.) lo que compromete la seguridad del sistema. Es por ello, que se utiliza el modelo de permisos que facilita al usuario conocer los riesgos antes de instalar una aplicación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación que se desarrolla en este artículo se fundamenta en evaluación de sistemas de información aplicados a la educación superior a través de métodos de evaluación multicriterios. Los ejes que se analizan en esta línea de investigación se describen a continuación. i) Tecnología de software y hardware aplicadas a la educación: la investigación se centró en realizar un estudio de las aplicaciones *mobile* en la educación superior; análisis de diferentes SOs en el mercado *mobile* y los IDEs para el desarrollo de aplicaciones.

ii) Métodos de evaluación: se estudiaron modelos de evaluación multicriterio ampliamente utilizados con el fin de evaluar eficiencia, usabilidad, seguridad que deben soportar las aplicaciones *mobile*.

iii) Seguridad en plataformas *mobile*: el equipo de investigación se ha centrado en el análisis de los sistemas de seguridad.

En la línea de investigación descripta en este trabajo, los autores emplean una aplicación [16] como caso de estudio. La aplicación ha sido desarrollada íntegramente por los autores de este artículo.

3. RESULTADOS OBTENIDOS / ESPERADOS

Los resultados de esta línea de investigación están orientados a:

i) El desarrollo de herramientas (aplicaciones) que utilizan la tecnología *mobile*. Estas herramientas permiten facilitar diversos aspectos de la educación superior, como por ejemplo: acceso inmediato a información relevante para alumnos como es horarios de cursada, becas, ubicación de aulas, contenidos de materias, salud estudiantil, etc.

ii) El estudio de métodos de evaluación multicriterios cuyo objetivo es de analizar, evaluar y diseñar aplicaciones *mobile* desarrolladas para el ámbito universitario. Se pretende obtener un ranking de la usabilidad de dichas aplicaciones; nivel de satisfacción de los módulos integrados en las aplicaciones; entre otros factores de importancia que hacen que estas aplicaciones sean ampliamente aceptadas y utilizadas por sus usuarios.

ii) Análisis de la seguridad de las aplicaciones *mobile* cuyo objetivo es determinar la seguridad en las aplicaciones.

A continuación se describen las tareas que se han llevado a cabo en esta línea de investigación.

i) Se desarrolló una aplicación *mobile*, para la educación de nivel superior, en la UNSL. El objetivo de esta aplicación se basa en los requerimientos de los usuarios (alumnos, docentes y no docentes) de la institución donde se accede a información de interés de forma inmediata.

ii) Se realizaron evaluaciones comparando la aplicación mencionada en el ítem anterior con aplicaciones similares. Estas comparaciones fueron realizadas empleando el método de evaluación LSP. En la Figura 1, se muestra el árbol de criterio utilizado para realizar las comparaciones mencionadas previamente.

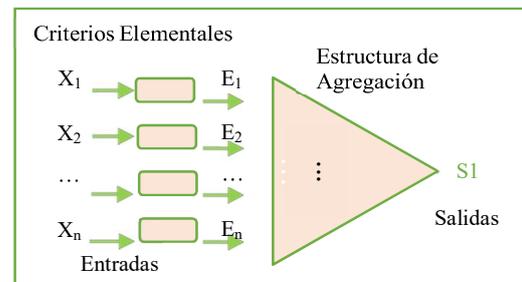


Figura 1: Vista del árbol LSP

Las entradas del árbol de criterio son características generales que se desean evaluar; como por ejemplo usabilidad, seguridad, robustez, entre otras. A continuación, se aplica una función que realiza la descomposición de las características mencionadas en atributos medibles. Finalmente, a través de un proceso de análisis/evaluación se obtiene como salida un número en el rango [1..10] que determina el nivel de satisfacción.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de profesionales de la UNSL que forman parte de la línea de investigación de este trabajo llevan adelante diferentes trabajos finales integradores de *Ingeniería en Informática*, *Ingeniería en Computación*, *Licenciatura en Ciencias de la*

Computación, y en un futuro próximo trabajos finales de especialización, tesis de maestría y doctorado. En particular, las investigaciones desarrolladas en este trabajo forman parte del lineamiento inicial de un trabajo final integrador para optar al grado de *Ingeniero en Informática* en la UNSL.

5. BIBLIOGRAFIA

- [1]. García, A. and Yussel, A., 2016. Sistema de almacenamiento en la nube para dispositivos móviles.
- [2]. Gubbi, J., Buyya, R., Marusic, S. and Palaniswami, M., 2013. Internet of Things (IoT): A vision, architectural elements, and future directions *Future generation computer systems*, 29(7), pp.1645-1660.
- [3]. Allen, S., Graupera, V. and Lun-drigan, L., 2010. Pro smartphone cross-platform development: iPhone, blackberry, windows mobile and android development and distribution. Apress.
- [4]. Bustos, M.A., Perez, N.B. and Berón, M., 2015, May. Plataformas para el desarrollo de aplicaciones móviles. *In XVII Workshop de Investigadores en Ciencias de la Computación* (Salta, 2015).
- [5]. Gronli, T.M., Hansen, J., Ghinea, G. and Younas, M., 2014, May. Mobile application platform heterogeneity: Android vs Windows Phone vs iOS vs Firefox OS. *In Advanced Information Networking and Applications* (AINA), 2014 IEEE 28th International Conference on (pp. 635-641). IEEE.
- [6]. Mantilla, M.C.G., Ariza, L.L.C. and Delgado, B.M., 2014. Metodología para el desarrollo de aplicaciones móviles. *Revista Tecnura*, 18(40), pp.20-35.
- [7]. Smyth, N., 2016. Android Studio 2 Development Essentials. eBookFren-zy.
- [8]. Vogel, L., 2014. Eclipse IDE tutorial.
- [9]. <https://developer.apple.com/xcode/>
- [10]. <https://www.visualstudio.com/>
- [11]. Jean-Pierre Brans and Bertrand Marschal. Promethee methods. In Multiple criteria decision analysis: state of the art surveys, pages 163–186. Springer, 2005.
- [12]. Enriquez, J.G. and Casas, S., 2014. Framework para implementar pruebas de usabilidad flexibles para aplicaciones móviles. *In XX Congreso Argentino de Ciencias de la Computación* (Buenos Aires, 2014).
- [13]. Miranda, E., Berón, M., Montejano, G., Pereira, M.J. and Henriques, P., 2013. NESSy: a new evaluator for software development tools. *In 2nd Symposium on Languages, Applications and Technologies* (SLATe'13) (Vol. 29, pp.21-38).
- [14]. Al-Qershi, F., Al-Qurishi, M., Rahman, S.M.M. and Al-Amri, A., 2014, January. Android vs. iOS: The security battle. *In Computer Applications and Information Systems* (WCCAIS), 2014 World Congress on (pp. 1-8). IEEE.
- [15]. Samawi, V.W., 2014. Security in Mobile Computing. *Handbook of Research on Threat Detection and Countermeasures in Network Security*, p.349.
- [16]. Bustos, M.A, Perez, N.B and Berón, M., 2016, November. Universitarios App: Aplicación Móvil para la UNSL. *In IV Congreso Nacional de Ingeniería en Informática/Sistemas de Información* (Co-NaIISI). (Buenos Aires, 2016).

Tecnología y Aplicaciones de Sistemas de Software: Calidad e Innovación en Procesos, Productos y Servicios

Ramon H. ¹, Ahmad H., Alonso N., Bendati N., Cicerchia B. ², Esnaola L., JatipN., Lencina P., López Gil F., Muscia A., Perez D., Piergallini R., Pompei S., Russo C. ³, Sarobe M., Schiavoni A., Smail A., Terzano J., Tessore J. ⁴

**Instituto de Investigación y Transferencia de Tecnología (ITT) -CIC⁵,
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires**

Sarmiento y Newbery, Junín (B) – TE: (0236) 4477050

{hugo.ramon, tamara.ahmad, nicolas.alonso, natalia.bendati, lucas.cicerchia, leonardo.esnaola, nicolas.jatip, paula.lencina, fernandolopezgil, alex.muscia, daniela.perez, sabrina.pompei, claudia.russo, monica.sarobe, antonela.schiavoni, ana.smail, julieta.terzano, juanpablo.tessore}@itt.unnoba.edu.ar

Resumen

El avance de las tecnologías es acompañado por el incremento de los sistemas de software donde la calidad y la innovación toman un lugar esencial tanto en el desarrollo de productos como de servicios. Actualmente al incluir al usuario en los procesos de innovación y ubicarlo en una posición centralizada se da lugar a la aparición de nuevos espacios de interacción y comunicación en los que el actor que consume productos y servicios adquiere un rol activo.

Con la premisa de incorporar al usuario en todas las fases del proceso de innovación surgen los Living Labs. Éstos

son entornos donde se expone a los usuarios a nuevas aplicaciones, mediante auténticas plataformas de innovación abierta de cocreación y codiseño con los usuarios, para experimentar sobre las implicaciones de una determinada tecnología, como computación ubicua, conectividad móvil, apropiación de las TIC en la educación, el comercio, el gobierno, entre otros, a fin de investigar y analizar, uso, experiencia de usuarios y oportunidades.

Nuestro propósito es estudiar cómo se modifican los procesos de desarrollo con la incorporación del usuario final en los procesos de innovación tecnológica desde los siguientes enfoques de investigación:

¹ Investigador Asociado CIC (Adjunto sin Director)

² Becario CIC

³ Investigador Asociado CIC (Adjunto sin Director)

⁴ Becario CIC

⁵ Instituto Asociado CIC

gestión de la innovación, calidad y cómputo ubicuo en gobierno/empresa/ciudades digitales.

Palabras clave: gestión de la innovación, living lab, cómputo ubicuo, calidad

Contexto

Las líneas descriptas en el presente documento están insertas en el proyecto de investigación "Tecnología y Aplicaciones de Sistemas de Software: Calidad e Innovación en procesos, productos y servicios", acreditado y financiado por la Universidad Nacional del Noroeste de Buenos Aires (UNNOBA) en la convocatoria "Subsidios de investigación bianuales" 2017. Las actividades de investigación de dicho proyecto tienen lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) de la UNNOBA.

1- Introducción

La calidad del software, de por sí compleja y subjetiva, es una cuestión clave a la hora de medir el éxito de un proyecto. Para abordarla hay dos aspectos importantes a tener en cuenta: identificar correctamente el software correcto a construir; y demostrar que ese software se ha logrado. [1]

DeLone identifica seis criterios para medir el éxito de un software: calidad del sistema, calidad de la información, uso de la información, satisfacción del usuario, impacto individual e impacto organizacional. [2] Por su parte, Deane basa la evaluación en la alineación de los resultados del proyecto con los requerimientos del cliente y las diferencias que pueden ocurrir entre ellos.

De allí es que se identifican cinco posibles diferencias (gap):

- Gap 1: Declaración ineficaz de los requerimientos de largo plazo de los clientes.
- Gap 2: Percepción incorrecta de los requerimientos del cliente.
- Gap 3: Traducción ineficaz de los requerimientos del cliente.
- Gap 4: Ejecución ineficaz del plan del proyecto.
- Gap 5: Comunicación ineficaz de los resultados. [3]

Atkinson, en cambio, define el modelo Square Route (SRM - Square Route Model) en donde muestra una perspectiva de éxito basado en múltiples perspectivas como una aproximación holística alternativa al triángulo de hierro (costo, tiempo y calidad). [4]

Uno de los cambios fundamentales en la generación de productos y servicios está dado por los ecosistemas digitales emergentes (ciudades digitales, agricultura de precisión, etc.) y los negocios que los alimentan, borrando los límites de la industria. Estos ecosistemas se basan entre otras cosas en el cómputo ubicuo, atribuido a Mark Weiser quien a fines de los '80 imaginó cómo evolucionaría la computación personal creando un escenario en el que cientos de dispositivos (de todos los tamaños) con capacidades de procesado y comunicación estarían integrados natural e invisiblemente en el entorno cotidiano de las personas "activando" la realidad. [5]

Una característica fundamental del software que soporta dispositivos ubicuos es la sensibilidad al contexto, donde se

requiere la capacidad de obtener información relevante para adaptar el comportamiento de la aplicación dinámicamente en función de su contexto. Esta relación entre la tecnología y el medio se da gracias a que los dispositivos móviles y los sensores pueden comunicarse entre sí. De hecho, esto es posible dado al surgimiento de tecnologías tales como el Bluetooth, la identificación por radiofrecuencia (RFID), los protocolos Zigbee y las redes WiFi, las cuales brindan la conectividad necesaria para la comunicación.

La construcción del software sensible al contexto es compleja, dado que además de la adquisición de la información relevante requiere el adecuado procesamiento de dicha información. Con el fin de facilitar dicho proceso se han desarrollado distintos frameworks para el prototipado rápido de aplicaciones dependientes del contexto y existen esfuerzos para definir estándares para el desarrollo de aplicaciones ubicuas. [6] [7] [8]

En el entorno de la empresa actual no existe tarea más vital para la dirección que la garantía de la calidad de sus productos y servicios y su constante innovación. Considerando luego que el software se ha convertido en un útil indispensable y fundamental de la economía de cualquier negocio, la gestión en cualquier proyecto tecnológico que involucre software en donde la calidad es más importante que la agenda de trabajo, es crítica.

Actualmente existen modelos de calidad específicos para la gestión orientada a procesos como también otros modelos de mejora de procesos para PyMEs. [9] [10] [11] Existen además otros modelos de normas y modelos

orientados a la calidad del producto y modelos de mejoras aplicables a equipos que integran hardware y software para la obtención de una certificación del producto completo (ISO/IEC 15408 Common criteria). Estos modelos aplican en los requerimientos de calidad definidos en la Ley 25922 (“Ley de Promoción de la Industria del Software”, 2005).

En cuanto a la innovación, esta supone la existencia de tres factores: una base de conocimientos tecnológicos y científicos; una necesidad económica y social, y un agente intermediario que transforme los conocimientos en bienes y servicios destinados a satisfacer la necesidad. El primer factor es el que proporciona la oportunidad de la innovación, el segundo el éxito y el tercero, el agente que lo realiza.

Respecto a la fuerza motriz para el cambio, existen tres grandes corrientes de pensamiento: aquellos que consideran que es la tecnología, que actualmente deriva directamente del esfuerzo de investigación y desarrollo, quien promueve la innovación (Technology push); otros que consideran que es el mercado con sus nuevas demandas el que motiva el cambio (Market pull); y, actualmente, es más aceptado que ambas fuerzas, el mercado y la tecnología, proporcionan la energía para el cambio que representa la innovación.

Medir la innovación parecería seguro desde los manuales basados en el consenso y los formularios de encuesta adoptados internacionalmente sobre todo considerando que las políticas de estímulo a la innovación están incorporadas en las agendas públicas desde hace bastante tiempo, sin embargo, se hace necesario responder algunas preguntas: ¿Cómo medir la eficacia de tales políticas y

establecer indicadores capaces de dar cuenta de procesos tan multifacéticos? ¿Las recomendaciones Internacionales o Latinoamericanas son aplicables a la región NOBA (Noroeste de la Provincia de Buenos Aires)? ¿Está esto valuado en las empresas de la región? [12]

Hoy en día la innovación intenta mejorar la posición competitiva de las empresas mediante la incorporación de nuevas tecnologías y conocimientos de distinto tipo. El proceso de innovación consiste así en una serie de actividades no solamente científicas y tecnológicas, sino también organizacionales, financieras y comerciales; acciones que, en potencia, transforman las fases productiva y comercial de las empresas.

Todo lo anterior conduce a que se esté ante un nuevo modelo emergente denominado *user-driven innovation* (innovación dirigida por el usuario). Dicho modelo, facilitado por la web (en sus versiones 2.0/3.0/4.0), la generalización de la “innovación abierta”, la universalización de la tecnología móvil y su conexión a Internet, y diversas formas de innovación social expresadas en forma de espacios de *co-working* (trabajo cooperativo), *crowdfunding* (financiamiento colectivo), y *P2P economy* (economía de pares), donde la característica clave es la participación de los propios “usuarios” o “ciudadanos” en el proceso de innovación; dando lugar al surgimiento de los *Living Labs* (laboratorios vivientes).

Estos laboratorios vivientes son entornos donde se expone a los usuarios a nuevas aplicaciones, mediante auténticas plataformas de innovación abierta de cocreación y codiseño, para experimentar sobre las implicaciones de una determinada tecnología, como

computación ubicua, conectividad móvil, apropiación de las TIC en la educación, el comercio, el gobierno, entre otros, a fin de investigar y analizar, uso, experiencia de usuarios y oportunidades.

Pero ante este nuevo modelo deberíamos respondernos ¿Cómo se modifican los procesos de desarrollo con la incorporación del usuario final (partícipe o representante de la sociedad) en los procesos de innovación tecnológica? El propósito de nuestro trabajo es estudiar cómo se modifican estos procesos de desarrollo desde tres líneas de investigación: gestión de la innovación, cómputo ubicuo y calidad.

2- Líneas de Investigación y Desarrollo

A fin de responder ¿Cómo se modifican los procesos de desarrollo con la incorporación del usuario final (partícipe o representante de la sociedad) en los procesos de innovación tecnológica? Se trabajará sobre tres ejes: gestión de la innovación, cómputo ubicuo y calidad.

Por un lado, se estudiarán los procesos metodológicos para abordar la innovación, la aplicación de *LivingLabs* y recomendaciones culturales. Se busca detectar casos de aplicación e indicadores de innovación. Se analizará la implicancia de la gestión de la innovación en los procesos de gestión de proyectos y actividades de Ingeniería de Software.

Por otra parte, se trabajará sobre el cómputo ubicuo en Gobierno/Empresa/Ciudades Digitales a fin de Investigar/Seleccionar/Adaptar/Desarrollar frameworks de desarrollo de aplicaciones. También se investigará sobre protocolos de redes de sensores

para dispositivos heterogéneos y su seguridad.

En cuanto al eje de calidad se busca Analizar/Adaptar/Implementar modelos de calidad aplicables a PyMES, evaluar modelos para la mejora de procesos de gestión y normas de certificación para productos tecnológicos y desarrollar modelos de calidad de procesos y productos que utilicen indicadores de innovación en PyMES de la región NOBA (Noroeste de la Provincia de Buenos Aires). También se pretende evaluar modelos en ambientes simulados para el entrenamiento de competencias, evaluar la viabilidad de utilización de documentos digitales y su legalidad en organizaciones de la región.

3- Resultados Esperados

Se pretende difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes revistas especializadas, congresos, jornadas y workshops, todos nacional es/internacionales relacionados con la mejora de los procesos, productos y servicios mediante la calidad y la innovación.

4- Formación de Recursos Humanos

El equipo de trabajo está compuesto por 11 investigadores formados (8 del área de Sistemas y 3 del área de Ciencias Económicas), 7 investigadores en formación, 3 becarios de posgrado y 8 becarios de grado.

Se espera que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su rol de investigadores o partícipes activos en equipos de investigación,

transferencia e innovación; fomentando la culminación de sus estudios superiores, promoviendo la redacción, exposición y defensa de Trabajos Finales de Grado y Postgrado, y la realización de Prácticas Profesionales Supervisadas.

Para los próximos dos años se espera contribuir al inicio y concreción de 2 (dos) Tesinas de Licenciatura en Sistemas, 3 (tres) Prácticas Profesionales Supervisadas de la Ingeniería en Informática, 1 (una) Tesis de Especialización, 4 (cuatro) Tesis de Magíster y 3 (tres) Tesis Doctorales.

5- Bibliografía

- [1] Holcombe, M. (2008). *Running an agile software development project*. John Wiley & Sons
- [2] DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, 3(1), 60-95.
- [3] Deane, R. H., Clark, T. B., & Young, A. P. (1997). Creating a learning project environment: aligning project outcomes with customer needs. *Information systems management*, 14(3), 54-60.
- [4] Atkinson, R. (1999). Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International journal of project management*, 17(6), 337-342.
- [5] Weiser, M. (1993). Hot topics- ubiquitous computing. *Computer*, 26(10), 71-72.
- [6] Fujinami, K., Yamabe, T., & Nakajima, T. (2004, November).

- Bazaar: a conceptual framework for physical space applications. In *International Symposium on Ubiquitous Computing Systems* (pp. 174-191). Springer Berlin Heidelberg.
- [7] Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-computer interaction*, 16(2), 97-166.
- [8] Modahl, M., Agarwalla, B., Abowd, G., Ramachandran, U., & Saponas, T. S. (2004, October). Toward a standard ubiquitous computing framework. In *Proceedings of the 2nd workshop on Middleware for pervasive and ad-hoc computing* (pp. 135-139). ACM.
- [9] Team, S. U. (2011). Standard CMMI Appraisal Method for Process Improvement (SCAMPI) A, Version 1.3: Method Definition Document.
- [10] Aguirre, A. F., Pardo, C., Pantoja, W. L., Mejía, M. F., & Pino, F. J. (2010). Reporte de experiencias de la aplicación de Competisoft en cinco mipymes colombianas. *Revista EIA*, (13), 107-122.
- [11] Calvo-Manzano, J. A., Garzás, J., Piattini, M., Pino, F. J., Salillas, J., & Sánchez, J. L. (2008). Perfiles del ciclo de vida del software para pequeñas empresas: los informes técnicos ISO/IEC 29110. *Innovación, Calidad e Ingeniería del Software*, 4(2), 96.
- [12] Albornoz, M. (2009). Indicadores de innovación: las dificultades de un concepto en evolución. *Revista iberoamericana de ciencia tecnología y sociedad*, 5(13), 9-25.

Tratamiento de Evidencias Digitales Forenses en Dispositivos Móviles

Marta C. Fennema, Liliana M. Figueroa, Graciela Viaña, Norma B. Lesca, Cecilia C. Lara

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
fennema@unse.edu.ar, lmyfigueroa@yahoo.com.ar; gv857@hotmail.com;
{norma.lesca, laraceciliacristina}@gmail.com

RESUMEN

En este artículo se presenta una propuesta de investigación atendiendo los requerimientos específicos planteados desde el ámbito judicial ante la implementación del Nuevo Sistema Procesal Penal en la Provincia de Santiago del Estero, respecto del proceso de obtención de evidencias digitales.

El proceso de adquisición de evidencias digitales debe ser legalmente aceptable, apoyándose en métodos científicos que permitan recolectar, analizar y validar las mismas, recurriendo entonces a la Informática Forense.

Es en este contexto que se propone el estudio y definición de un protocolo para la gestión de evidencias digitales forenses obtenidas de dispositivos móviles.

Otra cuestión a investigar es el diseño de un repositorio de evidencias digitales extraídas de dispositivos móviles, que permita almacenar, recuperar, distribuir y compartir las evidencias forenses (de manera abierta y segura), definiendo las herramientas de repositorios que son apropiadas para implementar el modelo de datos propuesto.

Palabras clave:

Informática Forense en dispositivos móviles, evidencias digitales, protocolo de extracción de datos, repositorio de evidencias.

CONTEXTO

La presente línea de investigación se encuentra inserta en el proyecto “Computación Móvil: desarrollo de

aplicaciones y análisis forense”, que propone una continuación del trabajo en el ámbito de la computación móvil iniciada en el año 2012, financiado por el Consejo de Ciencia y Técnica de la Universidad Nacional de Santiago del Estero [5].

La justicia moderna necesita ampliar la mirada a la hora de obtener evidencias y pruebas digitales, que sean legalmente aceptables y que ayuden a resolver conflictos apoyándose en métodos científicos que permitan recolectar, analizar y validar pruebas digitales.

Es en ese ámbito que se advierte la necesidad de trabajar en el estudio y definición de un protocolo para el análisis de evidencias forenses obtenidas de dispositivos móviles. En respuesta a ello, se propone investigar: protocolos de intervenciones forenses, equipamientos dedicados para la extracción de información y la gestión (acceso, almacenamiento, recuperación, seguridad) de la información obtenida.

1. INTRODUCCIÓN

La Informática Forense es una nueva disciplina que se desprende directamente de una serie de sucesos que han afectado a la sociedad globalizada e informatizada de fines del siglo XX y principios del XXI, en donde se observa el auge de una serie de delitos que están afectando diferentes áreas de la sociedad [3].

Esta disciplina de las ciencias forenses aplicadas en medios informáticos, considera las tareas propias asociadas con la evidencia, procura descubrir e interpretar la

información en los medios informáticos para establecer los hechos y formular las hipótesis relacionadas con el caso. Así también, es la disciplina científica y especializada que entendiendo los elementos propios de las tecnologías de los equipos de computación ofrece un análisis de la información residente en dichos equipos [1].

Está dedicada a la recolección de pruebas digitales para fines judiciales, mediante la aplicación de técnicas de análisis y de investigación. El propósito de esta disciplina es determinar los responsables de los delitos, así como también esclarecer la causa original de un ilícito o evento particular para asegurar que no se vuelva a repetir.

Según [1], la criminalística ofrece un espacio de análisis y estudio sobre los hechos y las evidencias que se identifican en el lugar donde se llevaron a cabo las acciones criminales, por lo tanto, es necesario establecer un conjunto de herramientas, estrategias y acciones que ayuden a identificar estos hechos y evidencias dentro del contexto informático. Por consiguiente, se hace indispensable la aplicación de procedimientos estrictos y cuidadosos, desde el momento en que se realiza la recolección de la evidencia, hasta que se obtienen los resultados posteriores a la investigación [2,8]. Si bien existe un modelo generalizado para realizar este tipo de investigaciones, cada región necesita adaptarse al Código Procesal Penal y al contexto tecnológico propio.

Para que una evidencia digital pueda ser usada en procesos judiciales, debe cumplir con las siguientes características:

- Admisibilidad: toda evidencia recolectada debe ajustarse a ciertas normas jurídicas para presentarlas ante un tribunal.
- Autenticidad: la evidencia debe ser relevante al caso, y el investigador forense debe estar en capacidad de representar el origen de la misma.
- Completitud: la evidencia debe contar todo en la escena del crimen y no una perspectiva en particular.

- Fiabilidad: las técnicas usadas para obtener la evidencia deben gozar de credibilidad y ser aceptadas en el campo en cuestión, evitando dudas sobre la autenticidad y veracidad de las evidencias.

- Entendimiento y Credibilidad: se debe explicar con claridad y pleno consentimiento, qué proceso se siguió en la investigación y cómo la integridad de la evidencia fue preservada, para que ésta sea comprensible y creíble en el tribunal.

Como en cualquier investigación forense, existen una variedad de enfoques que se pueden utilizar para la recolección y análisis de información. Un aspecto clave para ello es que el procedimiento que se siga, no modifique la fuente de información de ninguna manera, o que de ser esto absolutamente necesario, el analista esté en la capacidad de justificar por qué realizó esta acción [6].

En la actualidad, el empleo de dispositivos móviles se ha incrementado notablemente, principalmente por su facilidad de uso y la propiedad de mantener en contacto permanente a sus usuarios. A partir de esto, se ha generado un cambio significativo en la forma en que las personas se comunican, pero también por su proliferación, se ha incrementado su uso en actividades de orden delictivo.

A diferencia de la Informática Forense clásica, el análisis forense sobre dispositivos móviles, es un campo relativamente nuevo y los procedimientos y normas para su análisis aún se encuentran en desarrollo [11]. Un análisis forense que se lleve a cabo sobre un dispositivo móvil, puede ser admitido o no en un juicio dependiendo de lo que considere el juez y la formalidad con que se desarrolle el procedimiento de recolección, control, análisis y presentación de las evidencias.

En el ámbito de la Informática Forense, en nuestro país, se han desarrollado varios proyectos, entre los cuales se destacan:

- Facultad de Ingeniería de la Universidad FASTA:

a *Guía Integral de Empleo de la Informática Forense en el Proceso Penal: PAIF-PURI* [4]. Es una Guía de Actuación en Informática Forense para ser adoptada y promovida por el Ministerio Público de la Provincia de Buenos Aires como estándar oficial de trabajo, tanto para peritos como para investigadores judiciales, en base a lo establecido por el Proceso Unificado de Recuperación de Información. El Proceso PURI ha sido desarrollado por el Grupo de Investigación en Informática Forense y Sistemas Operativos, formalizando un proceso general que guía a peritos informáticos en la obtención de información digital que pueda ser considerada como evidencia válida por los operadores de justicia.

b *Proyecto INVESTIGA* [7]. Ambiente integrado de visualización y análisis de datos, tiene como objetivo el desarrollo de un sistema informático que permita la consolidación de datos provenientes de múltiples fuentes en un ambiente integrado que facilite su visualización y análisis. El sistema informático objetivo de este proyecto pretende reemplazar al software que actualmente utiliza el Ministerio Público con fines similares, ganando en flexibilidad e independencia tecnológica.

c *Proyecto FOMO* (Forensia en Equipos Móviles) [7]. Tiene como objetivo el desarrollo de un sistema informático que permita realizar la extracción forense de la información contenida en equipos de telefonía móvil. Se pretende, de esta forma, reemplazar al software privativo y extranjero que actualmente utiliza el Ministerio Público con fines similares; desarrollándose un sistema informático propio específicamente orientado a Smartphones Android y Nextel, ganando en flexibilidad e independencia tecnológica.

d *Proyecto GT-LIF* [7]: Guía Técnica para la Implementación de un laboratorio de Informática Forense Judicial, tiene como objetivo el desarrollo de una Guía

Técnica para la implantación de un laboratorio de Informática Forense para ser utilizada por el Ministerio Público de la Provincia de Buenos Aires y en el resto de las provincias, a través del Consejo Federal de Procuradores. Esta guía técnica complementa la “Guía integral de empleo de la Informática Forense en el proceso penal”. Se espera que una vez finalizado el proyecto se cuente con una guía que permita estimar y evaluar los aspectos claves de diseño de un laboratorio forense a nivel estratégico, institucional, edilicio, estructural y tecnológico.

- Facultad de Ingeniería en Sistemas, Universidad Abierta Interamericana, Sede Rosario. Se encuentra desarrollando el Proyecto “Análisis Digital Forense, Conceptos y Aplicaciones” [10], que tiene como principal objetivo generar nuevas herramientas y protocolos de análisis digital forense que sean pertinentes de ser utilizados en nuestro país.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Considerando la amplitud de los aspectos relacionados con la computación móvil, la línea de investigación se refiere a:

Informática Forense: protocolo y gestión de evidencias digitales obtenidas de dispositivos móviles.

A partir de ella, se proponen dos líneas de investigación derivadas, consideradas desde el ámbito de la justicia penal de Santiago del Estero:

- Protocolo de actuación para la extracción de evidencias digitales de dispositivos móviles.
- Modelo de datos para la gestión de las evidencias.

3. OBJETIVOS

El objetivo general de la investigación propuesta relacionada a esta línea de investigación es:

- *Contribuir al progreso del campo de la Computación Móvil mediante el análisis forense de dispositivos móviles.*

Los objetivos específicos que permitirán alcanzar el objetivo general son:

- *Definir un protocolo de actuación en la extracción de evidencias digitales de dispositivos móviles en el marco del nuevo Código Procesal Penal de la provincia de Santiago del Estero.*

- *Diseñar un repositorio de evidencias digitales extraídas de dispositivos móviles en el marco del proceso penal mencionado.*

Se plantea una investigación descriptiva-cualitativa, dado que si bien se puede definir una hipótesis que relacione variables, la misma no alcanzará a ser corroborada en el plazo de dos años que dura esta investigación.

La hipótesis planteada es la siguiente:

El uso de un protocolo preestablecido de Informática Forense para móviles y de un repositorio especializado, optimiza la gestión de evidencias digitales extraídas de los dispositivos móviles.

Como puede observarse en la misma, la variable a estudiar es la “optimización de la gestión de evidencias criminales obtenidas de dispositivos móviles”, la cual en futuras investigaciones podrá ser evaluada a través de indicadores cuantitativos que se pueden aplicar a casos de prueba especialmente diseñados.

Se espera generar nuevo conocimiento científico-tecnológico, plasmado en un protocolo para la recolección y tratamiento de evidencias digitales criminales extraídas de dispositivos móviles, acompañado de un modelo para la gestión óptima de dichas evidencias en el ámbito del Poder Judicial y del Ministerio Público Fiscal de la Provincia de Santiago del Estero, y de acuerdo a lo establecido en el nuevo Código Procesal Penal de la provincia.

Se considera que la obtención del mencionado protocolo traerá un beneficio

muy importante para la justicia santiagueña, dado que actualmente no existe un procedimiento claro y definido. Permitiría mejorar la calidad de las evidencias digitales y ayudará en la labor de los fiscales de la provincia.

4. FORMACIÓN DE RECURSOS HUMANOS

La Directora y Codirectora del proyecto pertenecen al Departamento de Informática de la Universidad Nacional de Santiago del Estero. Los asesores pertenecen a LIDI-FI-UNLP y FCE-UNSala. El resto de los integrantes son docentes investigadores de la Universidad Nacional de Santiago del Estero, con distintas categorías de investigación.

5. REFERENCIAS

1. CANO, J. (2006). Introducción a la informática forense: Una disciplina técnico-legal. Revista Sistemas, Asociación Colombiana de Ingenieros de Sistemas (ACIS). Vol.96, pp. 64-73. http://52.0.140.184/typo43/fileadmin/Revista_96/dos.pdf
2. CASTILLO, C., ROMERO, A., CANO, J. (2008). Análisis Forense Orientado a Incidentes en Teléfonos Celulares GSM: Una Guía Metodológica. Conf. XXXIV Conferencia Latinoamericana de Informática, Centro Latinoamericano de Estudios en Informática (CLEI). <http://www.clei2008.org.ar>.
3. DARAHUGE, M. (2011). Manual de Informática Forense. Buenos Aires. Errepar.
4. DI IORIO, ANA HAYDEE. [et al.] (2015). Guía Integral de Empleo de la Informática Forense en el Proceso Penal. Universidad FASTA. Mar del Plata. Argentina.
5. Herrera, Susana I., Najjar Ruiz P., Rocabado S., Fennema, C., Cianferoni, M. (2013). Optimización de

- la calidad de los sistemas móviles.
http://sedici.unlp.edu.ar/bitstream/handle/10915/27200/Optimizaci%C3%B3n_de_la_calidad_de_los_sistemas_m%C3%B3viles.pdf?sequence=1
6. HOOG, A. (2009). iPhone Forensics: Annual Report on iPhone Forensic Industry. Chicago Electronic Discovery.
 7. INFO-LAB. (2016). Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense. Universidad FASTA, Ministerio Público Fiscal de la Provincia de Buenos Aires y la Municipalidad de General Pueyrredon. Buenos Aires. Argentina.
 8. LEIGLAND, R. (2004). A Formalization of Digital Forensics. International Journal of Digital Evidence. University of Idaho. Volume 3, Issue 2.
<http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B8472C-D1D2-8F98-8F7597844CF74DF8.pdf>.
 9. MEYERS, M., ROGERS, M. (2004). Computer Forensics. The Need for Standardization and Certification. International Journal of Digital Evidence, CERIAS, Purdue University, Volume 3 Issue 2.
<http://ww.utica.edu/academic/institutes/ecii/publications/articles/A0B7F51C-D8F9-A0D0-7F387126198F12F6.pdf>.
 10. PADULA, EUGENIO J.; FOGLIATO, NELSON M.; CASCO, MARIA E. (2016). Análisis Digital Forense, Conceptos y Aplicaciones. Facultad de Ciencias Exactas, Ingeniería. Universidad Nacional de Rosario. Rosario, Santa Fe. Argentina.
 11. VARSALONE, J., KUBASIAK, R. (2009). Mac Os X, iPod and iPhone Forensic Analysis DVD Toolkit. Syngress Publishing, Inc, pp. 355-475.

Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web

Adriana MARTIN, Gabriela GAETAN, Viviana SALDAÑO, Analía PIRES, Gabriela MIRANDA, Silvia VILLAGRA, Alejandra CARRIZO, Claudia CARDOZO, Hernán SOSA

Grupo de Investigación y Formación en Ingeniería de Software (GIFIS)
Instituto de Tecnología Aplicada (ITA)
Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia (UNPA-UACO)

{amartin// ggaetan// vivianas // apires// gmiranda// svillagra// acarrizo}@uaco.unpa.edu.ar;
claudia_yoryi@yahoo.com.ar; hassio_09@hotmail.com

RESUMEN

Mejorar la experiencia del usuario (UX: “User eXperience”), ha cobrado protagonismo en el ámbito de las organizaciones que despliegan sus actividades en la Web, y desencadena una problemática nada trivial, relacionada con la importancia de identificar más completamente a los usuarios para poder reconocer sus expectativas y necesidades. Si bien, no son menores los esfuerzos y aportes realizados por las comunidades de investigación, lo cierto es que aún existe cierto vacío de propuestas integradoras a la que los equipos de desarrollo puedan recurrir a la hora de diseñar productos Web priorizando la UX.

Este Proyecto de Investigación (PI) está dirigido a desarrollar productos Web que atiendan las expectativas y necesidades de grupos de usuarios de interés, proponiendo, aplicando y validando un enfoque integrador de técnicas y herramientas basadas en la UX, a los efectos de:

(i) satisfacer a usuarios adultos mayores; (ii) incluir propiedades de Accesibilidad al ciclo de vida UX; y (iii) considerar la UX en el desarrollo Web móvil.

Palabras clave: *Enfoque Integrador | Experiencia de Usuario (UX) | Diseño y Evaluación | Interfaces de Usuario (UI) | Accesibilidad Web | Grupos de Usuarios de Interés.*

CONTEXTO

El Grupo de Investigación y Formación en Ingeniería de Software (GIFIS) perteneciente al

Instituto de Tecnología Aplicada (ITA), Universidad Nacional de la Patagonia Austral (UNPA), Unidad Académica Caleta Olivia (UACO), se encuentra ejecutando el Proyecto de Investigación (PI) N° 29/B194, Período: 2016-2018, Denominado: “Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web”, dirigido por la Dra. Adriana Martín y codirigido por la Mg. Gabriela Gaetán. Para desarrollar la problemática del presente PI, GIFIS cuenta con antecedentes y experiencia recabada desde 2010 a través de la ejecución de los siguientes Proyectos de Investigación: PI N° 29/B167, Período: 2014-2016, Denominado: “Identificación, Desarrollo y Uso de Soluciones Web Centradas en el Usuario” y, PI N° 29/B144, Período: 2012-2014, Denominado: “Diseño y Evaluación de Portales Web”.

En este contexto, se han enfocado los esfuerzos del presente PI N° 29/B194, Período: 2016-2018 en desarrollar productos Web proponiendo un enfoque integrador que incorporando herramientas conceptuales, tales como Estrategias de Contenido y propiedades de Usabilidad y Accesibilidad Web en un ciclo de diseño y evaluación de interfaces Web, permita mejorar la experiencia del usuario (UX) perteneciente a un perfil de usuario de interés.

1. INTRODUCCIÓN

Actualmente la Web se ha convertido en el instrumento preferido y adoptado por la mayoría de las organizaciones públicas y privadas, sin fines de lucro y de negocio. Esto es debido a que los productos con plataforma de despliegue

en la Web constituyen una herramienta muy poderosa a la hora de acercar propuestas de información, comunicación, productos y servicios a los ciudadanos.

Sin embargo, en general y como una conducta repetitiva y casi contradictoria, el desarrollo para la Web ha descuidado sistemáticamente al usuario destinatario de la evolución tecnológica, ya que no contempla en su justa dimensión las necesidades de los ciudadanos (usuarios destinatarios de los productos Web).

Numerosas áreas del conocimiento aportan propuestas y herramientas para mejorar el tratamiento de las necesidades de los usuarios, tales como: Diseño Centrado en el Usuario (“User-Centered Design”), Interacción Hombre-Computadora (“Human-Computer Interaction”), Estrategias de Contenido Web (“Content Strategy”), Usabilidad y Accesibilidad Web, entre otras.

Particularmente, en los últimos años y relacionado con esta necesidad de identificar más completamente a los usuarios Web para poder reconocer sus expectativas respecto a un producto desarrollado para la Web, el término “experiencia de usuario” (UX: User eXperience), ha cobrado protagonismo en los ámbitos de investigación y de los negocios. El término UX evoca un espectro mucho más amplio para atender a la diversidad humana, que el propuesto por las áreas del conocimiento antes mencionadas desde sus nichos de trabajo y aportes focalizados.

Diseñar para la experiencia del usuario, implica interacción social y cultural, diseño con valor sensitivo e impacto emocional, es decir, cómo la experiencia de acceso e interacción con el producto Web incluye alegría, comunicación emocional, impresión senso-perceptivas en el uso, etc., lo que en el diseño de interfaz de usuario, se denomina el “look and feel”.

Desde el Grupo de Investigación y Formación en Ingeniería de Software (GIFIS), se viene trabajando fuertemente en esta problemática desde 2012 y particularmente en los dos últimos años este trabajo ha permitido identificar valiosos aportes y recursos en el área, tales como [1] [2] [3] [4] [5] [6] [7] [8], entre muchos otros.

La experiencia recabada durante varios años de investigación, permiten aseverar, que cuando se trata de mejorar la UX en la Web, no sólo la Usabilidad es un componente clave, sino que también lo son otros factores de calidad, tales como la Accesibilidad Web, que es un recurso indispensable para aspirar a alcanzar una justa consideración de la diversidad de las personas destinatarias (usuarios Web), sus necesidades y características distintivas y diferentes. Con esta motivación, en el presente PI se definen los siguientes Objetivos:

Objetivo Direccional: “Desarrollar productos Web aplicando un enfoque integrador de técnicas y herramientas basadas en la experiencia del usuario (UX), que satisfaga expectativas y necesidades de perfiles de usuarios de interés.”

Objetivos Operacionales:

OO.1 Explorar requerimientos de información, comunicación y servicios de los usuarios pertenecientes al perfil de interés.

OO.2 Explorar necesidades de Usabilidad, Accesibilidad e interacción de los usuarios pertenecientes al perfil de interés.

OO.3 Seleccionar, extender y/o proponer técnicas y herramientas para satisfacer los requerimientos y necesidades de 1. y 2.

OO.4 Proponer enfoque integrador de las técnicas y herramientas de 3. que mejore la experiencia de los usuarios (UX) pertenecientes al perfil de interés.

OO.5 Aplicar el enfoque integrador propuesto en 4. en el desarrollo de productos Web que satisfagan expectativas y necesidades de perfiles de usuarios de interés.

OO.6 Validar los productos Web de 5. con los usuarios pertenecientes al grupo de interés.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El Proyecto de Investigación (PI) N° 29/B194, Período: 2016-2018, denominado “*Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web*”, cuenta con un Plan de Actividades en el que constan las siguientes Líneas de Investigación, las cuales comparten como hilo conductor los Objetivos Operacionales:

LI.1: Usuarios Adultos Mayores

Desde 2013, GIFIS ha estado vinculado al dictado de los Cursos de Computación en el marco del Convenio UPAMI (UNPA-PAMI). Y a partir de 2016, está a cargo de la Elaboración de la Propuesta y Dictado de Cursos de Computación a nuestros Abuelos de la región. Este espacio de intercambio e interacción ofrece el marco adecuado para ejecutar la LI.1.

Actividades de la LI.1:

A1.1 Definir un marco de experimentación para el análisis de las necesidades de usuarios adultos mayores.

A1.2 Evaluar la aplicabilidad de técnicas y herramientas que permitan satisfacer necesidades de usuario adultos mayores.

A1.3 Desarrollar y validar productos Web dirigidos a satisfacer necesidades de usuarios adultos mayores.

LI.2 Experiencia de Usuario (UX)

Entre 2013-2014, GIFIS ha estado trabajando en la identificación de problemas en los sitios Web Universitarios para los usuarios ciegos y en la aplicación de patrones de accesibilidad para asistir a la solución de estos problemas. A partir de 2015, se han recibido pedidos de desarrollo de productos Web que satisfagan las necesidades de los usuarios pertenecientes a comunidades específicas, tales como las del ámbito educativo primario y secundario. La necesidad de satisfacer las necesidades de los usuarios pertenecientes a dominios específicos de interés, ofrece el marco adecuado para ejecutar la LI.2.

Actividades de la LI.2:

A2.1 Definir un enfoque UX que considere propiedades de Accesibilidad.

A2.2 Desarrollar un producto Web que aplique el enfoque UX propuesto.

A2.3 Validar el producto Web considerando las necesidades de un grupo de usuarios de interés.

LI.3 Web Móvil

A los efectos de explorar las tendencias en el desarrollo de productos Web móviles considerando la UX, GIFIS ha decidido llevar a cabo la LI.3.

Actividades de la LI.3

A3.1 Identificar técnicas y herramientas para la Web móvil.

A3.2 Analizar las necesidades de grupos de usuarios de interés.

A3.3 Proponer un enfoque UX para desarrollo Web móvil que incluya la participación de los usuarios.

3. RESULTADOS OBTENIDOS/ ESPERADOS

En lo que se refiere a resultados obtenidos, GIFIS viene trabajando desde 2012 en el área del conocimiento relacionada con la problemática abordada por el presente PI y de este trabajo previo han surgido publicaciones, entre las cuales podemos citar algunas de las más recientes [9][10][11][12][13][14][15].

3.1. Resultados Esperados

Alineado a los objetivos (presentados en la Sección 1.), se espera alcanzar los siguientes resultados a la finalización del presente PI:

1. Desarrollar prototipos/ productos Web que satisfagan necesidades de grupos de usuarios de interés.
2. Aplicar un enfoque UX integrador al desarrollo de los productos Web en 1.
3. Interactuar con los grupos de usuarios de interés que permitan reforzar la validación de los prototipos/ productos Web desarrollados.

3.2. Resultados Obtenidos

En la línea de investigación orientada a Usuarios Adultos Mayores, se identificaron barreras a las que se enfrentan estos usuarios en el acceso e interacción con la Web, en particular con la red social Facebook [13][14][18]. Esta línea de investigación, se desarrolla aprovechando el contexto que ofrece el dictado de talleres de Extensión destinados a adultos mayores, por parte de los integrantes de GIFIS. Por otra parte, en referencia a la línea Experiencia de Usuario, se realizó un proceso de rediseño de un sitio Web universitario, siendo necesario identificar el tipo de contenido que permite mejorar la forma de comunicación del sitio con sus visitantes. Para ello, se evaluó el contenido del sitio Web de la Universidad Nacional de la Patagonia Austral, y también sitios de otras universidades. De acuerdo a los resultados obtenidos se propuso una guía de referencia para desarrollar una estrategia de contenidos completa [16].

Otros resultados a destacar consisten en la generación de proyectos de extensión y vinculación para llegar a la comunidad regional:

- Proyecto de Extensión y Vinculación (PEyV) “Mejoras en la Accesibilidad Web de un Portal Universitario: Evaluación con Usuarios Ciegos”. En este PEyV participan alumnos de grado y postgrado.
- Proyecto de Extensión “Abuelos 2.0: Taller de Redes Sociales para Adultos Mayores”

Los integrantes del grupo de investigación también participan activamente con ponencias en las ediciones anuales de eventos Locales, Nacionales e Internacionales:

- XIII Semana de la Ciencia, la Tecnología y el Arte Científico¹: charla “Principios Básicos de Internet para Adultos Mayores” y taller “Evaluación de la Experiencia de Usuarios en la Web”, ambas actividades abiertas a la comunidad en general.
- XVIII WICC, se presentó el artículo “Evaluaciones de Accesibilidad y Usabilidad en la WWW: Propuestas para Mejorar la Experiencia del Usuario” [15].
- 4º Encuentro de Investigadores UNPA, en que se presentaron los siguientes artículos: i) “Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web” [17], ii) “Los Adultos Mayores y la Utilización de Redes Sociales en Dispositivos Móviles” [18] y iii) “Un Enfoque de Desarrollo Basado en la Experiencia de Usuario (UX) para favorecer la Accesibilidad Web [19].
- INCISCOS 2016, International Conference on Information Systems and Computer Science. Se presentó el artículo: “Content Analysis: A Strategic Foundation to Improve the User Experience of a University Website” [16].

4. FORMACIÓN DE RECURSOS HUMANOS

Es importante señalar que GIFIS está enfocado en brindar el marco adecuado para que cada integrante alcance sus objetivos particulares de investigación y formación al contribuir con el objetivo del PI en curso.

En nuestro PI, contamos con 4 integrantes alumnas de posgrado, las cuales están cursando

la Maestría en Informática y Sistemas (MIS-UNPA):

- Una de las alumnas trabaja en su tesis, denominada “Un Enfoque de Desarrollo Basado en la Experiencia del Usuario (UX) para Favorecer la Accesibilidad Web”.
- Otra alumna está realizando el estudio preliminar para su tesis, la cual se enfocará en el área de diseño y mejora de interfaces de redes sociales para adultos mayores. Esta alumna obtuvo una beca para Alumnos de Postgrado UNPA.
- Otras 2 alumnas, están realizando sus Tutorías de investigación, cuyo plan se titula “Estudio de Técnicas y Herramientas para mejorar la Experiencia Web en Dispositivos Móviles”, y como resultado se elaborarán los respectivos Informes Científico-Técnicos (ICT).

Además, contamos con 1 alumno de grado que continúa su formación como investigador para mejorar la accesibilidad e interacción de los usuarios, y específicamente, se encuentra trabajando sobre sitios Web universitarios. En este marco, obtuvo una Beca de Estímulo a las Vocaciones Científicas del CIN, cuyo plan de trabajo se denomina “Rediseño de un sitio Web para mejorar la Experiencia del Usuarios”. El alumno está rindiendo las últimas materias de la Carrera de Ingeniería en Sistemas UNPA.

Los docentes dirigen también una Beca de Iniciación a la Investigación de 1 alumno de la carrera Analista de Sistemas.

Por otra parte, los investigadores docentes ofrecen cursos de capacitación para la formación de recursos humanos:

- En el marco de la edición XIV de la “Escuela de Informática UNPA-UACO 2016”, se dictaron 2 cursos de extensión y de postgrado denominados: “Diseño de Experiencia de Usuario (UX). Prototipo para una Aplicación Móvil” y “El Documento Científico-Técnico: Un Valioso Recurso para el Ámbito Educativo y Profesional”.

Durante 2016, una docente investigadora ha comenzado a cursar a distancia la Especialización “Diseño de Interacción para Dispositivos Móviles” con la Universitat Oberta de Catalunya (UOC).

¹< <http://www.semanadelaciencia.mincyt.gob.ar/>>

5. AGRADECIMIENTOS

A la UNPA por el soporte al Proyecto de Investigación (PI) N°: 29/B194, Período: 2016-2018, Denominado: “*Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web*”.

6. REFERENCIAS

- [1] Brajnik, G. Barrier Walkthrough: Heuristic Evaluation Guided by Accessibility Barriers http://users.dimi.uniud.it/~giorgio_brajnik/projects/bw/bw.html
- [2] Hartson, R., Pyla, P. S. The UX Book: Process and Guidelines for Ensuring a Quality User Experience. Morgan Kaufmann; 1st Edition, March 2012.
- [3] Vieritz, H., Schilberg, D., Jeschke, S. Merging Web Accessibility and Usability by Patterns, Computers Helping People with Special Needs. Lecture Notes in Computer Science; Volume 6179, 2010.
- [4] Suárez Torrente, M. SIRIUS: Sistema de Evaluación de la Usabilidad Web Orientado al Usuario y Basado en la Determinación de Tareas Críticas <<https://dialnet.unirioja.es/servlet/tesis?codigo=22190>>
- [5] Ihmig, S. Web-Accessibility Patterns. Thesis, University Hamburg, Informatic Department, 2007, 143 pages.
- [6] UWEN (Unified Web Evaluation Methodology) Metodología de Evaluación Europea de la Accesibilidad <<http://www.wabcluster.org/uwem1/>>
- [7] Redish, J. G. Letting Go of the Words: Writing Web Content that Works. Second edition. Ed: Morgan Kaufmann; 2012.
- [8] Halvorson, K.; Rach, M. Content Strategy for the Web. Ed: New Riders; 2012.
- [9] Gaetán, G., Martín, A., Saldaño, V., Sosa, H. Identifying Accessibility Barriers of a University Website: An Evaluation from the Perspective of Blind Users. Sixth International Conference on Advances in New Technologies, Interactive Interfaces and Communicability (ADNTIIC), 2015. ISBN: 978.88.96.471.43.2
- [10] Martín, A., Gaetán, G., Saldaño, V., Miranda, G., Sosa, H., Vilte, D., Pires, A., Molina, S. Hacia una WWW a la medida de las Personas: Identificación, Desarrollo y Uso de Soluciones Web Centradas en el Usuario. WICC 2015. ISBN: 978-987-633-134-0. Salta.
- [11] Moyano, A., Gaetán, G., Martín, A. Interfaz Móvil para el Sitio Web de la UACO. Un Prototipo centrado en el Usuario. ICT-UNPA-129-2015.
- [12] Sosa, H., Gaetán, G., Martín, A. Rediseño de un Portal Web Universitario aplicando Patrones de Accesibilidad. ICT-UNPA 2015, Vol.7, N°2, P137-166. ISSN: 1852-4516.
- [13] Miranda, M. G., Martín, A., Saldaño, V., Gaetán, G. Usabilidad y Accesibilidad en la Redes Sociales. Revista ICT-UNPA 2014, Vol.6, N°2, P118-140. ISSN: 1852-4516.
- [14] Vilte, D., Saldaño, V., Gaetán, G., Martín, A. Identificando Barreras en la Interacción con Facebook: Una Experiencia con Adultos Mayores de la Patagonia Austral. ICT-UNPA 2015, Vol.7, N°2, P249-268. ISSN: 1852-4516.
- [15] Martín, A., Gaetán, G., Saldaño, V., Miranda, G., Sosa, H., Pires, A., Nichele, E. Evaluaciones de Accesibilidad y Usabilidad en la WWW: Propuestas para Mejorar la Experiencia del Usuario. WICC 2016. ISBN: 978-950-698-377-2. Entre Ríos.
- [16] Gaetán, G., Martín, A., Saldaño, V. Content Analysis: A Strategic Foundation to Improve the User Experience of a University Website. INCISCOS 2016. ISBN: 978-9978-389-32-4. Ecuador. 2016.
- [17] Martín, A., Gaetán, G., Saldaño, V., Cardozo, C., Miranda, G., Sosa, H. Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web. 4° Encuentro Investigadores UNPA. Octubre 2016.
- [18] Cardozo, C., Saldaño, V., Martín, A., Gaetán, G. Los Adultos Mayores y la Utilización de Redes Sociales en Dispositivos Móviles. 4° Encuentro Investigadores UNPA.
- [19] Miranda, G., Gaetán, G., Martín, A., Saldaño, V. Un Enfoque de Desarrollo Basado en la Experiencia de Usuario (UX) para favorecer la Accesibilidad Web. 4° Encuentro Investigadores UNPA. Octubre 2016.

Un Modelo de Calidad Mixto como Soporte a la Mejora de los Productos Software con Impacto en los Procesos Organizacionales

Saldarini Javier*, Carrizo Claudio*, Salgado Carlos⁺, Sanchez Alberto⁺, Peralta Mario⁺

*Grupo de I+D Calidad de Software - Facultad Regional San Francisco Universidad Tecnológica Nacional

Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147

{saldarinijavier, cjcarrizo77}@gmail.com

⁺Departamento de Informática Facultad de Ciencias Físico-Matemáticas y Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina e-mail: {csalgado, mperalta, asanchez}@unsl.edu.ar

RESUMEN

Las organizaciones en general cuentan con infinidad de sistemas informáticos implementados y en funcionamiento, la pregunta es si estos sistemas en la actualidad representan, y en qué medida lo hacen, a los requerimientos de los usuarios y cómo responden a las necesidades del dominio que los contiene. Para ello será necesario contar con instrumentos que posibiliten la medición objetiva a la hora de llevar a cabo una evaluación sobre el software que da soporte a los procesos organizacionales, posibilitando la detección de requerimientos de los usuarios y/o necesidades del dominio no cubiertas.

Se propone la construcción de un modelo de calidad de software a través de una metodología establecida y que permita llevar a cabo una evaluación objetiva sobre un software perteneciente a una organización, partiendo de factores de calidad apropiados y deseables y que respondan a las demandas del dominio y a los requerimientos de sus usuarios.

Palabras Clave: *Calidad de Software, Modelos de Calidad, IQMC, ISO 25010*

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de Software con Calidad – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-031516. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

La calidad de los sistemas informáticos se ha convertido hoy en día en uno de los principales objetivos estratégicos de las organizaciones debido a que, cada vez más, sus procesos más importantes y, por lo tanto, la propia supervivencia de las organizaciones depende de los sistemas informáticos según se menciona en [1]. Cuando hablamos de sistemas informáticos debemos mencionar que los productos de software son un componente de importancia dentro del

sistema que los contiene, y la calidad del mismo afectará a la calidad del sistema en su conjunto, aquí entonces sería importante dar alguna definición sobre calidad de software, si bien en la literatura hay varias definiciones, podemos mencionar la de [2] donde define a la calidad de software como el cumplimiento de los requisitos de funcionalidad y desempeño explícitamente establecidos, de los estándares de desarrollo explícitamente documentados, y de las características implícitas que se espera de todo software desarrollado profesionalmente. También en [3] se la define como: “Grado en que el producto software satisface las necesidades expresadas o implícitas, cuando es usado bajo condiciones determinadas”.

La aplicación y/o modernización de cualquier software configura todo un desafío, dado que el software a ser implementado y/o adaptado debe dar soporte a los procesos organizacionales y satisfacer las necesidades de los usuarios.

Para saber si un software que está implementado y en funcionamiento en una organización realmente cumple con los objetivos organizacionales y satisface las necesidades de sus usuarios, se debería contar con herramientas y/o instrumentos que permitan realizar algún tipo de medición al respecto. El saber esto traería aparejado saber en cuánto ese software cubre las demandas organizacionales y las necesidades de los usuarios, y en función de ello proponer las mejoras necesarias para cubrir las necesidades no cubiertas por el software y posibilitar futuras evaluaciones a medida que el software vaya evolucionando.

Entonces, si se habla de organizaciones se debería saber que cada organización cuenta con un sinnúmero de singularidades que la hace diferente a otra, por cuanto, y como punto de partida

sería de gran utilidad contar con modelos conceptuales de los procesos organizacionales involucrados, a los cuales los productos de software dan soporte. Esto permitiría unificar la terminología, facilitar la comunicación y validación del modelo, entre otros beneficios. Es por ello que es de interés poder contar con modelos de procesos de negocios (MPN). Como se expresa en [4], se puede decir que un modelo de proceso de negocio muestra las actividades que se deben realizar para alcanzar los objetivos establecidos por la organización en su negocio. También en [4] se menciona que, en el campo del modelado de procesos de negocio se pueden encontrar numerosas propuestas de lenguajes de modelado, como IDEF0, IDEF3, UML, UML 2.0 y BPMN, por mencionar algunos.

Otro aspecto importante es que se pueda evaluar la calidad de un producto de software que da soporte a los procesos organizacionales, según se describe en [5], la evaluación de la calidad de un producto de software juega un papel fundamental, tanto en la selección de un componente de software (ej., un sistema legado o un componente a ser adquirido, ya sea de índole comercial -COTS- o libre y de código abierto -FOSS-), como la validación de un producto resultante de un proceso de construcción a la medida (sea este construido por una unidad interna de la organización o por una empresa externa subcontratada para este propósito). Lejos de ser un problema sencillo, esta tarea suele ser muy compleja de realizar. Los Modelos de Calidad del Software (MC), son artefactos específicamente diseñados y construidos para apoyar en estos procesos.

Según Carvallo en [6], y el estándar ISO 8402 [7] un modelo de calidad puede definirse como el conjunto de factores de calidad, y de relaciones entre ellos, que proporciona una base para la

especificación de requisitos de calidad y para la evaluación de la calidad de los componentes software. Los modelos de calidad se estructuran generalmente como una jerarquía (ya sea un árbol, ya sea un grafo dirigido), donde factores de calidad más genéricos, como eficiencia o usabilidad, se descomponen en otros más particulares, como tiempo de respuesta o facilidad de aprendizaje, probablemente en diversos niveles de descomposición. También menciona que las propuestas existentes de modelos de calidad se pueden clasificar según si tienen un enfoque de modelos de calidad fijos, a medida o mixtos.

Entre los modelos de calidad fijos se pueden observar los de McCall et al. (1997) [8], Boehm et al. (1978) [9], Keller et al. (1990) [10] y FURPS Grady y Caswell (1987) [11]. Para los modelos de calidad a medida existen diversas propuestas de métodos para crearlos entre las que podemos destacar a GQM (Goal-Question-Metric) de Basili[12] y la del estándar IEEE 1061 [13].

Para el caso de los modelos de calidad mixtos se pueden destacar el ADEQUATE Horgan[14], el modelo de Gilb[15] y el modelo propuesto en el estándar ISO/IEC 9126-1 [16], este último es actualizado y reemplazado por el estándar ISO/IEC 25010 [17].

Coincidiendo con lo expresado en [6] (Carvallo J.P., et al. 2010 Capítulo 10), la construcción de modelos de calidad viene dificultada por distintas circunstancias relacionadas con: (1) el equipo que realiza la construcción del modelo, en el caso de que este equipo no tenga experiencia en la construcción de modelos de calidad o bien en el contexto del dominio del componente objeto; (2) el dominio para el que se construye el modelo, para el que en muchas ocasiones no existe una terminología común; (3) factores metodológicos, ya que es difícil conocer el nivel de profundidad hasta el

que es necesario descomponer los modelos, y por tanto cuándo se puede decir que un modelo de calidad se ha finalizado.

También se expresa que la existencia de un método que proporcione unas pautas para la construcción de los modelos de calidad puede ayudar a paliar estas dificultades y a continuación los autores presentan el método Individual Quality Model Construction (IQMC), el cual proporciona un conjunto de guías y técnicas para la identificación de los factores de calidad apropiados que deben ser incluidos en un modelo de calidad que permita analizar la calidad de componentes pertenecientes a un cierto dominio de software.

IQMC adopta un enfoque de construcción mixto. El catálogo de partida original para la construcción de los modelos fueron los constituido por ISO/IEC 9126-1 [16] e ISO/IEC 9126-1-

NT. Si bien no hay demasiado desarrollo en la actualidad el catálogo de partida debería ser actualizado al estándar ISO/IEC 25010[17].

El método IQMC propone unos pasos para el refinamiento de dichos catálogos que conducen a la construcción de modelos de calidad para componentes software de un cierto dominio. Los modelos que se obtienen se deberían estructurar a través del estándar ISO/IEC 25010 [17].

Según se detalla en [18], los Modelos de Calidad del Software (MC), son artefactos específicamente diseñados y contruidos para soportar evaluación y selección de componentes de software. Permiten la definición estructurada de criterios de evaluación, la especificación de requerimientos, la descripción de componentes en relación a ellos y la identificación de desajustes de manera sistemática facilitando el proceso de evaluación y selección del software. Para

la construcción de éstos, es importante contar con una metodología que de soporte al proceso de construcción.

El objetivo de este trabajo es desarrollar un modelo que favorezca la evaluación de la calidad del software implementado y funcionando en una organización con el fin de detectar los requerimientos de los usuarios y/o necesidades del dominio no cubiertas, en pos de mejorar la eficiencia organizacional a través de la mejora en la calidad del software que da soporte a los procesos de la organización. El modelo que se propone se centrará en la calidad del producto de software, tomando como punto de partida el modelo conceptual de los procesos de la organización.

Es de gran importancia que las organizaciones puedan contar con modelos de calidad que partan del modelo conceptual de sus procesos y que les permita la evaluación objetiva y de manera sistémica de la calidad del software que da soporte a los procesos organizacionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación están asociados a:

- Evaluación de la calidad de productos de software.
- Estudio de modelos conceptuales aplicados a la calidad de productos de software.
- Estudio de estándares y metodologías aplicadas a la construcción de MC.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los objetivos que persigue esta línea de investigación son:

Objetivo general

- Lograr, a través de un MC, una

evaluación objetiva de un determinado software para saber si este cubre las demandas organizacionales y las necesidades de los usuarios, y en función de ello proponer las mejoras necesarias para cubrir las necesidades no cubiertas por el mismo.

Objetivos específicos

- Construir un modelo de calidad de software mixto a través de la metodología IQMC y la Norma ISO 25010 para un dominio de aplicación específico.
- Llevar a cabo la evaluación del software utilizando el MQ construido para tal fin.
- Elaborar un informe con los resultados de la evaluación, este informe contendrá las necesidades no cubiertas por el software, tanto a nivel organizacional como a nivel de usuarios.
- Difusión y divulgación de los resultados.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación se trabaja en lo referente a distintos métodos de evaluación de calidad con la participación del Grupo de Investigación "Calidad de Software" perteneciente a la UTN Facultad Regional San Francisco, Córdoba.

Se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) del Ing. Javier Saldarini, de la Facultad de Ciencias Físico- Matemáticas y Naturales, Universidad Nacional de San Luis. Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación y trabajos finales en el marco de la Ingeniería en Informática.

5. BIBLIOGRAFÍA

- [1]. Piattini Velthuis et al., Calidad de sistemas de información. 2ª edición: RaMa, 2011.
- [2]. R. Pressman, Ingeniería del Software. 6ª Ed: Mcgraw-Hill, 2005.
- [3]. ISO/IEC 25000 Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-System and software quality models.
- [4]. C. Salgado “MEMPN: Método para la Evaluación de Modelos Conceptuales de Procesos de Negocio”, Maestría, Facultad de Ciencias Físico Matemáticas y Naturales. Universidad Nacional de San Luis, San Luis. Argentina, 2013.
- [5]. A. Villalta, J.P. Carvallo “Modelos de calidad de software: Una revisión sistemática de la literatura” en Maskana, CEDIA 2015.
- [6]. C. Calero, M. Piattini, M. Moraga, Calidad del producto y proceso software: Ra-Ma, 2010.
- [7]. INTERNATIONAL STANDARDS ORGANIZATION (1986). ISO International Standard 8402: Quality Management and Quality Assurance-Vocabulary
- [8]. MCCALL, J.A., RICHARDS, P.K. y WALTERS, G.F. (1977) “Factors in Software Quality”. RADC TR-77-369, Vols I, II, III, US Rome Air Development Center Reports NTIS AD/A-049.
- [9]. BOEHM, B.W., BROWN, J.R., KASPAR, H., LIPOW, M., MACLEOD, G.J. y MERRITT, M.J. (1978). “Characteristics of Software Quality”. North Holland Publishing Company.
- [10]. KELLER, S., KAHN, L. y PANARA, R. (1990) “Specifying Software Quality Requirements with Metrics”. Systems and Software Requirements Engineering - IEEE Computer Society Press – Tutorial
- [11]. GRADY R.B. y CASWELL, D.L. (1987). “Software Metrics: Establishing a Company-Wide Program”. Prentice-Hall.
- [12]. BASILI, V.R., CALDIERA, G. y ROMBACH, H.D. (1994). “Goal Question Metric Paradigm”. En: Encyclopedia of Software Engineering 1, John Wiley & Sons.
- [13]. INSTITUTE OF ELECTRICAL ELECTRONIC ENGINEERING (1998). IEEE Std 1061-1998 IEEE Standard for a Software Quality Metrics Methodology – Description.
- [14]. HORGAN, G., KHADDAJ, S. y FORTE, P. (1999) “Aessential Views Model for Software Quality Assurance”. En: Project Control for Software Quality, Shaker Publishing.
- [15]. GILB, T. (1988). “Principles of Software Engineering Management”. Addison Wesley.
- [16]. ISO/IEC 9126-1. (2001). Software engineering — Product quality — Part 1: Quality model.
- [17]. ISO/IEC 25010:2011. Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models.
- [18]. J. Bermeo Conto, M. Sánchez, J. J. Maldonado, and J.P. Carvallo “Modelos de Calidad de Software en la Práctica: Mejorando su Construcción con el Soporte de Modelos Conceptuales”, CEDIA 1-abr- 2016.

Una Plataforma Tecnológica para el Seguimiento y Evaluación de Calidad de Proyectos Ágiles

Nicolás Tortosa; Blas Cabas Geat; Noelia Pinto

GICS (Grupo de Investigación en Ingeniería y Calidad del Software), UTN,
FRRe

French 414, Resistencia, Chaco
{nicotortosa; blasc147; ns.pinto}@gmail.com

RESUMEN

El uso de procesos ágiles por parte de la industria de software va ganando terreno día a día. Y en este sentido la determinación de la calidad con la que dichos procesos son llevados a cabo por parte de las empresas de software es fundamental para el éxito o fracaso de los productos de software que se fabriquen. Si bien existen estándares de calidad para procesos de software, estos no están pensados para evaluar la calidad de procesos ágiles de software, ni se ajustan a la realidad de las empresas de software del medio. Por ello, es notoria la necesidad de evaluar la calidad de los procesos de desarrollo ágiles, y verificar, entonces, si las buenas prácticas que estas metodologías ofrecen, están siendo cumplidas. En este trabajo se presenta la línea de investigación que se lleva adelante con el objetivo de implementar un framework que permita realizar el seguimiento de proyectos ágiles y la evaluación de calidad de los procesos asociados, evitando una mala administración de recursos, tiempos y costos.

Palabras clave: Calidad del Producto y del Proceso de Software, Metodologías Ágiles,

CONTEXTO

El trabajo que aquí se presenta está enmarcado en el proyecto “Evaluación de Calidad en Procesos Ágiles de Desarrollo de Software”, que es financiado por la UTN y

ejecutado en el Grupo de Investigación en Ingeniería y Calidad de Software (GICS) de la Facultad Regional Resistencia, con el código IAI4445TC.

Asimismo algunas actividades son compartidas con el Proyecto de Desarrollo Tecnológico y Social (PDTs) “Aporte a la competitividad de las empresas de desarrollo de Software del NEA”, también radicado en el GICS y aprobado por el Consejo Interuniversitario Nacional (CIN) en su Convocatoria 2014. Y se relaciona con actividades enmarcadas en el Proyecto “Herramientas y Métodos de soporte a la Ingeniería de Software: requerimientos, estrategias ágiles y calidad de procesos y productos”, radicado en la Facultad Regional Santa Fe y homologado como Proyecto de la Secretaría de Ciencia, Tecnología y Posgrado de la UTN, con código IPN4409.

1. INTRODUCCIÓN

En Argentina, la Industria del Software se compone mayoritariamente por PYMES donde la calidad del trabajo realizado, los bajos costos y las entregas oportunas son elementos esenciales para el incremento de las ventas internas y la proyección a nivel internacional.

Además este tipo de empresas trabaja continuamente en pos de mejorar sus niveles de competitividad forzándolas a la búsqueda de la mejora continua de sus procesos mediante una gestión exitosa que implica una correcta definición, ejecución, medición y control de sus actividades [1]. Para ello,

existen numerosas propuestas metodológicas que guían el ciclo del desarrollo de software y que inciden en distintas dimensiones del proceso.

Así, las metodologías más tradicionales se centran especialmente en una rigurosa definición de roles, de las actividades involucradas, los artefactos que se deben producir, y las herramientas y notaciones que se usarán [2].

Por otro lado las metodologías ágiles han surgido como una alternativa para maximizar el valor del negocio a través de pequeños equipos de trabajo auto-organizados utilizando tecnologías flexibles y la participación temprana de los clientes para mejorar iterativamente software [3].

En este contexto, existen varias herramientas de administración de proyectos que asisten a los equipos de trabajo en visualizar planes y progreso del proyecto. Sin embargo, estas herramientas no ofrecen información respecto a la calidad de los procesos asociados al proyecto ni ayuda a los administradores a tomar decisiones en pos de la mejora continua.

Se presenta, entonces, en este artículo la propuesta denominada AQF (Agiles Quality Framework) que asiste a los equipos de trabajo a lo largo de las distintas etapas del ciclo de vida del proyecto, y ofrece a los administradores la posibilidad de evaluar la calidad de sus procesos, de acuerdo a los componentes definidos en el modelo QuAM (Quality Agile Model) [4].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo general de la línea que aquí se describe es contribuir en la mejora de los procesos de desarrollo de software y en la medición cuali-cuantitativa de los productos obtenidos por las pymes del NEA. Para ello se trabaja en el diseño y desarrollo de una plataforma que incluya la propuesta de un nuevo modelo de calidad adaptado a proyectos

ágiles y su evaluación, junto a herramientas tecnológicas que permitan su implementación flexible e integrada en las empresas pymes de desarrollo de software del NEA.

Hasta el momento se ha logrado definir y validar el modelo QuAM [5], incluyendo el patrón de métricas que deberán evaluarse para obtener el nivel de calidad asociado al proyecto ágil. Sin embargo el proceso de seguimiento no puede ser cubierto solo con la definición del modelo, resulta necesario el diseño de una suite de aplicaciones que permita la gestión automatizada de los elementos de QuAM.

Así actualmente el equipo de trabajo ha iniciado el proceso de modelado y desarrollo de QuAGI, una aplicación web cuyo objetivo se enfocará en la administración de los proyectos a través de la visualización del plan asociado, informes respecto a estados del mismo, y el seguimiento de las actividades requeridas. Además esta aplicación permitirá brindar soporte continuo a los procesos de toma de decisiones asistiendo a los responsables mediante informes sobre evaluación de calidad del proyecto en cuestión y recomendaciones de ajustes para la mejora continua en caso de desviaciones que ocurran a lo largo del mismo.

Respecto a la arquitectura de QuAGI, el equipo realizó, en primer instancia y desde la perspectiva de lógica del negocio, el diseño de componentes basados en servicio REST para intercambio de toda aquella información solicitada entre diversas aplicaciones web asegurando escalabilidad y facilidad de uso. Además se estableció la utilización del framework de desarrollo Django, basado en un patrón MVT (Modelo-Vista-Template) el cual permite un patrón de diseño modular, facilitando la distribución y realización de las tareas en el equipo de desarrollo [6].

En cuanto a la organización del trabajo, este Proyecto se llevará adelante siguiendo prácticas ágiles adaptadas a la realidad del equipo y las características del desarrollo en cuestión. Así se establecieron las siguientes

actividades como transversales a todas las etapas del ciclo de obtención de QuAGI:

- Definición y priorización de todas las historias de usuarios, con sus correspondientes criterios de aceptación, que componen el Product Backlog.
- Definición del alcance de cada sprint y posterior asignación de tareas a cada uno de los integrantes del equipo de desarrollo. Estas tareas, a nivel de historia de usuario, involucran diseño e implementación de UI a nivel de front-end, desarrollo a nivel de back-end, incluido el servicio rest, y diseño de los test unitarios para verificar el correcto funcionamiento de cada una de las entidades definidas en el modelo.
- Para la puesta en producción de la aplicación web, se definen una serie de releases, incluyendo la configuración del servidor web, instalando todo lo necesario para el acceso de los usuarios desde el navegador.

Una vez obtenida una primer versión integrada de QuAGI, se establecerá un mecanismo de validación que implicará la delineación de una estrategia que permita vincular el trabajo real llevado a cabo por una PYME con el uso del framework que aquí se presenta.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Se presenta en este artículo el framework AQF con el objetivo de optimizar la gestión de proyectos de software basados en prácticas ágiles teniendo en cuenta la calidad del proceso y del producto final, a partir del uso de QuAGI como herramienta de soporte y basado en QuAM como modelo de calidad.

Actualmente el equipo abocado a este proyecto se encuentra en la fase de ajustes en el desarrollo de QuAgi. Este trabajo se realiza

teniendo en cuenta los resultados obtenidos en la experiencia de validación previa con empresas PYMES del NEA, que participaron de un proceso de encuestas a partir de las métricas definidas en QuAM [7].

Como trabajos futuros se pretende, además de continuar con el desarrollo de las funcionalidades correspondientes a las actividades de Seguimiento y Evaluación de Proyectos utilizando QuAgi, incorporar el diseño e implementación de un Asistente Virtual a la plataforma con el objetivo de conocer en base al nivel de calidad obtenido al momento, cuáles son las recomendaciones de ajustes para mejorar el valor alcanzado. Incluso se podrá visualizar mejoras que podrían obtenerse simulando nuevos parámetros asociados al Proyecto en cuestión.

De esta forma se prevé continuar con el desarrollo del framework AQF, integrando el Seguimiento de Proyectos y el enfoque inteligente que será provisto por el Agente Virtual que dará soporte a los usuarios de QuAgi, y a través de la implementación en al menos 2 empresas que permitan validar la aplicabilidad y correctitud de la plataforma.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación del Grupo de Investigación en Ingeniería y Calidad de Software (GICS) de UTN Facultad Regional Resistencia está integrado por el Director (Doctor, Categorizado en el Programa de Incentivos y como Docent Investigador de UTN), y tres Docentes Investigadores (Ingenieros, dos con Categoría "E" UTN y otro Categoría "G" UTN). Actualmente existen dos Tesis de posgrado radicadas en el Grupo, uno correspondiente a la Maestría en Ingeniería de Software y otro al Doctorado en Ciencias

Informáticas ambas carreras de Universidad Nacional de La Plata.

Además el equipo no sólo incluye investigadores formados sino que incorpora y forma jóvenes investigadores algunos de los cuales colaboran como docentes en diversas cátedras y otros se encuentran insertos laboralmente en Pymes de Software del NEA. Se cuenta, así, con un Becario BINID y un Becario alumno de Investigación provistos por la Secretaría de Ciencia y Tecnología de la Facultad.

REFERENCIAS

- [1] Mora, B., Garcia, F., Ruiz, F., Piattini, M., Boronat, A., Gomez, A., ... & Ramos, I. (2008). Software generic measurement framework based on MDA. *IEEE Latin America Transactions*, 6(4), 363-370.
- [2] Letelier, P., Penadés, P. (2006) "Metodologías ágiles para el desarrollo de software: eXtreme Programming (XP)" *Técnica Administrativa*, Buenos Aires. ISSN 1666-1680.
- [3] Lee, S., & Yong, H. S. (2013). Agile software development framework in a small project environment. *Journal of Information Processing Systems*, 9(1), 69-88.
- [4] Pinto, N., Acuña, C., & Cuenca Pletsch, L. R. (2016). Quality Evaluation in Agile Process: A First Approach. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [5] Pinto, N., et al (2016). Validación del diseño de componentes de QuAM: un Modelo de Calidad para procesos Ágiles. Publicado en *Anales del IV Seminario Argentina-Brasil de Tecnologías de la Comunicación y la Comunicación (SABTIC 2016)*. ISBN 978-987-3619-15-1
- [6] Información disponible en <https://docs.djangoproject.com/en/1.10/>
- Revista Española de Innovación, Calidad e Ingeniería del Software, Vol.1, No. 2
- [5] Pasini, A. C., Esponda, S., Bertone, R. A., & Pesado, P. (2008). "Aseguramiento de Calidad en PYMES que desarrollan software." XIV Congreso Argentino de Ciencias de la Computación.
- [6] Letelier, P., Penadés, P. (2006) "Metodologías ágiles para el desarrollo de software: eXtreme Programming (XP)" *Técnica Administrativa*, Buenos Aires. ISSN 1666-1680
- [7] Noelia Pinto, Gabriela Tomaselli , Liliana Cuenca Pletsch , Nicolás Tortosa , César J. Acuña. "Validación del diseño de componentes de QuAM: un Modelo de Calidad para procesos Ágiles". IV Seminario Argentina-Brasil de Tecnologías de la Comunicación y la Comunicación (SABTIC 2016). ISBN 978-987-3619-15-1

**Innovación en Educación
en Informática**

Agente Hornero. Ampliando las Posibilidades de Aprender a Programar

Pablo Kogan¹

Jorge Rodríguez'

Federico Amigone¹

email: {pablo.kogan, j.rodrig, f ede.amigone}@fi.uncoma.edu.ar

¹ *Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

Resumen

La apropiación de conceptos desarrollados en el campo de las Ciencias de la Computación constituye un recurso fundamental para comprender el mundo moderno. Por otro lado, aspectos fundamentales de la disciplina, como la comprensión de estructuras algorítmicas, suministran al estudiante un valioso recurso cognitivo para la resolución de problemas no computacionales y modelos útiles para entender el dinamismo de la sociedad en la que se encuentra inmerso.

De este marco emerge una iniciativa promovida por la Facultad de Informática, la implementación de un modelo de enseñanza y aprendizaje lúdico y colectivo a través de un Agente que soporta torneos de programación llamado Hornero. En éste contexto se presenta una Línea de Investigación que propone la ampliación funcional de determinadas características técnicas de Hornero, el estudio sobre enfoques metodológicos que se verifiquen en la práctica.

Resulta valioso comprender el impacto que determinadas características funcionales reportan sobre los procesos de enseñanza y de aprendizaje. La

articulación de la experiencia en el terreno de las actividades de extensión constituye una retroalimentación de gran importancia para la evaluación de los modelos y estrategias utilizadas.

Palabras Clave: ENSEÑANZA DE LAS CIENCIAS DE LA COMPUTACIÓN, ENSEÑANZA DE LA PROGRAMACIÓN, HORNERO, PROGRAMACIÓN POR BLOQUES, ENFOQUES BASADOS EN COMPETENCIA, APRENDIZAJE COLABORATIVO

Contexto

Esta propuesta se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial que buscan desarrollar conocimiento especializado en el área de la Inteligencia Artificial.

Así mismo, la línea se nutre del Convenio Marco de Colaboración firmado durante 2016 entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén. El objetivo de dicho convenio es el de contribuir recíprocamente al desarrollo de actividades de investigación, formación de recursos humanos y promoción de la enseñanza de las Ciencias de la Computación durante la escolaridad

obligatoria.

Este trabajo se presenta en el marco del proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica y por el Consejo Provincial de Educación en contexto del Convenio Marco de Colaboración. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

Las actividades concretadas en el ámbito de la investigación se plantean como articuladas al Proyecto de Extensión Universitaria *Torneos de Programación para Promover el Aprendizaje en la Escuela Media* [7], en ejecución durante 2017, con intención de construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el campo de la praxis.

Introducción

La enseñanza y el aprendizaje de conceptos propios a las Ciencias de la Computación son considerados prioritarios para la escolaridad obligatoria. La hipótesis de trabajo plantea que trabajar sobre este tipo de aprendizajes mejora las posibilidades de los estudiantes de comprender e intervenir el mundo que los rodea, aportando conceptos y estrategias que pueden ser transpuestas a otros contextos [9, 3, 6].

En este sentido se considera que la formación en aspectos fundamentales del campo de los Algoritmos y la Programación amplia las oportunidades de comprender como están contruidos los dispositivos de software con los que interactúan frecuentemente y desarrolla las capacidades para crear sus propias aplicaciones [8].

Por otro lado, la construcción de algoritmos y la programación favorece el

desarrollo del pensamiento computacional contribuyendo a la construcción de habilidades y estrategias para la resolución de problemas potencialmente transferibles a dominios no computacionales [10, 11].

En este marco, durante los últimos años, se intensifica la concreción de diversas iniciativas que buscan promover la enseñanza y el aprendizaje de la algorítmica y la programación en el ámbito de la educación obligatoria [1, 2].

Estos esfuerzos se ubican al menos en tres dimensiones que resultan convergentes en los proceso de promoción de enseñanza de la programación: el desarrollo de productos que buscan hacer accesible este tipo de conocimientos a mas estudiantes, la definición de enfoques metodológicos y el diseño de estrategias que buscan acercar el contenido disciplinar a la escolaridad obligatoria.

En este contexto, a partir de 2014, la Facultad de Informática desarrolla una Línea de Investigación y Desarrollo articulada a iniciativas en el ámbito de Extensión Universitaria que proponen la aproximación de estudiantes y escuelas secundarias a la programación a partir de la realización de torneos de programación gestionados por Hornero [5, 4].

Hornero es un agente diseñado para soportar competencias de programación en varios lenguajes posibilitando la participación de un amplio universo de estudiantes. Actualmente puede competir con Java, Python, Php, C, C++, Pascal, Javascript, C#, Ciao-Prolog, Perl, Bash, Lisp, Ruby, Smalltalk y PSeInt.

A las líneas de acción que se proponen en el presente trabajo, se suma la definición de un enfoque metodológico diseñado para favorecer los procesos de enseñanza y aprendizaje de la programación basados en competencia y aprendizaje colaborativo.

La competencia en entornos distendidos introduce a los estudiantes en una

trama de aprendizaje compartido, que construye un sentido de logro colectivo y superación individual.

En este trabajo se presenta una Línea de Investigación y Desarrollo que busca fortalecer las características funcionales de Hornero así como la puesta a punto de determinados enfoques metodológicos lúdicos y colaborativos, los cuales en suma, exploran la posibilidad de hacer accesible las Ciencias de la Computación a la generalidad de los estudiantes.

En este sentido se plantea trabajar en la incorporación de lenguajes de programación por bloques, mejorar el grado de integración con PSeInt y desarrollar interfaz web o móvil para resolver problemas en papel.

La convergencia entre las características técnicas de la aplicación y la instancia social de aprendizaje será materia de estudio en términos de búsqueda de mejores aproximaciones en los enfoques de enseñanza y de aprendizaje, sin desatender el estudio de las implementaciones que las posibilitan.

El desarrollo de la Línea de Investigación requiere necesariamente de la sostenida retroalimentación desde el campo de la praxis. En este ámbito las actividades de Extensión Universitaria resultan un componente necesario para avanzar en la construcción de conclusiones teóricas y productos tecnológicos.

Lineas de Investigación

En el marco de la Línea de Investigación y Desarrollo que se presenta en este trabajo se plantea el desarrollo de actividades en relación a temáticas vinculadas al ambiente de trabajo construido en torno a Hornero, tanto en lo referido al fortalecimiento de las características tecnológicas como a la formalización y mejora de los enfoques metodológicos.

Se prestará especial atención a la identificación del conocimiento producido en el campo de los Agentes Inteligentes y los Modelos Formales susceptibles a ser aplicados en el contexto específico de trabajo demarcado por Hornero.

Se proyecta abordar las siguientes líneas de trabajo:

- Desarrollos sobre Hornero: Integrar al menú del PSeInt. Incluir la comunicación con Scratch. Interfaz Web y/o móvil para resolver problemas en papel. Aspectos de seguridad para validar usuarios en torneos presenciales.
- Estudios en el contexto del enfoque metodológico y clasificaron del catalogo de problemas.
- Definición de modelo para la promoción de la enseñanza de la Ciencias de la Computación a partir de torneos.
- Medición de resultados en términos apropiación de conceptos de programación tanto en la dimensión grupal como en la individual

Estas líneas de trabajo se presentan como articuladas para mejorar las posibilidades de avanzar en la construcción de productos tecnológicos y enfoques teórico conceptuales consistentes e integrados.

Resultados Esperados

En el contexto de las actividades proyectadas se espera:

- Determinar modelos formales existentes que tengan implicancia en el contexto educativo y cuya aplicación sostenga un abordaje pedagógico acorde a las capacidades

de aprehensión de la base de estudiantes.

- Avanzar en la identificación y descripción de los elementos que permitan fortalecer el enfoque metodológico.
- Diseñar modelos conceptuales para la promoción de la enseñanza de las Ciencias de la Computación.
- Avanzar en el fortalecimiento de las características técnicas del Agente Hornero.

Actividades de Extensión relacionadas al Proyecto

El proyecto busca construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el ámbito de la Extensión Universitaria en relación a la aplicación de agentes al contexto educativo. La iniciativa de referencia se expresa en el siguiente Proyecto de Extensión Universitaria:

- *Torneos de Programación para promover el aprendizaje en la Escuela Media.* En este proyecto se propone organizar el proceso de aprendizaje de la programación a partir de la realización de torneos de programación soportados por Hornero [7].

Esta iniciativa extensionista se realiza en colaboración con cuatro escuelas secundarias de la región. En este contexto se espera desarrollar actividades de formación docente en el campo de la enseñanza de la programación y de formación a estudiantes secundarios en conceptos fundamentales del área de conocimiento.

El proyecto se ejecuta durante

2017 y está avalado por Resolución Fal 086/16.

Formación de Recursos Humanos

La ejecución del proyecto de investigación brindará un marco propicio para la iniciación y/o finalización de estudios de posgrado de los integrantes docentes. Asimismo, será un ámbito adecuado para la realización de tesis de grado de la carrera de Licenciatura en Ciencias de la Computación. Se espera, además, la consolidación como investigadores de los miembros más recientes del grupo de investigación.

El área científico - tecnológica principalmente impactada por la formación de recursos humanos en el marco de este proyecto será el área de Tecnologías de la Información y de las Comunicaciones.

Referencias

- [1] M. Bonello and H. Czemerinski. Program.ar: una propuesta para incorporar ciencias de la computación a la escuela argentina. 2015.
- [2] M. Borchardt and I. Roggi. Ciencias de la computación en los sistemas educativos de América Latina. 2017.
- [3] D. S. Chair, S. Carey, B. Fuschetto, I. Lee, D. Moix, D. O'Grady-Cunniff, B. Owens, C. Stephenson, and A. Verno. *K-12 Computer Science Standards*. The Computer Science Teachers Association, New York, 2016.
- [4] C. C. Fracchia, P. Kogan, A. C. Alonso de Armin°, I. Godoy, and L. M. López. Realización de torneos

- de programación como estrategia para la enseñanza y el aprendizaje de programación. In *XX Congreso Argentino de Ciencias de la Computación (Buenos Aires, 2014)*, 2014.
- [5] C. C. Fracchia, P. Kogan, and S. Amaro. Competir + motivar + hornero = aprender programación. *TE & ET*, 2016.
- [6] S . F u r b e r . *Shut down or restart?The way forward for computing in UK schools* . The Royal Society Education Section, 2012.
- [7] I. Godoy and P. Kogan. *Torneos de Programación para promover el aprendizaje en la Escuela Media*. 2017, avalado por Resolución FaI 086/16.
- [8] M. Resnick, J. Maloney, A. Monroy-Hernandez, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, et al. Scratch: programming for all. *Communications of the ACM*, 52(11):60-67, 2009.
- [9] M. Smith. Computer science for all. *The White House*, 2016.
- [10] J. M. Wing. Computational thinking. *Communications of the ACM*, 49(3):3335, 2006.
- [11] J . M . W i n g . Computational thinking and thinking about computing. *Philosophical transactions of the royal society of London A: mathematical, physical and engineering sciences*, 366(1881):3717-3725, 2008

Aplicaciones de Software con Tecnología NFC en Dispositivos Móviles, para Niños con Diversidad Funcional

Contreras, Victor (Dirección de Informática y Tecnología de las Comunicaciones)

García, Mariana (Docente de Psicología Evolutiva)

Moreno, Viviana (Vinculación Tecnológica)

Langer, Ariel (Director de Vinculación Tecnológica)

vcontreras, mgarcia, vemoreno ,alanger {@unpaz.edu.ar}

Universidad Nacional de José Clemente Paz (UNPAZ)

Leandro N. Alem 4731, José C. Paz (CP 1665), Buenos Aires

RESUMEN

El creciente desarrollo de las neurociencias, en articulación con la práctica clínica profesional e interdisciplinaria y los avances tecnológicos, han abierto y enriquecido las herramientas disponibles para el abordaje e intervención de las personas con TEA, como así también de otros trastornos del desarrollo, añadiendo nuevos recursos y herramientas de apoyo.

Es desde allí, que la incorporación del uso de la tecnología NFC, en dispositivos móviles, para niños con diversidad funcional, puede tornarse un aporte para terapeutas, docentes y familias. Nuestro enfoque teórico parte del manual diagnóstico DSM V y la Teoría de la mente de Barón-Cohen.

Siendo por lo general, las conductas de las personas con autismo de características rígidas, es que los cambios los pueden afectar significativamente. Por ello, este abordaje prevé el uso del recurso de secuencias temporales para atender a las necesidades de cada niño de forma anticipada y sistemática.

Asimismo, el uso de la tecnología NFC puede tornarse también un instrumento medidor y habilitante entre el niño, sus necesidades e intereses y el medio.

Finalmente, se prevé que esta experiencia pueda inaugurar el diseño de nuevas propuestas de actividades de transferencia tecnológica a otros campos de aplicación.

PALABRAS CLAVES: Innovación tecnológica - NFC- pictogramas animados - dispositi-

tivos móviles - Autismo - información de contexto.

CONTEXTO

La presente investigación, se enmarca en un proyecto más general, titulado: “Las interfaces gestuales en procesos de integración educativa y social. Innovaciones técnicas para la inclusión de niños con autismo y otros trastornos en escuelas especiales”. Su objetivo general, es fortalecer las actividades de investigación aplicada a partir del desarrollo de herramientas de software que utilicen interfaces naturales en Trastornos de Espectro Autista (TEA), analizando las posibilidades de transferencia tecnológica a escuelas especiales que trabajen con niños con estas u otras dificultades de aprendizaje.

Este, pretende, profundizar una línea de investigación financiada por la convocatoria PROCODAS, en la que se presentó una iniciativa de transferencia tecnológica a la Escuela Especial N° 501 obteniendo la aprobación del proyecto por Resolución Ministerial N° RESOL-2016-693-E-APN-MCT de fecha 23 de Octubre de 2016.

En este marco, prevemos el diseño de aplicaciones de software, con tecnología Near Field Communication (NFC) que favorezcan la comunicación, permitan crear procesos de anticipación, planeamiento y organización tanto de situaciones y conductas, como así también brindar información de contexto. Partimos de que la tecnología NFC en dispositivos móviles es un recurso posible a utilizar

tanto en niños con TEA como en otros con diversidad funcional, – a través de pictogramas animados-; así también, el que se puede propiciar y extender las actividades de transferencia tecnológica a otros campos de aplicación.

1. INTRODUCCIÓN

El proyecto *Aplicaciones de software con tecnología NFC en dispositivos móviles, para niños con diversidad funcional*, tiene por objetivo general, fortalecer las actividades de investigación aplicada a partir del desarrollo de aplicaciones de software, con tecnología NFC (mediante el uso de pictogramas animados en dispositivos móviles). El mismo, busca favorecer la comunicación, permitir planificar, organizar, y anticipar, así como brindar información de contexto, a niños con diversidad funcional. Así también, prevé extender las actividades de transferencia tecnológica a otros campos de aplicación.

Se espera que estas herramientas contribuyan a las terapias convencionales y favorezcan una mejor interacción social. Las actividades de las aplicaciones están pensadas para ser desarrolladas de manera tripartita, esto es, se fomentará la interacción entre el niño, las nuevas tecnologías y el profesional o tutor y se apunta a que sea transferible a distintos ámbitos de referencia del niño, principalmente familia y escuela.

Lo innovador de este proyecto radica en que no se han encontrado investigaciones o proyectos relacionados con la aplicación de herramientas de software que utilicen tecnología NFC - pictogramas animados -, para niños con TEA o diversidad funcional de otro tipo, a nivel regional, por lo que será un aporte respecto a nuevas prácticas de educación inclusiva.

Partimos de entender al autismo como un trastorno del neurodesarrollo caracterizado por una tríada de síntomas observables que consiste en la afectación en el desarrollo del lenguaje, conductas estereotipadas asociadas a intereses restringidos y trastorno en la interacción social (Rapin I, 1998).

Según estadísticas publicadas por Centers for Disease Control and Prevention se estima que 1 de cada 68 niños, nacidos a partir de 2002, son diagnosticados con Trastorno del Espectro Autista (TEA). Los niños con TEA pueden tener problemas significativos de tipo social, comunicativo y conductual.

Torras Virgili (2015), señala que tanto el Manual Diagnóstico y Estadístico de Desórdenes Mentales (DSM), que es el sistema de clasificación diagnóstica con mayor relevancia internacional, como la Organización Mundial de la Salud (OMS) consideran los Trastornos de Espectro Autista (TEA) como trastornos generalizados del desarrollo. Éstos se caracterizan por una amplia variedad de expresiones clínicas y conductuales que son el resultado de disfunciones multifactoriales de desarrollo del sistema nervioso central.

La autora, agrega que, pese a que el autismo no puede considerarse como un diagnóstico único, sino como un conjunto de trastornos, y que podemos encontrar grandes diferencias en cuanto a coeficiente intelectual o gravedad de los síntomas, existen una serie de características comunes: problemas de reciprocidad social y de comunicación (verbal y no verbal) y conductas restringidas y repetitivas. Este desorden suele manifestarse durante los tres primeros años de vida, aunque determinados tipos de autismo, como el síndrome de Asperger, al no existir ningún tipo de retraso intelectual ni rasgo físico que ayude a identificarlo puede diagnosticarse mucho más tarde.

Los trastornos del espectro autista no tienen cura, aunque sí tratamiento. Éste se basa en diversos abordajes terapéuticos, los cuales llevado adelante por profesionales idóneos, favorecen una mejoría sustancial.

De ahí que creemos que el uso de las TIC puede adquirir una gran relevancia en la estimulación de los niños y niñas con autismo y Down, en tanto puede contribuir activamente en su desarrollo y bienestar, independencia y el estilo de vida de las personas con diversidad funcional.

Cabe aclarar, que partimos de la postura de que el autismo no puede entenderse únicamente como un trastorno en el que todas las áreas del desarrollo están alteradas o en défi-

cits, de hecho, los niños con autismo suelen presentar habilidades destacables en ciertas áreas que superan a los niños de desarrollo tipo, como la memoria, capacidad para recolectar datos de su interés, reconocimiento icónico, entre otras. Nuestro abordaje, toma como apoyo dichas fortalezas, y procura favorecerlas y estimularlas en lo que respecta al manejo de datos, compresión y/o anticipación de situaciones y conductas.

Entendemos que la propuesta de actividades secuenciales puede propiciar un aprendizaje más avanzado en lo que respecta a hábitos de la vida cotidiana (tales como, alimentación, higiene, rutina diaria), enriquecido por la posibilidad de brindar información de contexto, lo cual permitirá ampliar el bagaje de información y flexibilizar progresivamente conductas o hábitos.

Moreno (2012), sobre el tema agrega, que las NFC tienen múltiples aplicaciones prácticas; entre los usos actuales de dichos dispositivos, se están investigando el uso de las NFC en la asistencia a los discapacitados visuales, en la búsqueda de productos en los supermercados. Así, el aspecto sobre el que se está indagando, es la aplicabilidad de pasar un smartphone por cada producto y que éste lea y brinde oralmente las características del producto a los discapacitados visuales.

En la misma línea, Gómez (2012) propone una innovación al proponer el uso de dos tecnologías complementarias, NFC y lectura de etiquetas de código de barra, para de este modo proporcionar una solución más eficaz para los usuarios –menor tiempo empleado en la búsqueda y obtención del producto y una mayor tasa de éxitos- con discapacidades visuales. En el caso del prototipo que vamos a desarrollar en este proyecto, la precisión no es un requerimiento apreciable.

NFC es una tecnología que tiene presencia desde hace más de una década, y que está adquiriendo importancia de cara al futuro de la tecnología y los dispositivos móviles.

La misma, presenta ventajas respecto a los tradicionales códigos QR debido a que no necesita realizar una fotografía o enfocar a algún lugar, ya que funciona por proximidad.

En este sentido, creemos que el desarrollo de las NFC, está sujeto a la incorporación de las mismas por parte de los fabricantes de teléfonos móviles, en sus dispositivos; Así, cuanto mayor sea el uso, más potencial podrá adquirir. La incorporación a la vida diaria de los usuarios de tecnología parece ser probable en los próximos años, por lo que la propuesta de este proyecto, podrá contribuir a la profundización del estudio del uso de estas aplicaciones de software, en tecnologías móviles.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En la actualidad, la tecnología se convirtió en un soporte esencial para múltiples terapias y ejercicios para la mejora de habilidades cognitivas. Sin embargo, existen ciertas limitaciones en las interfaces de usuario tradicionales entre el humano y el computador. Para intentar mitigar estas limitaciones se intenta hacer uso de interfaces más intuitivas y naturales (interfaces tangibles). Un tipo especial de esta interfaz se logra mediante el uso de tecnología móvil y NFC -Near Field Communication- (Igoe Tom, 2012).

NFC permite a un sistema informático detectar cuando un objeto se encuentra cerca de un dispositivo móvil. El objeto debe contar con una tarjeta NFC y el móvil con un lector NFC. El estándar internacional NFC define una nueva tecnología inalámbrica basada en radiofrecuencia que funciona en un radio de cobertura pequeño.

NFC es la evolución de RFID (Radio Frequency Identification). El RFID puede definirse como aquel sistema que tiene como principal función la identificación de determinados objetos a distancia, utilizando para tal efecto las ondas de radio. Para conseguirlo, esta tecnología proporciona soporte para el almacenamiento y la recuperación remota de datos, en etiquetas o tarjetas RFID que contienen la información necesaria para el reconocimiento.

Considerando las Etiquetas RFID, se puede decir que se componen principalmente por

una antena, un transductor de radio y un material encapsulado o chip (Campa Ruiz, 2011). El Near Field Communication brinda nuevas funciones a la tecnología RFID. Esto se debe a la combinación de un lector y una etiqueta RFID en un mismo equipo NFC. De esta manera se facilita la comunicación en ambas direcciones entre dos dispositivos y se quiebra la separación funcional de los esquemas RFID, es decir, el lector por un lado y las etiquetas por el otro (Chavarría, 2011).

Realizando una comparación entre NFC y RFID, encontramos que la primera posee la capacidad de cómputo necesaria para ejecutar operaciones, hecho que hace fácil su integración en dispositivos como teléfonos móviles (Tolsada Bris, 2012). Además, NFC provee una comunicación peer to peer, medio que permite intercambiar información entre dos dispositivos (Campa Ruiz, 2011).

Otra cuestión que tiene NFC es que admite rangos pequeños de comunicación, y, por ende, posee una seguridad inherente. Suceso que lo hace preferible para cuestiones de comunicación que requieren seguridad como los medios de pago o intercambio de información personal. Además, NFC no puede ser activado de forma remota, involuntariamente o por accidente. El teléfono obliga a que deba existir un acercamiento entre dispositivos antes de iniciar una comunicación (Campa Ruiz, 2011). En cuanto al sistema de corto alcance NFC se compone de dos elementos: un iniciador y un objetivo, en donde cualquier dispositivo con NFC puede adoptar las funciones o el comportamiento de una de estas partes. El NFC puede ser instalado en cámaras fotográficas, reproductores, televisores, teléfonos celulares y hasta en controles remotos (Chavira, et al, s/f).

Técnicamente, el NFC proporciona la comunicación inalámbrica de corto alcance mediante un campo magnético que permite el intercambio de datos, y opera en un ambiente en donde los dispositivos se encuentran separados una distancia de 20 cm. como máximo. El sistema se maneja en la banda de frecuencia no licenciada de $f_c = 13.56$ MHz, y un ancho de banda que oscila 7 KHz a cada lado

de fc. Las comunicaciones pueden ser half o full duplex. Se utiliza el esquema de modulación Amplitud Shift Keying (ASK) y la codificación Manchester. Finalmente, se disponen de tres velocidades de transmisión 106, 212 y 404 kbps que son fijadas por el dispositivo que inicia la conexión (Bueno et al, 2011).

3. RESULTADOS Y OBJETIVOS

El proyecto tiene estipulado 2 años de duración y cuenta con los siguientes objetivos: Desarrollar e implementar prototipos empleando la tecnología NFC (Near Field Communication), como complemento educativo, cognitivo y social en niños con diversidad funcional.

Generar conocimientos sobre la temática niños con diversidad funcional y los actores intervinientes y las metodologías vigentes para atender los mismos.

Indagar, clasificar y planificar las actividades ocio-educativas como instrumentos para el desarrollo de prototipo que abarque el conocimiento del medio físico, cognitivo y fomenten la inclusión.

Desarrollar actividades de modalidad tripartita (niño tutor dispositivo Kinect), que hagan uso de interfaces naturales

Implementar prototipos basados en modelos lúdicos, dinámicos y variados.

Diseñar aplicaciones de software, con tecnología NFC para niños con diversidad funcional en dispositivos móviles.

Mejorar los prototipos a partir de la implementación de las aplicaciones del software a fin de perfeccionar de forma continua las actividades de modalidad tripartita.

Formar e iniciar a los estudiantes avanzados de las carreras de UNPaz en investigación y transferencia.

Fortalecer la articulación entre la UNPaz y la comunidad.

Explorar las posibilidades de aplicación de NFC en otras instituciones educativas especiales y en trastornos alternativos.

4. FORMACIÓN DE RECURSOS HUMANOS

Director del Proyecto:

Mg. Ariel Langer. Director de Vinculación Tecnológica – Secretaría de Ciencia y Tecnología (SECyT) – Universidad Nacional de José Clemente Paz (UNPAZ).

Co-Director del Proyecto:

Lic. Víctor Contreras. Maestrando en Ingeniería de Tecnología Informática UAI. Director de Sistemas Informáticos- Dirección General de Informática y Tecnología de las Comunicaciones– UNPAZ.

Investigadores:

Lic. García, Mariana: Posgrado “Necesidades Educativas y Prácticas Inclusivas en Trastornos del Espectro Autista. FLACSO. Posgrado Autismo y Neurodesarrollo Universidad Favaloro. Argentina. Docente de Psicología Evolutiva, Licenciatura en Instrumentación Quirúrgica – UNPAZ.

Mg. Viviana Moreno. Doctoranda en Ciencias Sociales UNGS-IDES. Integrante del área de Vinculación Tecnológica-Secretaría de Ciencia y Tecnología – UNPAZ.

Ing. Daniel Fernandez. Maestrando en Ingeniería del Software UNLP. Docente universitario de UAI.

Campa Ruiz, A (2011) Desarrollo de una aplicación de pago a través de la tecnología NFC. Universidad Carlos III de Madrid.

Chavira, S. W. Nava, R. Hervás, J. Bravo, Carlos Sánchez, (s/f) Localización e Identificación: Una combinación RFID-NFC.

Chavarría, D (2011) TECNOLOGÍA DE COMUNICACION DE CAMPO CERCANO (NFC) Y SUS APLICACIONES. Universidad de Costa Rica.

Gómez (2012) Sistema Integral de guiado y localización de objetivos en interiores para personas con discapacidades visuales basado en dispositivos móviles libres. Módulo de localización e identificación de productos.

Igoe Tom, (2012) Getting Started with RFID. Primera Edición - Marzo 2012. O’Reilly

Moreno (2012), Aplicaciones prácticas de NFC. En, Revista de Investigación 3Ciencias. Área de Innovación y Desarrollo. SL.

RapinI, K. R. (1998). Neurology of autism. Ann Neurol, 43:7-14.

Tolsada Bris, L (2012) Desarrollo de una aplicación de transferencia de ficheros basada en NFC y Bluetooth. Universidad Carlos III de Madrid.

Torras Virgili (2015) “Estrategias Educativas para niños con Autismo”, En, Trastornos del Espectro Autista. Universidad Internacional de Valencia.

5. REFERENCIAS

Bueno Delgado, M, Pavón Mariño, P y de Gea García, A (2011) La tecnología NFC y sus aplicaciones en un entorno universitario. Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Cartagena.

Aproximando las Ciencias de la Computación a la Escuela Secundaria

Sonia Sommer² María Eda Cornejo² Jorge Rodríguez¹ Laura Cecchi¹

medacornejo@gmail.com, soniasommer@yahoo.com, {j.rodri, lcecchi}@fi.uncoma.edu.ar

¹ *Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

² *Consejo Provincial de Educación*
MINISTERIO DE EDUCACIÓN DE LA PROVINCIA DE NEUQUÉN

Resumen

La enseñanza de las Ciencias de la Computación en el ámbito de la educación obligatoria ha logrado altos niveles de consenso en los últimos años.

En Argentina convergen varias iniciativas promovidas por las Universidades, Program.ar y el Programa Nacional Conectar Igualdad que buscan mejorar la aproximación de las Ciencias de la Computación a la escuela secundaria.

La incorporación de forma sostenible y rigurosa de la enseñanza de las Ciencias de la Computación en la educación secundaria es aún un proceso en desarrollo en la mayoría de los países.

En este trabajo se presenta una Línea de Investigación que busca avanzar en la definición de modelos y enfoques metodológicos que impacten positivamente en los procesos de aproximación de las Ciencias de la Computación a la escuela secundaria.

Se presta especial atención a la innovación a partir de la integración de recursos didácticos disciplinares como robótica educativa y programación tangible.

Palabras Clave: ENSEÑANZA DE LAS CIENCIAS DE LA COMPUTACIÓN, RECURSOS DIDÁCTICOS EN CIENCIAS DE LA COMPUTACIÓN, ESCUELA MEDIA .

Contexto

Esta propuesta se ubica en el contexto de las iniciativas promovidas por el Grupo de Investigación en Lenguajes e Inteligencia Artificial de la Facultad de Informática y del Convenio Marco de Colaboración firmado durante 2016 entre la Facultad de Informática y el Ministerio de Educación de la Provincia del Neuquén.

El convenio persigue el objetivo de contribuir recíprocamente al desarrollo de actividades de investigación, formación de recursos humanos y promoción de la enseñanza de las Ciencias de la Computación durante la escolaridad obligatoria.

Este trabajo se desarrolla en el ámbito del proyecto de investigación *Agentes Inteligentes. Modelos Formales y Aplicaciones para la Educación (04/F015)* que está financiado por la Universidad Nacional del Comahue a través de la Secretaría de Ciencia y Técnica y por el Consejo Provincial de Educación en el contexto del Convenio Marco de Colaboración. El proyecto tiene prevista una duración de cuatro años a partir de enero del 2017.

Las actividades concretadas en el ámbito de la investigación se plantean como articuladas a un conjunto de proyectos de extensión [4, 15], en ejecución durante 2017, con intención de construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el campo de la praxis.

1 . Introducción

Introducir conceptos propios de la Ciencia de la Computación en las propuestas curriculares para las escuelas secundarias es una necesidad reconocida y con gran consenso entre gobiernos, organizaciones no gubernamentales, grupos de investigación y docentes en el campo de la enseñanza de la computación, tanto en países desarrollados como en desarrollo [3, 7, 16, 18].

La comprensión de conceptos fundamentales acerca de arquitectura y redes de computadoras, algoritmos, ingeniería de software, modelado de datos e inteligencia artificial, entre otros, resulta necesaria para mejorar las posibilidades de entender e intervenir el mundo que rodea a los estudiantes [2, 5, 16].

En línea con lo antedicho, en Argentina, a partir de 2013 convergen iniciativas como Program.ar y el Programa Nacional Conectar Igualdad que proponen la implementación de estrategias orientadas a aproximar las Ciencias de la Computación a la escuela secundaria [3].

En este marco el Consejo Federal de Educación declaró de importancia estratégica a la enseñanza y el aprendizaje de la Programación en todas las escuelas durante la escolaridad obligatoria, en agosto de 2015. [1].

Varios estudios muestran que, más allá de los acuerdos construidos y esfuerzos desarrollados, la incorporación de forma sostenible y rigurosa de la enseñanza de las Ciencias de la Computación en la educación secundaria es aún un proceso en desarrollo en la mayoría de los países [6, 9, 10].

En la provincia de Neuquén, la computación está presente en la mayoría de los planes de estudios vigentes para la escuela secundaria con una asignación horaria similar a otras disciplinas académicas como matemática, geografía o historia. Sin embargo, la diversidad de contenidos que se abordan es muy amplia y generalmente relacionados a la

alfabetización digital y al desarrollo de habilidades vinculadas al dominio de herramientas de oficina.

A partir del 2005, la Facultad de Informática de la Universidad Nacional del Comahue establece vínculos de colaboración con varias escuelas del nivel medio de la región con la intención de promover la inclusión progresiva y sostenida de contenidos relacionados a las Ciencias de la Computación en las propuestas de enseñanza [17, 14].

El contexto descripto evidencia la necesidad de desarrollar líneas de investigación, específicas al campo de la Educación en Ciencias de la Computación, tendientes a la producción de constructos teóricos, enfoques metodológicos y modelos de abordaje, que posibiliten la aproximación de la disciplina a la escuela secundaria de forma rigurosa y continua.

En este trabajo se presenta una Línea de Investigación que busca identificar y formalizar un conjunto de argumentaciones en torno a la necesidad de incorporar las Ciencias de la Computación como disciplina escolar troncal durante la escolaridad secundaria, contribuir al desarrollo de frameworks que orienten la construcción de diseños curriculares que propongan un abordaje transversal sobre las áreas de conocimiento del campo disciplinar y la definición de enfoques metodológicos que faciliten la concreción de instancias de formación continua docente que impacten positivamente en los procesos de aproximación de las Ciencias de la Computación al sistema educativo.

2 . Líneas de investigación y desarrollo

La inclusión de contenidos relacionados a las Ciencias de la Computación en las propuestas de enseñanza para la escuela secundaria plantea la necesidad de desarrollar líneas de investigación dedicadas específicamente al estudio de la temática.

Las líneas que se abordan en relación a este tema, en el contexto del proyecto, refieren a la construcción de un grupo de argumentos que justifiquen su inclusión, investigaciones que

contribuyan a definir una didáctica específica para la disciplina, a identificar recursos didáctico disciplinares adecuados y diseñar estrategias que promuevan y acompañen los procesos. Las líneas son:

- Enseñanza de las Ciencias de la Computación como una disciplina escolar
El conocimiento de los contenidos de la disciplina, ayuda a comprender mejor el mundo que nos rodea; contribuye a la democratización del conocimiento y colabora en los procesos de construcción de vocaciones. Conocer estos contenidos ofrece mejores oportunidades al momento de elegir sus futuros estudios. Por otro lado, por ser considerada disciplina STEM, la formación en este campo resulta fundamental para el desarrollo económico, productivo y social de los países.

Esta línea se centra en el estudio de los alcances, metas y objetivos formativos de las Ciencias de la Computación como una disciplina académica en la educación secundaria que logren articular una argumentación consistente para orientar procesos de diseño curricular. Considerar diferentes razones para enseñar la computación conduce a diferentes opciones en relación a los objetivos de aprendizaje [11].

- Contenidos de Ciencias de la Computación como disciplina
Los sujetos interactúan cotidianamente con un conjunto de dispositivos computacionales que median su relación con el mundo y con otros sujetos. La comprensión de conceptos y prácticas fundamentales de la disciplina, constituye un cuerpo de contenidos a abordar desde esta área curricular.

Es objeto de estudio de esta línea, el diseño de modelos conceptuales que consideren a la computación como

materia troncal durante la escolaridad obligatoria, con estructura y contenidos concretos orientados a plantear un recorrido amplio por las áreas de conocimiento de este campo disciplinar y con una carga horaria equivalente a las restantes materias troncales. Otros aspectos a considerar en este ámbito son las formas en que el conocimiento disciplinar se transforma en conocimiento didáctico disciplinar; cuáles son los recursos que hacen accesible a los estudiantes este tipo de saberes, por ejemplo la programación tangible, programación por bloques, robótica educativa y construcción de aplicaciones móviles; y cuáles son las prácticas para enseñar la disciplina, por ejemplo el par de programación, el aprendizaje colaborativo y el aprendizaje basado en proyectos [12, 13].

- Definición de Enfoques Metodológicos para el abordaje en la escuela
Más allá de la importancia de establecer los contenidos como disciplina escolar, es necesario pensar cómo estos puedan hacerse presente en las aulas. Es decir, investigar acerca de las formas efectivas para que los docentes desarrollen habilidades didáctico disciplinares para integrar satisfactoriamente este tipo de saber a las propuestas de enseñanza.

La propuesta se centra en la definición de enfoques metodológicos para la formación docente continua, destinada a profesores de estos espacios curriculares, estructurados en tres tramos o planos: el primero relativo a la formación disciplinar y didáctico disciplinar, el segundo a la puesta en práctica y por último la evaluación de la propuesta.

- Formación disciplinar y didáctico-disciplinar: en este tramo el acceso al conocimiento se produce mediante la interacción y el intercambio entre colegas, constituyéndose el grupo como facilitador de los aprendizajes

y los capacitadores como enseñantes, coordinadores, guías y acompañantes de los procesos y responsables de las situaciones de enseñanza.

- Prácticas asistidas: el equipo pedagógico acompaña a los docentes en la preparación y ejecución de talleres para sus estudiantes. Se construyen recursos para ser utilizados como material de apoyo didáctico durante el desarrollo del taller. Los talleres se ubican en el marco del framework "Computing for the Social Good" proponiendo la construcción de productos para el bien social [8].
- Taller de evaluación de experiencias: se desarrolla un taller para socializar las experiencias, debatir estrategias didácticas y generar materiales para un repositorio web que sirva como modelo a experiencias futuras.

Estas líneas no se presentan como aisladas, se considera que la investigación en estas áreas debe plantearse como integrada y estudiarse en forma conjunta para avanzar en la elaboración de modelos teórico conceptuales consistentes.

3. Actividades de Extensión relacionadas al Proyecto

El proyecto busca construir y ampliar conocimiento a partir de la revisión y análisis de resultados desarrollados en el ámbito de la Extensión Universitaria en relación a la aplicación de agentes al contexto educativo.

Las iniciativas de referencia se expresan en los siguientes Proyectos de Extensión:

- Vamos a la Escuela. Acercando las Ciencias de la Computación a la Escuela Media, la propuesta metodológica se estructura a partir de la concreción de talleres breves que se organizan en tres tramos: formación docente, taller con estudiantes secundarios, evaluación de

experiencias con docentes. Las temáticas a desarrollar son: Aprendizaje de conceptos sobre programación a partir de la construcción de Aplicaciones Móviles y la Robótica Educativa y sobre Arquitectura de computadora y sistemas embebidos a partir de la construcción de Robots Educativos [15]

- Agentes Robots. Divulgando Computación en la Escuela Media, en el ámbito de la robótica educativa se propone a los estudiantes del último año de la Escuela Media, la construcción colectiva de un producto de software que permite a un robot resolver un problema específico, por ejemplo "ejecutar un penal" [4].

Estos proyectos se implementan en colaboración con diez escuelas secundarias de la región. En este contexto se desarrollan actividades de formación docente en el campo de la enseñanza de las Ciencias de la Computación y de formación a estudiantes secundarios en conceptos fundamentales de la disciplina. Por otro lado se busca contribuir al mejoramiento de las propuestas de enseñanza a partir de la incorporación progresiva y rigurosa de contenidos curriculares.

4. Resultados esperados

En el contexto de trabajos realizados [14, 17] y de las actividades proyectadas se busca construir conclusiones que se ubiquen como referencias teóricas para los procesos de inclusión de contenidos de las Ciencias de la Computación a las propuestas de enseñanza para la escuela secundaria. Se espera:

- Avanzar en la identificación y formalización de una colección de argumentaciones.
- Diseñar modelos conceptuales que favorezcan la inclusión de contenidos disciplinares en el diseño curricular.
- Describir enfoques metodológicos destinados al desarrollo de dispositivos para la formación docente continua y situada que faciliten una incorporación estable y rigurosa.
- Definir criterios que orienten la utilización adecuada de recursos didácticos, como las

aplicaciones móviles, la robótica educativa o la programación tangible, entre otros, en la enseñanza de las áreas de conocimiento de las Ciencias de la Computación.

5. Formación de Recursos Humanos

Se aspira a la formación como investigadores de los miembros más recientes del grupo. En este sentido se buscará fortalecer la vinculación con otros grupos de investigación e instituciones abocados a la temática en el ámbito nacional e internacional.

Asimismo, se espera la iniciación del posgrado de uno de los autores de este trabajo.

Referencias

- [1] Consejo Federal de Educación- Resolución 263/15. Último acceso Marzo 2017, website <http://www.me.gov.ar/consejo/resoluciones/res15/263-15.pdf>.
- [2] M. Bonello and H. Czemerinski. Programar: una propuesta para incorporar ciencias de la computación a la escuela argentina. 2015.
- [3] M. Borchardt and I. Roggi. Ciencias de la computación en los sistemas educativos de américa latina. 2017.
- [4] L. Cecchi and G. Grosso. *Proyectede Extensión Agentes Robots: Divulgando Computación en la Escuela Media*. 2017, avalado por Resolución FaI 088/16.
- [5] ACM, Code.org, CSTA, Cyber Innovation Center, and National Math and Science Initiative. K–12 Computer Science Framework. 2016.
- [6] T. Crick and F. Moller. A national engagement model for developing computer science education in wales. 2016.
- [7] S. Furber. *Shut down or restart? The way forward for computing in UK schools*. The Royal Society Education Section, 2012.
- [8] M. Goldweber, J. Barr, T. Clear, R. Davoli, S. Mann, E. Patitsas, and S. Portnoff. A framework for enhancing the social good in computing education: a values approach. *A CM Inroads*, 4(1):5879, 2013.
- [9] Google and Gallup. *Searching for computer science: Access and barriers in U.S. K-12 education*. 2015.
- [10] Google and Gallup. Trends in the state of computer science in u.s. k-12 schools. 2016.
- [11] M. Guzdial. Learner-centered design of computing education: Research on computing for everyone. *Synthesis Lectures on Human-Centered Informatics*, 8(6):1-165, 2015.
- [12] M. C. Herring, M. J. Koehler, and P. Mishra. Handbook of technological pedagogical content knowledge (TPACK) for educators. Routledge, 2016.
- [13] M. Oliver. Handbook of technological pedagogical content knowledge (tpck) for educators, 2011.
- [14] J. Rodríguez, G. Grosso, R. Zurita, and L. Cecchi. Intervención de la facultad de informática en la enseñanza de ciencias de la computación en la escuela media basada en robótica educativa. In XI Congreso de Tecnología en Educación y Educación en Tecnología, 2016.
- [15] J. Rodríguez and R. Zurita. Proyecto de Extensión Vamos a la Escuela: Acercando las Ciencias de la Computación a la Escuela Media. 2017, avalado por Resolución FaI 087/16.
- [16] M. Smith. Computer science for all. The White House, 2016.
- [17] S. Sommer, J. E. Sznek, and J. Rodríguez. Divulgando temáticas computacionales-internet segura. In X Congreso Tecnología en Educación y Educación en Tecnología, 2015.
- [18] C. Wilson, L. A. Sudol, C. Stephenson, and M. Stehlik. Running on Empty: The Failure to Teach K-12 Computer Science in the Digital Age. ACM and The Computer Science Teachers Association, 2010.

Articulación de Contenidos Curriculares de las Disciplinas Sistemas Operativos y Sistemas de Tiempo Real con Sistemas Embebidos en la Carrera Ingeniería en Informática

Juan Pablo Moreno¹; Carola Victoria Flores¹; Marcos Aranda¹; Enrique Miranda²; Gabriel Vilallonga^{1,2}

¹Departamento Sistemas de la Facultad de Tecnología y Ciencias Aplicadas de la UNCa Maximio Victoria N° 55 - C.P: 4700 - San Fernando del Valle de Catamarca
Teléfono: 03834- 435112 – int 168

²Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis. Ejercito de los Andes 950. San Luis
Tel: +54 (0266) 4520300. Int. 2127

jomoreno@tecno.unca.edu.ar, carolaflores@tecno.unca.edu.ar, mdaranda@tecno.unca.edu.ar, eamiranda@unsl.edu.ar, gvilallo@tecno.unca.edu.ar/unsl.edu.

Resumen

La formación de profesionales de las carreras de ingeniería en informática demanda conocimientos correctamente integrados por áreas afines de forma articulada para que los alumnos logren aprehender los conceptos por medio de prácticas interdisciplinarias.

El desarrollo de software para sistemas de tiempo real (STR) es realizado por ingenieros de software especializados en el área. Si bien en todo desarrollo de sistemas el tiempo es una variable importante, existen aquellos sistemas donde el no cumplimiento de una meta temporal puede provocar fallas catastróficas. Esto promueve el uso de técnicas de ingeniería de desarrollo para estos tipos de sistemas.

La gestión de software para STR involucra el conocimiento de diversas áreas, entre ellas los sistemas operativos (SO). Una de las principales áreas de estudio son los SO de tiempo real (SOTR), y su base teórica, que permite garantizar la correcta gestión de las tareas. Por ser el SO la tarea de más alta prioridad, en un STR, es que su correcta elección, por sus características, debe ser estudiada con minuciosidad.

Con una visión ingenieril, el desarrollo de sistemas con restricciones temporales debe involucrar desde la adquisición de las

especificaciones hasta la instalación del software en su ambiente de ejecución. El software es embebido en una placa electrónica donde el mismo correrá. La placa electrónica cuenta con diversos dispositivos y componentes electrónicos, acorde a los requerimientos del STR.

Desde el año 2011 se están llevando a cabo distintos talleres y escuelas de sistemas embebidos donde se ha promovido la actualización y capacitación de los docentes que se encuentran trabajando en el área de los sistemas embebidos (SE), incluso se llegó a desarrollar la Computadora Industrial Abierta Argentina (CIAA) y su versión educativa EDU-CIAA, la cual es una placa electrónica que permite la instalación de distintos SOTR y luego la implantación de software que desarrollará tareas de tiempo real.

Las distintas Universidades Nacionales que aportan al desarrollo de la CIAA poseen un stock de las mismas y van incorporando el manejo de la misma al dictado de sus cátedras. Esto posibilita contar con el material adecuado para realizar prácticas sobre las EDU-CIAA tanto para la instalación de SOs, como así también la implantación de software de tiempo real.

En la presente línea de investigación se propone confeccionar un marco de trabajo para la articulación de los contenidos

curriculares de las disciplinas Arquitectura de Computadores, Sistemas Operativos y Sistemas de Tiempo Real, de la carrera ingeniería en Informática, que por medio de la utilización de las EDU-CIAA permitan a los alumnos en primer lugar instalar un SO adecuado, y luego sobre este embeber el software desarrollado y validado por técnicas formales de la ingeniería de software. Esta experiencia permitirá al alumno la integración efectiva de conocimientos de STR.

Palabras claves: Ingeniería de Software (IS) Sistemas de Tiempo Real (STR), Sistemas Operativos (SO), Sistemas Embebidos (SE).

Contexto

El proyecto de investigación *“Fortalecimiento de la Calidad y la Productividad en Ingeniería de Software”*, posee una línea de investigación dedicada al análisis de implementación de buenas enseñanzas de la ingeniería que tiene como objetivo mejorar la formación de ingenieros de software.

El mencionado proyecto es evaluado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca y se desarrolla y ejecuta en ámbitos del Departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCa y en el Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis donde los integrantes se desempeñan como docentes de la carrera de Ingeniería en Informática (UNCa), e Ingeniería en Informática y Licenciatura en Ciencias de la Computación (UNSL) respectivamente.

Esta línea de investigación, inserta en el proyecto, establece como prioridad la formación científico-técnico de los integrantes con la premisa de proceder a la transferencia de resultados a los medios informáticos.

Introducción

Los STR se encuentran en un gran número de elementos de la vida cotidiana. Desde los

sistemas de control de lavavajillas, sistema de control de reproductores de DVD, CD, etc hasta los sistemas de antibloqueo de los frenos de los automóviles, pasando por el control de tracción y el climatizador del aire. Los controladores realizan el análisis del medio en el cual el controlador está embebido y activan las operaciones pertinentes en fracciones de segundo. También se encuentran en lugares más sensibles como los sistemas de navegación y posicionamiento de aviones, y en lugares ultrasensibles como las centrales termonucleares.

En los últimos años tanto el hardware con el software están siendo embebidos en la mayoría de los STR para monitorear y controlar sus operaciones. Sin lugar a dudas la corrección de estos sistemas es de suma importancia debido que en muchos de estos sistemas la falla parcial o total puede acarrear consecuencias incómodas en el mejor de los casos, hasta catastróficas en otro extremo.

El desarrollo de estos sistemas implica el conocimiento de casi todas las teorías generales de los sistemas de computación, desde ingeniería de software hasta la teoría de SOTR, pasando por la arquitectura de la computadora hasta el conocimiento de lenguajes de programación con habilidades para el manejo de conceptos de tiempo real.

El estudio profundo de los STR implica un amplio conocimiento de varias áreas temáticas de las ciencias de la computación.

Los SO brinda una interfaz amigable entre el hardware y el usuario, permitiendo al desarrollador de software abstraerse del manejo mecánico/electrónico del dispositivo o componente electrónico que desea utilizar y centrarse solamente en la lógica del problema que desea solucionar. Para lograr este grado de abstracción, los SO ejecutan diversos procesos en paralelo, los cuales son administrados por un planificador, el cual se encarga de asignar tiempo de procesamiento a cada uno de ellos y de mediar entre procesos cuando dos o más procesos compiten por un mismo recurso.

Los SOTR poseen características especiales que los diferencian de los SO tradicionales, estas características están ligadas a los tiempos de respuesta, asignación de

prioridades y tolerancia a fallos. El tiempo de respuesta de

un SOTR debe estar dentro de valores preestablecidos, tanto para responder a interrupciones que puedan presentar, como así también a la comunicación interprocesos. La asignación de prioridades permite al usuario indicar al planificador del SOTR cuáles son los procesos a los cuales se debe asignar una prioridad mayor por encima de otros o viceversa. Por último, la tolerancia a fallos donde el SOTR intentará mitigar o corregir las consecuencias que pueda ocasionar una falla dentro del sistema, para ello es necesario conocer cada uno de los distintos escenarios que puedan presentarse a lo largo de su ciclo de vida.

En un SO tradicional si algo falla o un proceso demora más de lo normal no genera más que un inconveniente o la necesidad de reiniciar el sistema, en contraste con un SOTR la ocurrencia de este tipo de problemas puede llevar a que la vida de una o varias personas corra riesgo.

Los SE son una combinación de hardware y software que son desarrollados para satisfacer un requerimiento particular, es decir, poseen una funcionalidad determinada, diferenciándolos de los sistemas de propósito general. Una PC puede ser utilizada para procesar texto, diseño gráfico, programación, edición de sonido, etc., lo que la define como un sistema de propósito general, el hardware que posee está orientado para satisfacer todas esas necesidades. Al contrario, un SE solamente va a tener el hardware necesario para satisfacer los requerimientos específicos que posee, tratando de abaratar los costos de su construcción y además logrando un grado de determinismo beneficioso para el sistema.

Tanto en las áreas curriculares de la carrera de electrónica como la de informática y sistemas es imposible cubrir todos los aspectos teórico-prácticos de tiempo real. El aspecto teórico es muy rico pero no menos complejo desde el punto de vista de la enseñanza. La verificación y testeado de estos sistemas, previo a su embebido en el medio de ejecución, involucra el estudio que va desde los modelos de sistemas hasta distintos tipos de lógicas,

temporales y no temporales, necesarios para producir sistemas formalmente verificados.

Para lograr una correcta formación el alumno debe aprehender los conocimientos teórico-prácticos de la producción de software para STR. La experiencia de implantación del software en su medio de ejecución es algo que todo alumno de la carrera ingeniería en informática debe ser adquirida, especialmente en el área de STR.

Para lograr el objetivo es primordial que el alumno desarrolle prácticas de implantación de software necesarias para la efectiva adquisición e integración de los conocimientos teóricos.

La demanda de recursos humanos altamente calificados para el correcto desempeño en proyectos de desarrollo de software, donde las especificaciones temporales son una variable fundamental, hace que esta área de la ingeniería de software, tome real importancia y donde un ingeniero informático debe poseer las herramientas necesarias para un correcto desempeño en grupos de trabajo abocados a estos tipos de proyectos.

La formación de los alumnos de Ingeniería en Informática por medio de las disciplinas SO y STR, junto con la disciplina de Software Embebido permiten el desarrollo de los conocimientos teóricos y la habilidad en el manejo de herramienta de tiempo real por parte de los futuros profesionales.

Esta línea de investigación permite abrir una perspectiva multidisciplinar que tiene como efecto sinérgico el trabajo con sistemas reales, que permiten la aplicación de las herramientas informáticas utilizadas en el desarrollo de sistemas complejos.

Estas actividades exigen de la realización de cursos acerca de la temática, como así también el estudio de material bibliográfico y de publicaciones científicas.

Es de hacer notar que estos grupos de trabajos están coordinados por docentes investigadores de las distintas universidades.

Resultados y Objetivos

El objetivo principal de esta línea de investigación es contribuir de manera efectiva y mensurable en la concreción de aportes en

la formación de los alumnos de la carrera ingeniería en informática por medio de la adquisición de habilidades en el desarrollo de STR con especial énfasis en la implantación en su medio de ejecución.

Esto conlleva la revisión, o nuevas propuestas, técnicas, y metodologías de enseñanza que asistan al desarrollo de software de tiempo real embebido.

El efecto deseado, también, es incidir significativamente en las actividades académicas de formación de recursos humanos, y la transferencia al medio.

Estos objetivos están siendo alcanzado gracias al trabajo conjunto entre los equipo de la UNSL-UNCA.

Formación de Recursos Humanos

Esta línea de investigación está en una etapa intermedia, donde integrantes del proyecto son docentes de las disciplinas SO y STR con formación de posgrado en Ingeniería de Software.

Se ha procedido a la incorporación de alumnos de los últimos años con la finalidad de incluirlos en actividades de investigación y desarrollo en las áreas específicas del proyecto

El proyecto prevé un programa de capacitación y formación de recursos humanos, que contempla las siguientes actividades:

- Dirección de tesinas de grado de la carrera de Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA integrados al proyecto, como así también en la UNSL.
- Participación de los integrantes del proyecto en cursos de actualización y posgrado en el área de estudio.

Para garantizar la capacitación y actualización del equipo de investigación, así como la difusión de los avances y resultados logrados, se propuso la participación en eventos nacionales e internacionales de la especialidad, como congresos, simposios, seminarios y cursos.

Bibliografía

[1] Real Time Systems Design and Analysis. Laplante Phillips. IEEE PRESS 3o edn. 2004. ISBN-10: 0471228559, ISBN-13: 978-0471228554

[2] Sistemas de tiempo real y lenguajes programación. Alam Burns & Andy Welings. Pearson Educación, 2003. ISBN: 8478290583, 9788478290581

[3] Real-Time Systems: Scheduling, Analysis and Verification. Albert M. K. Cheng. John Wiley & Sons, Inc., Hoboken, New Jersey. 2002. ISBN-10: 0471184063, ISBN-13: 978-0471184065.

[4] Sistemas Operativos Modernos - 3ra edición. Andrew Tanenbaum, Pearson Educación. 2009. ISBN: 9780136006633.

[5] Fundamentos de sistemas operativos - 7ma edición. Abraham Silberschatz, Peter Baer Galvin & Greg Gagne, McGraw-Hill. 2006. ISBN: 9788448146412.

[6] Programacion de Sistemas Embebidos en C. Gustavo Galeano, Alfaomega. 2009. ISBN- 10: 9586827704 ISBN-13: 978-9586827706.

[7] Diseño y Programación de Sistemas Embebidos con el Núcleo Microblaze: Fundamentos, Conceptos y Método del Cómputo de Propósito Específico. Olmo Alonso Moreno Franco, Editorial Académica Española. 2012. ISBN-10: 3848450593 ISBN- 13:978-3848450596.

Avances en Torno a la Formación en Accesibilidad Web

Sonia I. Mariño, María V. Godoy, Pedro L. Alfonzo

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura.

9 de Julio 1449. CP: 3400. Corrientes. Argentina.

Universidad Nacional del Nordeste.

simarinio@yahoo.com, mvgg2001@yahoo.com, plalfonzo@hotmail.com

Resumen

Se exponen los desarrollos concretados en torno a la AW, como un aspecto de la calidad del software centrada en actividades de I+D+i. Particularmente los conocimientos adquiridos y los desarrollos tecnológicos favorece la formación de recursos humanos en la temática y su desarrollo en otras áreas académicas.

Palabras clave: Accesibilidad Web, métodos y herramientas, formación de recursos humanos, transferencia de conocimientos.

Contexto

En el marco de proyectos de I+D+i acreditado por la Secretaría General de Ciencia y Técnica (UNNE), se indaga y aplican métodos y herramientas para evaluar la accesibilidad según pautas WCAG como un aspecto de la calidad de la Ingeniería del Software.

Introducción

La Accesibilidad Web (AW) referencia el acceso universal a este servicio de Internet, independientemente del tipo de hardware, software, infraestructura de red, idioma, cultura, localización geográfica y capacidades de los usuarios [1], [2].

En el marco del proyecto de "Sistemas de Información y TIC: métodos y herramientas" y el TI en los Sistemas de Información: modelos, métodos y herramienta (continuación del primero) se avanza en la indagación de métodos y herramientas y su aplicación con miras a aportar a la inclusión de los ciudadanos en el uso de herramientas informáticas en este siglo, y contribuir que las Tecnologías de la Información y Comunicación (TIC) estén al servicio de la comunidad para mejorar su calidad de vida.

Es Responsabilidad Social de los actores de la Universidad contribuir a la consolidación de la sociedad de la Información. Así involucrar desde el grado a los futuros profesionales, se considera de relevancia.

Los avances sobre el tema de responsabilidad social universitaria proponen una meta contextual de análisis del espacio para la creación, el desarrollo, la transferencia y la innovación del conocimiento. Orientar en los aspectos mencionados para lograr impactos particulares en la sociedad, permite crear y establecer la relación Universidad- Empresa-Estado [3].

Desde la perspectiva de la Ingeniería del Software (IS) es notable determinar la calidad de los productos software en proceso de elaboración, siendo la Accesibilidad Web una medida aplicable desde etapas tempranas de desarrollo y tratada como un requerimiento no funcional [4].

Otro aspecto de relevancia en la IS es el mantenimiento de sus productos para lograr mayor funcionalidad en la comunidad de usuarios.

Para el equipo es fundamental el estudio teórico y la definición de procesos orientados a aplicar la AW a lo largo de desarrollos tecnológicos.

Cabe aclarar que la iniciativa de estudio de AW, desarrollada por este equipo de trabajo de la UNNE, coincide con el tratado en otras universidades ([5] [6] [7] [8] [9] [10]).

Líneas de Investigación, Desarrollo e Innovación

Con el propósito de extender conceptos teóricos y desarrollos empíricos en la temática se fomenta la construcción de sistemas informáticos que respondan a estándares

internacionales de accesibilidad como los establecidos por la WCAG 2.0 se:

- relevan, estudian, analizan y aplican métodos para el tratamiento de la AW basados en las pautas WCAG 2.0 [11].
- seleccionan, analizan, estudian y aplican herramientas informáticas utilizadas para la medición de accesibilidad, entre las que se mencionan: Brailersurf [12], Lynx Browser [13], TAW [14], HERA [15], EXAMINATOR [16], TextAloud [17], NVDA [18], WEBBIE [19], entre otras CheckMyColours [20] WCAG Contrast checker [21].
- utilizan diversos dispositivos móviles, teléfonos celulares, tablets que responden a diversas configuraciones para evaluar el nivel de accesibilidad de software basado en la web desde los mismos.
- aplican las pautas WCAG 2.0 [11], en etapas tempranas del desarrollo de sitios web, particularmente se desarrollo el sitio del Congreso Educar Para Incluir, 2016 [22].

Resultados y Objetivos

La RedUNCI [23] ha incluido la Accesibilidad como un descriptor recomendado. El abordaje de la AW en el proyecto y su transferencia en espacios de educación superior -asignaturas Trabajo Final de Aplicación y Proyecto Final de Carrera- y la formación de recursos humanos, contribuye al desarrollo de la Industria del Software con compromiso y responsabilidad social.

Por otra parte dado que la asignatura Trabajo Final de Aplicación corresponde al plan de estudios anterior, se asegura que los alumnos de ambos planes de estudios accedan a contenidos y estrategias didácticas similares.

Los resultados concretados en el año 2016, se plasman en producciones teóricas y experiencias con énfasis en la transferencia de conocimientos y productos.

Se han realizado presentaciones con el propósito de difundir los resultados ([22], [24], [25]).

Se compilaron diversas experiencias en la temática en un documento el que será difundido en la comunidad de práctica e interesados en el tema [26].

En el análisis y evaluación de la accesibilidad se aplicaron procedimientos híbridos sustentados en medios automáticos y semi-automáticos y se procedió a la revisión manual.

Formación de Recursos Humanos

El Sector de Servicios y Sistemas Informáticos se encuentra en desarrollo, motivo por el cual desde el ámbito académico de la región NEA se tiene un fuerte compromiso de aportar. Estos desarrollos tecnológicos construidos con recursos humanos formados y en formación permiten fortalecer la temática y su concreta realización.

La formación de recursos humanos en la temática, dado que se avanzó en:

- la revisión y profundización en métodos y herramientas de accesibilidad, y su introducción desde etapas tempranas del ciclo de vida de sistemas de información, actividad realizada con investigadores, becarios de grado [27] y postgrado de la SGCyT – UNNE.
- se concretó la defensa de un Proyecto Final de Carrera [28] y un Trabajo Final de Aplicación [29], que abordó temas de investigación y aplicación de la AW.
- se concluyó en el desarrollo de un sistema informático para la gestión de AW en proyectos software [24]. El mismo se presentó como beca de pregrado [27] y fue nominado como mejor trabajo [25] y seleccionado para de las XXIV Jornada de Jóvenes Investigadores – AUGM 2016.
- en una asignatura de programación del tercer año de la carrera para la elaboración de aplicaciones web se incorporaron aspectos relacionados con la accesibilidad desde etapas temprana del desarrollo. Se mencionan el uso correcto de los códigos HTML y CSS. Se fomentó la comprobación del código de forma

automática, a través de los servicios de validación suministrados por W3C.

- en la asignatura Trabajo Final de Aplicación y Proyecto Final de Carrera, espacios donde se generan las tesinas para la obtención del título de grado, Licenciado en Sistemas de Información, se fomentó el abordaje de la temática considerando cómo un aspecto clave en soluciones web. Se enfatizó en su tratamiento como un aspecto integrado al modelo de ciclo de vida.

Actualmente se realiza una evaluación longitudinal de diversos desarrollos software diseñados para entornos web a fin de estudiar el grado de aplicación de las normas W3C. La información sistematizada se utilizará para fundamentar capacitaciones y construir una base de conocimiento sobre la temática

Como líneas futuras de trabajo se menciona desarrollar asesorías con la finalidad de aportar, desde la producción de software en la conformación de la sociedad inclusiva centrada en los sujetos consumidores de la tecnología.

Referencias

- [1] Consorcio World Wide Web (W3C). Disponible en: <http://www.w3c.es/>
- [2] Oficina Española. (2008). Word Wide Web - Guía Breve de Accesibilidad Web. Disponible en: <http://www.w3c.es/divulgacion/guiasbreves/accesibilidad>
- [3] Hernández-Arteaga, R. I., Alvarado-Pérez, J. C. y Luna, J. A. (2015). Responsabilidad social en la relación universidad-empresa-Estado. Educación y Educadores. Vol. 18, No. 1, 95-110.
- [4] Mariño, S. I., Godoy, M. V., Alfonzo, P., Acevedo, J., Gómez Solis. L, Fernández Vázquez, A. (2012) Accesibilidad en la definición de requerimientos no funcionales. Revisión de herramientas. Multiciencias, Vol. 12, N° 3, 2012 (305 - 312).
- [5] Varas, V. D., Agüero, A. L., Guzmán, A. Elena, Martínez, M. (2015). Importancia y beneficios de la Accesibilidad Web para todos, X Congreso sobre Tecnología en Educación & Educación en Tecnología, p. 357-366.
- [6] Gallardo, C., Funes, A., Ahumada, H. (2016). Modelo Integral para la Evaluación de la Calidad de la Accesibilidad al Contenido Web. XVIII Workshop de Investigadores en Ciencia de la Computación, WICC 2016.
- [7] Rossi, B., Ortiz, C., Chapetto, V. (2016). Accesibilidad de la Información en Sitios Web argentinos. XVIII Workshop de Investigadores en Ciencia de la Computación, WICC 2016.
- [8] Martín, A., Gaetán, G., Saldaño, V., Miranda, G., Sosa, H., Pires, A., Nichele, E. (2016). Evaluaciones de Accesibilidad y Usabilidad en la WWW: Propuestas para Mejorar la Experiencia del Usuario. XVIII Workshop de Investigadores en Ciencia de la Computación, WICC 2016.
- [9] Rossi, B., Chapetto, V., Curti, M. (2015). Accesibilidad de la información en los sitios web de entidades bancarias públicas y privadas de la República Argentina. 44 JAIIO. Jornadas Argentinas de Informática e Investigación Operativa.
- [10] Fernández Reuter, B., Duran E. 2016, Modelos de usuario como una alternativa para mejorar la Accesibilidad, STS 2016, 3° Simposio Argentino sobre Tecnología y Sociedad, 45 Jornadas Argentinas de Informática e Investigación Operativa. 81-99.
- [11] Web Content Accessibility Guidelines (WCAG) 2.0. Disponible en: <http://www.w3.org/TR/WCAG20/>
- [12] Brailersurf. Disponible en: <http://www.snv.jussieu.fr/inova>
- [13] Lynx Browser. Disponible en: <http://invisibleisland.net/lynx/>
- [14] TAW. Disponible en: <http://www.tawdis.net/>

- [15] HERA. Disponible en: <http://www.sidar.org/hera/>
- [16] EXAMINATOR. Disponible en: <http://examinator.ws/>
- [17] TextAloud. Disponible en: http://www.nextup.com/files/manualtrans/Spanish_TextAloudManual.html
- [18] NVDA. Lector de pantalla. Disponible en: <http://nvda.softonic.com/descargar>
- [19] WEBBIE. Explorador web para usuarios ciegos y deficientes visuales. Disponible en: <http://www.webbie.org.uk/es/index.htm>
- [20] Ckecksmicolours. Disponible en: <http://www.checkmycolours.com/>.
- [21] WCAG Contrastchecker. Complemento para el navegador Mozilla Firefox. Disponible en: <https://addons.mozilla.org/es/firefox/addon/wcag-contrast-checker/>.
- [22] Rodriguez, M., Mariño, S. I., Godoy, M. V., Alfonzo, P. L. (2016). Desarrollo de sitio web del Congreso.
- [23] Red UNCI. (2014). Documento de Recomendaciones Curriculares de la RedUNCI 2014-2015. Disponible en: <http://redunci.info.unlp.edu.ar>,
- [24] Cavalieri, J., Mariño, S. I., Alfonzo, P. L., Godoy, M. V. (2016). Herramienta Software para Administrar el Desarrollo de Sitios Web Accesibles según Pautas WCAG 1.0. Revista Latinoamericana de Ingeniería de Software, 4(3): 159-166.
- [25] Cavalieri, J., Alfonzo, P. L., Mariño, S. I. (2016). Una herramienta software para asistir la evaluación de las pautas WCAG 1.0. XXIII Reunión de Comunicaciones Científicas de la UNNE. 14 -15/06/2016. Campus Universitario Corrientes. Corrientes. Argentina.
- [26] Mariño, S. I., Godoy, M. V. Alfonzo, P. L., Compilación de experiencias del equipo en torno a la Accesibilidad Web. Inédito.
- [27] Cavalieri, J. I., Alfonzo, .P, Mariño, S. I. (2015). Sistema informático como soporte al análisis de las pautas de accesibilidad al contenido web. Beca de Pregrado otorgada por la Ciencia y Técnica de la Universidad Nacional del Nordeste.
- [28] Cavalieri, J. (2016). Herramientas de apoyo para el desarrollo de sitios web accesibles de acuerdo a las WCAG 1.0. Proyecto Final de Carrera. Carrera Licenciatura en Sistemas de Información. FaCENA. UNNE. Prof. Orientador: P. L. Alfonzo.
- [29] Blanco, V. (2016). Sistema web de gestión de turnos para una clínica de ojos. Trabajo Final de Aplicación. Carrera Licenciatura en Sistemas de Información. FaCENA. UNNE. Prof. Orientador: P. L. Alfonzo.

Catedra de Proyecto Fin de Carrera una Gestión de Calidad

Roberto Eribe, Jorge Eterovic, Silvana Ardanaz, Juan Martin Hernández

Departamento de Ingeniería e Investigaciones Tecnológicas
 Universidad Nacional de la Matanza
 Florencio Varela 1903, (B1754JEC) San Justo, Prov. de Buenos Aires, Argentina.
 Tel: (54 11) 4480-8900

eri_10@yahoo.com, jorge_eterovic@yahoo.com.ar, silvanaardanaz@yahoo.com.ar,
juanmartinhernandez@hotmail.com

Resumen.

La asignatura proyecto fin de carrera tiene como objetivo fundamental completar la formación académica y profesional de los futuros ingenieros.

En este trabajo se describen las acciones abordadas por los docentes de la cátedra desde el año 2004 con el propósito de que los alumnos, logren integrar y aplicar los conocimientos adquiridos durante sus años de estudio en el desarrollo de un producto de software.

El objetivo principal de este estudio se basó en la necesidad de incrementar la calidad de ese producto desde el punto de vista de su creatividad e innovación, como así también desarrollar un modelo de gestión de la materia tal que los alumnos terminen su proyecto en tiempo y forma. A tales efectos se desarrollaron dos líneas de trabajo, una basada en la presentación de los proyectos en congresos nacionales e internacionales, con el fin de que los mismos sean evaluados más allá de la evaluación interna de la cátedra. La otra realizando una encuesta a los alumnos al finalizar la cursada, con la finalidad de conseguir información sobre la problemática de estos al cursar la materia y aspectos relevantes de la misma. Esto nos permitió crear un modelo de gestión de calidad para la materia, para cumplir con los objetivos planteados. En esta presentación solamente se hará referencia a las métricas obtenidas desde el punto de vista de la evaluación del

plantel docente, piedra angular del éxito para el cumplimiento de los objetivos buscados.

Los resultados han sido excelentes, y han permitido desarrollar un modelo de gestión para la cátedra, que en función de los buenos resultados obtenidos produjo una alta motivación en los docentes para continuar con este trabajo.

Palabras Clave.

Proyecto Fin de Carrera. Metodología. Gestión de calidad. Gestión de proyecto. Desarrollo de software.

Contexto.

La Universidad debe formar ingenieros con capacidad creativa, actitud crítica y disposición a la actualización permanente, así como con una formación teórica y práctica.

La formación recibida en los últimos años de estudio debe asegurar que el futuro ingeniero se adapte a la diversidad de situaciones que se presenten en su desempeño profesional.

1 Introducción

Esta materia, Proyecto Fin de Carrera, aglutinante y convergente de conocimientos, debe integrar los conceptos adquiridos del proceso de desarrollo de software, a saber: análisis, diseño, proyecto de ingeniería, metodología, responsabilidad legal y ética profesional, planificación, factibilidad temporal, cálculo del retorno de

la inversión, optimización, innovación y perspectiva tecnológica, el diseño por evolución, confiabilidad y falla, ensayo y evaluación, mantenibilidad, calidad del producto y del proceso y documentación [1]. [2]., para completar la formación del futuro ingeniero, mediante la investigación personal de aspectos teórico prácticos actuales, en lo referente al gerenciamiento de proyectos informáticos y su traslado a la práctica mediante el desarrollo e implementación de un proyecto informático de envergadura, adecuado en tiempo y forma.

El acelerado avance de las tecnologías de la información y las comunicaciones (TICs) y los consecuentes requerimientos de un cada vez más alto grado de especialización por parte de las empresas, hace que el dominio de la aplicación sea un recurso indispensable para asumir los cambios y participar en la creación de nuevos conocimientos [3].

El objetivo general será completar la formación académica y profesional de los alumnos, posibilitando la integración y utilización de los conocimientos adquiridos durante sus años de estudio para la resolución de problemas de índole profesional o científico [4] y [5].

En los últimos años, en foros y congresos académicos se han expuesto con gran preocupación dos problemas que impactan fuertemente en esta materia. El primero, la falta de creatividad e innovación en los proyectos que proponen los alumnos, y el segundo que los equipos de desarrollo de los proyectos no logran finalizar el mismo en tiempo y forma. Desde el año 2010, en [6] [7] y [8], un grupo de trabajo se ha abocado a detectar y analizar las causas de los retrasos en la realización del proyecto final de la carrera Analista en Computación. Ante lo expuesto, esta cátedra ha estado trabajando desde hace doce años en esta problemática, con el fin de encontrar una estrategia para una gestión de calidad que permita ir mejorando año tras año los

productos desarrollados y encontrar la mejor metodología de gestión para la

materia a fin de conducir los proyectos con éxito. Entendiéndose por éxito, que los mismos sean desarrollados en los tiempos estimados y que se cumplan los requerimientos funcionales establecidos.

Planteados los problemas y después de varias reuniones de cátedra la solución a la que se llega por parte de los docentes tiene fundamentalmente dos vistas.

La primera, en cuanto a evaluar la calidad de los proyectos, desde la visión de la creatividad e innovación, la solución que se ha adoptado a fin poder medir la calidad de los trabajos más allá de la evaluación interna de la cátedra, fue exigir que los mismos sean presentados en exposiciones y congresos nacionales e internacionales. Esta medida se aplicó desde el año 2012, realizándose presentaciones en Imagine Cup Microsoft 2012, CoNaIISI, 2014, 2015 y 2016, obteniéndose en todos los casos puestos de relevancia. Lo que nos permite inferir que estamos a la altura de las circunstancias que la realidad nos exige.

En segundo lugar, y con el criterio de considerar que la piedra angular del éxito de los proyectos son los docentes, y sus técnicas de coaching, en función del acompañamiento que estos hacen a lo largo de la cursada, fue darles a los alumnos la posibilidad de expresarse respecto al equipo de profesores, en forma anónima, por medio de un cuestionario al finalizar el ciclo lectivo. Con el objetivo de conocer las fortalezas y debilidades de la cátedra.

Mariño y Herrmann [9] describieron la función principal del plantel docente, orientada al asesoramiento, seguimiento y tutorización en el diseño y desarrollo del plan de trabajo y las condiciones a cumplir por el alumno, para regularizar la asignatura.

2 Líneas de investigación y desarrollo Para llevar a cabo esta investigación se utilizó un proceso sistemático en el que se consideró el uso de técnicas simples y económicas, pero que permitieron obtener

información confiable y relevante al trabajo.

En primer lugar, se incluyó exploración bibliográfica en soporte tradicional y en Internet para el estudio de teorías que permitan enmarcar con actualidad el estado del conocimiento de los temas tratados en el ámbito de esta investigación.

De esta forma se llegó a elaborar un cuestionario de 25 preguntas cerradas y de respuestas múltiple choice.

Cabe destacar, que esta encuesta abarca varios temas relacionados con la cátedra, en esta exposición solo serán expuestos a modo de ejemplo los resultados relacionados con la gestión de calidad de los docentes.

Con respecto a la muestra, se decidió hacer la encuesta sobre el total de la población, dado que se estudiaron las cohortes dando como resultado que en ningún caso superaría los 150 alumnos en promedio, con lo cual sería totalmente factible llevarla adelante. El porcentaje de respuestas a través de los años fueron entre un 60/70 por ciento de la población total.

La encuesta, debería ser realizada vía Internet, teniendo carácter anónimo. Después de realizar una investigación sobre los productos ofrecidos en la web, se decidió utilizar Create Survey [10], producto de origen ruso, alojado en servidores residentes en EEUU.

Este software nos permite crear la encuesta, con sus respectivas preguntas y tipos de respuestas.

Luego se procedió a enviar el formulario de la encuesta vía web al correo electrónico de cada alumno.

Este procede a completar el formulario y lo envía a través de Internet, incorporándose la respuesta a la base de datos de resultados, a la cual se puede acceder por medio de un login y password para imprimir los resultados o hacer el análisis de los mismos.

A continuación, se procede a la revisión, validación y organización del material recolectado para hacer el análisis y la generación de conclusiones.

Este proceso no fue puramente lineal, sino que se trabajó en forma iterativa incremental año tras año, a medida que surgieron nuevos temas o categorías de análisis.

Estos análisis se realizaron con una continua reflexión de la evolución de la problemática en estudio, a la luz del surgimiento de nuevos conceptos desde los datos examinados.

Indudablemente la validez y confiabilidad de un estudio cuantitativo están definidas por la precisión y pertinencia de los resultados.

Al ser los datos de tipo cuantitativo, el reporte de los resultados se hizo mediante gráficos y descripciones del proceso completo y comparativo a lo largo de los años.

Principalmente se buscó credibilidad y utilidad de los resultados, basados en la coherencia y confiabilidad.

3 Resultados Obtenidos / Esperados.

Como se mencionó en párrafos anteriores, se presentan en esta exposición los resultados obtenidos entre los años 2004 / 2015 en la cátedra de proyecto fin de carrera.

Solo se expondrá a modo de ejemplo dos cuadros de variables relacionadas con la gestión de calidad de los docentes de la cátedra.

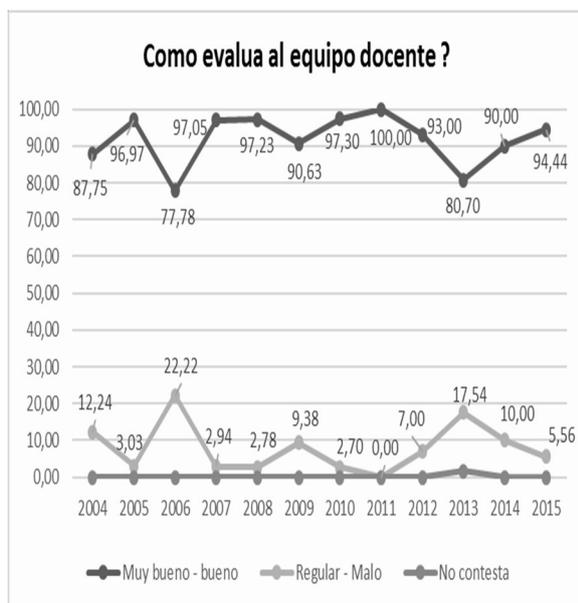


Figura 1. Evaluación del equipo docente. (en %). Fuente, alumnos de la Cátedra de Proyecto fin de Carrera.

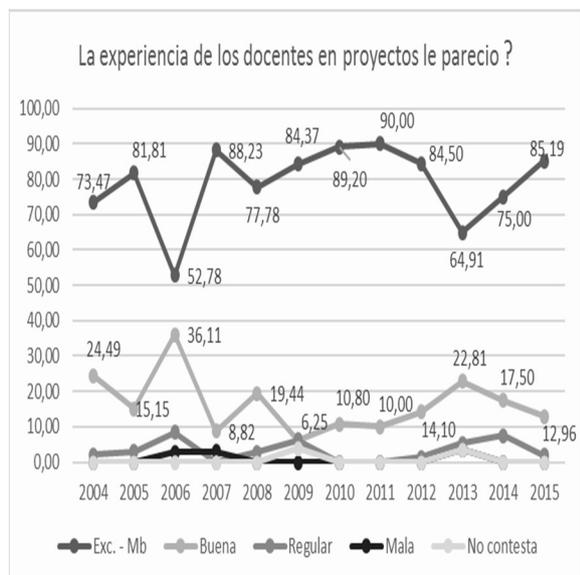


Figura 2. La experiencia de los docentes en proyectos. (en %). Fuente, alumnos de la Cátedra de Proyecto fin de Carrera.

3. 1 Discusión.

A lo largo de estos años, hemos logrado un avance significativo respecto de la calidad de los proyectos, no solo desde el producto final, como un producto de software, sino también en la gestión de los proyectos. Habiendo podido bajar la gestión de la cátedra al análisis cuantitativo, los números y los hechos nos permiten afirmar que estamos en el camino correcto, buscando en todo momento el objetivo tan anhelado y

nunca totalmente alcanzable de la excelencia académica.

Esta presentación no tiene el ánimo de constituirse en una receta, está muy lejos de haber sido concebida como una prescripción normativa. Solo nos motiva el afán de realizar una sugerencia práctica y comprobable abierta a la comunidad académica.

Para poder mantener los valores de las variables de análisis propuestas entre un 70% y un 100%, es necesario mantener actualizados y en permanente discusión las nuevas tecnologías emergentes, con todos los docentes de la cátedra, independientemente de su categoría.

Cabe destacar la caída en los valores en los años 2006 y 2013, producida por la incorporación de una nueva camada de ayudantes que distorsiono el habitual funcionamiento de la cátedra.

Cada profesor, desde su cargo, debe conocer en profundidad que y como debe hacer su trabajo y eso será responsabilidad de los jefes de cátedra.

Tan bien recibida ha sido esta investigación por parte de alumnos y profesores de la cátedra, y tanto ha servido al mejoramiento de su gestión, que la misma será continuada, profundizándose su análisis año tras año.

3.2 Conclusión.

En los últimos años, en foros y congresos académicos se han expuesto con gran preocupación dos problemas que impactan fuertemente en la materia proyecto fin de carrera. El primero, la falta de creatividad e innovación en los proyectos que proponen los alumnos, y el segundo que los equipos de desarrollo de los proyectos no logran finalizar el mismo en tiempo y forma.

A fin de paliar esta situación, la cátedra de proyecto ha venido trabajando desde el año 2004 en desarrollar un modelo de gestión de calidad que elimine o disminuya a su mínima expresión esta problemática.

En tal sentido con todos los profesores de la cátedra se trabajó en dos sentidos.

Primero se dispuso, que los proyectos sean evaluados, independientemente de la evaluación interna de la cátedra, presentándolos en congresos nacionales e internacionales.

En segundo lugar, se dispuso armar una encuesta de carácter anónimo para el alumno, con 25 preguntas que abarcan desde la evaluación de los docentes de la cátedra, del grado de satisfacción del alumno por el producto desarrollado, el grado de conflictos internos dentro del grupo de desarrollo del proyecto, etc.

Los resultados han sido excelentes, y permitió desarrollar un modelo de gestión de cátedra, que en función de los buenos resultados obtenidos produjo una alta motivación en los docentes para hacer de esta investigación un proceso continuo.

4 Formación de Recursos Humanos

La estructura de los recursos humanos está dada por el director, el co director, y dos ingenieros ayudantes de cátedra involucrados en forma directa en el armado y captura de datos a través de la encuesta.

Por otro lado, cabe destacar que las conclusiones obtenidas y las decisiones tomadas en consecuencia derraman sobre la totalidad de la cátedra, 11 profesores que también integran el proyecto.

Referencias.

[1] Ian Sommerville, Ingeniería de software, novena edición. (2011).

[2] Roger S. Pressman. Ingeniería del software: Un enfoque práctico, 7ma Edición. (2014).

[3] Adell, J. 1997. "Tendencias en educación en la sociedad de las tecnologías de la información".

EDUTECH, Revista Electrónica de Tecnología Educativa, n° 7, ISSN: 1135-9250.

[4] Mariño, S. I. y Herrmann, C. F.: Experiencias curriculares en la asignatura Trabajo Final de Aplicación en la FACENA. Anales del I Congreso de Tecnología en Educación y Educación en Tecnología (2006).

[5] Mariño, S. I. y Herrmann, C. F.: Los trabajos finales de aplicación en informática. Una alternativa de promoción de docencia, investigación, extensión y transferencia. V Encuentro Nacional y II Latinoamericano La universidad como objeto de investigación (2007).

[6] Marcela Daniele, Fabio Zorzan, Paola Martellotto, Mariana Frutos, Marcelo Uva, Ariel Arsaute, F. Brusatti, J. Guazzone, S. Angeli. Estimación y Planificación de Proyectos de Software versus duración de proyectos finales en la carrera Analista en Computación. Secretaría Académica y de Ciencia y Técnica, Universidad Nacional de Río Cuarto. RR No /11. (01/02/2011, 31/12/2012).

[7] Marcela Daniele, Fabio Zorzan, Paola Martellotto, Marcelo Uva, Ariel Arsaute, Mariana Frutos. Causas que producen que los estudiantes de Computación retrasen la culminación de su Trabajo Final. Secretaria Académica y de Ciencia y Técnica, Universidad Nacional de Río Cuarto. Presentado en abril de 2013. En evaluación.

[8] Fabio Zorzan, Mariana Frutos, Ariel Arsaute, Marcela Daniele, Paola Martellotto, Marcelo Uva, Carlos Luna Delayed Completion of Final Project of the Career Computer Analyst: Seeking its Causes. XX Congreso Iberoamericano de Educación Superior (CIESC 2012), en el Marco de la XXXVIII Conferencia Latinoamericana en Informática – CLEI 2012 - octubre 1 al 5 de 2012 -Medellín, Colombia. ISBN 978-1-4673-0792-5.

[9] Mariño, S. I. y Herrmann, C. F.: Innovaciones en el desarrollo de trabajos finales de aplicación en una carrera informática. Cohortes 2003-2007". Revista Latinoamericana de Tecnología Educativa - RELATEC, Vol 8(1):141-148 (2009).

[10] <http://www.createsurvey.com/contact.htm>.

El Pensamiento Computacional en el **Ámbito Universitario**

María V. Rosas, Mariela E. Zúñiga, Jacqueline M. Fernández, Roberto A. Guerrero

Laboratorio de Computación Gráfica, Departamento de Informática, FCFMyN Universidad Nacional de San Luis

Ejército de los Andes 950 (San Luis), Teléfono: 266-4520300 (2120)

{ mvrosas, mezuniga, jmfer, rag}@unsl.edu.ar

RESUMEN

Para comprender y desempeñarse en el mundo actual es fundamental desarrollar la habilidad de resolver problemas. El Pensamiento Computacional se basa en resolver problemas haciendo uso de conceptos básicos de la informática.

Por ello, es fundamental discutir el proceso educativo y el lugar de la Informática en el conocimiento humano reconociendo al pensamiento computacional como un objetivo a lograr, necesario en la educación actual.

Para favorecer su desarrollo resulta imperiosa la inclusión de los fundamentos computacionales de la Informática en todos los niveles educativos, incluyendo a la Universidad.

No obstante, la deserción y desgranamiento universitario junto con el poco interés de los jóvenes por estudiar carreras de informática representan serios problemas a resolver en el ámbito académico. Es necesario que desde la universidad surjan estrategias que permitan afrontarlos y convertirlos en soluciones viables, visibles y comprensibles.

En este contexto, las iniciativas de reforma implican un cambio de perspectiva profundo y complejo de los diferentes actores involucrados en los procesos de enseñanza y aprendizaje, que resultará más significativo mientras más temprano se lleve a cabo, favoreciendo en particular la formación de profesionales informáticos capacitados para desenvolverse en los desafíos del mundo actual.

Palabras clave: Pensamiento Computacional,

Enseñanza de la programación, Deserción y desgranamiento universitario.

CONTEXTO

La línea de investigación presentada en este trabajo, está inserta en el proyecto de investigación “*Estrategias para la Mejora de la Enseñanza de la Programación a Alumnos Ingresantes de las carreras de Ciencias e Ingeniería*” del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el Laboratorio de Computación Gráfica (LCG) de la Facultad de Ciencias Físico, Matemática y Naturales de la Universidad Nacional de San Luis (UNSL). Este proyecto está acreditado y financiado por la UNSL.

1. INTRODUCCIÓN

El Pensamiento Computacional (PC) se define como “*procesos de pensamiento involucrados en formular problemas y encontrar sus soluciones de manera que estén representadas de forma tal que puedan llevarse a cabo por un agente que procesa información (máquina o humano)*” [1, 2, 4].

El PC se compone del pensamiento crítico, relacionado con el modo de pensar y las estructuras inherentes del acto de pensar y la computación. El poder de la computación refuerza y desarrolla al pensamiento crítico apoyándose en procesos de pensamientos como la abstracción y la descomposición de problemas que pueden aplicarse a cualquier área del conocimiento [13].

La abstracción representa el proceso de

pensamiento de orden superior más importante en el PC. Junto con la descomposición de problemas, el reconocimiento de patrones y la definición de algoritmos, conforman los cuatro pilares fundamentales del PC [3, 14].

El PC se hace concreto cuando se aprende a programar. La programación es una disciplina que requiere del uso simultáneo de la creatividad, un conjunto de conocimientos técnicos asociados y la aptitud de trabajar con abstracciones, tanto simbólicas como mentales. La creatividad necesaria para programar no difiere demasiado de aquella utilizada, por ejemplo, para producir textos. El empleo de un conjunto de conocimientos técnicos asociados a la operación de las computadoras es lo que la convierte en una actividad especial. Sin embargo, al poseer una naturaleza ligada a la resolución de problemas del mundo real, se requiere de una capacidad de abstracción que permita trabajar sin que los conocimientos técnicos representen un obstáculo para el programador, donde el límite a la creatividad está dado por la imaginación permitiendo crear mundos virtuales sin las restricciones del mundo físico [13].

La sociedad y los sistemas de producción, necesitan de profesionales calificados en las industrias de la información. A pesar de que el sector está en crecimiento, existe una gran demanda de profesionales para abastecer la Industria del Software y la cantidad de graduados del sistema universitario argentino en carreras de la disciplina informática, universitarias y terciarias, no satisfacen a la demanda [6, 7, 8, 12].

Esta problemática obedece a diversas razones, pero la cuestión de fondo radica en tres puntos fundamentales: pocos potenciales alumnos de la universidad eligen estudiar estas carreras, el alto porcentaje de alumnos que no aprueban el ingreso y el significativo nivel de deserción y desgranamiento en las carreras de Ciencias e Ingenierías [8, 9].

Afrontar el problema implica hacerlo desde diversas ópticas considerando los diferentes actores involucrados y sus características. La profundidad y complejidad en el cambio de perspectiva hacen necesario, entre otras cosas,

abordar la enseñanza de las capacidades que favorecen el PC de manera temprana dando el tiempo suficiente para la comprensión y maduración cognitiva imprescindible. Por tratarse de una característica fundacional de estas disciplinas, es necesario favorecer su desarrollo en el primer año de las carreras de Ciencias e Ingenierías como así también promover y difundir su desarrollo en el ámbito de las escuelas secundarias [5, 10, 11].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Con la intención de contribuir al corpus teórico a partir del mundo empírico, el énfasis de la investigación está circunscripto en un enfoque cuantitativo con alcance exploratorio-descriptivo. Exploratorio, dado que aborda una temática sobre la que se está haciendo camino, fundamentalmente en Argentina; y descriptivo, pues se busca especificar las características y los aspectos relevantes del fenómeno objeto de estudio.

Dentro del proyecto que da marco a la investigación presentada en este trabajo se han definido tres grandes líneas:

- Línea 1: Desde el punto de vista del alumno. Indagar qué habilidades relacionadas al PC ponen en juego los alumnos que ingresan a la UNSL en lo que respecta al aprendizaje de la programación en carreras de Ciencias e Ingeniería.
- Línea 2: Desde el punto de vista del docente. Conocer las teorías implícitas de los docentes acerca del desempeño académico de los alumnos en el aprendizaje de la programación, fundamentalmente de los primeros años de las carreras de ciencias e ingeniería.
- Línea 3: Desde el punto de vista de la didáctica. Analizar las estrategias de enseñanza y aprendizaje que favorecen el desarrollo del PC en el aprendizaje de la programación.

3. RESULTADOS OBTENIDOS/ESPERADOS

Durante la primera etapa del proyecto antes mencionado, el propósito general ha consistido en capturar los rasgos generales de los alumnos que desean ingresar al sistema universitario y de los que se encuentran cursando el primer año de su respectiva carrera.

El trabajo de campo se ha realizado con potenciales alumnos a carreras de Informática e Ingeniería de la Facultad de Ciencias Físico, Matemática y Naturales de la UNSL, tanto del ingreso 2016 como del 2017, y con alumnos de primer año de dichas carreras.

De la observación y análisis de las distintas problemáticas que se plantean y a partir de la interpretación de los primeros resultados obtenidos, se ha decidido realizar diversas acciones con la intención de experimentar diferentes estrategias que permitan atenuar los conflictos más importantes encontrados para, finalmente, poder determinar el impacto que provocan.

En consonancia con las líneas de investigación definidas se destacan las siguientes acciones:

Línea 1

- El fortalecimiento del vínculo universidad-nivel medio, posibilitando a través de diversas actividades, acercar a los alumnos experiencias propias del trabajo del profesional informático e incorporar las nociones básicas de programación.
- La formulación y presentación de proyectos nacionales, en particular, en respuesta a las convocatorias realizadas por la Fundación Sadosky en los años 2015 y 2016. En el año 2015 la UNSL quedó seleccionada por la Convocatoria A, y en el año 2016 por las Convocatorias A y B.
- La planificación de talleres orientados a los alumnos del curso de ingreso elaborados a partir del desarrollo de actividades lúdicas que permiten in-

corporar las nociones básicas de programación.

Línea 2

- El dictado de cursos de capacitación a docentes del nivel medio durante los años 2015 y 2016 y en el marco del convenio firmado con la Fundación Sadosky.
- La planificación y formulación de métodos exploratorios que caractericen las teorías implícitas que los docentes de las carreras de informática e ingeniería ponen en juego en su desempeño cotidiano.

Línea 3

- La selección y determinación de herramientas que permitan reformular las estrategias actuales de enseñanza favoreciendo sustancialmente el aprendizaje de los alumnos.
- La planificación de talleres orientados a los docentes involucrados.

Los principales resultados de las actividades desarrolladas son:

- El dictado de un curso de inicio a la programación como apoyo al ingreso. Experiencia piloto no vinculante destinada a alumnos del ingreso 2017. Se encuentra en análisis y evaluación la posibilidad de establecer dicho curso como requisito del ingreso y los posibles cambios curriculares que serían necesarios realizar.
- La definición de una red de escuelas que incorporen al PC como estrategia para mejorar el rendimiento académico de los alumnos. Esto surge como corolario de la información relevada durante el desarrollo de actividades con las escuelas del medio.
- La firma de convenios de asistencia técnica con otras instituciones, en particular con la Fundación Sadosky.

4. FORMACIÓN DE RECURSOS HUMANOS

En el equipo de trabajo participan cuatro docentes investigadores con formación de grado y posgrado en carreras relacionadas a la Informática y la Educación Superior. Desde el proyecto se han generado propuestas de temas para ser abordadas como trabajos finales de carreras de grado o posgrado, tanto afines a la informática como a las de índole educativo. Por consiguiente, actualmente se están desarrollando: un trabajo de doctorado y dos trabajos de especialización en Educación Superior, además de los planes de tesis de los propios integrantes del proyecto.

De la misma manera, desde el proyecto se ha consolidado un espacio de interacción con otros grupos de investigación para propiciar actividades interdisciplinarias que promuevan nuevas estrategias didácticas en la enseñanza de la programación.

Para este año se ha iniciado el proceso de la incorporación de pasantes para la investigación en las distintas líneas pertenecientes al proyecto.

5. BIBLIOGRAFÍA

- [1]. J. M. Wing, *Computational Thinking and Thinking about Computing*, de Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 2008, pp. 3717-3725.
- [2]. M. Prensky, *Digital natives, digital immigrants*, USA: On the Horizon, 2001.
- [3]. G. Stager, *En pro de los computadores*, 13 enero 2004. [En línea]. Available: <http://www.eduteka.org/ProComputador.es.php>.
- [4]. J. M. Wing, *Computational thinking*, Commun. ACM, 2006, p. 49(3):33-35.
- [5]. A. Szpiniak y G. Rojo, *Enseñanza de la programación*, TE&ET: Revista Iberoamericana, 2006.
- [6]. Secretaría de Políticas Universitarias Argentina- 2013. ISSN 1850-7514.
- http://informacionpresupuestaria.siu.edu.ar/DocumentosSPU/Anuario_2013.pdf
- [7]. Instituto Nacional de Educación Tecnológica. *Demandas de capacidades 2020*. Ministerio de Educación. Presidencia de la Nación. Junio 2016.
- [8]. M. F. Botta, L. Dughera, G. Yansen, M. Zukerfeld. *Y las mujeres... ¿dónde están? Informe final*. Fundación Sadosky. Universidad Maimónides. 2013.
- [9]. I. Palou, G. Utges. *Teorías implícitas de docentes universitarios sobre la problemática de deserción y desgranamiento en carreras de Ingeniería. Un estudio contextualizado*. Revista Argentina de Enseñanza de la Ingeniería. Año I. N° 1. Marzo 2012.
- [10]. J. Wing *Computational Thinking: What and Why?*. Noviembre 2010. Disponible en <http://www.cs.cmu.edu/~CompThink/resources/TheLinkWing.pdf>.
- [11]. G. Simari. *Los fundamentos computacionales como parte de las ciencias básicas en las terminales de la disciplina Informática*. VIII Congreso de Tecnología en Educación y Educación en Tecnología. junio 2013. UNSE. Santiago del Estero.
- [12]. M. Zapata-Ros. *Pensamiento computacional: Una nueva alfabetización digital*. RED-Revista de Educación a Distancia, N° 46. Septiembre 2015.
- [13]. J. L. Zapotecatl Lopez *Pensamiento Computacional*. <http://www.pensamientocomputacional.org/>. (2014).
- [14]. Martínez López, P. (2014) *Las base conceptuales de la Programación. Una nueva forma de aprender a programar*. ISBN 978-987-33-4081-9. Obtenido de: <http://www.gobstones.org/bibliografia/Libros/BasesConceptualesProg.pdf>.

Experimentación en la Enseñanza de Requerimientos

Alejandro Oliveros

UNTREF – Departamento de Ciencia y Tecnología

Valentín Gómez 4752, Caseros, Buenos Aires, Argentina

aoliveros@untref.edu.ar

Resumen

Se propone desarrollar una serie de experimentos en el curso de Ingeniería de Requerimientos de la carrera de Ingeniería en Informática. De esta forma se propone ampliar el campo ya establecido en las investigaciones experimentales de enfoques de enseñanza de ingeniería de requerimientos en prácticas innovadoras para obtener los requerimientos del usuario. La primera de las dos líneas básicas del proyecto se orientan a un reprocesamiento de los datos hasta ahora obtenidos con el objetivo de establecer la calidad de los productos y la segunda se enfoca en evaluar la utilización de buenas prácticas para obtener el Léxico Extendido del Lenguaje.

Palabras clave: Elicitación de requerimientos, experimentación, enseñanza.

Contexto

En el contexto del proyecto de investigación en Ingeniería de Requerimientos [1], se viene desarrollando una línea de investigación en la enseñanza de Ingeniería de Requerimientos en la carrera de Ingeniería Informática de la Universidad Argentina de la Empresa. Dicha línea se ha construido sobre la base de actividades en la asignatura dedicada a la Ingeniería de Requerimientos.

En [2] se encuentra una detallada explicación del curso mencionado. El curso *Ingeniería de Requerimientos* forma parte del plan de estudio de la carrera Ingeniería en Informática, es una asignatura cuatrimestral de segundo año de la carrera. Es el primer curso que encara temas de Ingeniería de

Software y tiene un enfoque hacia los sistemas socio-técnicos. En el curso se brindan herramientas para comprender los requerimientos del usuario, definir los requerimientos de sistemas y documentarlos adecuadamente. El contenido abarca los procesos de la Ingeniería de Requerimientos con diverso grado de profundidad. El libro de texto que siguen los alumnos es el de Wiegers [3].

En esta asignatura se han desarrollado varias actividades de investigación encarando los temas de la enseñanza de las prácticas de la Ingeniería de Requerimientos. En [1], [2], [4], [5] y [6], se han informado los resultados obtenidos. Estas investigaciones se continúan con el proyecto acá presentado.

Introducción

El proceso de elicitación de requerimientos es el proceso de la Ingeniería de Requerimientos mediante el que se obtiene el conocimiento necesario para producir la especificación de requerimientos. La dificultad de la elicitación de requerimientos radica fundamentalmente en la incertidumbre derivada del proceso de transferencia de información desde una o varias fuentes, no siempre consistentes, hacia el ingeniero de requerimientos.

En la obtención de los requerimientos de un sistema software, el conjunto de documentos vinculados con el sistema constituye una fuente relevante. Un subconjunto de estos son los *documentos del usuario*, esto es: aquellos documentos accesibles directamente por el usuario sin necesidad de intermediación técnica y que contienen una gran parte de la información necesaria para formular los requerimientos del sistema. Estos documentos pueden ser los

disponibles de los actuales sistemas, por ejemplo el *manual del usuario* o documentación de procesos anteriores de elicitación.

También deben considerarse normas, disposiciones legales, estándares de la industria, información de la competencia, notas internas de la organización, planes de la organización, formularios de procedimientos, etc.

Estos documentos están escritos en lenguaje natural (la principal fuente de conocimiento de los requerimientos) y su análisis requiere una serie de técnicas para poder establecer claramente su contenido y poder identificar información que alimentará la formulación de los requerimientos. De hecho la mayor parte de la información necesaria para formular los requerimientos se encuentra en el cuerpo de documentación del usuario.

Esta fuente de requerimientos es la de uso más antiguo para obtención de requerimientos y muy habitualmente utilizada. Permite subsanar algunos déficits de disponibilidad de tiempo del usuario. El analista puede desarrollar una idea muy precisa del trabajo a ejecutar en la fase de requerimientos antes de comenzar a interactuar directamente con la organización. Obtener estos beneficios requiere que el analista maneje técnicas de análisis eficaces en la obtención de información y eficientes en el uso de su tiempo.

Líneas de investigación y desarrollo

Como se explicó anteriormente los documentos (en diversos formatos) constituyen una fuente fundamental de información para el proceso de elicitación de requerimientos. Por ello es que el objetivo establecido consiste en desarrollar experimentos en la enseñanza de técnicas de análisis de documentos para elicitación de requerimientos.

El objetivo general del proyecto consiste en desarrollar experimentos orientados al

proceso de enseñanza con procesos y productos para la elicitación de requerimientos.

Este objetivo general se refinó en dos objetivos específicos del proyecto:

1. Evaluar la calidad de las especificaciones de requerimientos producidas en cursos anteriores de la asignatura Ingeniería de Requerimientos en un contexto de usuario real
2. Evaluar la calidad de los LEL (Léxico Extendido de Lenguaje) producidos por los estudiantes de la asignatura Ingeniería de Requerimientos en un contexto de usuario real.

Objetivo específico 1.

Con los trabajos que se han desarrollado en el proyecto de enseñanza de requerimientos, se produjo una serie de material que ya ha sido procesado para distintas publicaciones.

Ese material se generó en cursos de la carrera de Ingeniería en Informática y fueron especialmente analizados en términos de enseñanza de técnicas. Ahora se tratará de analizarlos en términos de la calidad del producto obtenido por los alumnos en los trabajos desarrollados en la asignatura Ingeniería de Requerimientos (ver referencias del punto “Contexto”).

Se apunta especialmente a utilizar técnicas de análisis que permitan establecer patrones de calidad en los distintos grupos de productores de las especificaciones.

Objetivo específico 2.

Como base conceptual se tomará un enfoque elaborado de mejores prácticas en la elaboración del Léxico Extendido de Lenguaje (LEL) [7], [8]. Se capacitará a los alumnos del curso de Ingeniería de Requerimientos para producir un LEL a partir de entrevistas a un usuario real. Con ese material se desarrollará un análisis de los resultados para establecer características de calidad del producto obtenido por los grupos

de alumnos. Básicamente se trata de establecer la calidad de las entradas y la completitud del conjunto a partir de un análisis de las entradas producidas por los alumnos en relación a las mejoras prácticas definidas en los trabajos de Antonelli *et al* citados más arriba.

Resultados y objetivos

Con relación al **Objetivo Específico 2**, plantean los siguientes objetivos detallados.

1. desarrollar material para ser utilizado por los alumnos de cátedra en la realización de un caso;
2. contar con un caso que compartan todos los cursos de la cátedra;
3. homogeneizar los puntos de vista de los miembros de la cátedra mediante la realización de talleres;
4. las entrevistas se desarrollarán siguiendo el esquema ya utilizado en el proyecto [6] asegurando la filmación de las mismas.

Los pasos principales a desarrollar son los siguientes:

Capacitación en el LEL.

Con el enfoque propuesto de proceso de construcción del LEL [7] se capacitará a los alumnos. El enfoque será el de presentación teórica del tema y realización de ejercicios con el modelo de trabajo práctico habitual en las asignaturas.

Entrevista en la cámara de Gesell.

Desarrollada por un equipo de alumnos dentro de la cámara de Gesell con los usuarios. Los restantes alumnos observan desde fuera de la cámara con la posibilidad de sugerir preguntas.

Elaboración del LEL.

A partir del conocimiento obtenido en la entrevista con el usuario, producir el LEL

Corrección por los docentes.

Corrección del trabajo realizado por los docentes de la cátedra.

Revisión de las buenas prácticas, depuración de los símbolos.

Revisión de lo desarrollado y ajuste al estándar derivado del documento de buenas prácticas. En cuanto a esto debe tenerse en cuenta que el proceso de evaluación por los docentes puede tolerar errores (que reducirán la calificación del alumno). Sin embargo la construcción del modelo del dominio no puede tolerar esos errores. Ese es el cometido de esta etapa

Informe de investigación.

Producción de un documento detallado con la recopilación de la experiencia, los productos obtenidos, los resultados alcanzados y las conclusiones del proceso.

Formación de recursos humanos

En el proyecto participan todos los docentes de la cátedra en diferente medida. En la actividad final de evaluación de las entrevistas mediante los videos y las minutas participan junto con los investigadores en el mismo nivel. Ello se integra con el plan de incrementar la participación de los docentes en la actividad de investigación.

Los estudiantes tendrán una doble participación en el proceso: como parte del objeto de estudio y en el cierre del proyecto se hará una evaluación mediante un cuestionario.

Referencias

- [1] A. Oliveros, J. Zuñiga, R. Wehbe, S. Rojo, and J. Rousselot, "Requerimientos para Aplicaciones Web," presented at the Workshop de Investigadores en Ciencias de la Computación, Rosario, Argentina, 2011.
- [2] A. Oliveros, J. Zuñiga, R. Wehbe, S. Rojo, and S. Martinez, "La enseñanza de elicitación de requerimientos," in *Actas del Congreso Argentino de Ciencias de la Computación (CACIC2012)*, Bahía Blanca, Argentina, 2012.
- [3] K. Wiegers, *Software Requirements*, 2nd ed. Microsoft Press, 2003.

- [4] A. Oliveros, J. Zuñiga, S. Corbo, S. Rojo, P. Forradelas, and S. Martinez, “Enseñanza de técnicas de elicitación de requerimientos,” in *Libro de Actas de Congreso Argentino de Ciencias de la Computación (CACIC2013)*, Mar del Plata, 2013.
- [5] A. Oliveros, J. Zuñiga, R. Wehbe, S. Rojo, and S. Martinez, “Enseñanza de elicitación de requerimientos,” presented at the WICC 2012 - XIV Workshop de Investigadores en Ciencias de la Computación, Posadas - Misiones, 2012.
- [6] S. Martinez, A. Oliveros, J. Zuñiga, S. Corbo, and P. Forradelas, “Aprendizaje de la elicitación y especificación de requerimientos,” in *XX Congreso Argentino de Ciencias de la Computación*, La Matanza, Argentina, 2014.
- [7] L. Antonelli, G. Rossi, J. C. S. do P. Leite, and A. Oliveros, “Buenas prácticas en la especificación del dominio de una aplicación,” in *Memorias del XVI Workshop de Ingeniería en Requisitos WER 2013*, Montevideo, Uruguay, 2013, pp. 80–92.
- [8] L. Antonelli, J. C. S. do P. Leite, G. Rossi, and A. Oliveros, “Deriving requirements specifications from the application domain language captured by Language Extended Lexicon,” in *Proceedings of Workshop on Requirements Engineering*, Buenos Aires, Argentina, 2012.
- [9] P. Runeson, “Using Students as Experiment Subjects? An Analysis on Graduate and Freshmen Student Data,” in *Proceedings of the 7th International Conference on Empirical Assessment in Software Engineering*, Keele, UK, 2003, pp. 95–102.

***Framework* para la Construcción de Estrategias Didácticas para la Enseñanza Inicial de la Programación de Computadoras**

Ana M. Piccin, Diana Cicinelli

Teoría de la Computación / Facultad de Ingeniería y Tecnología de la Información /
Universidad de Belgrano/CONICET
ana.piccin;diana.cicinelli@comunidad.ub.edu.ar

Resumen

El proyecto que se describe tiene por objetivo principal la definición de un *framework* de herramientas didácticas que facilite la construcción de estrategias de enseñanza para los cursos introductorios de programación. Entendemos por herramientas didácticas a todo concepto, técnica, instrumento o procedimiento que permita hacer efectiva la enseñanza. Por estrategia de enseñanza, a la planificación específica, que hace el docente, del procedimiento a seguir y herramientas a utilizar para lograr los objetivos pedagógicos.

El *framework* en cuestión estará compuesto por herramientas didácticas, tanto conceptuales como tecnológicas, un conjunto de requerimientos para el diseño de nuevas herramientas y ejemplos de su aplicación para el diseño de estrategias de enseñanza de contenidos y desarrollo de competencias específicos.

De entrevistas a docentes, se obtendrán descripciones de estrategias y métodos que serán categorizados y organizados en estilos de enseñanza, así como recursos conceptuales y tecnológicos utilizados en la práctica de la enseñanza; se elicitarán requerimientos para el diseño de nuevas herramienta. Los resultados permitirán construir un modelo conceptual sobre el que se fundará el diseño del *framework*.

Se construirá un prototipo que permita validar los resultados de este proyecto.

Palabras clave: enseñanza de la programación; elicitación de requerimientos; *framework*; herramientas didácticas; modelo conceptual.

Contexto

Este proyecto se incluye en la línea de investigación institucional de la Universidad de Belgrano “Abordaje de problemas de la educación” y se encuadra en área Teoría de la Computación, Programación y Lenguajes de la Facultad de Ingeniería y Tecnología Informática. El proyecto está fuertemente vinculado con las carreras de grado, Ingeniería en Informática y Licenciatura en Sistemas de Información, y con la Tecnicatura en Programación.

La Facultad de Ingeniería y Tecnología informática se ha propuesto formar un laboratorio para el estudio de herramientas para la enseñanza de la programación en los cursos introductorios, habida cuenta de la importancia del conocimiento de esta disciplina y de las competencias asociadas para un buen desempeño académico y profesional y de las dificultades que su aprendizaje ofrece a los estudiantes. El proyecto que se

presenta es el punto de inicio de este proceso.

1. Introducción

Las publicaciones en congresos especializados sobre fracasos en el aprendizaje [McC2001] [Boo2001], y dificultades en la enseñanza [Rob2003] son numerosas. Del material revisado rescatamos las siguientes cuestiones:

- *Con respecto a la calidad de la publicación científica y de sus resultados:*

Los estudios sobre la programación en cuanto a disciplina, o que investigan esta disciplina desde la tarea docente, son notablemente escasos y de orden biográfico. [Ber2010] [Pea2007]

Los trabajos académicos sobre el tema no ofrecen los resultados esperados: “[...] a pesar de que la comunidad docente ha ensayado distintas estrategias, los resultados no cumplirían con las expectativas”. “[...] es posible concluir que cuanto más se cambian las cosas, más permanecen iguales”. [Ast2004] Y la rigurosidad de los estudios está frecuentemente puesta en duda. “[...] los docentes suelen basar sus argumentaciones, con respecto a las dificultades en la enseñanza, sobre conjeturas y anécdotas y sus propuestas difícilmente llegan a constituirse en investigaciones genuinas”. [Lis2003],

- *Con respecto a la influencia que pueda tener la publicación científica sobre la práctica docente*

“La mayoría de los académicos de ciencias de la computación llevan una doble vida, la de la investigación y la de la enseñanza, según dos mentalidades diferentes”. [Lis2006] En referencia a que los docentes publican, asisten a congresos, intercambian puntos de vista, construyen unos en función de las publicaciones de los otros. Pero que, en la práctica docente, en la enseñanza, se

orientan por sus opiniones y teorías personales, y que en, las discusiones especializadas, tienden a expresarse desde una “posición personal, a la que se llegó por intuición e introspección”. [Lis2006]

La didáctica general ya ha descrito este último fenómeno, que es común a todos los docentes: los docentes fundamentamos nuestra práctica en nuestras experiencias personales, y en las apreciaciones que tenemos de éstas. Se tratan de experiencias de toda la vida, además de la experiencia recogida durante su formación docente. La enseñanza está plagada de tomas de decisiones y cuando una decisión debe tomarse de modo inmediato, sin que pueda mediar un tiempo de necesaria reflexión, nuestros preconceptos y teorías implícitas se impondrían sobre la teoría. [Bas2013] [Rol2009] Lo mismo aplica al diseño estrategias para la enseñanza.

La didáctica no es una disciplina monolítica. Ofrece un conjunto de métodos y hace recomendaciones de aplicación dependientes del contexto en que se esté enseñando. Y a ese contexto pertenecen también las teorías implícitas de los docentes, su formación pedagógica, los alumnos, los conocimientos y las competencias en juego. El docente diseñará diferentes estrategias didácticas teniendo en cuenta estas variables. Para cada oportunidad una estrategia, para cada estrategia un conjunto de métodos y herramientas. [Cam2013]

En un contexto en el que los problemas de la enseñanza son revisados reiteradamente en busca de alternativas, la Joint Task Force for Computing Curricula, en la formulación de recomendaciones para la enseñanza inicial de la programación, induce a analizar la realidad desde otra perspectiva: “Notamos que más que algún paradigma o lenguaje en

particular haya salido favorecido a lo largo del tiempo [como resultado de los trabajos académicos sobre enseñanza de la programación], la última década sólo logró ampliar la lista de lenguajes de programación que en la actualidad son utilizados exitosamente en los cursos introductorios”. [The2013] A continuación, el informe enumera y presenta como posibles a distintos métodos, descritos o mencionados en los aportes académicos. los acepta como aplicables. Son estrategias y métodos utilizados por los docentes. Y los docentes conciben la enseñanza de la programación de distintas maneras. Según este informe [The2013], temas tan discutidos como el paradigma, o el lenguaje para la enseñanza inicial de la programación, no influirían significativamente sobre la enseñanza, sino cómo y de qué manera se utilizan para construir estrategias para la enseñanza efectiva, exitosa, en los cursos introductorios. No son fines en sí mismos, son herramientas. Los recursos didácticos existen, los docentes elegirán el que mejor se adapte a sus circunstancias. Esta es la idea fuerza que guía este proyecto.

2. Líneas de Investigación, Desarrollo e Innovación

En [The2013] se ofrece un ángulo diferente para el abordaje del tema. Se trata de tomar al conjunto de supuestos y teorías, desde las cuales los docentes de programación planifican la enseñanza, como un dato de la realidad. Recurrimos al campo de la Didáctica para avanzar en el de la Ingeniería de Software. Un análisis metódico de las formas en que los docentes ven la enseñanza de la programación dará por resultado un modelo conceptual que permitirá definir el perfil del *framework* que se desea construir. Existen estudios de este tipo sobre docentes de

tecnología en general, pero no de programación en particular. La especificidad de este trabajo permitirá la formulación de herramientas sobre contenidos particulares de la disciplina con características de diseño que permitirán adecuarse a las diversas modalidades de enseñanza vigentes. Este conocimiento será de utilidad para investigaciones sobre didáctica y gestión educativa, en particular para carreras de Informática.

El diseño del *framework* y la construcción de herramientas constituyen el segundo eje de desarrollo: la Ingeniería de Software. Esta línea permite arraigar al proyecto, con agilidad, en las carreras de grado y tecnicaturas que ofrece la Facultad de Ingeniería y Tecnología Informática. Favorecerá la formación de estudiantes avanzados en prácticas de investigación. La necesidad de incluir o mejorar herramientas en el *framework* por lo menos dará lugar a trabajos finales de carrera.

Los resultados de este estudio, que podrán ser aplicados en el aula, también servirán para fundamentar proyectos para la mejora de la programación

3. Resultados Obtenidos/Esperados

Objetivo General:

Formular un *framework* orientado a facilitar el diseño de estrategias para la enseñanza de la programación, en particular en los cursos introductorios. Este *framework* deberá tener por lo menos las siguientes características: a) Un modelo conceptual de la enseñanza de la programación, b) un conjunto de requerimientos para la construcción de herramientas, c) recomendaciones para el uso de dichas herramientas asociadas a modalidad de enseñanza preferida por el docente y las competencias o contenidos en cuestión.

Resultados previstos

La ejecución de las sucesivas etapas del proyecto proveerán los siguientes resultados:

1. Descripción de distintas modalidades de enseñanza de la programación.
2. Modelo conceptual para la enseñanza inicial de la programación que incluya un repertorio de enfoques de la enseñanza.
3. Repertorio de los requerimientos identificados para el diseño de aplicaciones informáticas didácticas específicas para la programación.
4. Diseño del núcleo para el desarrollo del *framework* didáctico con la siguiente estructura: a) modelo conceptual, b) guía didáctica, incluye herramientas y su aplicación y c) conjunto de requerimientos para el diseño de aplicaciones didácticas.
5. Prototipo.

Se ha revisado la literatura específica en forma sistemática. El equipo de trabajo está realizando la clasificación y depuración de los textos. Se está obteniendo la información necesaria para la formulación de los instrumentos de recolección de datos.

Los resultados que se planifica obtener redundarán en la constitución de las bases para un *framework* dinámico. Ni las herramientas, ni los métodos ni las estrategias se agotan en este estudio, por lo que éste abrirá el campo a nuevas investigaciones. Entre los futuros desarrollos que podrán derivar de este estudio consideramos:

- Desarrollo de *frameworks* para diferentes áreas del conocimiento.
- Normas de calidad para la producción de herramientas didácticas.

- Reformulación/adaptación de herramientas a nuevos contextos tecnológicos.
- Estudios orientados a actualizar la enseñanza de la programación.

4. Formación de Recursos Humanos

Equipo de investigación:

CC Ana Piccin, docente investigadora, Prof. Titular, Directora del proyecto.

CC Diana Cicinelli, docente investigadora Prof. Adjunta I, integrante de equipo, colaboradora de investigación.

La Prof. Piccin es docente de la asignatura Lenguajes de Programación de las carreras de grado, Licenciatura en Sistemas e Ingeniería en Informática y posee formación de posgrado en Gestión Educativa. Ha participado como expositora en los Seminarios de Metodología organizados por la unidad académica.

La Prof. Cicinelli es docente en los cursos iniciales de programación de la Licenciatura en Sistemas y la Ingeniería en Informática y de la Tecnicatura en Programación de Computadoras.

Otros colaboradores:

Se ha solicitado un ecario de tiempo parcial para incorporar como ayudante de investigación.

La duración del proyecto es de dos años. El plan de trabajo prevé la incorporación de tesis de grado a partir del segundo semestre. Se prevé la formación en investigación de seis estudiantes, que estén desarrollando su trabajo final de carrera, a lo largo del proyecto.

Planificación de cursos:

Son cursos dirigidos a estudiantes de las carreras, docentes de grado y de

enseñanza media y para desarrolladores de software.

- “Desarrollo de habilidades docentes: aplicaciones y tecnologías de apoyo para la enseñanza”.
- “La enseñanza introductoria de la programación: Cómo es concebida por los docentes de educación superior a la enseñanza de la programación”.
- “Estrategias, técnicas y métodos utilizados en la enseñanza de la programación”.
- “Taller para la elicitación de requerimientos orientado al diseño de aplicaciones didácticas”.
- Taller para el desarrollo de aplicaciones didácticas para la enseñanza de la programación.
- Diplomatura: “Desarrollo de aplicaciones didácticas”.

5. Bibliografía

- [Ast2004] Astrachan, Owen (2004) “Why I Care about Programming and How to Teach It”, en *SIGCSE'04*, Marzo 3-7, 2004, Norfolk, Virginia, USA , (235). Extraído de <http://www.acm.org> el 10 de noviembre de 2008.
- [Bas2013] Basabe, Laura y Colls, Estela (2013) “La enseñanza”, en Camilloni, A. (Comp.) *El Saber Didáctico*, Paidós, Buenos Aires.
- [Ber2010] Berglund, A., Lister, R (2010), “Introductory Programming and the Didactic Triangle”, en *Conferences in Research and Practice in Information Technology*, Vol. 103. Tony Clear and John Jamer Eds.
- [Boo2001] Booth, Shirley (2001), “Learning Computer Science and Engineering” in *Context, Computer Science Education*, Vol. 11, Nor. 3, pp. 160-188.
- [Cam2013] Camilloni R. W. de (2013) “Los profesores y el saber didáctico”, en Camilloni, A. (Comp.) *El saber didáctico*, Paidós, Buenos Aires.
- [Lis2003] Lister, Raymond (2003) “A Research Manifesto, and the Relevance of Phenomenography”, en *Inroads:The SIGCSE Bulletin*, Vol. 35, Num. 2, (15-16).
- [Lis2006] Lister, R., Berglund, A., Clear, T., Bergin, J., Garvin Doxas, K., Hanks, B., et al. (2006). “Research Perspectives on the Objects-Early Debate”. *SIGCSE Bulletin Inroads*, 38(4), 173 - 192
- [McC2001] McCracken, Michael et al (2001) *A multi-national, multi-institutional study of assessment of programming skills of first-year CS students*, ACM SIGCSE Bulletin, v.33 n.4 Extraído de <http://www.acm.org> el 10 de noviembre de 2008.
- [Pea2007] Pears, A. et al (2007) “A Survey of Literature on the Teaching of Introductory Programming” en *ITiCSE-WGR '07: Working group reports on ITiCSE on Innovation and technology in computer science education*, SIGCSE Bulletin Volume 39 Issue 4, December 2007
- [Rob2003] Robins, A. et al (2003) “Learning and Teaching Programming: A Review and Discussion”, en *Computer Science Education 2003*, Vol. 13, No. 2, 137–172.
- [Rol2009] Rolandsson, L. (2009) *Teachers' perceptions about learning programming*, The Netherlands, August 2009.
- [The2013] The Joint Task Force on Computing Curricula (2013) *Computer Science Curricula 2013 Final Report*, Association for Computing Machinery IEEE-Computer Society

Grupo de Investigación para el Desarrollo de Videojuegos dentro de Carreras de Informática

Romina Stickar, Rodrigo René Cura, Nahuel Defossé, Carlos Buckle

Departamento de Informática, Facultad de Ingeniería,
Universidad Nacional de la Patagonia San Juan Bosco.

Puerto Madryn, Chubut, Argentina. +54 280-4883585 – Int. 116

romistickar@gmail.com, rodrigo.renecura@gmail.com, nahuel.defosse@gmail.com,
carlos.buckle@gmail.com

RESUMEN

Este proyecto busca consolidar el recién formado Grupo de Desarrollo de Videojuegos (GDV) de la Facultad de Ingeniería de la Universidad Nacional de la Patagonia San Juan Bosco, sede Puerto Madryn (UNPSJB-PM). El mismo tiene como objetivo general estudiar la forma de mejorar y aplicar la adquisición de conocimientos relacionados a las áreas de Algoritmos y Lenguajes, Ingeniería de Software y Teoría de la Computación mediante el desarrollo de juegos, entendiendo a éste como un proyecto integral de construcción de software.

Palabras clave: Videojuegos, juegos serios, gamification, interfaces no convencionales.

CONTEXTO

Este grupo de investigación se lleva a cabo dentro de la Facultad de Ingeniería de la UNPSJB - PM.

Actualmente, el grupo está formado por alumnos y docentes que participan de forma voluntaria en un esquema de reuniones semanales presenciales para la discusión de objetivos, planificación de

tareas y formación en las técnicas y herramientas de desarrollo. Además, el grupo funciona como un núcleo para todas las actividades relacionadas a los videojuegos realizadas en la universidad.

Tanto el equipo docente como varios de los alumnos participantes ya cuentan con experiencia en actividades afines al desarrollo de videojuegos, entre las que se destacan: El vínculo entre la universidad y la escuela secundaria, por medio de los talleres de programación de videojuegos que se dictan desde 2013 [Barry 2014]. En 2016 se comenzó a dictar el taller "Alice/Sadosky" en relación con la Secretaría de Ciencia y Técnica de la provincia de Chubut y se mantiene una participación constante de la universidad en eventos abiertos con el fin de difundir las carreras de informática.

1. INTRODUCCIÓN

La formación de este grupo se debe a la necesidad de los alumnos en contar con un espacio para investigar e implementar videojuegos. A lo largo de varios años, los mismos han manifestado la falta de integración de esta rama de la computación con los conocimientos formales presentados en la carrera.

Los videojuegos fomentan la motivación, el aprendizaje de contenidos y tareas, los procedimientos y destrezas y las actitudes

como la toma de decisiones y cooperación [Marques 2000].

Hay varios enfoques que serán fundamentales en nuestra investigación.

Entre ellos están:

Juegos serios: el término juego serio tiene su inicio en la década del 60 por Clark Abt [Abt 1970], que hacía referencia a juegos que simulaban eventos de la Primera Guerra Mundial y recreaban estrategias de guerra dentro del aula. Actualmente juego serio se refiere a videojuegos cuya finalidad es la formación y educación, el entrenamiento de habilidades específicas de procesos complejos en un entorno simulado, controlado y lúdico.

Gamification: es el uso de técnicas, elementos y dinámicas propias de los juegos y el ocio en actividades no recreativas con el fin de potenciar la motivación, así como de reforzar la conducta para solucionar un problema u obtener un objetivo [Deterding 2011].

Interfaces no convencionales: con la aparición de nuevas tecnologías, las interfaces deben ser capaces de percibir el contexto en el que se encuentra el usuario y poder asistirlo de manera proactiva [Harris 2005]. El objetivo es poder lograr una mejor comprensión de las capacidades humanas (cognitivas, motoras, de comunicación y perceptivas), combinada con profundos avances en técnicas de síntesis y reconocimiento de lenguaje, visión por computadora, animaciones gráficas, e interfaces hápticas, entre otros. Los videojuegos son un excelente escenario para realizar este tipo de experiencias debido a que muchas plataformas ya han innovado en las formas con las cuales el jugador interactúa con el juego, por ejemplo, los acelerómetros de los dispositivos móviles, la kinect de XBox, pantallas táctiles, etc.

Según [Johnson & Johnson 1994] para que un grupo de jugadores trabaje de manera colaborativa, debe tener estas características:

- Interdependencia positiva.
- Interacción cara a cara.
- Responsabilidad individual y de grupo.
- Aprendizaje de habilidades sociales.
- Revisión del proceso del grupo.

Así como se presentan dentro del juego, consideramos que estas mismas pueden ser aplicadas en el desarrollo de los mismos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El grupo funciona como un espacio semi-estructurado donde los alumnos puedan desarrollar sus inquietudes relacionadas al diseño y desarrollo de videojuegos, capacitarse y dar continuidad a este tipo de actividades.

En particular, las principales líneas de trabajo a seguir son:

- Técnicas de ludificación (gamification) destinadas a aplicaciones educativas y ciencia ciudadana.
- Juegos serios.
- Interfaces no convencionales, como por ejemplo, Kinect, Leap Motion, sensores incorporados a dispositivos móviles como acelerómetro, giroscopio o interacción multitáctil, etc.
- Estudio de la experiencia del usuario (UX) en la interacción con

los juegos y la transferencia de la misma a aplicaciones no lúdicas.

- Estudio de diferentes plataformas de desarrollo Open Source.
- Estudio de técnicas y estrategias para la construcción de prototipos rápidos.

Además, como tareas transversales a estos ejes es requerido:

- Investigar formas de mejorar la gestión e integración de módulos de terceras partes para acelerar el ciclo de desarrollo con el fin de aumentar la productividad, potenciar la reusabilidad de código y mejorar, por lo tanto, la calidad del proceso y los resultados en general.
- Favorecer el trabajo colaborativo y cooperativo por medio de la aplicación práctica de metodologías modernas de desarrollo de software.
- Aplicar técnicas de integración continua (CI) y entrega continua (CD) para acelerar los ciclos de prueba de los juegos. La generación de entregables está asociada a los eventos de cambios en el código fuente, sin necesidad que los desarrolladores o artistas tengan que realizar operaciones complejas, inherentes al desarrollo multiplataforma.

- Fomentar la expresión artística y creativa en la construcción de software más allá de los requerimientos funcionales.

3. RESULTADOS OBTENIDOS/ESPERADOS

El grupo ha finalizado una primera etapa en la que logró una participación continua de los integrantes y la formación de un equipo estable. Ya se avanzó en el estudio y la incorporación de metodologías ágiles (SCRUM) y actualmente se está evaluando el desempeño del grupo y la manera de adaptar estas formas de trabajo a la naturaleza del mismo. También se logró un progreso significativo en la formación de los integrantes en los aspectos relacionados específicamente al desarrollo de videojuegos, y se espera en la etapa actual formalizar las estrategias de aprendizaje y facilitar la incorporación de nuevos integrantes en un nivel mínimo de capacidades mediante un taller de inducción. En los objetivos planificados a mediano plazo se espera poder concretar un conjunto de juegos básicos y publicar los mismos en la web.



Figura 1: Cubit, primer juego desarrollado por GDV

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está integrado por 3 docentes del Departamento de Informática de la Facultad de Ingeniería de la UNPSJB sedes Puerto Madryn y Trelew y 5 alumnos de diferentes años de la carrera de Licenciatura en Informática.

Es intención del proyecto que aquellos alumnos que estén pronto a recibirse puedan realizar sus tesinas en temáticas afines al mismo, así como trabajo finales de materias.

Se diseñará como parte del proyecto un taller de inducción destinado a proveer los conocimientos mínimos a los nuevos integrantes.

5. BIBLIOGRAFÍA

1. Barry, Stickar, Defosse, Salvatierra. (2014). Programación de videojuegos en el nivel medio.
2. MARQUÉS, P. (2000), "Las claves del éxito", Cuadernos de Pedagogía
3. Abt, C.C. (1970) Serious Games. Viking, New York.
4. (Harris, 2005) Harris, R., Voice Interaction Design: Crafting the New Conversational Speech Systems. Morgan Kaufmann Series in Interactive Technologies ISBN 1- 55860.768-4
5. Johnson, D. W. & Johnson, R.T. (1994). Sharan, S. (ed.). Learning Together. Handbook of Cooperative Learning Methods. Connecticut: Greenwood Press.
6. Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). ACM.

Identificando Fortalezas y Debilidades en la Formación de Graduados de Analista y Licenciatura en Computación para Mejorar la Enseñanza en el Marco de Nuestras Asignaturas

Ariel Arsaute, Fabio Zorzan, Marcela Daniele, Marcelo Uva, Daniela Solivellas, Mariana Frutos, Franco Bursatti, Guillermo Frascchetti y Sandra Angeli

Departamento de Computación, Facultad de Ciencias Exactas, Físico-Químicas y Naturales, Universidad Nacional de Río Cuarto

Ruta 36 Km. 601 –CP 5800 - Río Cuarto – Córdoba - Argentina Tel. (0358) 4676235

{arsaute, fzorzan, marcela, uva, dsolivellas, mfrutos, fbrussatti, gfrascchetti, sangeli}@dc.exa.unrc.edu.ar
asignaturas involucradas, están

Resumen

La industria del software crece y se expande a nuevos ámbitos permanentemente, generando nuevas metodologías de desarrollo y de trabajo, adopción de herramientas Open Source para automatizar y mejorar las actividades de gestión, planificación, diseño, implementación y automatización de pruebas que hacen a la producción del software.

Se hace necesario entonces una constante revisión y reflexión sobre las prácticas de enseñanza y los procesos de aprendizaje que se llevan a cabo en las carreras de computación de la Universidad.

Este trabajo involucra a docentes y asignaturas compartidas por el último año de Analista en Computación, y tercer año de la Licenciatura en Ciencias de la Computación. En ellas se abordan contenidos curriculares desde el punto de vista teórico, práctico, y actitudinal, que contribuyen al desarrollo de estrategias y habilidades para el trabajo grupal e individual en la industria del desarrollo de software.

Nos interesa conocer si los contenidos, actividades, estrategias y tecnologías que se proponen en las

contribuyendo a facilitar u obstaculizar la integración laboral de nuestros graduados en las empresas de desarrollo de software, para consolidar buenas prácticas de enseñanza y pensar en nuevas innovaciones que atiendan a problemas y necesidades actuales de formación.

Palabras clave: Ingeniería de Software, Universidad-Industria, Graduados, Inserción Laboral

Contexto

La línea de investigación presentada en este trabajo se desarrolla en el marco de la convocatoria de financiamiento a Proyectos de Innovación e Investigación para el Mejoramiento de la Enseñanza de Grado (PIIMEG) de la secretaría de Ciencia y Técnica de la Universidad Nacional de Río Cuarto. Para la formulación del proyecto se tuvieron en cuenta las pautas presentadas en [1] y [2].

1. Introducción

El constante crecimiento y expansión de la informática en los más

diversos ámbitos, deja a la luz la imperiosa necesidad de disponer de recursos humanos calificados [3].

La industria de desarrollo de software crece y se expande a nuevos ámbitos permanentemente, esto genera nuevas metodologías de desarrollo, nuevas modalidades de trabajo (como por ejemplo desarrollo outsourcing y freelance), adopción de herramientas Open Source para automatizar y mejorar las actividades de administración, gestión, planificación, diseño, implementación, automatización de pruebas, seguimiento y control de las tareas que hacen a la producción del software [4].

En este contexto, para que una empresa resulte competitiva necesita adaptarse a los cambios constantemente, por consiguiente demanda profesionales que puedan integrarse rápidamente a equipos con diferentes características, con una sólida formación que les permita adaptarse rápidamente a: cambios tecnológicos, diferentes modalidades de trabajo, nuevas metodologías de desarrollo, diferentes lenguajes tanto de codificación como de modelado, etc.

Paralelo a esto, la relación entre la Universidad y las empresas dedicadas al desarrollo de software es cada vez más estrecha. Pequeñas, medianas y grandes empresas se acercan continuamente a las universidades en busca de graduados en Ciencias de la Computación con perfiles profesionales que faciliten una rápida integración y adaptación a los procesos de trabajo. Por su parte, la universidad demanda a las empresas espacios de prácticas contextualizadas, de pasantías para la formación de sus estudiantes.

A partir de todo lo descrito se vuelve necesario una constante revisión y reflexión sobre las prácticas de enseñanza y los procesos de aprendizaje que llevan a cabo en las carreras de

computación que ofrece nuestra Universidad.

En ese marco se plantea este trabajo, que involucra a tres asignaturas compartidas por el último año de la carrera Analista en Computación, y por el tercer año de la Licenciatura en Ciencias de la Computación, y que están a cargo de los equipos docentes participantes en el proyecto. En estas asignaturas se abordan contenidos curriculares desde el punto de vista teórico, práctico, y actitudinal, que contribuyen: al desarrollo de estrategias y habilidades para el trabajo grupal e individual, y para el liderazgo; al conocimiento y aplicación de metodologías de desarrollo de software [5] [6]; a la selección y uso de herramientas que da soporte y automatizan diversas actividades del proceso desarrollo; a la comprensión y uso de diferentes lenguajes de modelado y de programación; al desarrollo de la comunicación oral y escrita. El equipo docente, a partir de la participación en diferentes proyectos de innovación e investigación para el mejoramiento de la enseñanza de grado (PIIMEG) a lo largo de los últimos 10 años, ha propuesto cambios para mejorar la enseñanza en las asignaturas, vinculados a formas de abordar los contenidos, incorporación de nuevos lenguajes y herramientas, tipo de actividades a desarrollar, bajo la hipótesis de que estos cambios contribuyen a la construcción de perfiles profesionales que pueden adecuarse a las necesidades cambiantes del mercado laboral.

Luego de esa trayectoria de innovación recorrida, y teniendo en cuenta los actuales escenarios de inserción laboral de nuestros graduados, se nos plantea la inquietud de conocer cómo las propuestas educativas de las asignaturas involucradas en este proyecto, están

contribuyendo a la formación de los estudiantes, y a facilitar su integración en las empresas de desarrollo de software

2. Líneas de Investigación, Desarrollo e Innovación

Con este trabajo nos interesa conocer si los contenidos, las actividades, las estrategias, y las tecnologías que se proponen en las asignaturas involucradas, están contribuyendo a facilitar u obstaculizar la integración laboral de nuestros graduados en las empresas de desarrollo de software.

Nos interesa conocer la mirada de nuestros egresados sobre: cuáles son los aprendizajes que realizaron durante el cursado de las asignaturas que ellos consideran que le han facilitado su integración y desempeño laboral; y sobre cuáles son las necesidades de formación que han evidenciado a partir de las actividades que realizan, y que actualmente desde las asignaturas en cuestión o de otras no estamos cubriendo.

También nos interesa recuperar la voz de los responsables de las áreas o empresas donde nuestros egresados trabajan, para conocer su mirada sobre: cuáles son las

fortalezas y las debilidades que ellos identifican actualmente en la formación de nuestros egresados, a partir de su desempeño; y sobre cuáles creen que son los aprendizajes que un estudiante de computación debe hacer durante su formación de grado, para poder integrarse y desarrollarse laboralmente en el ámbito de la industria del software.

Por otra parte, nos interesa conocer la opinión de otros docentes de las carreras, en particular de aquellos que participan en asignaturas con articulaciones

explícitas o implícitas en sentido vertical u horizontal en el plan de estudio, sobre lo que se debería

abordar en las asignaturas involucradas de acuerdo al perfil profesional que ellos entiende deberíamos contribuir a formar desde nuestra universidad.

Creemos que poder reunir esta información, analizarla, confrontar miradas, y reflexionar sobre lo que hemos estado y estamos haciendo, nos va ayudar a consolidar buenas prácticas de enseñanza, a pensar en nuevas innovaciones para dar solución a problemas que podamos detectar, y a identificar tendencias en necesidades de formación que de cara al futuro nuestras carreras deberían considerar.

Este trabajo se enmarca dentro del ámbito de la investigación cualitativa. Se propone realizar una investigación diagnóstica para la educación desde un enfoque de evaluación comprensiva. Para su desarrollo nos proponemos realizar las siguientes acciones:

- Búsqueda y lectura de algunos materiales que nos aporten elementos para el tipo de investigación que se quiere hacer.
- Sistematización de información de contactos de graduados y de responsables de las empresas en las que tenemos conocimiento que trabajan nuestros egresados, para conformar la muestra con la que se trabajará en este proyecto.
- Elaboración de los instrumentos para la recolección de datos y aplicación [7]: una encuesta semiestructurada, con preguntas de respuesta abiertas y de respuestas cerradas para los graduados que se implementará mediante un formulario web que se hará llegar por correo electrónico; un guion de entrevista para los responsables de las empresas, que en la medida de lo posible se realizará de

manera personal, y en caso de imposibilidad por distancia geográfica, a través de un chat o videoconferencia; y un guion de entrevista para otros docentes que se desempeñan en asignaturas que se pueden relacionar con las materias involucradas en el proyecto, que se realizarán de manera personal en la misma universidad.

- Sistematización de las respuestas obtenidas y análisis de las mismas: se elaborarán documentos escritos conteniendo toda la información obtenida, y se organizarán tratando de identificar algunas categorías que faciliten su análisis.
- Validación de la información: se hará un proceso de triangulación de las miradas de los distintos actores (graduados, responsables de empresas, otros docentes de las carreras) para identificar coincidencias, divergencias, tendencias.
- Elaboración de proyecciones: a partir del conocimiento obtenido con la realización del proyecto se pensará sobre posibles acciones de innovación en las asignaturas involucradas en el PIIMEG.
- Escritura de artículos e informes para socializar los resultados del proyecto.

3. Resultados y Objetivos

Esta propuesta tiene como objetivos:

- Conocer cómo las propuestas de enseñanza de las asignaturas involucradas, pueden estar contribuyendo a facilitar u obstaculizar la integración laboral de nuestros graduados en las empresas de desarrollo de software.
- Indagar sobre las percepciones que tienen los graduados de nuestras carreras respecto de aprendizajes que realizaron en la universidad y que les facilitan la adopción de

nuevas metodologías de trabajo, de nuevas tecnologías, el trabajo individual y en equipo, la resolución de problemas, y la realización de nuevos aprendizajes.

- Identificar las principales problemáticas a las que se enfrentan nuestros egresados al momento de ingresar a una empresa de desarrollo de software, que puedan estar vinculadas a ausencias en su formación de grado.
- Conocer las necesidades actuales de formación desde la perspectiva de los responsables de las empresas de software.
- Sistematizar información que pueda constituirse en un aporte para la toma de decisiones en ámbito de las Comisión Curricular Permanentes de las carreras de computación.
- Movilizar a los equipos docentes a realizar una evaluación permanente de los procesos de formación.

Pensar en posibles acciones para enriquecer la enseñanza y ofrecer una sólida formación académica que le permita a nuestros egresados adquirir nuevos conocimientos y desarrollarse en su ámbito laboral

4. Formación de Recursos Humanos

Durante el desarrollo de esta línea de investigación han logrado obtener el título Magister en Ingeniería de Software integrantes del grupo de trabajo. Otros integrantes están actualmente trabajando en su tesis de Magister y Doctorado.

También, se están formando ayudantes de segunda en las asignaturas de Análisis y Diseño de Sistemas, Ingeniería de Software, Base de Datos.

Los temas abordados en esta línea de investigación brindan un fuerte aporte al

proceso de perfeccionamiento continuo de los autores de carreras de computación en Universidades Nacionales como del exterior

5. Bibliografía

- [1] Macchiarola, Viviana (2016) La investigación evaluativa de proyectos de innovación pedagógica. Seminario Taller: La formulación y evaluación de proyectos de innovación pedagógica. UNRC.
- [2] Manccini, Alicia y otros. (2016) Presentación del primer encuentro del Seminario taller de formación: La formulación y evaluación de proyectos de innovación pedagógica. UNRC
- [3] Cohen, Corina Eliana (2013) La problemática de la escasez de profesionales IT en el marco del creciente desarrollo de las tecnologías de la información. Estrategias empresariales para la captación, capacitación y desarrollo de los profesionales IT. Tesis de la Maestría en Recursos Humanos. Facultad de Ciencias Económicas. UBA.
- [4] Estayno, M. y Otros (2011). Evaluación de calidad de software, formación de recursos humanos y políticas públicas para la promoción de la industria del software en la región NEA. Universidad Nacional de Lomas de Zamora, Universidad Nacional del Nordeste, y Universidad Tecnológica Nacional de Resistencia.
- [5] Pressman, R. (2015) Software engineering apractitioner's approach. 8th Edition
- [6] Sommerville Ian. (2005) Ingeniería del Software, 7th edition, Addisson Wesley. Pearson Education.
- [7] Ander-Egg, Ezequiel (2003) Métodos y Técnicas de investigación IV. Técnicas para la recogida de datos e información. Grupo editorial Lumen Hvmanitas. Buenos Aires. Argentina.

Influencia de la Educación Inversa en el Aprendizaje y Adquisición de Competencias Transversales

Fernández, Teresa Ema; Aubin, Verónica Inés; Guatelli, Renata Silvia; Bellani, Marcela Elisabet; Blautzik, Leonardo José

*Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de la Matanza
Florencio Varela 1903 (B1754JEC)- San Justo, Buenos Aires, Argentina*

terefernandez62@gmail.com; vaubin@unlam.edu.ar; renata.guatelli@gmail.com;
mebellani@yahoo.com.ar; leoblautzik@gmail.com

RESUMEN

Este proyecto integra en un marco pedagógico las asignaturas de matemática discreta y programación. La integración no sólo se centra en los contenidos sino también en las competencias. Las competencias sociales, científico-técnicas y metodológicas, tienen una participación fundamental en el incremento de las habilidades del estudiante, en su desempeño académico y profesional. Propone la implementación de metodologías activas, siguiendo un modelo de aprendizaje centrado en el estudiante, con un enfoque pedagógico basado en crear situaciones que provoquen la necesidad de compartir y colaborar.

Este cambio de enfoque se logra con una estrategia innovadora en la UNLaM, basada en el uso de la clase invertida (FC), donde los estudiantes realizan fuera del aula las tareas menos activas, y se reserva para el trabajo en clase las actividades que requieren mayor participación e interacción. Los docentes, obtienen así un tiempo valioso para llevar a cabo otro tipo de técnicas que facilitan la adquisición de los conocimientos. Los materiales didácticos utilizados en este modelo, están basados en formatos multimedia e interactivos.

Se propone elaborar una propuesta didáctica basada en el modelo FC. Diseñar, desarrollar e implementar materiales didácticos que permitan la adquisición de competencias junto con los conocimientos académicos de cada asignatura.

Palabras Claves: *Flipped Classroom, trabajo colaborativo, metodologías activas, competencias.*

CONTEXTO

Esta investigación tuvo como punto de partida el Proyecto Estratégico de Ingeniería para Ciencias Básicas (PEICB) del Departamento de Ingeniería e Investigaciones Tecnológicas de la universidad Nacional de La Matanza (UNLaM), en el marco del Programa de fortalecimiento de las carreras de Ingeniería de la Secretaría de Políticas Universitarias (2012-2016).

En el PEICB se propuso el desarrollo de estrategias pedagógicas superadoras que favorezcan procesos de aprendizaje significativos y la incorporación de enfoques basados en la formación por competencias tendientes a una mayor retención del alumnado. Postulando a las estrategias activas de enseñanza (aprendizaje basado en problemas, aprendizaje por proyectos, trabajo colaborativo, entre otras) como metodologías pedagógicas centrales en el desarrollo de las asignaturas.

El tiempo de realización de este proyecto es de dos años, desde comienzos de 2016 a finales del 2017.

La implementación del proyecto se realiza desde el inicio del mismo y se seguirá desarrollando luego de la finalización de este trabajo de investigación.

INTRODUCCIÓN

Hace algunos años que las tendencias educativas a nivel mundial establecen como centro de la planificación y el desarrollo didáctico, el aprendizaje por competencias [1][2] y las metodologías activas [3][4][5].

Por otra parte es poco usual en la carrera de Ingeniería en Informática de la Universidad Nacional de La Matanza, la integración de asignaturas de diferentes áreas aun cuando existan temáticas compartidas. Por lo general se identifican estrategias de integración entre asignaturas correlativas.

El aprendizaje estandarizado en un sistema de enseñanza tradicional no es compatible con las exigencias del mundo globalizado. En contraposición, las metodologías activas promueven una enseñanza centrada en el estudiante capacitándolo en competencias y adaptándose a sus necesidades individuales. Conciben el aprendizaje como un proceso constructivo y no receptivo. Fomentan la participación y reflexión continua de los estudiantes a través de actividades que incentivan el diálogo, la colaboración, la cual conduce a un aprendizaje entre pares donde se promueve el intercambio de conocimientos entre compañeros [6][7].

El modelo de Flipped-Classroom o Clase Invertida (FC) [8] [9] favorece el cambio metodológico. Los estudiantes asumen una mayor responsabilidad sobre su propia educación. El profesor asume el rol de guía que facilita el proceso de enseñanza-aprendizaje, apoyándose en diferentes recursos tecnológicos y aplicando metodologías activas. Todos los involucrados en el proceso de enseñanza - aprendizaje deben asumir el cambio que implica esta nueva metodología, pero esto no es siempre fácil, ni automático. Puede generar resistencia dado que están habituados a una educación tradicional, donde el profesor es el responsable de enseñar y transmitir los conocimientos y el alumno es el receptor.

Aunque la idea de la FC es muy sencilla, su aplicación requiere que el docente, dedique suficiente tiempo a investigar e identificar en qué tema de cada clase funcionará mejor [10]. Asimismo necesita desarrollar nuevas

habilidades en el uso de tecnologías, generar materiales educativos para motivar y animar a los alumnos y maximizar el tiempo de clase [11][12]. Al trasladar una parte de la instrucción fuera de la clase, los estudiantes tendrán acceso a los contenidos de una forma más flexible, en cualquier tiempo y lugar [13], pero podrán ver comprometido su tiempo libre y sentir que ellos realizan todo el esfuerzo.

Metodología del trabajo desarrollado

El tipo de estudio que enmarca esta investigación es cuali-cuantitativo, dependiendo esto de cada etapa de desarrollo del presente proyecto.

Cualitativo en cuanto a la ponderación del diagnóstico realizado y aplicación de las distintas metodologías a aplicar.

Cuantitativo, al final del trabajo, donde mediante encuestas y otros instrumentos se medirán los resultados obtenidos.

También es una investigación-acción, ya que es participativa, colaborativa, se trabaja para mejorar las prácticas de los estudiantes, creando grupos autocríticos en todas las fases del trabajo, procediendo a cambios más grandes a medida que se va avanzando. Es un proceso sistemático de aprendizaje, orientado a la praxis, induce a teorizar sobre la práctica. Somete a prueba las prácticas, las ideas y las suposiciones. Así mismo implica registrar, recopilar, analizar los propios juicios, reacciones e impresiones en torno a lo que ocurre con los estudiantes y con los docentes. Se utilizarán los cursos de la cátedra de Matemática Discreta, Programación y Programación Avanzada, para aplicar las metodologías propuestas y medir sus resultados.

Los datos usados para el análisis, se obtendrán de los docentes a cargo de las cátedras, de sus alumnos y de los registros o evidencias que se obtengan de: diagnósticos previos, parciales, finales, guías de TTPP, proyectos grupales de aula y aplicaciones prácticas.

Las consultas serán realizadas en base a encuestas, formularios y análisis de las evidencias, paneles, entrevistas y otras técnicas de prospectiva.

El procedimiento para el análisis de los datos obtenidos se centrará en mediciones objetivas aplicando diferentes métricas, según cada cátedra involucrada y mapeándolas a las competencias seleccionadas.

2. LÍNEAS DE DESARROLLO

El presente proyecto de investigación, da continuidad a diferentes proyectos y trabajos realizados en el campo de la enseñanza en la educación superior.

Proyectos de investigación desarrollados en el DIIT de la UNLaM:

C135. Reorganización disciplinar y didáctica de Matemática Discreta para las carreras de ingeniería. (2011-2012).

C137. Uso de nuevas métricas orientadas a las competencias en la gestión curricular. (2012-2013).

C167. Mejoras en el proceso de enseñanza-aprendizaje de programación utilizando metodologías propias de la industria del software como caso particular de las metodologías activas. (2014 – 2015).

C191. Fortalecimiento de competencias transversales de trabajo en equipo. (2016-2017).

Matemática discreta: guía de estudio y ejercicios [Bellani Fernández ISBN/ISSN: 978-987-1635-84-9].

3. RESULTADOS OBTENIDOS/ESPERADOS

La primera etapa del proyecto de investigación consistió en el análisis y evaluación, considerando la facilidad de uso y versatilidad, de diferentes herramientas para elaboración de videos interactivos que permitan implementar la Flipped Classroom. Como resultado de este análisis se ha escrito el artículo [14] “Evaluación de herramientas para la creación de materiales didácticos interactivos para implementar la clase invertida”.

Un cuestionario en línea ha sondeado diferentes aspectos de la forma de trabajo de los alumnos dentro y fuera del curso. A partir de este sondeo se presentan metodologías para la asignación de roles y de manejo de conflictos dentro de los grupos de trabajo y de estudio. En base a esta encuesta de escribió el

artículo [15] “Fortalecimiento de Competencias de Trabajo en Equipo en la Universidad”

El artículo [16] “Congruencia y TIC” presenta los inicios de una experiencia de cátedra, sobre teoría de números, involucrando las TIC y el trabajo tradicional, utilizando el aprendizaje colaborativo y metodologías activas de aprendizaje. La misma, pretende marcar una línea de enseñanza, que intentará conseguir relacionar la asignatura con el entorno del estudiante, y que se mantenga en el tiempo, concordando con las innovaciones educativas. Con la finalidad de identificar, evaluar y analizar el grado competencial inicial de los alumnos recursantes de la asignatura Matemática Discreta para las distintas carreras de ingeniería se realizó una evaluación diagnóstica de competencias matemáticas específicas. A partir de los resultados, se propone diseñar un conjunto de actividades en clase, basándose en metodologías activas de enseñanza-aprendizaje. Esto se documentó en el artículo [17] “Análisis competencial inicial en Matemática Discreta”

Todos estos artículos fueron presentados en el CONAIIISI 2016.

Al final del segundo cuatrimestre del 2016, se realizó otra encuesta de opinión entre los estudiantes de Matemática Discreta, sobre la aplicación de Flipped en la unidad de Grafos y Digrafos. Los resultados obtenidos serán el punto de partida de la siguiente etapa del proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

Algunos de los docentes - investigadores alcanzaron los títulos de Magister en Informática, Licenciada en Matemática Aplicada, Especialistas en Educación Superior y TIC. Otros se encuentran desarrollando diversas carreras de postgrado.

Así mismo, se han dirigido varias tesinas, tesis de postgrado y actividades de cátedras relacionadas con las herramientas involucradas.

5. BIBLIOGRAFÍA

- [1] Biggs, J.B. (2005). Calidad del aprendizaje universitario. Madrid. Nancea
- [2] Perrenoud, P. (2005). Diez nuevas competencias para enseñar - *Educatio Siglo XXI*, 23.
- [3] Metodologías activas y corrección por rúbricas en el proceso de enseñanza-aprendizaje de programación Aubin, V.; Blautzik, L., Guatelli, R.; Pafundi, F.; Salica, M.; Nisi, Z.; Zanga, M.
- [4] Fernández, A. (2006): Metodologías activas para la formación de competencias. En *EDUCATIO SIGLO XXI*, N° 24. Revista de la Facultad de Ciencias de la Educación de la Universidad de Murcia. ISSN 1699-2105.
- [5] Alfaro Rocher, I., Apodaca Urquijo, P., Arias Blanco, J., García Jiménez, E., & Lobato Fraile, C. (2006). Metodologías de enseñanza y aprendizaje para el desarrollo de competencias: orientaciones para el profesorado universitario ante el Espacio Europeo de Educación Superior. M. de Miguel Díaz (Ed.). Madrid: Alianza editorial.
- [6] Martín-Moreno, Q. (2004) Aprendizaje colaborativo y redes de conocimiento. Actas de las IX Jornadas Andaluzas de Organización y Dirección de Instituciones Educativas, 55-70.
- [7] Panitz T. "Collaborative versus cooperative learning-a comparison of the two concepts which will help us understand the underlying nature of interactive learning". 2001.
- [8] Tourón, J., Santiago, R., & Diez, A. (2014). *The Flipped Classroom: Cómo convertir la escuela en un espacio de aprendizaje*. Grupo Océano.
- [9] Bishop, J. L., & Verleger, M. A. (2013, June). The flipped classroom: A survey of the research. In *ASEE National Conference Proceedings*, Atlanta, GA (Vol. 30, No. 9).
- [10] Gorman, M. (2012, July 18). Flipping the classroom...a goldmine of research and resources keep you on your feet. URL (último acceso 26/03/2017): <https://21centuryedtech.wordpress.com/2012/07/18/flipping-the-classroom-a-goldmine-of-research-and-resources-to-keep-you-on-your-feet/>
- [11] Landau, M. (2007) Análisis de Materiales Digitales. En Posgrado del Proyecto en Educación y Nuevas Tecnologías, Flacso Argentina
- [12] Odetti, V. (2013). El diseño de materiales didácticos hipermediales para los niveles medio y superior: experiencias incipientes en Argentina. I Jornadas de jóvenes investigadores en Educación, FLACSO-Argentina. URL (último acceso 26/03/2017): <http://www.pent.org.ar/institucional/publicaciones/disenomaterialesdidacticoshipermedialesparanivelesmediosuperiore>
- [13] Burbules, N. C. (2012). El aprendizaje ubicuo y el futuro de la enseñanza Ubiquitous Learning and the Future of Teaching. *Encounters*, 13, 3-14.
- [14] Aubin, V.; Guatelli, R.; Blautzik, L.; Fernández, T. E.; Bellani, M. E. (2016) - Evaluación de herramientas para la creación de materiales didácticos interactivos para implementar la clase invertida. CONAIISI 2016 - IV Congreso Nacional de Ingeniería en Informática / Sistemas de Información. - Universidad Católica de Salta (UCASAL). ISSN- 2347-0372. URL (último acceso 26/03/2017): <http://ucasal.edu.ar/conaiisi2016/book/memorias.html>
- [15] Blautzik L.; Cabrera J. L., Videla L., Aubin V., Guatelli R., Crispino J., Otegui C. y Zanga M. (2016) - Fortalecimiento de Competencias de Trabajo en Equipo en la Universidad: Uso de Nuevas Tecnologías y Manejo de Conflictos. CONAIISI 2016 - IV Congreso Nacional de Ingeniería en Informática / Sistemas de Información. - Universidad Católica de Salta (UCASAL).

ISSN- 2347-0372. URL (último acceso 26/03/2017):

<http://ucasal.edu.ar/conaiisi2016/book/memorias.html>

[16] Fernández, T. E.; Aubin, V.; Guatelli, R. (2016). Congruencia y TIC. CONAIISI 2016 - IV Congreso Nacional de Ingeniería en Informática / Sistemas de Información. - Universidad Católica de Salta (UCASAL). ISSN- 2347-0372. URL (último acceso 26/03/2017):

<http://ucasal.edu.ar/conaiisi2016/book/memorias.html>

[17] Bellani, M. E.; Aubin, V.; Guatelli, R. (2016). Análisis competencial inicial en Matemática Discreta. CONAIISI 2016 - IV Congreso Nacional de Ingeniería en Informática / Sistemas de Información. - Universidad Católica de Salta (UCASAL). ISSN- 2347-0372. URL (último acceso 26/03/2017):

<http://ucasal.edu.ar/conaiisi2016/book/memorias.html>.

La Tecnología como Vehículo de Articulación

Nivel Medio / Universidad

Fernanda Beatriz Carmona, Alberto Eduardo Riba, Fernando Emmanuel Frati, Claudia Isaia, Alejandro Cruz, Jorge Tejada, Matías Pérez, Patricia Manriques

Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina

{fbcarmona, ariba, fefrati, cisaia, acruz, jtejada}@undec.edu.ar, {mataguper20, patriciamanriques311995}@gmail.com

Resumen

Esta línea de I+D pretende fortalecer la articulación Universidad – Nivel Medio, estimulando el diseño de mecanismos que permitan alcanzar un diagnóstico compartido y planes de trabajo, destinados a disminuir la brecha que existe entre estos dos niveles.

Se propone una estrategia pedagógica / tecnológica como soporte de los procesos de enseñanza y aprendizaje a ser utilizados en las Instituciones Educativas (IE) de nivel medio para despertar vocaciones tempranas vinculadas a las carreras tecnológicas, utilizando la Robótica Educativa (RE), el aprendizaje colaborativo y por descubrimiento guiado y el estímulo basado en competencias. Los temas abordados son transversales a varias áreas, educación, algoritmos, lógica, programación y arquitectura de computadoras.

El desarrollo de esta línea conlleva, a la generación de otros proyectos relacionados con el uso de la tecnologías en la articulación Universidad - Nivel Medio considerando la amplia cobertura de la temática a desarrollar y la imperiosa necesidad de articulación existente entre estos niveles educativos, como así también, en la gestación de líneas de investigación relacionadas con nuevas estrategias didácticas en Robótica y Educación, la enseñanza y aprendizaje de la programación, la utilización y programación de microcontroladores en otros ambientes de trabajo y la utilización de lenguajes interpretados.

Palabras clave: Educación, capacitación, TIC, enseñanza-aprendizaje, robótica educativa, aprendizaje experimental, aprendizaje inductivo, programación, juego.

Contexto

Esta línea de I/D/I corresponde al desarrollo de proyectos que fortalecerán la inserción de la UNdeC en la comunidad y especialmente su articulación con los demás niveles educativos. Incluye el proyecto “La robótica como introducción a la formación tecnológica” SPU 2013-2015 y la colaboración de otras instituciones del país y del extranjero a través de los proyectos “Red para la Integración de Universidades en el uso de TIC para la Inclusión en la Educación Superior”, año 2013-2014, el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” 2016-2017, ambos aprobados en la Convocatoria a Redes Internacionales, SPU, el proyecto “Mejora de la Enseñanza de las Ciencias” convocatoria La Universidad y la Escuela Secundaria, tercera etapa 2015-2017, SPU, el Proyecto “Repositorios Digitales con Contenidos Orientados a las Necesidades de Escuelas Rurales (ER)” por Resolución CE N° 1055/15 convocatoria de Proyectos de Desarrollo Tecnológico y Social (PDTS), 2016-2018 y el proyecto "Programación Colaborativa" aprobado por resolución N° EXP-S01:0001042/2016, 2016, "Universidad, Cultura y Sociedad", SPU.

Introducción

Con la constante evolución de las Tecnologías de la Información y la Comunicación (TIC) se hace necesaria la incorporación, integración y aprovechamiento pedagógico de éstas en el Sistema Educativo. Por otro lado, vivimos una época conocida como la “Era de la Información”, caracterizada por una gran demanda a

nivel nacional e internacional por profesionales calificados en el uso de TIC [1, 2,3]. En la Argentina esta realidad ha sido identificada como clave para la innovación, incremento de las exportaciones, actividades de investigación, emprendedurismo y producción de valor agregado para todos los sectores de la sociedad. Evidencia de ello son la existencia de distintas políticas de estado que promueven el desarrollo del sector, como ser la Ley de Software (Ley N° 26692), becas para jóvenes profesionales TIC [4] o la puesta en funcionamiento de la iniciativa Program.AR [5], que promueve la enseñanza de la computación en todas las escuelas argentinas. La implementación del programa del Estado Nacional denominado Modelo 1 a 1 a través del programa “Conectar Igualdad” al cual la provincia de La Rioja complementó con el plan “Joaquín V. González”, han alcanzado una alta cobertura, cumpliendo con la entrega de notebooks a estudiantes y docentes de escuelas secundarias. Sea cual sea el nivel de integración de las TIC en los centros educativos, los docentes y estudiantes necesitan una “alfabetización digital” y una actualización didáctica.

Es necesario entonces fortalecer la articulación Universidad / Nivel Medio, planteando nuevos escenarios donde los estudiantes no solo sean usuarios de las tecnologías sino participantes activos, generando una adecuada planificación y realización conjunta de acciones que favorezcan la inserción de los estudiantes en la Universidad para prevenir y disminuir las causas de deserción y estancamiento en los primeros años de las carreras universitarias. Esta línea aborda dos aristas del mismo problema, la escasa matrícula de ingresantes que optan por una carrera TIC y el marcado nivel de deserción en los primeros años de estudio de quienes lo hicieron.

La robótica y la programación como introducción a la formación tecnológica

Los procesos de desarticulación y segmentación educativa desarrollados en los

últimos años en nuestro país y en especial en la región, dificultan el pasaje, ingreso y permanencia, de los alumnos del nivel medio a los estudios universitarios [6]. Las instituciones universitarias requieren de sus alumnos aprendizajes vinculados a las destrezas y estilos de pensamiento de las culturas disciplinares, y el desarrollo de una personalidad autónoma y crítica para su desempeño social [7].

Pese a la marcada demanda por profesionales en carreras TIC, es notable que el número de estudiantes que eligen estas carreras sea muy inferior al de estudiantes que eligen carreras más clásicas. Aunque existen múltiples motivos que pueden explicar esta situación (vocación, popularidad de las carreras, entre otras) se destaca particularmente algunos “prejuicios” asociados a las carreras TIC: son carreras exclusivas para jóvenes sobresalientes, la tarea de programar es difícil, aburrida y/o repetitiva, hay que saber mucho inglés, entre otras.

Con el fin de acortar la brecha existente en la articulación Universidad / Nivel Medio se deben profundizar contenidos teóricos - prácticos y su aplicación en un proceso de formación integral que reúna conocimientos, habilidades, destrezas y actitudes propios de los modos de producción en los diferentes campos disciplinares, acordes a los requerimientos sociales.

Conforme a estas tendencias, consideramos oportuna la intervención de la UNdeC como nexo de articulación entre Nivel Medio y Universidad, dado que la universidad es el medio adecuado y pertinente para fortalecer los procesos de enseñanza y aprendizaje sobre dicha tecnología, con el objetivo final de despertar en los estudiantes de Nivel Medio la vocación por las carreras tecnológicas, particularmente en Informática.

La RE se concibe como un contexto de aprendizaje que se apoya en las tecnologías digitales y en los procesos de mediación pedagógica para que los estudiantes creen prototipos o simulaciones robóticas que surgen a partir del ingenio, la creati-

vidad y puesta en práctica de lo aprendido [8]. Es de interés plantear a la robótica como vehículo de aprendizaje con materiales concretos, motivando a los jóvenes a construir, diseñar y explorar nuevas formas de hacer las cosas, a través del aprendizaje experimental, el trabajo en equipo y el desarrollo de su confianza y habilidades innovadoras, brindando un espacio que les permita no sólo ser usuarios de las tecnologías, sino que, a partir de conocimientos matemáticos, mecánicos, físicos y lógicos, logren resolver problemas significativos.

En Argentina y en Latinoamérica se están implementando proyectos, como propuestas de enseñanza de la programación en los primeros años de la carrera, que incorporan entre sus estrategias el aprendizaje basado en problemas y el aprendizaje colaborativo, otros incluyen, además, a la robótica como una opción para la profundización y gestación de habilidades cognitivo-creativas [10, 11, 12].

Algunas de estas propuestas promueven la construcción de robots que compiten de acuerdo a reglas internacionales y por categorías, otras usan la robótica como recurso de apoyo en el estudio de habilidades básicas en matemáticas, ciencias o física y construyen y programan modelos que ayudan a representar con elementos externos esos conceptos. Otros promueven la construcción de robots que ejecutan tareas y funciones particulares o que se comportan de cierta manera ante variables del ambiente [13, 14].

Aunque todas ellas sirven de inspiración, aplicamos un enfoque diferente: utilizar actividades colaborativas y lúdicas relacionadas con la programación de un robot para la resolución de problemas, como estrategia de enseñanza de la programación destinada especialmente a estudiantes del Nivel Medio [15].

No buscamos crear un curso completo de programación, sino generar un espacio de acercamiento a la tecnología y al mundo de la programación que resulte atractivo para los jóvenes antes de que ingresen a la

universidad. Por otro lado la experiencia adquirida en los últimos años como organizadores de la sede Chilcito del Torneo Argentino de Programación (TAP) nos ha permitido poner en valor para el aprendizaje de la programación la motivación obtenida por las competencias y el reconocimiento entre pares: aquellos estudiantes que han participado del torneo han mejorado su desempeño académico, se muestran más dispuestos a compartir lo que saben y participan en las propuestas académicas extracurriculares.

Se está desarrollando una comunidad virtual de aprendizaje colaborativo de la programación. El soporte tecnológico para esta comunidad será dado por una red social diseñada y desarrollada por el equipo de trabajo y orientada a fomentar el aprendizaje de la programación entre pares. Los miembros de esta comunidad en lugar de ocupar jerarquías formales establecidas por los roles de los participantes, adquirirán una “reputación” dentro de la comunidad basada en distintas actividades realizadas a través de la plataforma. La reputación de los miembros permitirá construir un ranking similar al que se tiene en un juego basado en la competencia. Las actividades para adquirir créditos o aumentar la reputación consistirán en resolver problemas de programación usando distintos lenguajes, proponer nuevos problemas, ayudar a otro miembro con dificultades o describir soluciones y técnicas de programación. Se busca fomentar el aspecto lúdico de la competencia donde la participación es de carácter individual, complementado con la colaboración entre pares, donde se premia el desarrollo de habilidades grupales. Interactuando jóvenes de los dos niveles educativos con la finalidad de reducir la deserción en los primeros años de la carrera.

Líneas de Investigación, Desarrollo e Innovación

- Impacto de las TIC en el proceso de enseñanza-aprendizaje del Nivel Medio.

- Desarrollo de Objetos de Aprendizaje.
- Articulación Universidad - Nivel Medio.
- Programación colaborativa.
- Robótica Educativa.

Resultados y Objetivos

El desarrollo de esta línea permitirá:

- Fortalecer la relación entre Universidad e IE del Nivel Medio.
- Promover el desarrollo de habilidades de resolución de problemas de programación, trabajo colaborativo y vinculación entre estudiantes de nivel secundario y estudiantes de nivel universitario en carreras TIC.
- Desmitificar los prejuicios clásicos asociados a la programación entre los estudiantes de nivel secundario y universitario en carreras TIC.
- Generar un ambiente de aprendizaje que permita a los alumnos integrar distintas áreas del conocimiento, adquiriendo habilidades generales y nociones científicas.
- Constituir un equipo interdisciplinario para la investigación y desarrollo de contenidos educativos y estrategias didácticas en TIC, Robótica y Educación.
- Estimular las prácticas de estudio colaborativo entre pares, la competencia sana y el reconocimiento del saber académico.
- La expansión de las TIC en las Escuelas de Nivel Medio del departamento Chilecito y de la región.
- Disminuir del índice de deserción en el primer año de las carreras incluidas como oferta académica de la UNdeC.

Actualmente se ha diseñado y construido el kit de RE compuesto por un módulo principal que aloja al microcontrolador Arduino [9] UNO R3, un sistema de alimentación, un conjunto de sensores que componen el sistema de percepción del robot y actuadores que permiten modificar las variables del entorno controladas. El kit puede ser armado y desarmado de manera simple, siendo ideal para la experimentación y comprensión de conceptos

matemáticos, físicos e informáticos. Para su construcción se utilizaron piezas de bajo costo y disponibles en el mercado nacional para que en caso de pérdida o daño puedan ser reemplazadas fácilmente. Se incluye una librería con funciones, programadas por el equipo de trabajo, que se anexan al entorno de desarrollo Arduino, con el objeto de brindarle al estudiante un nivel de abstracción superior que le permita programar en un lenguaje más natural e intuitivo.

Se adquirió una impresora 3D con la cual se diseñan y construyen los distintos componentes del kit de RE.

Durante el año 2015 y 2016, se desarrollaron talleres con la participación de estudiantes de IE de nivel medio. Cada taller está constituido por tres módulos, un módulo permite que los estudiantes programen cada uno los componentes del kit experimentando el comportamiento de estos antes los estímulos incluidos, otro módulo muestra el funcionamiento de la impresora 3D y las aplicaciones para el diseño de las diferentes piezas, y el tercero permite que los estudiantes experimentan con el robot programando la resolución de un juego. En grupo de a diez, los estudiantes recorren cada módulo.

Avances de la línea fueron expuestos en el artículo "Robótica educativa: una estrategia para despertar vocaciones tempranas en Informática." aprobado y presentado en el IX Congreso sobre Tecnología en Educación y Educación en Tecnología, Chilecito (La Rioja) 2014; en las III y IV Jornadas Científicas de Estudiantes Investigadores, UNdeC 2014 / 2015 / 2016 con la participación de un alumno de grado; y en la Sesión de Demos Educativos en el X Congreso de Tecnología en Educación y Educación en Tecnología, Corrientes 2015.

Formación de Recursos Humanos

El equipo de investigación de esta línea de trabajo está compuesto por seis docentes de la Universidad Nacional de Chilecito y seis alumnos de grado. De los docentes:

uno es doctor en ciencias de la computación especializado en cómputo paralelo y tecnología grid, dos docentes que se encuentran desarrollando su tesis de Maestría en Informática uno en el área de Objetos de Aprendizaje y el otro en Mejora del Posicionamiento de Satélites y dos auxiliar docente actualmente egresados de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC. También participan 2 alumnos avanzados de grado que se encuentran definiendo su trabajo final para la obtención del título de Ingeniero en Sistemas.

Los integrantes son docentes de las asignaturas Algoritmo y Estructuras de Datos, Arquitecturas de Computadoras, Arquitecturas Paralelas, Programación.

Referencias

- [1] NIDIVERSIDAD. Las TIC, su importancia en la actualidad y el mercado laboral, 2016. <http://www.unidiversidad.com.ar/las-tic-su-importancia-en-la-actualidad-y-el-mercado-laboral>
- [2] El Mundo. Faltan 900.000 profesionales TIC | Economía | EL MUNDO. (s.f.) 2016. <http://www.elmundo.es/economia/2015/05/31/5568a4a1268e3e9e518b4592.html>
- [3] El Espectador. «Hay déficit de 15.000 ingenieros» para industria TIC: Mintic, 2014. <http://www.elespectador.com/noticias/economia/hay-deficit-de-15000-ingenieros-industria-tic-mintic-articulo-503625>
- [4] MINCYT. Apertura de Becas Jóvenes Profesionales TIC 2016. <http://www.agencia.mincyt.gob.ar/frontend/agencia/post/2311>
- [5] Program.AR | Acercando a los chicos a la computación. (s. f.). <http://program.ar/>
- [6] Josfie Joaquifin Brunner and Rocío Ferrada Hurtado, "Educación superior en Iberoamerica", Eds. Santiago, Chile: RIL R editores, Oct. 2011.
- [7] SPU, "Articulación escuela secundaria educación superior," Blog: Portal de Educación, 2007.
- [8] Ana Lourdes Acuña, Marfiá Dolores Castro, and Diana Matarrita Obando, "Desarrollo de capacidades para el diseño e implementación de proyectos de robótica educativa en América Latina y el Caribe," Fondo. Reg. para la Inov. Digital en América Latina y el Caribe y la Fund. Omar Dengo, Informe Final de investigación, 2011.
- [9] Massimo Banzi, "Getting Started with Arduino", 2nd ed. O'Reilly Media, 2011.
- [10] A. H. González and M. C. Madoz, "Utilización de TIC para el desarrollo de actividades colaborativas para la enseñanza de la programación," Jul. 2013.
- [11] R. Coppo, J. Iparraguirre, G. Feres, G. Ursua, and A. Cavallo, "Sistema didáctico para la enseñanza de la programación con metodologías de aprendizaje basado en problemas," 2011, eje: Tecnología informática aplicada en educación.
- [12] L. C. De Giusti, F. Leibovich, M. Sanchez, F. Chichizola, M. Naiouf, and A. E. De Giusti, "Desafíos y herramientas para la enseñanza temprana de concurrencia y paralelismo," Oct. 2013, WIEI - II Workshop de innovación en educación en informática.
- [13] Cristian Rigano and Juan Vivanco, "MHO: un robot de sumo - la primera experiencia en robótica con un LOGO!" Grupo de Robótica y Simulación, Departamento de Ingeniería Eléctrica, UTN-FRBB, Bahía Blanca, Reporte, 2006.
- [14] M. A. Junco Rey, R. Swain Oropeza, A. Aceves López, and J. Ramírez Uresti, "RoboCup: el reto tecnológico de monterrey campus estado de México," Arequipa, Perú, 2002.
- [15] Fernanda B. Carmona, et al. "Robótica educativa: una estrategia para despertar vocaciones tempranas en Informática." IX Congreso sobre Tecnología en Educación & Educación en Tecnología. La Rioja, 2014.

Pautas de Diseño de Actividades Formativas para el Desarrollo de Competencias Específicas de Orden Superior y Genéricas

Elizabeth Jiménez Rey (ejimenezrey@yahoo.com.ar)

Patricia Calvo (pat_calvo@yahoo.com)

Arturo Servetto (aserve@gmail.com)

Departamento de Computación, Facultad de Ingeniería, Universidad de Buenos Aires

RESUMEN

La línea de desarrollo que se propone se encuadra en el área de Innovación en Educación Informática, en lo que concierne a la estandarización curricular, y evaluación y aseguramiento de calidad de planes de estudio y currículos de asignaturas, en línea con los objetivos del Proyecto ALFA Tuning América Latina. Se asume un enfoque de competencias para el diseño curricular, y la graduación de éstas en cinco niveles de intensidad en base a las taxonomías de Bloom, Anderson y Krathwohl, y Churches. En base a la taxonomía adoptada, ya se ha propuesto un estándar para la expresión de objetivos de asignaturas, y se trabaja en la especificación de pautas de diseño de actividades formativas y métodos de evaluación en planificaciones de las mismas, para garantizar el desarrollo de competencias de niveles superiores, así como de competencias actitudinales y genéricas.

Palabras clave: competencias, asignaturas, actividades, currículos, métricas.

CONTEXTO

La línea de desarrollo se inserta en el Plan Integral de Mejora de Ingenierías para el Siglo XXI (FIUBA-Calidad Educativa), con financiamiento del Plan Estratégico de Formación de Ingenieros (PEFI – SPU), en el eje “Propuestas de Mejora de la Enseñanza y el Aprendizaje”.

1. INTRODUCCIÓN

Las prospectivas nacionales para la educación superior se enmarcan en el Programa de Internacionalización de la Educación Superior y Cooperación Internacional [1], que tiene como

uno de sus objetivos específicos impulsar la inserción de las instituciones de educación superior argentinas en los procesos de internacionalización, integración y desarrollo local y regional. En lo que respecta al ámbito nacional, la función principal del Programa radica en la canalización de las demandas y necesidades del Sistema Universitario Argentino en materia de internacionalización de la educación superior y la cooperación internacional, y articularlas con los proyectos y acciones impulsadas desde el Programa; este enlace se realiza principalmente a través del Consejo Interuniversitario Nacional (CIN), en particular por intermedio de la REDCIUN (Red de Cooperación Internacional de Universidades Nacionales). En lo que refiere a la cooperación multilateral, el eje estratégico se encuentra identificado en el MERCOSUR y la UNASUR; de esta manera el Programa participa en representación de Argentina en las reuniones de la Comisión Regional Coordinadora de Educación Superior del Sector Educativo del MERCOSUR y en las reuniones del Consejo Suramericano de Educación, Cultura, Ciencia, Tecnología e Innovación de UNASUR (COSECCTI) a través del Subgrupo de Educación Superior. Otros de los espacios multilaterales de la región donde el Programa tiene participación es el Espacio Iberoamericano del Conocimiento, que tiene como una de sus líneas de acción el fortalecimiento de los sistemas nacionales y regionales de acreditación y evaluación de la calidad de la educación superior, y el Proyecto ALFA Tuning América Latina, cuyo objetivo general es contribuir a la construcción de un Espacio de Educación Superior en América Latina a través de la convergencia curricular.

Los objetivos específicos del Proyecto ALFA Tuning América Latina en los que se enfoca la línea de desarrollo son:

- Avanzar en los procesos de reforma curricular basados en un enfoque en competencias en América Latina, completando la metodología Tuning.
- Promover la construcción conjunta de estrategias metodológicas para desarrollar y evaluar la formación de competencias en la implementación de los currículos que contribuyan a la mejora continua de la calidad, incorporando niveles e indicadores.

Para el diseño curricular de carreras según estándares internacionales, se impone considerar niveles de intensidad en las competencias que deben adquirir los estudiantes en áreas específicas de cada disciplina [2]. Los objetivos y contenidos de cada asignatura deberían poder asociarse fácilmente a intensidades y a áreas según su especificación, para facilitar la mejora y el control de calidad de los planes de estudio.

En orden a desarrollar elementos que faciliten el diseño sistemático de currículos de carreras y de asignaturas, se debe considerar que todo diseño curricular enuncia objetivos y contenidos, y para currículos de asignaturas, también actividades y formas evaluativas. Sin embargo, ocurre a menudo que tanto la redacción de los objetivos como la enunciación de las actividades del currículo de una asignatura no permiten establecer la profundidad real del aprendizaje correspondiente.

Los objetivos enunciados corresponden a las competencias a adquirir por el estudiante, entendiendo como competencia la integración de habilidades, conocimientos y actitudes en un contexto específico, y también la capacidad para responder a demandas complejas y llevar a cabo tareas diversas de forma adecuada. De lo cual se desprende que el sólo enunciado de los temas no es en absoluto suficiente; es necesario desarrollar un marco de referencia que exprese la graduación de competencias que adquieren los estudiantes y habilite comparaciones.

El enunciado de los objetivos debería hacer referencia al nivel de profundidad que se espera que el alumno adquiera sobre cada contenido, expresando claramente las competencias que se

espera que el estudiante haya adquirido al finalizar el trayecto de aprendizaje correspondiente. Esto se puede lograr a través de un vocabulario adecuado a la graduación esperada.

Pero para evaluar la intensidad de contenidos, además de ponderar los objetivos, es imprescindible analizar los currículos de las asignaturas que los cubren para verificar que las actividades planificadas y sus sistemas de evaluación sean coherentes con los mismos, por lo que un diseño curricular consistente debería comprender la planificación de cursos de asignaturas que vincule actividades con objetivos y un sistema de evaluación que permita efectivamente constatar la consecución de éstos [3].

Las competencias que deben adquirir los estudiantes al superar trayectos curriculares no sólo deben ser cognitivas (saber), procedimentales/instrumentales (saber hacer), y actitudinales (ser), sino también transversales o genéricas (aquellas que se adquieren en lo que suele denominarse “currículo oculto”), que no se asocian a contenidos sino a actividades, por ejemplo:

- capacidad para la autoorganización y planificación del trabajo individual y del proceso de aprendizaje,
- capacidad para el trabajo en grupo,
- habilidad para la comunicación oral y escrita, y
- motivación por la calidad del resultado.

Así, las actividades son puentes de unión entre el aprendizaje académico y el aprendizaje natural, que crean condiciones para el aprendizaje estratégico, activo, autoconsciente, motivado e intencional [4].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se propone desarrollar especificaciones de diseño de actividades formativas para tres niveles de intensidad, en función de una taxonomía de competencias de cinco niveles de intensidad, siendo los dos primeros **comprender** y **aplicar**, que implican las competencias cognitivas y procedimentales básicas que desarrollan los alumnos al cursar cualquier asignatura de tecnología básicas o aplicada y resolver guías de ejercicios. Para desarrollar las competencias de

mayor nivel, **analizar**, **sintetizar** y **crear**, se requieren actividades formativas más elaboradas y complejas, que pueden contemplar simultáneamente el desarrollo de competencias actitudinales específicas del tipo de conocimiento de una asignatura o transversales o genéricas, así como sistemas de evaluación que contemplen estos aspectos o dimensiones.

Se propone entonces caracterizar actividades formativas para los tres niveles superiores de intensidad, y proponer ejemplos de actividades de cada nivel, empleando actividades y recursos disponibles en el campus virtual (Moodle), y eventualmente incorporando el uso de alguna herramienta libre externa, que se pueda vincular al campus.

Según el nivel de intensidad de competencias que se pretenda, la actividad puede representar un trabajo complementario a las evaluaciones parciales, o un conjunto de actividades con evaluación constante, también complementarias a las evaluaciones parciales; también dicha actividad o actividades podrían implicar una alternativa para promocionar una asignatura.

La competencia de **analizar** corresponde a la capacidad de separar o diferenciar las partes del objeto de estudio en sus componentes para expresar su estructura organizativa, cómo se organizan o se relacionan entre sí las partes. Para poder analizar se presupone la capacidad de diferenciar, organizar y atribuir funciones a las componentes. Los requerimientos de cualquier actividad o secuencia de actividades que se propongan desarrollar esta competencia deberían expresarse en términos de verbos tales como: analizar, comparar, sondear, investigar, examinar, categorizar, diferenciar, contrastar, detectar, encuestar, clasificar, deducir, experimentar, escudriñar, descubrir, inspeccionar, discriminar, separar; y en orden a considerar las tecnologías actuales, se agrega hacer ingeniería inversa, recopilar información, hacer mapas mentales, encuestar, usar bases de datos, analizar mediante métodos de planeación estratégica, entre otras posibilidades.

La competencia de **sintetizar** se relaciona con la capacidad de poner las componentes juntas en la

disposición adecuada para formar un nuevo todo coherente, y también con la capacidad de hacer juicios en función de criterios. Para aspirar a que los estudiantes desarrollen esta competencia, las exigencias de las actividades formativas deberían plantearse en términos de: componer, diseñar, criticar, predecir, planear, formular, generalizar, desarrollar, organizar, construir; a esto se agrega formular hipótesis, experimentar, juzgar, probar, detectar, monitorear, moderar, validar, debatir.

La competencia de **crear** implica el nivel más profundo de cobertura de los temas, y se vincula con la capacidad de planificar, generar y reunir elementos para formar un todo coherente y funcional, pudiendo reorganizar estos elementos para generar también variantes del modelo o nuevas estructuras. Las actividades que pueden asociarse a este nivel pueden comprender: evaluar comparativamente, concluir, estimar, deducir, establecer método de elección, considerar validez, evaluar valor, inferir; también puede considerarse producir, dirigir, planear y elaborar.

Paralelamente al desarrollo de competencias de cualesquiera de estos niveles, se puede establecer modalidades de desarrollo y de evaluación que impliquen en simultáneo el desarrollo de competencias actitudinales (por ejemplo mediante rúbricas para la evaluación que comprendan aspectos como claridad y prolijidad, análisis y justificación de eficiencia, etc.), así como también genéricas (por ejemplo mediante métodos de trabajo en grupo, presentación de resultados con diapositivas o material audiovisual, rúbricas que contemplen la dinámica de trabajo grupal, etc.).

3. RESULTADOS OBTENIDOS/ESPERADOS

Se propuso una jerarquización de objetivos y se planteó la interrelación entre contenidos, áreas de conocimiento, objetivos, y actividades curriculares y sistemas de evaluación como base para definir métricas para evaluar la intensidad de las áreas de conocimiento en el currículo de una carrera.

Se propuso aplicar las taxonomías de Bloom, y de Anderson y Krathwohl [5] para establecer verbos a utilizar en la redacción de objetivos curriculares de modo de ofrecer una base para

estandarizar su escritura y medir el nivel de intensidad de contenidos.

Se ejemplificó la definición de objetivos para asignaturas iniciales de las áreas “Elementos de Programación” y “Algoritmos y Complejidad” [6]. En este planteo, una de las modificaciones fue asociar la capacidad de sintetizar con la posibilidad de crear, ubicando en consecuencia dicha competencia en un nivel superior. Si bien esta revisión incorpora elementos eficaces para el análisis y diseño de programas, cabe señalar que no cubre aspectos contemporáneos asociados a los procesos propios de la aplicación de las nuevas tecnologías de información y comunicación. Es por eso que se han ido desarrollando otras taxonomías, como la de Churches [7], la cual, considerando las características propias de la era digital, incorpora no sólo elementos del ámbito cognitivo (como verbos apropiados para el contexto tecnológico) sino también diversas herramientas (como actividades digitales).

Como resultado de esta línea de desarrollo se espera incidir en la estandarización de diseños curriculares y planificaciones de asignaturas de carreras de informática de la FIUBA de manera que, expresados los objetivos en términos de competencias, se pueda evaluar si las actividades formativas y métodos de evaluación de las planificaciones son coherentes con los mismos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I/D presentada se estructura con:

- un profesor adjunto con dedicación exclusiva,
- una profesora adjunta con dedicación parcial, y
- una jefa de trabajos prácticos con dedicación exclusiva.

4. BIBLIOGRAFÍA

[1] Ministerio de Educación, Secretaría de Políticas Universitarias, Subsecretaría de Gestión y Coordinación de Políticas

Universitarias:

<http://portales.educacion.gov.ar/spu/cooperacion-internacional/>

- [2] Calvo, Patricia; López, Gustavo; Servetto, Arturo; Echeverría, Adriana; Cabrera, José L.; Jeder, Ismael (2014): “Bases para el diseño curricular de carreras de Ingeniería en Informática / Sistemas de Información”. Tecnología y Ciencia, revista de la Universidad Tecnológica Nacional, CABA, año 13, noviembre 2015, Edición Especial CoNaIISI 2014: 2º Congreso Nacional de Ingeniería Informática / Sistemas de Información, San Luis, Argentina, 13 y 14 de noviembre de 2014. ISSN 1666-6917. http://conaiisi.unsl.edu.ar/Revista_UTN_2015/003-33a38.pdf.
- [3] Calvo, Patricia; Echeverría, Adriana; Servetto, Arturo; Cabrera, José L. (2015): “Bases para la estandarización de currículos de carreras de Ingeniería en Informática / Sistemas de Información”. CoNaIISI 2015: 3er Congreso Nacional de Ingeniería Informática / Sistemas de Información, 19 y 20 de noviembre, UTN, Facultad Regional Buenos Aires. ISBN: 978-987-1896-47-9.
- [4] Fernández Batanero, José M. (2004): “La transversalidad curricular en el contexto universitario: un puente entre el aprendizaje académico y el natural”. Revista Fuentes 5. Universidad de Sevilla.
- [5] Anderson, Lorin W. y Krathwohl, David R. (2001): “A Taxonomy for Learning, Teaching and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives”. Longman, New York.
- [6] Servetto, Arturo; Jiménez Rey, Elizabeth; Calvo, Patricia (2015): “Diseño Curricular de un Primer Curso de Algoritmos y Programación”. CoNaIISI 2015: 3er Congreso Nacional de Ingeniería Informática / Sistemas de Información, 19 y 20 de noviembre, UTN, Facultad Regional Buenos Aires. ISBN: 978-987-1896-47-9.
- [7] Churches, Andrew (2009): “Taxonomía de Bloom para la Era Digital”. Educational Origami.

PROGRAMAR en la Escuela: Nuevos Desafíos en las Aulas

Claudia Queiruga, Claudia Banchoff Tzancoff, Sofia Martín, Vanessa Aybar Rosales, Fernando López, Isabel Miyuki Kimura y Noelia Soledad Gómez

Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI). Facultad de Informática. Universidad Nacional de La Plata
50 y 120. La Plata

{claudiaq, cbanchoff, vaybar, sgomez}@info.unlp.edu.ar, {smartin, flopez, ikimura}@linti.unlp.edu.ar

RESUMEN

En un mundo cada vez más conectado, educar en tecnología es imprescindible en la formación de nuestros/as niños/as y jóvenes. Para esto es necesario que la Informática, y en especial la “programación”, se incorpore a los diseños escolares en todos los niveles educativos. En nuestro país han surgido iniciativas y políticas federales que sitúan la enseñanza de “programación” en el aula de la escuela.

El LINTI trabaja en esta área a partir del desarrollo de proyectos de investigación y extensión. El equipo de trabajo ha realizado varias experiencias replicables, sistematizado actividades y materiales, además de desarrollar software pensado para la enseñanza de contenidos de programación en escuelas. En la actualidad participa, junto a otras universidades nacionales y la Fundación Sadosky¹, en dos proyectos cuyo objetivo es colaborar en la incorporación de la disciplina Informática en las escuelas.

Luego de lograr afianzar el grupo y generar experiencias replicables en cuanto al acercamiento de la programación a la escuela, actualmente se trabaja en la formulación de estrategias y herramientas que permitan evaluar el desempeño en las dimensiones del *pensamiento computacional* de niños/as y jóvenes de escuelas primarias, secundarias y de nivel inicial.

Palabras clave: enseñanza de programación, robótica educativa, pensamiento computacional, programación en bloques.

CONTEXTO

La línea de investigación “PROGRAMAR en la escuela: nuevos desafíos en las aulas” presentada en este trabajo, está inserta en el proyecto de investigación “Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro” del Programa Nacional de Incentivos a docentes-investigadores, que se desarrolla en el LINTI de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). Este proyecto está acreditado por la UNLP y financiado por partidas del presupuesto nacional.

1. INTRODUCCIÓN

La enseñanza de Informática en los sistemas educativos nacionales es actualmente una preocupación global. Evidencia de ello es la incorporación desde 2014 de la enseñanza de programación en el Sistema de Educación Escolar del Reino Unido a partir de los 5 años; Estonia, una de las economías europeas que más apuesta al desarrollo tecnológico, viene impulsando a través de diferentes iniciativas estatales la enseñanza de la programación en la escuela; varios estados de los EEUU también han implementado políticas activas en respuesta

¹ Fundación Sadosky:
<http://www.fundacionsadosky.org.ar/>

al respaldo de la industria tecnológica a través movimiento “LearnToCode”, liderado fundamentalmente por las iniciativas globales code.org (<http://code.org>) y codecademy (<http://www.codecademy.com>). Asimismo, Israel, Australia, Japón y Finlandia son otros países que cuentan con iniciativas similares en relación con la incorporación de la enseñanza de programación en las escuelas (National curriculum in England: computing programmes of study, 2013) (Obama, 2013).

En los últimos años en nuestro país, un conjunto de iniciativas y políticas federales, han contribuido a consolidar una perspectiva que fortalece las razones por las cuales resulta necesario situar a la enseñanza de la “programación” en la escuela. Ejemplo de ello es el surgimiento de Programas de Políticas Educativas, tales como el proyecto “Program.AR”, la creación del “Programa Conectar Igualdad”, el desarrollo del “Plan Nacional de Telecomunicaciones Argentina Conectada”, la implementación en el territorio nacional del Programa “Primaria Digital”, que involucró en su primera instancia a las Escuelas PIIE (Programa Integral para la Igualdad Educativa) y el programa “Alfabetización digital” en la provincia de Buenos Aires (PAD), entre otros, que permiten situar social e históricamente, la necesidad de pensar en clave educativa el desarrollo, avance y sinergia en este campo. Asimismo el Consejo Federal de Educación declaró el aprendizaje de “programación” como una herramienta de “importancia estratégica para el sistema educativo argentino”, que será enseñada durante el ciclo de escolaridad obligatoria en todas las escuelas de la Argentina (Resolución CFE N° 263/15).

Algunas iniciativas están relacionadas con el acceso universal de las TIC en los procesos educativos mediante la entrega de equipos informáticos, en este caso netbooks y “pisos tecnológicos” que brindan el acceso a Internet, a alumnos/as, docentes y equipos directivos, con el objetivo de reducir la primera brecha digital: el acceso a las tecnologías.

Actualmente, nos enfrentamos a otra brecha digital que es la que Tim Berners-Lee (Bell, S., 2013) denominó segunda brecha digital y es la que separa a las personas que saben programar de aquellas cuyas habilidades informáticas se limitan simplemente a usar instrumentalmente las computadoras y manejar fluidamente aplicaciones, sin embargo no cuentan con las competencias suficientes para crear sus propias aplicaciones y adaptarlas a sus necesidades. Programar no significa simplemente codificar en un lenguaje de programación, saber programar abre el camino al *pensamiento computacional* y esto implica aprender a resolver problemas aplicando conceptos propios de la Informática como abstracción y descomposición (Wing J, 2006). Las competencias que se adquieren mediante el pensamiento computacional pueden impactar positivamente en otras áreas disciplinares independientes de la informática (NRC, 2012) y convertirse en un habilitador para entender e intervenir en un mundo cada vez más tecnológico, permitiendo crear artefactos tecnológicos (Brennan K, et al, 2012). Las iniciativas gubernamentales proveen el marco necesario para incorporar la enseñanza de la programación en el aula de la escuela, sin embargo se requiere de un acompañamiento a los docentes, dado que rara vez los mismos tienen una formación en esta temática.

Recuperando las experiencias realizadas, durante estos años de trabajo con las escuelas secundarias, encontramos que los docentes, en su mayoría no poseen formación en Informática, además de notar una diferenciación sobre esta cuestión, entre escuelas medias y escuelas técnicas. En las escuelas secundarias técnicas, la formación de los docentes no suele ser un problema, dado que muchos de los docentes sí cuentan con formación en el área. En las escuelas secundarias medias, donde el único espacio curricular cercano es la materia “Nuevas Tecnologías de la Información y la Conectividad” (Dirección General de Cultura y Educación de la provincia de Buenos Aires,

2010) no todos los docentes a cargo de esta asignatura están formados en la disciplina Informática. Esto se agudiza aún más, en la escuela primaria, donde las maestras y maestros no sólo no cuentan con la formación específica sino que tampoco cuentan con un espacio curricular donde desarrollarlo.

Esta situación referente a la formación de los docentes de los diferentes espacios y las diferentes escuelas, constituye una realidad cotidiana de las instituciones educativas, que emerge con la llegada de cada uno de los proyectos.

Desde el año 2006 el LINTI trabaja en proyectos de alfabetización digital, desarrollando actividades que incorporan a las TIC en procesos de aprendizaje, tanto en escuelas realizando actividades con docentes y alumnos/as, como en organizaciones civiles a los que concurren niños y niñas de sectores vulnerables de sociedad, como comedores barriales y centros comunales (Díaz J. et al, 2007). Desde el año 2008, se comienza a trabajar en una línea de investigación cuyo objetivo es abordar la enseñanza de la programación en el aula de escuela, desde una perspectiva cercana a la de nuestros niños/as y adolescentes (Díaz J. et al, 2009) (Díaz J. et al, 2014) (Aybar Rosales V, 2015) (Queiruga C. et al, 2016).

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta línea de investigación trabaja en el diseño de enfoques de intervención en relación a la incorporación del *pensamiento computacional* en la escuela a través de la programación y, en la evaluación del desempeño de los estudiantes en las dimensiones del *pensamiento computacional*.

El enfoque adoptado consiste en la incorporación de actividades y herramientas didácticas cercanas al mundo de los/as niños/as y adolescentes. Los lenguajes de programación visuales basados en bloques y la manipulación de objetos físicos facilitan la incorporación de

conceptos, prácticas y perspectivas propias del *pensamiento computacional*. Se han desarrollado herramientas y materiales que facilitan la apropiación de los elementos constitutivos de la programación, ejemplo de ello es RITA² y MIT APP Inventor³ como punto de partida para enseñar programación en la escuela secundaria. Son dos herramientas de programación visual basadas en bloques que evitan enfrentarse a cuestiones sintácticas propias de los lenguajes de programación textuales (como podrían ser JAVA, Python, JavaScript entre otros) y las instrucciones o comandos del lenguaje son más próximos al “lenguaje humano”. Los resultados de los programas construidos con estas herramientas pueden ser rápidamente visualizados y ciertas prácticas propias de la programación como el testeo y la depuración, pueden resultar procesos relativamente simples de comprender. Estas herramientas también permiten a los estudiantes crear juegos y *apps* propias y de esta manera expresar sus ideas y tener su propia perspectiva sobre qué se puede hacer con la tecnología. Manipular objetos físicos es otro elemento motivador en el universo de los/as niños/as y adolescentes, más aún si es posible programarlos. Desde una perspectiva incremental, en una etapa más avanzada del proceso de enseñanza de programación en la escuela, se introduce la programación de robots simples en lenguaje de programación Python. Se desarrollan actividades desde una perspectiva lúdica a través de un lenguaje de programación real, como lo es Python y el uso de robots físicos inalámbricos.

En el nivel primario e inicial se trabaja con la herramienta de programación basada en bloques Scratch⁴ y el curso inicial⁵ de code.org. Actualmente se están desarrollando

² RITA (Robot Inventor to Teach algorithms): disponible en <https://github.com/vaybar/RITA>

³ MIT APPInventor: disponible en <http://appinventor.mit.edu/explore/>

⁴ Scratch: <https://scratch.mit.edu/>

⁵ Curso 1 de code.org: <https://studio.code.org/s/course1>

herramientas para abordar los primeros pasos en la programación con niño/as pequeño/as.

3. RESULTADOS Y OBJETIVOS

Enmarcadas en esta línea de investigación se han desarrollado herramientas tanto en el marco de tesinas de grado como en trabajos de cátedra y proyectos propios del LINTI.

En el año 2016 se presentó “DROPSY: eDucational ROBot Programming SYstem”⁶, tesina de grado desarrollada por los alumnos Matías Fuentes y Diego Fernández, que extiende la tesina XRemoteBot (López, 2015). Ambos trabajos desarrollan herramientas que permiten la programación remota de los robots físicos: XRemoteBot lo hace a través de clientes para lenguajes textuales (Ruby, Python y Javascript) y Dropsy propone una interfaz basada en bloques, brindando una forma más adecuada para abordar esta temática con niños y niñas más pequeño/as. Respecto a la manipulación de los robots, integrantes de esta línea de investigación desarrollaron DuinoBotSocks⁷, una aplicación móvil pensada para niños/as pequeños/as que permite manipular a los robots desde una interfaz icónica.

RITA es una herramienta desarrollada en el LINTI que propone a los estudiantes programar, con bloques, juegos de robots virtuales que compiten en un campo de batalla y cuyo desafío es programar estrategias de robots ganadores. Actualmente se encuentra en proceso de testeo una versión de RITA en RED, que facilita la implementación de competencias de robots en el aula y promueve el trabajo en equipos.

Para la implementación de estas herramientas en el aula de la escuela se desarrolla material didáctico para docentes y estudiantes que acompañan la puesta en acción. Asimismo para la implementación de MIT APPInventor se

desarrolló material didáctico propio, adecuado a los intereses de nuestra cultura.

En esta línea de investigación, se viene trabajando con más de 15 escuelas, de las cuales se mantiene una relación continua en 10 de ellas. En las mismas se han desarrollado talleres que llevan la programación a las aulas y en varias de ellas los contenidos trabajados en estos talleres se han incorporado a la currícula escolar. Con respecto a la inserción de la programación en la escuela primaria se realizó una primera experiencia con niños/as de 6to. año de la Escuela Graduada “Joaquín V. González” de la UNLP que se extenderá en el actual ciclo lectivo a 4 escuelas más.

Los investigadores de este grupo integran el equipo de trabajo que participa en dos proyectos impulsados por la Fundación Sadosky cuyo objetivo es aportar desde una visión federal a la incorporación de la disciplina Informática en las escuelas, estos proyectos son:

- La elaboración de un manual de Informática destinado a docentes del segundo ciclo de nivel secundario.
- El diseño curricular de una especialización en “Didáctica de las Ciencias de la Computación” en el ámbito de Institutos Superiores de Formación Docente, destinada a docentes de nivel medio de la provincia de Buenos Aires.

Estos proyectos, dan un marco para indagar sobre el impacto de la incorporación de esta disciplina en las actividades escolares y trabajar en la formación de docentes.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo que lleva adelante estas líneas forma parte del LINTI.

Asimismo, se han formulado varias tesinas, tesis de postgrado y actividades de cátedras relacionadas con las herramientas involucradas.

⁶ Disponible en <https://github.com/dropsy-unlp>

⁷ Disponible en <https://github.com/Robots-Linti/DuinoBotSocks>

5. BIBLIOGRAFÍA

- Bell, Stephen (2013). Programming ability is the new digital divide: Berners-Lee. *Computerworld*.
- Brennan, Karen y Resnick, Mitchel (2012). New frameworks for studying and assessing the development of computational thinking. *Annual American Educational Research Association meeting, Vancouver, Canada*.
- Díaz Javier, Harari Ivana, Harari Viviana y Banchoff Tzancoff Claudia (2007). UNLP experiences to reduce the digital gap. *Anales del congreso IECC, International Conference on Engineering Education*. Coimbra. Portugal.
- Díaz Javier, Harari Ivana, Harari Viviana y Banchoff Tzancoff Claudia (2009). An attractive way to teach programming. *ICEER International Conference on Engineering Education & Research*. Seoul, Korea.
- Díaz Javier, Banchoff Tzancoff Claudia, Queiruga Claudia y Martín Sofía (2014). Experiencias de la Facultad de Informática en la Enseñanza de Programación en Escuelas con Software Libre. *Publicado en Memorias del Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación*. Buenos Aires, Argentina. Recuperado de: <http://www.oei.es/congreso2014/memoriactei/1426.pdf>
- Dirección General de Cultura y Educación de la provincia de Buenos Aires (2010). Diseño Curricular para la Educación Secundaria Ciclo Superior. ES4: Nuevas Tecnologías de la Información y la Conectividad. Coordinado por Claudia Bracchi. -1a ed.- La Plata, 2010. Disponible en <https://goo.gl/jHI8Wx>
- López Fernando (2015). XRemoteBot: un servicio para programar robots en forma remota. Disponible en <http://sedici.unlp.edu.ar/handle/10915/51032>
- National curriculum in England: computing programmes of study (2013). Recuperado de <https://www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study>
- NRC (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. *The National Academies Press*.
- Obama (2013). Barack Obama. Discurso difundiendo la hora del código. Computer Science Education Week 2013 Disponible en <https://www.youtube.com/watch?v=6XvmhE1J9PY>
- Queiruga Claudia, Banchof Tzancoff Claudia, Martín Sofía, Aybar Rosales Vanessa y López Fernando. PROGRAMAR en la escuela. *Libro de actas de WICC 2016*. ISBN: 978-950-698-377-2, p. 403-407. <http://sedici.unlp.edu.ar/handle/10915/53013>
- Aybar Rosales Vanessa, Queiruga Claudia, Kimura Isabel Miyuki, Barnetche Matías Brown y Gómez Soledad (2015). Enseñando a programar con RITA en escuelas secundarias. *Libro de actas de CACIC 2015*, ISBN 978-987-3724-37-4, p. 1201-1211.
- Wing Jeannette (2006). Computational thinking. *Communications of ACM*. Vol 49, N° 3. Mar. 2006, 33-35

Promoción del Pensamiento Computacional para Favorecer la Formación de Recursos Humanos en Disciplinas STEM

Gladys N. Dapozo, Cristina L. Greiner, Raquel H. Petris, María C. Espíndola, Ana M. Company

Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura Universidad Nacional del Nordeste. 9 de Julio N° 1449. Corrientes. Argentina.

{gndapozo, cgreiner, rpetris}@exa.unne.edu.ar, mcespindola@yahoo.com, anamacom@hotmail.com

Resumen

Este proyecto de investigación busca aportar información a la problemática de la formación de recursos humanos en disciplinas académicas relacionadas con las ciencias, las tecnologías, las ingenierías y las matemáticas, conocidas como STEM. Se plantean tres líneas de investigación, una orientada a la incorporación de innovaciones basadas en tecnologías (juegos, robots) en el campo de la didáctica de la programación al interior de las carreras de Informática, otra orientada a promover el pensamiento computacional a través de la formación de formadores y articular acciones entre la universidad y otros niveles educativos no universitarios para estimular vocaciones en STEM. Paralelamente a estas actividades, se propone también indagar sobre las causas que influyen en la elección de la formación profesional de las mujeres. Finalmente, se busca también contribuir a mitigar los índices de desgranamiento y de abandono estudiando las dificultades que se presentan en la enseñanza de la Matemática en el primer año de las carreras de Ciencias Exactas.

Palabras clave: Didácticas específicas, herramientas lúdicas, pensamiento computacional, participación de la mujer.

Contexto

Las líneas de I/D corresponden un proyecto nuevo presentado a la Secretaría de Ciencia y Técnica de la Universidad Nacional del Nordeste (UNNE), en etapa de evaluación.

Introducción

Las dificultades en la formación en STEM (Science, Technology, Engineering and Mathematics), es un problema global. Diversos países a nivel mundial se ven afectados por una marcada disminución de estudiantes en carreras TIC, y la escasez de recursos humanos calificados. Un informe del Consejo Presidencial de Asesores en Ciencia y Tecnología (PCAST) en EE.UU. señala que se requerían, en la próxima década, aproximadamente 1 millón de graduados universitarios en los campos de STEM., Además, en este país menos del 40% de los estudiantes que ingresan a la universidad con la intención de especializarse en un campo STEM lo logran [1].

La consultora Everis publicó un estudio [2] del cual se desprende que en 2019 habrá un 40% menos de ingenieros y técnicos informáticos en España. América Latina no es ajena a esta situación, y se llevan adelante diversas políticas públicas, conscientes de esta “revolución digital”, en el que los grandes datos (big data) y la información en tiempo real, son nuevas fuentes de creación de valor. [3].

Sin embargo, en un estudio reciente [4] Katz señala que es evidente la producción limitada de graduados en STEM respecto de las necesidades del campo productivo y, la falta de masa crítica en capacidades para trabajar en la innovación digital. Destaca que, si bien se observa un avance significativo de la digitalización en la región, persiste una carencia estructural de capital humano de TIC. Considera que

esta carencia tiene efectos negativos, debido a que *“jóvenes que han transitado dos o tres años de carrera tienden a abandonarlas tentados por los altos salarios del mercado”* [4].

Propuestas de soluciones

Ante este problema de carácter global, existen numerosos programas que proponen soluciones, que buscan mejorar la educación en estos temas desde los primeros niveles de educación (*“Some STEM for All”*), concentrar los esfuerzos sólo en los interesados o en los destacados (*“All STEM for Some”*) [5], fortalecer la formación de los docentes de STEM y promover el incremento de los mismos mediante beneficios adicionales. [6].

En este contexto, la identificación de las vocaciones se torna esencial para detectar las causas del alto índice de deserción y desgranamiento de los estudiantes. *“Es indudable la íntima relación existente entre el bajo rendimiento, la deserción, el fracaso y la despersonalización del estudiante en el transcurso de su vida universitaria”* [7].

Por estos motivos desde la UNNE se realizan diferentes esfuerzos tendientes a mejorar el desempeño y lograr la permanencia, en la carrera Licenciatura en Sistemas de Información se presentó un conjunto de estrategias, orientadas a mejorar esta realidad descrita en [8].

Nueva alfabetización

Un debate en esta época de transformaciones, atravesada por el exponencial desarrollo de las tecnologías, es sobre la noción de alfabetización: cuáles son los conocimientos necesarios para la vida cotidiana y de qué forma deben ser transmitidos en contextos educativos. Actualmente se requiere la apropiación de habilidades ligadas al desarrollo de formas de gestión y procesamiento de la información. *“Se trata de dimensiones que involucran el ejercicio de una nueva ciudadanía, la del siglo XXI, con nuevas formas de participación, reglas y valoraciones”* [9].

Las Ciencias de la Computación (CC) reúnen los conceptos básicos de esta nueva alfabetización necesaria para que los estudiantes de todos los niveles puedan comprender, por ejemplo, los principios algorítmicos sobre los que funciona el buscador que utilizan cotidianamente o para decidir qué datos publicar en las redes sociales en que participan y con quiénes están dispuestos a compartirlos.

Las habilidades y competencias propias de las CC que pueden ser aplicadas en otros campos de estudio, tales como la descomposición de problemas complejos en subproblemas, el diseño de algoritmos como estrategia de solución, factibles de ser implementados y ejecutados por un dispositivo digital, la capacidad de abstracción como la representación de un problema de manera conceptual y general, la generalización o modelización y formalización, traducida en la operación de plantear una solución de modo que permita resolver múltiples problemas que comparten características comunes, aun cuando se denen contextos diferentes y cambiantes; el reconocimiento de estrategias de solución que han sido útiles para resolver determinados problemas y podrían servir para resolver otros diferentes o bien combinarse de nuevas maneras para solucionar otros en el futuro; la identificación de patrones en la organización de datos y la resolución de problemas[3].

Promoción del pensamiento computacional

En la última década, la web 2.0, los dispositivos móviles, la industria del videojuego y el resurgir del movimiento DIY (Do It Yourself) han puesto en un primer plano las necesidades y carencias en la formación sobre pensamiento computacional y programación. Como respuesta a esta demanda gobiernos, empresas y organizaciones sin ánimo de lucro o instituciones educativas, han desarrollado proyectos y adoptado

decisiones relacionadas con el fomento de la programación en la educación.

El concepto de pensamiento computacional es una competencia compleja de alto nivel relacionada con un modelo de conceptualización específica de los seres humanos que desarrolla ideas y está vinculada con el pensamiento abstracto-matemático y con el pragmático-ingenieril que se aplica en múltiples aspectos de la vida diaria. No es sinónimo de capacidad para programar un ordenador, puesto que requiere pensar en diferentes niveles de abstracción y es independiente de los dispositivos. Se puede desarrollar pensamiento computacional sin utilizar ordenadores (basta papel y lápiz). Por otra parte, es una competencia básica que todo ciudadano debería conocer para desenvolverse en la sociedad digital, pero no es una habilidad rutinaria o mecánica, ya que es una forma de resolver problemas de manera inteligente e imaginativa (cualidades humanas que no poseen los ordenadores). Además, posee las características de combinar abstracción y pragmatismo, puesto que se fundamenta en las Matemáticas, un mundo de ideas, y se desarrolla a partir de proyectos de ingeniería que interactúan con el mundo real.

La programación no es solo una competencia cognitiva que se utiliza para diseñar códigos. Es también una competencia social y cultural que se usa para participar en grupos. Este aprendizaje conectado es algo más que lenguaje de programación porque permite comprender cómo funciona la tecnología y cómo el diseño puede incorporar nuevas posibilidades y soluciones a problemas de la vida cotidiana [10].

Con frecuencia se atribuye a la tecnología un nivel de complejidad que solamente expertos o especialistas podrían abordar. Tradicionalmente la matemática, los principios de la física, los conceptos de biología o química son abordados en todas las etapas del sistema educativo

formal con diferentes niveles de profundidad y de complejidad. Del mismo modo, los fundamentos de las CC pueden ser incorporados de manera espiralada a lo largo de los niveles inicial, primario y secundario. Esta iniciativa se vería facilitada por un gran número de herramientas lúdicas actuales que favorecen la enseñanza de la programación en todos los niveles. La clave sigue estando en la formación de profesionales de la educación, orientadas tanto para los docentes en formación como para los que están en ejercicio. Y esto, sin dudas, demandará cambios en las instituciones formadoras y en las estrategias de desarrollo profesional docente.

Vocaciones TIC. ¿Por qué los jóvenes no eligen carreras TIC?

En nuestro país, la CESSI (Cámara de Software y Servicios Informáticos) realizó en 2006 un estudio que aporta algunas causas, que se reportan en [5], a los que se suma la experiencia personal de los autores. Entre las causas que se mencionan, se destacan: Los jóvenes tienden a asociar a los informáticos con la imagen del “nerd”, personas aisladas, enfrascadas en la computadora, Imaginan a las carreras TIC como difíciles, en parte por deficiencias en formación matemática, en parte por prejuicio puro. Atendiendo a esta situación, en la UNNE, se decidió acompañar la iniciativa de la Fundación Sadosky de promover el estudio de carreras vinculadas con la Informática, y en este marco, indagar sobre el perfil de los alumnos del nivel medio en relación a su formación y habilidades tecnológicas, su interés por la elección de carreras vinculadas con la Informática y los factores que influyen en esta elección, a fin de aportar información que contribuya al objetivo de acercar a los jóvenes a la Informática, como actividad profesional.

Un párrafo aparte merece la cuestión de género en TIC. Según Pavez [11] las mujeres están más expuestas a quedar

rezagadas en el mundo de las TIC porque deben sortear mayores barreras que sus pares masculinos para acceder y aprovechar las ventajas ofrecidas por el mundo digital.

La iniciativa Program.AR

La Iniciativa Program.AR sostiene que el mundo moderno no puede comprenderse cabalmente sin contar con un conjunto de herramientas que permitan decodificar la lógica de la tecnología que media en buena parte los vínculos que establecemos con el mundo en el que estamos insertos.

Los diseñadores de esta propuesta exponen las razones que fundamentan la enseñanza de la programación en la escuela argentina. Para comprender y problematizar los saberes vinculados a la tecnología no alcanza con simplemente conocerlos, nombrarlos o estar en constante contacto con dispositivos tecnológicos. La lógica de funcionamiento de la tecnología no se revela con su mero uso [12].

La problemática en la UNNE

En un estudio realizado en la UNNE [13] se muestra la evolución de la cantidad de alumnos inscriptos en la carrera Licenciatura en Sistemas de Información de la UNNE, por año académico y sexo. Se puede observar una marcada tendencia decreciente en el ingreso pasando de 518 inscriptos en el año 2001, a 210 en el 2014. Se observa una mejora a partir del 2015 (274), 2016 (257) y 2017 (304). La misma podría estar relacionada con las diferentes acciones de promoción que se vienen llevando a cabo desde el 2013, como por ejemplo las Visitas a las Escuelas, en el marco de las acciones de la Fundación Sadosky. En la caracterización de los alumnos se nota también una preponderancia de varones. El porcentaje más alto de mujeres se dio en el 2001 con un 41%, siendo decreciente en los años siguientes, llegando al porcentaje más bajo (12%) en el año 2012. Esta problemática se ve agravada por índices

desfavorables de desgranamiento, deserción y lento avance en la carrera, que influye en el bajo número de egresados, entre 20 y 30 alumnos por año (Estadísticas UNNE. www.unne.edu.ar). En síntesis, la problemática que dio origen a este proyecto se puede resumir en las siguientes cuestiones:

Líneas de investigación y desarrollo

En todos los casos, se tomará como fuente de datos la información de alumnos de la Facultad de Ciencias Exactas de la Universidad Nacional del Nordeste (FaCENA-UNNE) y la información recabada en las actividades de articulación con el nivel medio.

Para el logro de los objetivos del proyecto se propone trabajar en las siguientes líneas:

- Estudiar la aplicación de metodologías innovadoras en el campo de la didáctica de la programación, mediante la incorporación de herramientas lúdicas y robots.

Actividades:

- a) Diseñar y aplicar estrategias de enseñanza de la programación que incorporen software lúdico, como PilasBloques, Scratch, Alice, Ligthbot, entre otros, en el primer año de la carrera Licenciatura en Sistemas de Información (LSI).
 - b) Evaluar si las estrategias implementadas mejoraron la comprensión de los conceptos de programación y los resultados finales del cursado.
- Promover el pensamiento computacional a través de actividades orientadas a la formación de formadores.

Actividades:

- a) Estudio, evaluación y selección de software para diseñar y desarrollar actividades orientadas a consolidar el pensamiento computacional en la enseñanza de la Química, Física, Biología y Matemática.

- b) Capacitación de docentes y estudiantes avanzados de los profesorados de Ciencias Experimentales (Química, Física, Biología) y Matemática que se dictan en la FaCENA, en herramientas computacionales que faciliten la resolución de problemas en las distintas áreas.
 - c) Análisis del grado de apropiación del pensamiento computacional de los participantes a través de la evaluación de las actividades diseñadas para ser implementadas en el aula.
- Estudio de las dificultades que se presentan en la enseñanza de la Matemática en el primer año de las carreras de Ciencias Exactas.

Actividades

- a) Diseño de un instrumento de recolección de información que permita detectar las dificultades de aprendizaje de conceptos propios de las asignaturas de Matemática de primer año.
 - b) Sistematización de la información, análisis y clasificación de las dificultades.
 - c) Evaluación de métodos o estrategias que permitan mejorar la comprensión de los conceptos detectados como dificultosos para los alumnos.
- Articular acciones entre la universidad y otros niveles educativos no universitarios para estimular vocaciones en STEM, en particular, la formación en Ciencias de la Computación.

Actividades

- a) Dictado de talleres de robótica e introducción a la programación con herramientas lúdicas.
- b) Evaluar el grado de interés de los alumnos por la formación en carreras de Informática, y en particular, el interés de las mujeres.

- Análisis de los estilos de aprendizaje su impacto en el aprendizaje de la programación.

Actividades

- a) Diseño de un instrumento que permita determinar los estilos de aprendizaje de los alumnos de primer año de la LSI.
 - b) Análisis de la relación entre el estilo de aprendizaje y las dificultades en el aprendizaje de la programación.
- Estudiar la problemática de la cuestión de género en las carreras STEM

Actividades

- a) Recabar información sobre los factores que influyen en la decisión de las mujeres respecto de su formación profesional, en las distintas actividades que se realicen en el marco de este proyecto.

Resultados obtenidos

Integrantes del grupo de investigación formaron parte de los equipos de asistencia técnica a la Fundación Sadosky para la implementación de talleres de Alice en los colegios secundarios de la ciudad de Corrientes y la capacitación de docentes en didáctica de la programación. Los resultados de estas experiencias se publicaron en [13]. Se muestran los resultados de las acciones de capacitación en programación, que permitieron ajustar la propuesta incorporando actividades motivadoras para las chicas, enfatizando la relevancia que representa una titulación universitaria para el desarrollo personal, y de la sociedad en su conjunto. En [14] se presentaron los resultados de la experiencia de formación en Didáctica de la Programación destinada a docentes de los niveles no universitarios de la ciudad de Corrientes, (2015 y 2016), en la cual se pudo detectar un elevado porcentaje (82%) de los docentes que manifiestan estar de acuerdo con la incorporación de la programación en las escuelas.

Formación de recursos humanos

En el Grupo de Investigación GI-TIPC (Tecnologías Informáticas y Pensamiento Computacional) están involucrados ocho docentes investigadores y dos tesis de posgrado.

Referencias

[1] President's Council of Advisors on Science and Technology. (2012). "Report to the president. Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics". Disponible en: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

[2] Consultora EVERIS. (2012) "La falta de ingenieros TIC: situación actual y perspectiva". Barcelona. Disponible en: <http://www.everis.com/spain/WCLibraryRepository/La%20falta%20de%20ingenieros.pdf>

[3] Borchardt, M.; Roggi, I. (2017) Ciencias de la Computación en los Sistemas Educativos de América Latina. SITEAL. <http://www.siteal.iipe-oei.org>

[4] Katz, R. (2016). TIC, digitalización y políticas públicas. En Entornos Digitales y Políticas Educativas. IPEE-UNESCO.

[5] Atkinson, R.D.; Mayo, M. "Refueling the U.S. innovation economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education". Disponible en: <http://www.itif.org/files/2010-refueling-innovation-economy.pdf>

[6] Committee Members. "Rising above the gathering storm, revisited". Disponible en: <http://www.sandia.gov/NINE/documents/RisingAbove.pdf>

[7] Romero, H.; Pereyra, A. (2003). "Elección vocacional e ingreso a la universidad". III Coloquio internacional sobre gestión universitaria en América del Sur.

[8] Dapozo, G.; Greiner, C.; Godoy Guglielmone, M. (2011). "Retención de alumnos y culminación de estudios. Análisis y propuestas para la Licenciatura en Sistemas de Información de la UNNE". TE&ET 2011. ISBN 978-987-633-072-5. Pp. 424-433.

[9] Brito, A., (2015). Nuevas coordenadas para la alfabetización: debates, tensiones y desafíos en el escenario de la cultura digital, SITEAL TIC. Disponible el 26/03/2017 en: http://tic.siteal.org/sites/default/files/stic_publicacion_files/tic_cuaderno_alfabetizacion.pdf

[10] Valverde Berrocoso, J.; Fernández Sánchez, M.R.; Garrido Arroyo, M. del C.. El pensamiento computacional y las nuevas ecologías del aprendizaje. RED-Revista de Educación a Distancia, 46(3). 2015. Disponible en: http://www.um.es/ead/red/46/valverde_et_al.pdf

[11] Pavez, I. (2015). Niñas y mujeres de América Latina en el mapa tecnológico: una mirada de género en el marco de políticas públicas de inclusión digital, SITEAL TIC. Disponible en: http://tic.siteal.org/sites/default/files/stic_publicacion_files/tic_cuaderno_genero_20160210.pdf

[12] Factorovich, P.; Sawady O'Connor, F. Actividades para aprender a ProgramAR. Volumen 1. Edición: Ignacio Miller. E-Book. - (ProgramAR). ISBN 978-987-27416-1-7.

[13] Dapozo, G.; Greiner, C.; Pedrozo Petrazzini, G.O.; Chiapello, J. (2014) "Investigación para fortalecer actividades de promoción y retención de alumnos en carreras de informática". Anales del XX Congreso Argentino de Ciencias de la Computación. ISBN 978-987-3806-05-6.

[14] Dapozo, G.; Petris, R.; Greiner, C. (2016). "Programación en las escuelas. Experiencia de formación docente en el Nordeste Argentino". III Encuentro Internacional de Educación y Cultura en Ambientes Virtuales. Guadalajara. México.

Robótica Aplicada como Estrategia de Enseñanza de Programación

Lilia Palomo, Norma Lesca y Juan Mulki

Facultad de Matemática Aplicada - Universidad Católica de Santiago del Estero

lilia.palomo@ucse.edu.ar – norma.lesca@gmail.com – jmulki@ucse.edu.ar

RESUMEN

La importancia que han tomado en la actualidad las TIC (Tecnologías de la Información y de la Comunicación) en el campo educativo, ha permitido el desarrollo de diferentes herramientas que han contribuido a mejorar las estrategias de enseñanza.

La programación de robots, permite asimilar conceptos de programación en forma intuitiva y entretenida, de manera tal que se propicie el aprendizaje a través de experiencias basadas en casos reales, mediadas por tecnología, lo que puede aportar al desarrollo de personas creativas, capaces de resolver problemas, articulando la teoría y la práctica.

Este trabajo tiene como finalidad presentar estrategias didácticas para la enseñanza de la programación a través de experiencias en este entorno tecnológico, que permita al alumno comprobar cómo se aplica al diseño, a la implementación y a la prueba del software; y a la vez, comprender los conceptos abstractos.

Palabras clave: Robótica educativa, estrategias didácticas, programación, motivación, enseñanza.

CONTEXTO

La facultad de Matemática Aplicada de la Universidad Católica de Santiago del Estero impulsa la implementación de proyectos que apuntan a incentivar la investigación desde las cátedras promoviendo la interacción vertical y horizontal entre ellas, y posibilitando a los docentes obtener resultados que puedan ser aplicados en las aulas.

Por ello es que, en el año 2016, desde las asignaturas de Programación I, Estructuras de Datos y Análisis Numérico surge el proyecto “Robótica Aplicada a la Enseñanza de la Programación” con el objetivo de promover la investigación aplicada, la formación de recursos humanos, la innovación de los contenidos de las cátedras y de las prácticas profesionales.

El presente artículo presenta una línea de investigación que surge de ese proyecto, ante la necesidad de determinar estrategias didácticas a utilizar para la enseñanza de la programación en la educación superior, bajo prototipos robóticos que se han seleccionado, con el fin de incrementar el interés para aprender e investigar, además de potenciar el aprendizaje basado en problemas.

1. INTRODUCCIÓN

Actualmente, las nuevas tecnologías influyen en todas las esferas de la actividad humana, entonces, no es extraño que también las universidades las incorporen en su quehacer renovando su modelo docente [3]. En nuestro país, esta afirmación es una realidad a la que nos encaminamos

Por su parte, la robótica es un área de conocimiento que rápidamente va cobrando presencia en el sector educativo, tanto escolar como universitario.

La utilización de la tecnología como medio en la docencia no se limita a la utilización del ordenador y a los medios multimedia clásicos, sino se están utilizando otros mecanismos robóticos como herramienta educativa, ganando cada vez más popularidad [5].

Como señalan, Angulo [2] y Carmona [4], los profesores de programación y materias afines

de los primeros cursos universitarios enfrentan diferentes problemas relacionados con los altos porcentajes de deserción y la dificultad creciente para que sus alumnos comprendan las técnicas y criterios fundamentales que deben seguir en el diseño y desarrollo de los programas informáticos.

La idea principal que plantea Márquez [6], es que, el proceso de enseñanza-aprendizaje se fortalece cuando se aplica la robótica, donde el alumno se apropia del conocimiento cambiando su manera de acercarse a la ciencia motivado por la imaginación y la inventiva, con convergencia hacia la aplicación del método científico. Durante este proceso de trabajo en equipo, la ciencia y la ingeniería se combinan obteniendo un producto final operativo y funcional que involucra al estudiante con los saberes formales y ciertos valores asociados al trabajo científico, como la responsabilidad, el compromiso, la organización y el cumplimiento de tareas.

Por su parte, Odorico [7] explica que las situaciones de aprendizaje mediado por tecnología más idóneas son las que permiten que el alumno lleve a cabo una actividad estructurada, en colaboración con sus compañeros y guiada por el docente, donde se centre la atención en un contenido determinado de la materia y se expliciten los objetivos de manera clara para aprovechar el potencial del medio tecnológico utilizado.

En este sentido, un aspecto interesante de la robótica educativa es que permite la integración natural de conocimientos provenientes de diferentes áreas como las matemáticas, las ciencias experimentales, la tecnología y las ciencias de la información y la comunicación.

Tal como plantean Acuña [1] y Ramírez [8], antes de implementar estrategias de este tipo es necesario considerar factores pedagógicos, administrativos y técnicos, prestando especial atención a los aspectos de apropiación del conocimiento que se desean promover y las habilidades que se desean fortalecer a partir de la experiencia. Es necesario contar con recursos materiales y humanos calificados y

realizar adaptaciones específicas en la currícula para incorporar la robótica educativa, donde la propuesta pedagógica y metodológica contemple los contenidos a abordar, las estrategias de aprendizaje, los recursos físicos mínimos y los criterios de evaluación a aplicar.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación del presente trabajo tienen como eje central las estrategias de enseñanza usando procedimientos o recursos utilizados para promover el aprendizaje significativo de los conceptos abstractos de programación.

Entre los supuestos que guían el proyecto se encuentran:

- La robótica educativa emerge como una rama promisoriosa en la que el alumno puede experimentar en la construcción de conocimientos partiendo de lo concreto a lo abstracto
- La formulación de procesos de análisis y síntesis, comprendiendo las relaciones causa-efecto.
- El desarrollo de la creatividad y el despertar de intereses vinculados al área científico-tecnológica.
- La aplicación sencilla y sistemática de los principios tecnológicos para la resolución de problemas.
- La disponibilidad de recursos tecnológicos con diferentes potencialidades, que se pueden seleccionar según las competencias que se desee fortalecer.

Se trabajará en la planificación de la asignatura Programación I, teniendo en cuenta la articulación entre los programas de las cátedras intervinientes y la integración de la práctica utilizando los nuevos recursos que proveen los robots.

Se investigó, algunos productos de robótica educativa como Lego Mindstorms, Arduino,

Scriblers, E-puck, Multiplo, y Probots, en referencia a las teorías y técnicas que posibilitan el aprovechamiento práctico del conocimiento científico y las distintas posibilidades de aplicación, todo ello en relación a su implementación como estrategia didáctica, que comprenda una transferencia de tecnología acorde a las necesidades, recursos y características de las cátedras involucradas en el proyecto.

3. RESULTADOS ESPERADOS

El objetivo general de este trabajo es proponer estrategias didácticas para la enseñanza de la programación a través de experiencias en el entorno tecnológico de los robots.

Este objetivo general se desarrollará mediante los siguientes objetivos específicos:

- Promover el aprendizaje significativo de los conceptos abstractos de programación usando procedimientos y recursos de del mercado.
- Permitir al alumno comprobar cómo se aplica al diseño, a la implementación y a la prueba del software.

Para alcanzar el objetivo planteado, se llevarán a cabo las actividades siguientes:

- Identificación los puntos de articulación entre los programas de las cátedras intervinientes.
- Determinación de los contenidos de ambos programas donde sea apropiada la integración de la práctica utilizando los nuevos recursos.
- Selección de estrategias didácticas en el proceso de enseñanza.
- Diseño y elaboración de ejercicios prácticos.
- Evaluación del impacto de la utilización de las herramientas y ajustes.

Se espera que los resultados de esta investigación se incorporen a los contenidos de las cátedras relacionadas y al espacio curricular correspondiente a través de

procedimientos o recursos de enseñanza para promover aprendizaje significativo, a fin de generar cambios estructurales en la organización y desarrollo de dichas asignaturas, enfocados en la adquisición de habilidades y actitudes en los estudiantes, quienes no deben estar limitados a metodologías tradicionales de enseñanza, sino que el aprendizaje de conocimientos en el aula debe ser dinámico a través de estrategias de formación que impulsen el deseo de estudiar y aprender, incrementando niveles de motivación y gusto por la investigación.

Como resultados indirectos se espera la consolidación del grupo de investigación, la formación de nuevos investigadores y la motivación y entrenamiento en investigación de los estudiantes de grado.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está conformado por tres docentes de la carrera de Ingeniería en Informática, con dedicación simple.

El grupo hace difusión y formación de recursos humanos desde las cátedras: Programación I, Estructura de Datos y Análisis Numérico.

Asimismo, se considera de gran interés la incorporación de becarios, para motivar a los alumnos de la carrera de Ingeniería en Informática a realizar su trabajo final de grado en el área de este proyecto.

5. BIBLIOGRAFIA

- 1] ACUÑA, A. (2009) La robótica educativa: un motor para la innovación. http://www.fod.ac.cr/robotica/descargas/roboteca/articulos/2009/motorinnova_articulo.pdf
- 2] ANGULO, J. & ANGULO, I. (2005) Robots para la enseñanza de la informática. Revista Informática en ingeniería.

- 3] BREMMER, F. (1998) Networked Economy. Online courses get high marks. Usa Today digital, 22 de noviembre, 1998.
- 4] CARMONA, F. (2014) Robótica educativa: una estrategia para despertar vocaciones tempranas en Informática. En IX Congreso sobre Tecnología en Educación & Educación en Tecnología (La Rioja, 2014). http://sedici.unlp.edu.ar/bitstream/handle/10915/38429/Documento_completo.pdf?sequence=1
- 5] DRUIN, A., & HENDLER, J. (2000). "Robots for kids: Exploring new technologies for learning". San Diego, CA: Academic Press
- 6] MÁRQUEZ, D. (2014) Robótica Educativa aplicada a la enseñanza básica secundaria. Didáctica, innovación y multimedia, 2014, no 30, p. 1-12.
- 7] ODORICO, A. (2004). Marco teórico para una robótica pedagógica. Revista Informática Educativa y Medios Audiovisuales, 2004, vol. 1, no 3, p. 34-46. <http://laboratorios.fi.uba.ar/lie/Revista/Articulos/010103/A4oct2004.pdf>
- 8] RAMÍREZ, P. & SOSA, H. (2013). Aprendizaje de y con robótica, algunas experiencias. Revista Educación, 2013, vol. 37, no 1, p. 43-63. <http://revistas.ucr.ac.cr/index.php/educacion/article/view/10628/10298>

TASA – Tecnología Asistiva Mediante Sistema Arduino

Manuel Danieli y Franco Faggiolani

Facultad de Tecnología Informática, Universidad Abierta Interamericana.

Av. Ovidio Lagos 944, Rosario, Santa Fe, Argentina

mdanieli@outlook.com

francofaggiolani@hotmail.com

RESUMEN

El objetivo del proyecto es desarrollar un dispositivo tecnológico (prototipo) fundado en los beneficios que proporciona la Tecnología Asistiva con Sistema Arduino (TASA), como herramienta remediadora y compensatoria de niños con discapacidad motriz y/o motora. El proyecto consiste en la creación de una botonera controlada mediante una placa Arduino que permite el envío de SMSs a un número de teléfono móvil previamente configurado. El contenido de dichos mensajes serán del tipo “Quiero comer”, “Quiero ir al baño”, o, personalizarlo de acuerdo a los requerimientos del usuario.

Palabras clave—motricidad, comunicación, integración, arduino, tecnología.

CONTEXTO

La Tecnología Asistiva es cualquier herramienta utilizada para permitir a individuos con discapacidades o necesidades especiales verse involucrados en las mismas actividades de sus pares sin discapacidad. Dentro de esta categoría podemos estar hablando de herramientas como el uso de sillas de ruedas, el uso de computadoras o el uso de un dispositivo como el TASA.

A continuación, se hace referencia a diferentes proyectos y dispositivos que fueron tomados como referencia y punto de partida para este trabajo.

El trabajo número 1 hace referencia a dispositivos de no tecnología, corresponde a la utilización especial que se le da a objetos o métodos de uso cotidiano. Por ejemplo, comunicarse indicando objetos, figuras, carteles con instrucciones, entre otras cosas.

El trabajo número 2 hace referencia a aplicaciones móviles para celulares y/o tablets. Por ejemplo, TippyTalk[2].

TippyTalk es una aplicación móvil creada por RobLaffan para facilitar la comunicación con su hija Sadie, la cual posee autismo no verbal.

El trabajo número 3 hace referencia a dispositivos semejantes al trabajo actual. Por ejemplo BigMack Comunicador[4].

BigMack Comunicador es una botonera que permite el grabado de mensajes en cada uno de sus botones.

Tomando como referencia los trabajos anteriores expuestos, a continuación, se presenta una tabla comparativa buscando los ítems más significativos y referenciales para llevar a cabo el proyecto.

Tabla 1. Comparación de dispositivos y trabajos anteriores con el proyecto actual.

TRABAJOS	C1	C2	C3	C4
Asistencia a través de dispositivos manuales de NO tecnología (por ejemplo; objetos, dibujos, etc).		X		X
Asistencia a través de dispositivos móviles como celulares y/o tablets.	X		X	
Asistencia a través de dispositivos similares (por ejemplo; BIG Mack Comunicador)	X	X	X	
Trabajo actual	X	X	X	X

C1: Alta tecnología. C2: Personalización. C3: Automatización. C4: Bajo costo.

1. INTRODUCCIÓN

Quienes habitamos este planeta, contamos con diferentes capacidades y el derecho a la inclusión e integración desde entornos

accesibles que garanticen el desarrollo sostenido de nuestras potencialidades y propicien recursos remediativos y compensatorios de diferentes necesidades específicas, no sólo en relación a la usabilidad, sino en la utilidad de los mismos a la medida de sus usuarios.

Ante la presencia de disfunciones motrices y/o motoras, especialmente en niños, la afección no solo se centra en lo físico, sino que impacta también en lo emocional y social, principalmente en la comunicabilidad, ansiedad e independencia, la atención y el procesamiento de la información del entorno que los rodea.

António Damásio, neurocientífico portugués que desenvuelve su labor en el sur de California, dice: *“En la construcción de nuestro mundo, el cuerpo es la herramienta de mapeación y percepción del mismo, realizando este proceso de acuerdo con las capacidades neuronales desarrolladas y a las que el propio cuerpo contribuye en su desarrollo, en un sistema de retroalimentación complejo y en gran medida aún desconocido”*.

De ahí la importancia de una **buena selección** y aplicación de aquellos recursos o herramientas con tecnología asistiva, de modo que otorguen múltiples medios de representación, de acción y expresión particularmente, en niños con discapacidad motriz y/o motora.

Situación Problemática

La igualdad pensada desde las diferencias, es un punto de partida para crear posibilidades y oportunidades en la calidad de vida de niños con discapacidad motriz y/o motora, por ello: ¿Qué aportes brinda un dispositivo de Tecnología Asistiva con Sistema Arduino (TASA) a niños con discapacidad motriz y/o motora en la comunicabilidad inmediata de sus necesidades básicas a su familia, acompañantes terapéuticos o personas con quienes conviven en su hogar?

Objetivos específicos

- Analizar antecedentes sobre la aplicación de este Sistema en productos tecnológicos utilizados en personas con discapacidad.

- Diseñar un Dispositivo (prototipo) de Tecnología Asistiva con Sistema Arduino (TASA) que facilite la comunicabilidad de necesidades básicas en niños con discapacidad motriz y/o motora dentro del hogar.
- Propiciar una herramienta de accesibilidad y utilidad a la medida de sus pequeños usuarios.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los niños con algún tipo de discapacidad motriz y/o motora pueden encontrar en dispositivos de Tecnología Asistiva con Sistemas especializados, del tipo Arduino, el apoyo perfecto para llevar una vida más plena e independiente en vista a la comunicabilidad de sus necesidades y demandas básicas.

Para lo cual, estos recursos con tecnología asistiva deben enmarcarse dentro de un Diseño Universal, cuya convergencia tecnológica cumpla, preferentemente, con tres características fundamentales: UBICUIDAD para que se pueda utilizar en cualquier momento y lugar; INVISIBILIDAD para que, dentro de lo posible, esta tecnología pase desapercibida en el entorno y ADAPTABILIDAD para que responda a las necesidades de la persona.

A su vez, es importante tener presente que son varios los aspectos a tener en cuenta al momento de pensar y crear dispositivos tecnológicos asistivos, remediativos y compensatorios para niños con discapacidad motriz y/o motora:

- Las capacidades motrices condicionantes, aquellas capacidades energético-funcionales del rendimiento, que se desarrollan producto de las acciones motrices conscientes del individuo (fuerza rápida, máxima; velocidad de traslación, de reacción; resistencia de corta, media o larga duración; entre las principales)
- Las capacidades coordinativas, aquellas capacidades que se realizan conscientemente en la regulación y dirección de los movimientos, con una

finalidad determinada. Dependen de las aptitudes físicas del niño y en su enfrentamiento cotidiano con el medio (regulación y dirección del movimiento; orientación y equilibrio; anticipación; diferenciación; agilidad; movilidad activa o pasiva; entre otras)

Lo que se propone como solución es un dispositivo que permita el envío de mensajes de texto a un número de teléfono específico. Dicho dispositivo cuenta con botones donde cada uno de estos tiene configurado un mensaje diferente, es decir, el individuo que adquiera el producto podrá enviar SMSs como ser: “Quiero ir al baño”, “Quiero comer”, o los que él solicite, dependiendo las necesidades de la persona.

A su vez, el dispositivo posee un switch/on/off el cual permite deshabilitar/habilitar los diferentes botones para poder capacitar a la persona que lo necesite, entrenando a esta, utilizando botón a botón. Se propone también que aquella persona que adquiera el dispositivo pueda agregarle texturas o colores a los diferentes botones para lograr una mejor identificación de cada uno de estos.

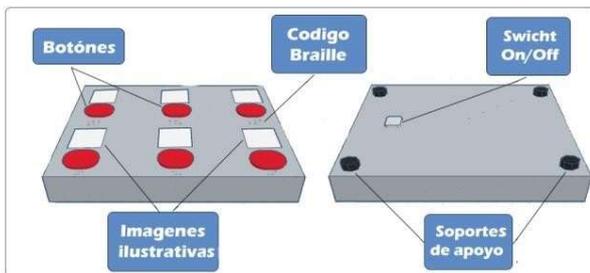


Fig 1: Esquema del dispositivo TASA

Para interiorizarse aún más sobre lo que es la motricidad y su vínculo con el movimiento y el conocimiento, Vítor Da Fonseca, lo resume claramente en la siguiente frase: “...Cuanto más compleja es la motricidad, más complejo es el mecanismo que la planifica, regula, elabora y ejecuta. La motricidad conduce a esquemas de acción sensoriales que a su vez son transformados en patrones de comportamiento cada vez más versátiles y disponibles. La motricidad retrata, en términos de acción, los productos y los procesos funcionales creadores de nuevas acciones sobre acciones anteriores.

Por la motricidad utilizadora, exploratoria, inventiva y constructiva, el Hombre y el niño, humanizando, esto es, socializando el movimiento, adquirirán el conocimiento.” (Vítor Da Fonseca, 1989).

En cuanto a la implementación, el dispositivo TASA se encuentra diseñado sobre una plataforma open-hardware denominada Arduino basada en una sencilla placa con entradas y salidas (E/S), analógicas y digitales, y en un entorno de desarrollo que implementa el lenguaje Processing/Wiring.

Dicha plataforma cuenta con una unidad de procesamiento del tipo Arduino Mega 2560 la cual es la encargada de interactuar con los diferentes componentes que integran TASA.



Fig 2: Arduino Mega 2560

El Arduino Mega está basado en el microcontrolador ATmega2560. Tiene 54 pines de entradas/salidas digitales (14 de las cuales pueden ser utilizadas como salidas PWM), 16 entradas analógicas, 4 UARTs, cristal oscilador de 16 Mhz, conexión USB, jack de alimentación, conector ICSP y botón de reset. Incorpora todo lo necesario para que el microcontrolador trabaje; simplemente se conecta a la PC por medio de un cable USB o con una fuente de alimentación externa.

Características:

Microcontroller	ATmega1280
Operating Voltage	5V
Input Voltage	7-12V
Input Voltage	6-20V
Digital I/O Pins	54
Analog Input Pins	16

DC Current per I/O Pin	40 mA
DC Current for 3.3V Pin	50 mA
Flash Memory	128 KB of which 4 KB used by bootloader
SRAM	8 KB
EEPROM	4 KB
ClockSpeed	16 MHz

Otro de los componentes principales es una plaqueta ultra compacta GPRS SIM900 la cual permite acceder a redes de telefonía celular.



Fig3: Shield GPRS/SIM900

Características:

Totalmente compatible con Arduino
Conexión con el puerto serial
Quad-Band 850/ 900/ 1800/ 1900 Mhz
GPRS multi-slot clase 10/8
GPRS mobile station clase B
Compatible GSM fase 2/2+
Clase 4 (2 W (AT) 850 / 900 MHz)
Clase 1 (1 W (AT) 1800 / 1900MHz)
TCP/UP embebido
Soporta RTC
Consumo de 1.5 mA (susp)

La conexión con la unidad de procesamiento principal es del tipo UART "Universal Asynchronous Receiver-Transmitter" utilizando comandos del tipo AT.

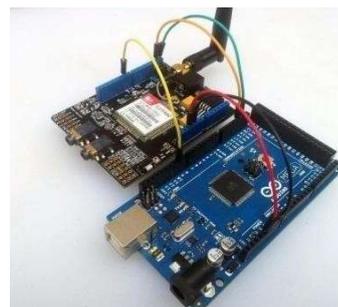


Fig4: Ejemplo de comunicación entre ambos componentes.

Distribución de Pines:

Pin	Nombre de Pin	Función de Pin	Notas
1	RST	Reset	
2	CE	Chip Selection (Selección de chip)	
3	DC	Data/Commandschoice	
4	DIN	Serial data in	
5	CLK	Serial clock	
6	VCC	Positive powersupply (Alimentación positiva)	2.7V a 3.3V
7	LIGHT	LED backlightsupply	Conectar a GND para max brillo
8	GND	Ground (Tierra)	

Para integrar los diferentes componentes se diseño una plaqueta electrónica con un circuito impreso.

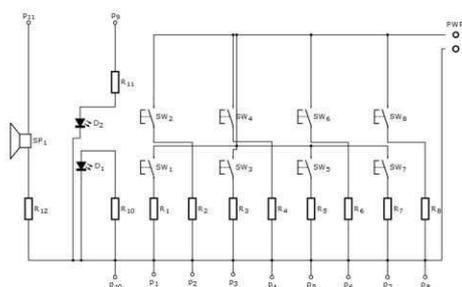


Fig6: Circuito electrónico integrador.

Los componentes electrónicos utilizados son los siguientes:

- 2 diodos leds (rojo y verde)
- 8 pulsadores

- 8 resistencias 220ohm
- 1 switch DIP de 8 canales

3.RESULTADOS OBTENIDOS/ESPERADOS

“Las necesidades y experiencias que solicita el niño son las que necesita el cerebro para desarrollarse, por lo que más que preocuparnos por la estimulación, debemos preocuparnos por satisfacer las necesidades y esto nos permitirá corresponder con las solicitudes de la familia y las sociedades cada vez más exigentes con el futuro,” (Terre Camacho, 2014)

Con el uso de la tecnología y la innovación cualquier persona puede realizar cualquier actividad pese a poseer alguna discapacidad. Lo que falte físicamente se suple mecánicamente y lo que no se puede realizar se hace con ayuda de herramientas controladas por una computadora. Cuando se pierde una capacidad el mecanismo de defensa de la persona busca suplirla de alguna manera, busca cómo seguir adelante, consigue herramientas que le permitan continuar con la normalidad de su vida.

De acuerdo con los resultados evidenciados de nuestro dispositivo una vez puesto en funcionamiento, se observa que este presenta grandes beneficios ya que permite satisfacer una necesidad muy importante como es la comunicación en personas y por otro lado lograr cierta independencia por parte del usuario respecto a sus padres/tutores.

Además, la modularidad del dispositivo es una gran ventaja ya le que permite ser modificado fácilmente de acuerdo a las necesidades del usuario.

4.FORMACIÓN DE RECURSOS HUMANOS

El proyecto está dirigido por los ingenieros Pedro López y Pablo Audoglio, ambos docentes de la Universidad Abierta Interamericana. Además, integran el proyecto los alumnos Manuel Danieli y Franco Faggiolani, ambos

cursando el último año de Ingeniería en Sistemas en la UAI sede Rosario.

5.BIBLIOGRAFÍA

- [1] Gonzáles, Rosa M. (2012). Informe sobre el Uso de las Tecnologías de Información y Comunicación (TIC) en la educación para Personas con Discapacidad. Disponible en: <http://unesdoc.unesco.org/images/0021/002163/216382s.pdf>.
- [2] Rob Laffan (2016). Tippy Talk - Instant Communication for Non Verbal People. Disponible en: <http://tippy-talk.com>
- [3] Alberoni, Orlando (2012). De las ayudas técnicas a la tecnología asistiva. Disponible en: <http://blog.ucc.edu.ar/tecnologiaasistiva/files/2012/07/DE-LAS-AYUDAS-T%C3%89CNICAS-A-LA-TECNOLOG%C3%8DA-ASISTIVA1.pdf>
- [4] BIGmack (2017). AbleNet - Solutions for Individuals with Disabilities. Disponible en: <http://www.ablenetinc.com/bigmack>
- [5] Berruezo, Pedro Pablo (2005). Los contenidos de la Psicomotricidad. Disponible en: <https://www.um.es/cursos/promoedu/psicomotricidad/2005/material/contenidos-psicomotricidad-diapositivas.pdf>
- [6] Barreto, Mario Alberto (2010). Sistemas de Tecnología Asistida para personas con discapacidad motriz. Disponible en: <https://tecasistidaud.wikispaces.com/file/view/SISTEMA+DE+TECNOLOG%C3%8DA+ASISTIDA+PARA+PERSONAS+CON+DISCAPACIDAD+MOTRIZ.pdf>
- [7] Campos, Luís R. (2016). Desarrollos Tecnológicos Especiales. Disponible en: <http://www.centrocamac.org/ayudas-tecnicas.pdf>
- [8] Belloch Orti, Consuelo (2004). Tecnologías de ayuda: Sistemas Alternativos de Comunicación. Disponible en: <http://www.uv.es/bellochc/pdf/pwlogo5.pdf>
- [9] Guenaga, María Luz , Barbier, Ander , Eguiluz, Andoni (2007). La accesibilidad y las tecnologías en la información y comunicación. Disponible en: <http://trans.uma.es/presentacion.html>

**Innovación en
Sistemas de Software**

Agentes Inteligentes y Web Semántica: Hacia la Verbalización de un Subconjunto de UML en una Herramienta Gráfica Web

Matías Garrido¹

Germán Braun^{1,2,3}

Sandra Roger¹

email: {roger,german.braun}@fi.uncoma.edu.ar

¹Grupo de Investigación en Lenguajes e Inteligencia Artificial
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²Laboratorio de I&D en Ingeniería de Software y Sistemas de Información
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

RESUMEN

El proyecto de investigación Agentes Inteligentes y Web Semántica, financiado por la Universidad Nacional del Comahue, tiene como objetivo general la generación de conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la Web, es decir, lo que se ha llamado la Web Semántica.

El objetivo general del trabajo de investigación es la extensión de una herramienta de modelado ontológico, denominada crowd, mediante la verbalización de un subconjunto del lenguaje de modelado conceptual UML. Esta integración permitirá generar especificaciones en Lenguaje Natural a partir de un diagrama de clases.

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

Palabras Clave: Verbalización, Generalización de Lenguaje Natural, UML, Ontologías.

CONTEXTO

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el contexto del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014), por la Universidad Nacional del Sur a través del proyecto de investigación Integración de Información y Servicios en la Web (24/N027) y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral con duración de 5 años, finalizando en abril de 2019.

1. INTRODUCCIÓN

La fase de requerimientos es la más problemática en un proceso de desarrollo de software [1]. Estos problemas incluyen las dificultades de elicitar correctamente los requerimientos del usuario, en su entendimiento y en la transformación de estos mismos dentro de un modelo computacional que puede ser semiformal, usualmente referidas a una notación gráfica tal como los modelos Orientados a Objetos (OO) [2] los cuales usan el Lenguaje Unificado de Modelado (UML) [3] o lenguajes de

especificación formal tales como Vienna Development Method (VDM) [4] o Z [5].

La verbalización es el proceso de escribir la semántica capturada en una teoría lógica (relaciones entre entidades y sus restricciones) en sentencias en lenguaje natural.

En la fase de análisis del desarrollo de los sistemas de información, es importante que el esquema conceptual sea validado por el experto del dominio, para asegurar que el esquema modela con precisión los aspectos relevantes del dominio del negocio.

Una manera efectiva de facilitar esta validación es verbalizar el esquema en un lenguaje fácilmente comprensible por el experto del dominio, quien puede no contar con el conocimiento técnico adecuado. Dicha verbalización también puede ser usada como una manera de integrar a los usuarios en los procesos de chequeo de consistencia cuando los cambios son realizados en el diseño o en la implementación. Como consecuencia, esta brecha de comunicación entre los modeladores y los expertos en el dominio es minimizada.

En este sentido, existen algunas investigaciones a tener en cuenta que se relacionan con sistemas de Generación de Lenguaje Natural (GLN) [6]. Uno de ellos, es el sistema ModEx (Model Explainer) [7], que genera lenguaje natural desde descripciones de modelos de software OO. Sin embargo, ModEx no verifica semánticamente la salida final y asume que dicha verificación es realizada por los usuarios, comparando el diagrama con el sistema de especificación generado mediante la GLN. En general, el sistema funciona con éxito para los modelos que cumplen con las suposiciones de como las clases y las relaciones deben ser llamadas.

Con el fin de extender las capacidades del sistema anterior y considerando una convención de nomenclaturas más amplia, surgió GenLangUML [8]. Se trata de un sistema que propone la generación de una especificación de lenguaje natural en inglés a partir de diagramas de clases UML. Utiliza WordNet, una ontología lingüística, para realizar el análisis sintáctico de los nombres de entrada y la verificación de las sentencias generadas. La validación de GenLangUML fue

realizada extrayendo convenciones de nomenclaturas de libros académicos. Sin embargo, esto podría impactar negativamente en su aplicación en la industria, debido a que las organizaciones podrían tener diferentes convenciones para definir sus diagramas de clases.

Hay enfoques establecidos con respecto a la verbalización multilingüe. Uno de ellos es DogmaModeller [9], basado sobre ORM 1 (Object-Role Modeling) [10], para la generación automática de verbalizaciones en lenguaje pseudo-natural. Es una herramienta de ingeniería de ontologías basada en los principios de ORM e implementa totalmente la verbalización multilingüe. Para cada uno de estos idiomas de salida, DogmaModeller posee un template que contiene estructuras determinadas por la sintaxis para cada tipo de restricción de ORM. Cada estructura contiene etiquetas para referenciar a los tipos de objetos y los roles que forman parte de dichas restricciones. Por otro lado, DogmaModeller es extensible al poder crear nuevos templates de verbalización para otros idiomas. Sin embargo, para hacer que las sentencias verbalizadas sean gramaticalmente correctas en cualquier lenguaje natural, se requiere un tratamiento más complejo a través de un análisis morfológico automatizado para cada idioma. Este es un área de investigación activa en NLG.

Otra herramienta de software de verbalización automatizada, que soporta modelos ORM de segunda generación (ORM2) es NORMA (Natural ORM Architect) [11]. En esta herramienta, la verbalización de elementos individuales en el modelo ORM principal se genera utilizando una transformación XSLT aplicada a un archivo XML. En este proceso se identifican rápidamente diferentes patrones de verbalización y posteriormente se describe como deben combinarse las frases para producir la verbalización en un inglés legible.

En este contexto, el objetivo del presente trabajo es extender la arquitectura de nuestra herramienta gráfica de modelado ontológico crowd [12, 13], para soportar la verbalización multilingüe de un subconjunto del metamodelo

de UML referido a los diagramas de clases. crowd es un prototipo cliente-servidor, actualmente en desarrollo por nuestros grupos de investigación, en respuesta a la complejidad inherente al modelado conceptual y ontológico, además de explotar las bondades de los sistemas basados en Lógicas Descriptivas (DL) [14].

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección 3 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

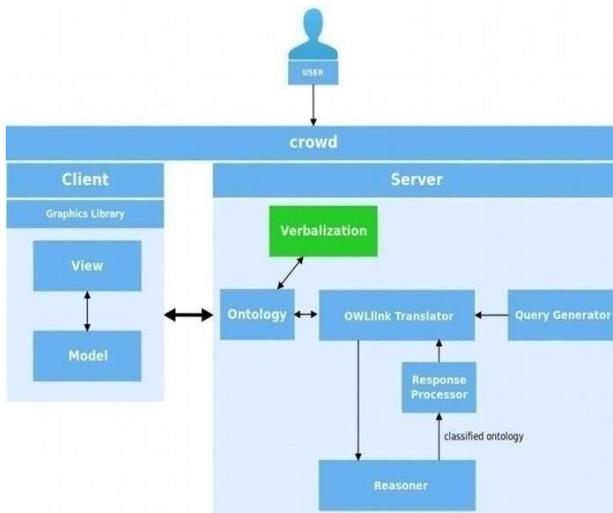


Figura 1. Arquitectura de crowd

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación Agentes Inteligentes y Web Semántica tiene como objetivo general generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica.

Por otro lado, en el proyecto de investigación Integración de Información y Servicios en la Web se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web,

fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado conceptual.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la Representación del Conocimiento, las Lógicas Descriptivas, [15], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento.

En los trabajos [12, 13] se presentó la arquitectura inicial de crowd (ver Figura 1) y un prototipo implementado que permite, en primer instancia, determinar la consistencia de un modelo gráfico representando una ontología. El front-end permite al usuario modelar de forma gráfica usando diagramas de clases UML, mientras que el back-end trabaja del lado del servidor con un razonador capaz de inferir posibles restricciones implícitas en los modelos. Los módulos en el servidor traducen el modelo inicial en uno lógico basado en DL, como propone [16]. La comunicación entre el cliente y el servidor es a través del protocolo OWLlink [17].

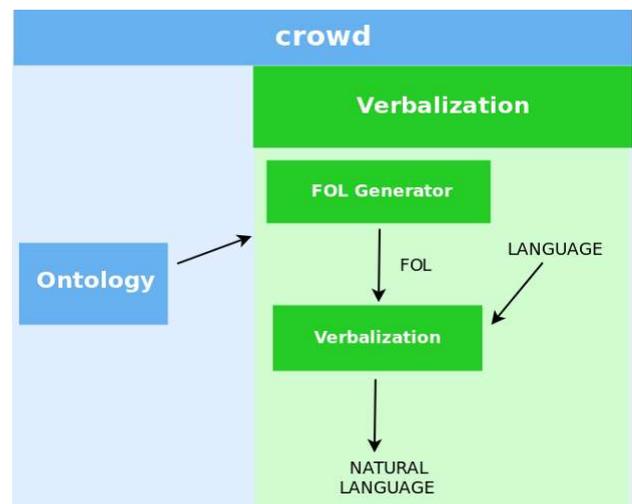


Figura 2. Arquitectura de Verbalización

La Figura 2 muestra la arquitectura básica del proceso de verbalización planteada e incorporada a la arquitectura mostrada anteriormente. La entrada a dicho proceso es

un subconjunto del metamodelo de UML, que es traducido a un lenguaje intermedio basado en lógica de primer orden (FOL), por el módulo FOL generator. La verbalización es multilingüe, por lo cual el usuario puede elegir el idioma para traducir las sentencias FOL. En una primera etapa, los idiomas de traducción serán inglés y español y, posteriormente, planeamos extender el conjunto de idiomas posibles.

3. RESULTADOS OBTENIDOS Y TRABAJOS FUTUROS

Inicialmente, se diseñó una primera versión de la arquitectura cliente-servidor, incluyendo entre otros módulos la generación de consultas, librerías gráficas y un traductor para OWLlink.

Con el fin de extender esta herramienta para validación de requerimientos del usuario, se incorporó un módulo de verbalización. El mismo fue ideado con el objetivo de ser multilingüe y, para facilitar esta traducción, se decidió utilizar una representación intermedia en lógica de primer orden de los diagramas de clases UML. Finalmente, para abordar la complejidad inherente a la generación de las sentencias en el idioma destino elegido, el proceso de verbalización procederá con la creación de patrones de escritura generales, basados sobre algunas herramientas lingüísticas existentes [18, 19, 20].

Actualmente, nos encontramos en la fase de diseño de los módulos de verbalización en crowd, y próximos a iniciar la implementación de los mismos. Asimismo, se pretende estudiar diferentes técnicas para la validación de nuestro prototipo.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante la realización de este sistema se espera lograr, como mínimo, la culminación de 2 tesis de grado dirigidas y/o codirigidas por los integrantes del proyecto. Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

Finalmente, es constante la búsqueda hacia la consolidación como investigadores de los miembros más recientes del grupo.

BIBLIOGRAFÍA

- [1] Michael Christel and Kyo Kang. Issues in requirements elicitation. Technical Report CMU/SEI-92-TR-012, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 1992.
- [2] Grady Booch, Robert Maksimchuk, Michael Engle, Bobbi Young, Jim Conallen, and Kelli Houston. Object-oriented Analysis and Design with Applications, Third Edition. Addison-Wesley Professional, third edition, 2007.
- [3] Grady Booch, James Rumbaugh, and Ivar Jacobson. Unified Modeling Language User Guide. Addison-Wesley Professional, 2005.
- [4] Jones, Cliff B. Systematic Software Development using VDM. Prentice-Hall, Upper Saddle River and NJ 07458 and USA, 1990.
- [5] J. M. Spivey. The Z Notation: A Reference Manual. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
- [6] Ehud Reiter and Robert Dale. Building Natural Language Generation Systems. Cambridge University Press, New York, NY, USA, 2000.
- [7] Benoit Lavoie, Owen Rambow, and Ehud Reiter. The modelexplainer, 1996.
- [8] Farid Meziane, Nikos Athanasakis, and Sophia Ananiadou. Generating natural language specifications from UML class diagrams. *Requir. Eng.*, 13(1):1–18, 2008.
- [9] Mustafa Jarrar. Towards Methodological Principles for Ontology Engineering. PhD thesis, Vrije Universiteit Brussel, Brussels, 5 2005.

- [10] Terry Halpin and Tony Morgan. *Information Modeling and Relational Databases*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2 edition, 2008.
- [11] Matthew Curland and Terry A. Halpin. The NORMA software tool for ORM 2. In *Information Systems Evolution - CAiSE Forum 2010*, Hammamet, Tunisia, June 7-9, 2010, Selected Extended Papers, pages 190–204, 2010.
- [12] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una Arquitectura Cliente-Servidor para Modelado Conceptual Asistido por Razonamiento Automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [13] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *Proc. of the 2nd Simposio Argentino de Ontologías y sus Aplicaciones (SAOA) colocated at Jornadas Argentinas de Informática (JAIIO) - to appear*, 2016.
- [14] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: theory, implementation, and applications*. 2003.
- [15] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [16] Daniela Berardi, Diego Calvanese, and Giuseppe De Giacomo. Reasoning on UML class diagrams. *Artif. Intell.*, 168(1-2):70–118, 2005.
- [17] Thorsten Liebig, Marko Luther, Olaf Noppens, and Michael Wessel. OwlLink. *Semantic Web*, 2(1):23–32, 2011.
- [18] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, 2014.
- [19] Edward Loper and Steven Bird. Nltk: The natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics - Volume 1*. Association for Computational Linguistics, 2002.
- [20] Hamish Cunningham, Diana Maynard, Kalina Bontcheva, Valentin Tablan, Niraj Aswani, Ian Roberts, Genevieve Gorrell, Adam Funk, Angus Roberts, Danica Damjanovic, Thomas Heitz, Mark A. Greenwood, Horacio Saggion, Johann Petrak, Yaoyong Li, and Wim Peters. *Text Processing with GATE (Version 6)*. 2011.

Agricultura de Precisión y Data-driven Agriculture en Regiones de Climas Áridos

Alberto Eduardo Riba⁽¹⁾, Jorge Damián Tejada⁽¹⁾, Fernanda Beatriz Carmona⁽¹⁾, Fernando Emmanuel Frati⁽¹⁾, Matías Agustín Pérez⁽¹⁾, Emmanuel Alejandro Portugal Murcia⁽¹⁾, Nelson Acosta^(2,3), Juan Manuel Toloza^(2,3)

¹ Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina
{ariba, jtejada, fcarmona, fefrati}@undec.edu.ar
{emmanuel.portugal.91, mataguper}@gmail.com

² Universidad Nacional del Centro de la Provincia de Buenos Aires
General Pinto 399, Tandil, Buenos Aires, Argentina
{nacosta, jmtoloza}@exa.unicen.edu.ar

³ Universidad Nacional de Tres de Febrero
Mosconi 2736 - Sáenz Peña (B1674AHF), Buenos Aires, Argentina

Resumen

La tecnología es la impulsora de lo que hoy se denomina agricultura de precisión, los avances en las herramientas de sensado y la generalización del uso de los sistemas de posicionamiento globales, permiten mejorar la productividad agrícola mediante el uso intensivo de los datos generados. Esta tendencia llamada data-driven agriculture no solo implica la digitalización y almacenamiento de información sino también desplegar políticas de gestión de los datos enfocadas a su análisis, explotación y protección.

Esta línea de I+D+i de gran interés regional, se enfoca en el problema de la racionalización y uso eficiente de los recursos agrícolas en regiones de climas áridos, con el objetivo de incrementar la productividad. Estos problemas se detectaron en empresas privadas y organizaciones del medio que se beneficiarán con los resultados obtenidos.

Los temas abordados son transversales a varias áreas, como programación,

probabilidad y estadística, investigación operativa, bases de datos y análisis de sistemas.

Cabe destacar que en los proyectos participan alumnos de grado avanzados y que los desarrollos propuestos convergen en tesinas de la Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC.

Palabras clave:

Agromática, Red de Sensores, Agricultura de Precisión, Data-Driven Agriculture, Series de Datos Agrometeorológicos.

Contexto

Esta línea de investigación se presenta como continuidad de los proyectos “Programación y generación de pronósticos de riego presurizado para cultivo de olivo en regiones áridas, utilizando TIC” y como complemento del proyecto “Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para

cultivo de olivo en regiones áridas”, aprobados en los años 2009 y 2012 respectivamente, por la Secretaría de Ciencia y Tecnología (FICyT - UNdeC).

A partir de los cuales surgen los proyectos “Utilización de métodos de diseño de software para desarrollar un sistema automatizado de riego”, 2011 – 2013 y “Red de Sensores Inalámbricos basado en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”, 2013-2015 aprobados por la Secretaría de Ciencia y Tecnología, convocatoria para estímulo y desarrollo de la investigación científica y tecnológica (FICyT - UNdeC).

Por otro lado, se coopera activamente con una empresa del medio dedicada al cultivo de olivo, y se espera que otras empresas de la región se beneficien de los resultados de esta línea de I+D.

Introducción

La actividad agrícola en la región sólo es posible con la ayuda del riego artificial, utilizándose las tierras para cultivos como el nogal, olivo, vid y frutales. La agronomía es la ciencia aplicada que rige las prácticas agrícolas y es considerada una ciencia espacial. En un lote cultivado es posible encontrar sectores de alta productividad, muy próximos a sectores menos productivos. Esta variación espacial suele estar asociada a factores como la pendiente del suelo, la permeabilidad, el tipo del suelo y la fertilidad. Sin embargo, los agricultores manejan el cultivo de forma homogénea, aplicando dosis de fertilizante o irrigando de manera uniforme todo el lote.

La agricultura de precisión utiliza complejas fórmulas y modelos matemáticos para el análisis de los grandes volúmenes de datos geo- espaciales generados por las distintas tecnologías de sensado, convirtiéndose en

un sistema de control en donde la retroalimentación de la información permite a los productores diseñar tratamientos específicos situados para incrementar la eficiencia en el uso de los recursos involucrados [1].

El sector agrícola es el mayor consumidor de agua, por el aumento de la superficie irrigada y por la escasa eficiencia en los sistemas de riego. Por este motivo es de vital importancia lograr un equilibrio hidrológico que asegure el abastecimiento de agua a la población y al sector agrícola.

En regiones de climas áridos el costo final de explotación es afectado en un alto porcentaje por factores imputables al riego. Dentro de este costo se considera la inversión inicial del sistema (detección de las napas, perforación del pozo y tendido del sistema de distribución) y el costo energético para su extracción y distribución (energía eléctrica y combustibles necesarios).

El sistema de irrigación más utilizado por los agricultores es el riego presurizado. Desde el punto de vista agronómico se denominan riegos localizados porque humedecen un sector de volumen de suelo suficiente para un buen desarrollo del cultivo.

Para incrementar la eficiencia en los sistemas de riego es necesario estudiar las distintas variables (requerimientos hídricos de los cultivos, características del suelo, condiciones meteorológicas, propiedades y limitaciones del sistema de riego) y como se relacionan para determinar el uso adecuado del recurso. Estas variables son de naturaleza heterogénea y algunas de ellas pueden ser capturadas automáticamente mediante el uso de sensores para permitir su posterior tratamiento y análisis para una correcta programación del riego.

Monitorización de riego en cultivos
Para la monitorización y programación del riego se debe tener en cuenta:

Factores ambientales:

- la medición o estimación de variables ambientales: temperatura, presión, radiación solar, evapotranspiración, entre otras.
- el monitoreo en el crecimiento y desarrollo del cultivo o fitomonitorio: tamaño del fruto, tamaño del tallo, flujo de savia, índice de estrés hídrico.
- la medición de variables en el suelo: temperatura, humedad, conductividad, constante dieléctrica.
- la medición de variables propias del agua aportada: salinidad, alcalinidad, etc.

Factores artificiales:

- Limitantes propios del sistema de riego instalado (cañerías, presiones, válvulas, aspersores, goteros). Entre ellos perfil-umbral (capacidad máx. diaria), coeficiente de uniformidad, etc.
- Aspectos económicos relativos al consumo energético utilizado por las bombas. Los aranceles de Kwatt/hora varían según el momento del día en que son consumidos.

Actualmente las parcelas de la región capturan la información relacionadas con algunas de estas variables de manera aislada e independiente por tipo de sensor (de temperatura ambiente, de humedad de suelo, dendrómetros, estaciones meteorológicas, dataloggers, etc.), utilizando para el análisis e interpretación de datos la interface de software suministrada por el fabricante del dispositivo. Para otras variables el método de registro es totalmente manual utilizando planillas en papel y dicha información luego es volcada en planillas de cálculo.

En trabajos anteriores sobre esta línea [2, 3, 4] se implementó un sistema de registro para el procesamiento de esta información, para la cual la carga de datos se realiza manualmente. Trabajos actuales pretenden automatizar la captura de la información de los distintos sensores en cada sector de la plantación utilizando redes de sensores inalámbricos [5]. Esta práctica ha sido implementada con éxito en diversos ámbitos como detección de incendios forestales [6] o la monitorización de viñedos [7]. Cada nodo de la red estará compuesto por un dispositivo inalámbrico autónomo y un conjunto de sensores para la recolección de datos de distinta naturaleza. Estas redes se caracterizan por su escalabilidad, ausencia de cableado, bajo consumo y gran variedad de magnitudes físico/químicas medibles, lo que las vuelve muy interesantes para aplicaciones en agricultura, ya que serían más costosas y complejas de implementar con otras tecnologías.

Se espera que a partir de la automatización de la recolección de información de los sensores y la integración con el sistema actual, se optimice la generación de estimaciones estadísticas y gráficos, lo que contribuirá al proceso de toma de decisiones relacionadas con el pronóstico y la planificación del riego.

Riego automatizado inteligente

Existe una gran cantidad de sistemas para la determinación, control y automatización del riego que permiten un consumo óptimo de agua.

En nuestra región, los equipos de riego localizado empleados en la mayoría de las explotaciones agrícolas no poseen automatismos, o en algunos casos, estos se encuentran en desuso. La apertura y cierre de las válvulas de las diferentes subunidades y sectores de riego se realiza en forma manual, acotando los tiempos de

riego, basándose principalmente en la experiencia de quien lo programa o en recomendaciones surgidas de mediciones de humedad de suelo. La fertirrigación sufre de las mismas limitaciones.

Ajuste de series meteorológicas

La automatización de la recolección de información de sensores permitió notar que las series agrometeorológicas resultantes presentan problemas de completitud, veracidad y exactitud.

La calidad de estas series depende de dispositivos electro-mecánicos (sensores, estaciones meteorológicas, dataloggers), de redes de transmisiones inalámbricas, de personal de mantenimiento y de baterías. Cualquiera de estos factores resulta un punto de fallo ineludible que degrada la calidad [8]. Para mitigar estos se pueden establecer políticas y estándares de funcionamiento de hardware y software que tiendan a prevenirlos, minimizarlos y ante su aparición, controlarlos. A pesar de la implementación de tales políticas existen casos en los que simplemente los fallos ocurren, y se deben tomar medidas ante estas situaciones.

Todos estos motivos inclinaron al estudio, análisis y desarrollo de métodos y técnicas de detección, corrección y ajuste de datos de series agrometeorológicas para mejorar su calidad.

Líneas de Investigación, Desarrollo e Innovación

- Sistemas de Información y Base de Datos
- Sistemas de Tiempo Real
- Diseño de dispositivos de automatización
- Comunicaciones inalámbricas
- Redes de sensores
- Ajuste de series de datos

Resultados y Objetivos

Objetivos

- Analizar y estudiar las diferentes tecnologías de microcontroladores, módulos de conexiones inalámbricas y tipos de sensores (de temperatura ambiente, de humedad de suelo, dendrómetros) disponibles en el mercado.
- Monitorizar en forma centralizada, remota y en tiempo real las variables capturadas por los diferentes sensores.
- Mejorar la capacidad de administración y planificación de los recursos hídricos destinados al riego, a través del análisis de las variables obtenidas de los nodos instalados en sectores con diferentes características de suelo y clima.
- Configurar adecuadamente los algoritmos de control y aplicar estrategias de riego que optimicen la relación kg. producido por m³ de agua aplicada.
- Evaluar distintos esquemas de control, comparar los resultados e inferir en la elaboración de nuevas estrategias de riego.
- Aumentar la uniformidad y la eficiencia de la aplicación del agua del riego.

Reducir los costos asociados de consumo del agua.

Resultados

Como resultados se puede mencionar el trabajo de tesis de grado de uno de los integrantes, denominado “Sistema de Gestión de Riego y Fertilización”, presentado en 2014.

En el marco de otro trabajo final de grado (fecha de defensa junio 2017), el desarrollo de una herramienta, que combina dispositivos de hardware y aplicaciones de software para la monitorización en tiempo real de la

medición de diferentes variables que influyen en la necesidad hídrica del cultivo, presentado evento 45 JAIIO – Concurso de trabajos Estudiantiles 2016, titulado “Red de sensores inalámbricos basados en microcontroladores para la monitorización del riego presurizado en plantaciones de olivo”.

Formación de Recursos Humanos

El equipo de trabajo está formado por seis docentes de las carreras Ingeniería en Sistemas, Licenciatura en Sistemas e Ingeniería Agronómica de la UNdeC (acreditadas por CONEAU), de la carrera Ing. Sistemas de la UNICEN y UNTREF y dos alumnos de la UNdeC.

De los docentes: 1 es posdoctorado en Informática, 2 son doctores en Informática; 2 maestrandos que presentarán su tesis en la Universidad Nacional de San Juan; un doctorando que presentará su tesis en la Universidad Nacional del Centro de la Provincia de Buenos Aires.

Los alumnos de grado se hallan realizando su trabajo de tesina final en esta línea de I+D.

Los integrantes son docentes de las asignaturas Arquitecturas de computadoras II, Arquitecturas paralelas, Programación I y II y Agromática I y II. Estas asignaturas fomentan la participación en proyectos de investigación, por lo que pueden surgir nuevos trabajos en esta línea.

Referencias

- [1] E. W. Schuster, S. Kumar, S. E. Sarma, J. L. Willers and G. A. Milliken (2011) "Infrastructure for data-driven agriculture", IEEE 8th International Conference & Expo on Emerging Technologies for a Smarter World, New York, NY
- [2] F. B. Carmona, A. Riba, A. Sfeir, and F. E. Frati, “Pronósticos de riego presurizado para cultivo de olivo en regiones áridas,” (2008), eje: Ingeniería de Software y Base de Datos. [Online]. available: <http://hdl.handle.net/10915/20640>
- [3] F. B. Carmona, E. A. Riba, A. Sfeir, and F. E. Frati, “Pronósticos de riego presurizado para cultivo de olivo en regiones áridas,” (2010), eje: Innovación en Sistemas de Software. [Online]. Available:<http://hdl.handle.net/10915/19571>
- [4] E. A. Riba, F. B. Carmona, F. E. Frati, J. D. Tejada, N. Acosta, and J. M. Toloza, “Integración y administración de índices de sensado en la programación y pronósticos de riego presurizado para cultivo de olivo en regiones áridas,” (2012), eje: Innovación en sistemas de software. [Online]. Available: <http://hdl.handle.net/10915/19218>
- [5] C.-Y. Chong and S. Kumar, “Sensor networks: evolution, opportunities, and challenges,” Proceedings of the IEEE, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.
- [6] Javier Solebera, “Detecting forest fires using wireless sensor networks,” Sep. 2010. [Online]. Available: <http://www.libelium.com/wireless-sensor-networks-to-detect-forest-fires/>
- [7] Alberto Bielsa, “Smart agriculture Project in Galicia to monitor a vineyard with waspmote,” Jun. 2012. [Online]. Available: <http://www.libelium.com/smartagriculture-vineyardsensorswaspote/>
- [8] M. N. Khaliq, T. B. M. J. Ouarda (2007) “Short communication on the critical values of the standard normal homogeneity test (SNHT)” International Journal of Climatology, Vol. 27 681-687p

Análisis e Implementación de Nuevas Tecnologías para la Web Móvil

Rocío A. Rodríguez, Pablo M. Vera, M. Roxana Martínez,
Fernando A. Parra Beltrán, Jeff Alcidor

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{rocioandrea.rodriguez, pabломartin.vera, roxana.martinez} @uai.edu.ar
{Adolfo.ParraBeltran, Jeff.Alcidor} @alumnos.uai.edu.ar

RESUMEN

La presente línea de investigación y desarrollo tiene por objetivo analizar los nuevos estándares de la web móvil para su implementación. Para lo cual es necesario analizar los estándares del W3C y generar documentación necesaria para su uso, o bien tomar las nuevas APIs y con ellas crear bibliotecas de funciones (mediante JavaScript) que brinden un valor agregado y faciliten su uso para los desarrolladores web. Esto permitirá traer innovación a la web pudiendo facilitar el acceso a componentes de hardware de los dispositivos móviles, así como a datos del dispositivo que puedan ser consumidos desde la web.

Palabras clave: Web Móvil, Dispositivos, API, JavaScript

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Laboratorio de Algoritmos y Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado y una duración de 2 años habiendo iniciado el año pasado.

1. INTRODUCCIÓN

La alta inserción de los dispositivos móviles (entre ellos el teléfono celular) hacen necesario tomar en consideración

particularidades al momento de diseñar aplicaciones, motivo por el cual el W3C cuenta con guías de buenas prácticas [1], [2] y diversos autores han puesto el foco de atención sobre el usuario de la web móvil destacándose algunos artículos basados en la usabilidad [3], [4], [5].

Los teléfonos celulares, traen cada vez mayor equipamiento, lo cual permite realizar aplicaciones más ricas haciendo uso de sus componentes y sensores. Con estos avances es necesario que las aplicaciones puedan incorporar estándares creados para facilitar el acceso al hardware. El W3C (Consortio Web a nivel internacional) tiene grupos de trabajo los cuales se centran en la creación de estándares [6], la figura 1 muestra las distintas etapas de estandarización. No obstante, los navegadores no incorporan rápidamente los estándares, así mismo las aplicaciones web construidas para dispositivos móviles tampoco incorporan mediante APIs [7] [8] el aprovechamiento de los componentes o sensores de los dispositivos. Por otra parte, si bien el W3C provee APIs para acceso al hardware deben realizarse pruebas que permitan conocer el funcionamiento en diversos equipos dado que el comportamiento varía. Cabe destacar que existen APIs implementadas que no son utilizadas por lo general en las aplicaciones, como es el caso de la APIs de vibración es un estándar consolidado desde inicios del 2015 sin embargo las aplicaciones no usan esta API como por ejemplo para poder brindar un feedback al usuario.



Figura 1. Etapas de estandarización del W3C

Existen distintas estrategias al momento de diseñar aplicaciones móviles. La web ofrece algunos beneficios como la portabilidad y reduce la brecha que existía entre la web y las aplicaciones móviles; mediante la llegada de HTML 5 [9]. La construcción de aplicaciones puede realizarse para la web o en forma nativa tal como se muestra en la tabla 1.

Tabla 1. Comparativa entre aplicaciones nativas y aplicaciones web

	Aplicaciones	
	Nativas	Web
Acceso al Hardware	Completo	Limitado, pero se encuentra en crecimiento de ampliar el acceso al mismo.
Conectividad	Online y Offline	Mayormente Online, pero se puede trabajar desconectado
Portabilidad	Se encuentra restringida por el sistema operativo	Puede ser utilizado en cualquier dispositivo con

		cualquier sistema operativo compatible con los estándares web
Gráficos	Puede utilizar todas las capacidades del hardware del dispositivo para crear aplicaciones importantes a nivel gráficos de juegos.	Limitado a la visualización que soporte el navegador web que se esté utilizando.
Interfaz de usuario	El acceso a los controles nativos permite la creación de aplicaciones más estándares.	Con el uso de frameworks de JavaScript, los controles nativos pueden ser simulados en los navegadores

Como puede observarse, las aplicaciones web que tenían como característica principal la

portabilidad, actualmente suman otros beneficios que antes se podían obtener sólo con aplicaciones nativas como ser: el acceso al hardware, una interfaz de usuario con controles nativos, etc. Los nuevos estándares permiten que la brecha que separa a las aplicaciones web móviles de las nativas se reduzca. Teniendo actualmente aplicaciones visualmente similares y que ofrecen las mismas prestaciones.

Por ello es importante contribuir a la implementación de dichos estándares y su uso estratégico en distintos tipos de aplicaciones.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Análisis de estándares para la web móvil.
- Diseño de pruebas de nuevas tecnologías, análisis de compatibilidad con distintos navegadores.
- Analizar ventajas de encapsular funcionalidades en APIs de alto nivel.
- Analizar ámbitos de aplicación e impacto de uso.
- Generación de Guías de Uso y Buenas Prácticas para los estándares analizados.

3. RESULTADOS OBTENIDOS/ESPERADOS

El avance de los estándares impulsado por el W3C se efectúa rápidamente, sin embargo, la implementación de estos estándares se ve rezagada en los navegadores actuales. La prueba de dichos estándares y la construcción de documentación junto a APIs que permitan ponerlos en funcionamiento, mejorará las aplicaciones existentes aportándole los beneficios que incorporan los nuevos estándares. En base a esta premisa, en el primer año del proyecto de investigación fue posible trabajar con distintas tecnologías generando documentación que facilite su uso para los desarrolladores web:

- 1) Rendimiento: Se documentan dos APIs del W3C [10] [11] comparándose entre sí y estableciendo sus escenarios de uso.
- 2) Manejo de Interfaz: Eventos Touch [12] analizando la posible identificación de gestos en la pantalla del navegador para facilitar tareas habituales.
- 3) Acceso al Hardware: Acceso al sensor de proximidad, se construye una API que encapsula las cuestiones básicas de acceso al sensor de proximidad, permitiendo identificar gestos en el aire y que los mismos desencadenen una acción particular.

Producto de este tercer ítem se realizaron pruebas con diversos sensores de proximidad notándose diferencias en la velocidad de detección y en los valores arrojados para indicar cercanía o lejanía de un objeto frente al sensor. Todas estas consideraciones hicieron que sea necesario realizar ajustes a una API realizada, la cual fue publicada en una revista indexada en el IEEE [13].

En este segundo año del proyecto de investigación se continúa avanzando con el análisis de nuevas tecnologías y la consideración de sus posibles usos.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 8 personas.

- 3 Docentes (2 de Postgrado y 1 de Grado).
- 5 (2 Alumnos de Posgrado, 3 Alumnos de Grado).

En el área de dispositivos móviles se encuentran en realización 2 tesis de maestría en la UAI (Universidad Abierta Interamericana) y una de doctorado en la UNLP (Universidad Nacional de La Plata).

5. BIBLIOGRAFIA

- [1] W3C, "Mobile Web Best Practices 1.0", 2008
<https://www.w3.org/TR/mobile-bp/>
- [2] W3C, "Mobile Web Application Best Practices", 2010
<https://www.w3.org/TR/mwabp/>
- [3] Nielsen, Jacob. "Jakob Nielsen's Alertbox", September 26, 2011.
<http://www.useit.com/alertbox/mobile-usability.html>
- [4] Nielsen Jacob, Budiu Raluca. "Mobile Usability", Editorial: New Riders. Estados Unidos. 2012
- [5] Neil Theresa. "Mobile Design Pattern Gallery", Editorial: O'Reilly. Estados Unidos. 2012
- [6] W3C, "Standards for Web Applications on Mobile: current state and roadmap", 2015
<https://www.w3.org/Mobile/mobile-web-app-state/>
- [7] W3C, "JavaScript APIs Current Status", 2016
https://www.w3.org/standards/techs/js#w3c_all
- [8] Fling Brian, "Mobile Design and Development: Practical Concepts and Techniques for Creating Mobile Sites and Web Apps". Editorial: O'Reilly. Estados Unidos. 2009.
- [9] Firtman Maximiliano, "HTML5 compatibility on mobile and tablet browsers with testing on real devices", 2015.
<http://mobilehtml5.org/>
- [10] W3C, User Timing, 2013
<https://www.w3.org/TR/2013/REC-user-timing-20131212/>
- [11] W3C, Navigation Timing, 2012
<https://www.w3.org/TR/2012/REC-navigation-timing-20121217/>
- [12] W3C, Touch Events, 2013
<https://www.w3.org/TR/2013/REC-touch-events-20131010/>
- [13] Vera Pablo, Rodríguez Rocío, Creating and Using Proximity Events on Mobile Websites. IEEE Latin America Transactions, Volumen: 14, Número: 11, Diciembre 2016
<http://ieeexplore.ieee.org/document/7795832/>

Aumento de la Precisión Posicional Empleando Técnicas y Algoritmos para el Tratamiento del Error en Receptores GNSS de Bajo Costo

Alberto Eduardo Riba⁽¹⁾, Jorge Damián Tejada⁽¹⁾, Fernando Emmanuel Frati⁽¹⁾, Nelson Acosta^(2,3), Juan Manuel Toloza^(2,3)

¹ Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito 9 de Julio 22, Chilecito, La Rioja, Argentina
{ariba, jtejada, fefrati}@undec.edu.ar

² Universidad Nacional del Centro de la Provincia de Buenos Aires General Pinto 399, Tandil, Buenos Aires, Argentina
{nacosta, jmtoloza}@exa.unicen.edu.ar

³ Universidad Nacional de Tres de Febrero Mosconi 2736 - Sáenz Peña (B1674AHF), Buenos Aires, Argentina

Resumen

Los sistemas GNSS (Global Navigation Satellite System) son utilizados en innumerables ámbitos para georreferenciar objetos, existen tecnologías implementadas como NAVSTAR-GPS y GLONASS, y otras en vías de implementación como GALILEO y COMPASS.

En la actualidad existe una gran variedad de receptores GNSS y su costo varía en función de su precisión. Los receptores de bajo costo proveen posiciones con precisiones de una decena de metros (aplicables a navegación, seguimiento de fauna y flotas, turismo y ocio, etc.). Otros, más caros y precisos son capaces de entregar localizaciones con precisión de centímetros e incluso milímetros, que se utilizan para estudios y aplicaciones de carácter técnico y científico (agricultura de precisión, topografía, fotogrametría, conducción autónoma, replanteo en ingeniería civil, etc.).

La presente línea de I+D apunta al estudio, diseño y desarrollo de algoritmos, técnicas y métodos que

permitan disminuir el error en la posición entregada por receptores GNSS de bajo costo para mejorar la precisión del posicionamiento.

Los temas abordados son transversales a varias áreas, como algoritmos, lógica, programación, arquitectura de computadoras, sistemas de posicionamiento. Se espera que de esta línea surjan futuras tesis de grado de las carreras relacionadas de la UNDeC, UNICEN y UNTREF.

Palabras clave:

GNSS, GPS, DGPS. Posicionamiento de precisión, GNSS diferencial de bajo costo, minimización de errores de sensores.

Contexto

Forma parte de esta línea de investigación el proyecto "Incremento de la precisión posicional relativa utilizando receptores GPS de bajo costo" presentado en la Secretaría de Ciencia y Tecnología convocatoria para estímulo y desarrollo de la investigación científica y

tecnológica (FICyT – UndeC - convocatoria 2013-2014), aprobado en junio de 2015 actualmente en desarrollo. Cabe destacar que esta línea se inicia con el proyecto de tesis doctoral “Algoritmos y técnicas de tiempo real para el incremento de la precisión posicional relativa usando receptores GPS estándar” del Dr. Juan Manuel Toloza becario CONICET. Y actualmente se encuentra en desarrollo el trabajo de tesis de Maestría titulado “Técnicas de análisis de la geometría de los satélites para el incremento de la precisión posicional relativa utilizando receptores GPS monofrecuencia L1” de Alberto Eduardo Riba.

Introducción

El posicionamiento de un objeto sobre la corteza terrestre ha sido objeto de innumerables estudios. Los mercantes usaban los cuerpos celestiales para ubicarse en su navegación por los océanos [1, 2]. En la actualidad los nuevos sistemas autónomos utilizan micro-dispositivos embebidos con distintos tipos de sensores, que resuelven en cuestión de segundos la posición actual de un objeto para asistirlo en su navegación [3].

Los sistemas GNSS constan de tres partes fundamentales: los satélites en órbita alrededor de la Tierra, las estaciones terrestres de seguimiento y control, y los receptores propiedad de los usuarios [4].

Los receptores necesitan recibir la señal de al menos cuatro satélites para resolver su hora local precisa y sus coordenadas tridimensionales de latitud, longitud y altitud. Para ello utilizan técnicas de trilateración basada en la medición de retardo de la señal [5, 6].

Existe una amplia oferta de receptores GNSS de bajo costo. En nuestro país los

más utilizados son los que implementan la tecnología NAVSTAR-GPS aunque en los últimos años es posible adquirir dispositivos que soportan más de una constelación como por ejemplo GLONASS y la emergente GALILEO.

Cuando se trabaja con este tipo de receptores no se puede asegurar de obtener posiciones con una precisión mayor a los 15 metros el 95% de las veces [7, 8].

Existe un gran número de aplicaciones que utilizan esta tecnología y pueden llevar a cabo su tareas sin inconvenientes en cuanto a la precisión obtenida, pero alguna áreas como la agricultura de precisión, la aeronavegación, la navegación marítima, los desarrollos aeroespaciales, entre otras, necesitan una mayor precisión.

Varios desarrollos de empresas, gobiernos e instituciones afrontan a diario este desafío de encontrar nuevas técnicas para mejorar la precisión del posicionamiento, muchos de ellos con resultados exitosos y comprobables [9]. Pero estos desarrollos no alcanzan a la totalidad de los usuarios que los necesitan, en algunos casos las razones son económicas y en otros porque la región de residencia no se ve beneficiada por el servicio. Adquirir la infraestructura para montar un sistema de posicionamiento preciso, puede requerir de grandes inversiones. En otros casos puede implicar el abono oneroso de un servicio mensual. Aun así, y más allá de contar con la capacidad económica, en ciertas regiones no es posible acceder a las señales de corrección por características del terreno o por estar fuera del área de cobertura. Por este motivo es muy importante desarrollar soluciones tecnológicas que cubran estas necesidades.

Algunos requisitos operacionales para los que las constelaciones GPS y

GLONASS no se elaboraron, se enuncian a continuación:

- **Exactitud:** diferencia entre la posición estimada y la real.
- **Integridad:** confianza sobre la información total proporcionada.
- **Continuidad:** funcionamiento sin interrupciones no programadas.
- **Disponibilidad:** es la parte del tiempo durante la cual el sistema presenta simultáneamente la exactitud, integridad y continuidad requeridas.

Los sistemas de aumentación permiten mejorar la precisión posicional y garantizan que los GNSS actuales cumplan con los requisitos antes expuestos. Para superar estas limitaciones inherentes, la tecnología GPS ha diseñado y normalizado tres sistemas de aumentación: el sistema basado en aeronave (Aircraft Based Augmentation System – ABAS), el basado en tierra (Ground Based Augmentation System - GBAS), y el basado en satélite (Satellite Based Augmentation System – SBAS) [5]. Para aplicaciones en tiempo real, las correcciones de los parámetros de cada satélite de las constelaciones GNSS existentes (GPS y GLONASS) deberán ser transmitidas a los usuarios a través de equipos de radio VHF (GBAS) o si se requiere una amplia cobertura a través de satélites geoestacionarios que emitan pseudocódigos con información de corrección (SBAS) [10].

Esta línea de investigación pretende reducir la brecha existente entre estos desarrollos y los usuarios finales que necesitan realizar tareas con mayor precisión posicional que la entregada originalmente por los sistemas de posicionamiento, utilizando receptores de bajo costo. Además los sistemas GNSS están generando nuevas áreas de investigación, algunas previstas desde sus

comienzos, otras como resultado de una corriente innovadora en este campo que ha mostrado que estos sistemas pueden dar respuesta a muchos problemas que originalmente no fueron considerados.

Líneas de Investigación, Desarrollo e Innovación

- Posicionamiento de precisión
- Sistemas de Tiempo Real
- Protocolo NMEA
- Geometría satelital
- Sistemas Multiconstelación
- Posicionamiento diferencial de bajo costo
- Desarrollo de métricas para evaluar el error posicional

Resultados y Objetivos

Resultados

Como resultados se puede mencionar el trabajo de tesis doctoral de uno de los integrantes, que desarrolló una herramienta donde se implementa un conjunto de técnicas y algoritmos para el tratamiento de información que proviene de los receptores para incrementar la exactitud del posicionamiento.

Esta herramienta es totalmente configurable y posee portabilidad de manera que funcione en cualquier región donde no se cuenta con servicios de aumentación para mejorar la precisión.

Durante 2015 y 2016 se adquirió el equipamiento necesario para realizar nuevos experimentos que permitieron analizar nuevos parámetros relacionados con la mejora de la precisión posicional.

Para ello el equipo desarrolló una herramienta que permite la manipulación de los grandes volúmenes de datos obtenidos de los receptores GNSS que facilita el tratamiento y estudio de posibles indicadores y métricas

relacionadas con el error posicional. En CACIC 2016 WISS se presentó esta herramienta en el artículo “Procesamiento de sentencias NMEA-0183 para el análisis de la geometría satelital utilizando receptores GPS de bajo costo”.

Además se logró fortalecer el grupo de trabajo, como así también incentivar la producción científica en la UNdeC.

Objetivo General

Desarrollar un conjunto de técnicas, métodos y algoritmos de mejora a la precisión del posicionamiento utilizando receptores GNSS de bajo costo en un prototipo de GNSS diferencial.

Objetivos Específicos

Analizar técnicas y algoritmos de mejora de la precisión del posicionamiento para receptores GNSS de bajo costo. Analizar y diseñar técnicas para el análisis geométrico de la posición de los satélites. Estudiar el impacto que tiene el factor de dispersión espacial de los satélites en el error de posicionamiento de un GNSS de bajo costo. Definir métricas que permitan analizar y evaluar este error. Analizar técnicas y algoritmos de la integridad de los datos recibidos por el dispositivo. Analizar técnicas de aumentación de la precisión.

Formación de Recursos Humanos

El equipo de investigación de esta línea de trabajo en posicionamiento está compuesto por 5 docentes.

De los docentes: 1 es posdoctorado en Informática, 2 son doctores en Informática; un maestrando que presentará su tesis en la Universidad Nacional de San Juan; un doctorando que presentará su tesis en la Universidad Nacional del Centro de la Provincia de Buenos Aires.

Referencias

[1] Rao (2010) Global Navigation Satellite Systems. Tata McGraw-Hill Education, 478 pp.

[2] Misra P. & Enge P. (2010) Global Positioning System: Signals, Measurements, and Performance. New York, Ganhga-Jamuna Press, 590 pp.

[3] Enrique David Martí, David Martín, Jesús García, Arturo de la Escalera, José Manuel Molina and José María Armingol. (2012) “Context-Aided Sensor Fusion for Enhanced Urban Navigation”. Open Access Sensors, Article.

[4] <http://www.gps.gov/> Información oficial del Gobierno de los Estados Unidos relativa al Sistema de Posicionamiento Global y temas afines.

[5] Elliott D. Kaplan, Christopher J. Hegarty (2006) Understanding GPS Principles and Applications 2° Edition, pp. 26-28

[6] Gleason S. & Gebre-Egziabher D. (2009) Gns Applications and Methods. Artech House, 508 pp.

[7] Zandbergen P. A. & Arnold L. L. (2011) Positional accuracy of the Wide Area Augmentation System in consumer-grade GPS units. Computers & Geosciences Volume 37 Issue 7, Elsevier, pp. 883-892.

[8] Garmin International, Inc.: GPS 18x Technical Specification. (2008).

[9] Tolosa Juan Manuel. (2012) “Algoritmos y técnicas de tiempo real

para el incremento de la precisión posicional relativa usando receptores GPS estándar”. SEDICI, Universidad Nacional de La Plata.

[10] Alberto Riba, (2015) “Incremento de la precisión posicional relativa utilizando receptores GPS de bajo costo,” in IV Jornadas Científicas de Estudiantes Investigadores (IV-JCEI), UNDeC.

Completitud de los Métodos de Acceso a Datos Basado en Ontologías: Enfoques, Propiedades y Herramientas

Sergio Alejandro Gómez, Pablo Rubén Fillottrani

Laboratorio de I&D en Ingeniería de Software y Sistemas de Información (LISSI)
 Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur,
 San Andrés 800 - Campus de Palihue, (8000) Bahía Blanca, {sag,prf}@cs.uns.edu.ar
 Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

RESUMEN

Exploramos el acceso a datos basado en ontologías (OBDA). Una ontología brinda una visión conceptual de un repositorio de datos relacional. La instancia relacional está representada con una base de datos relacional mientras que el esquema de la base está representado en el lenguaje OWL2. OWL2 tiene tres *perfiles* que hacen ciertas concesiones en la capacidad expresiva para salvaguardar la eficiencia computacional de las operaciones pero comprometiendo la completitud del razonamiento. Estudiaremos cómo es posible extender tales capacidades de representación para mejorar la completitud de los métodos de OBDA.

Palabras clave: Acceso a datos basado en ontologías, Description Logics, Representación de conocimiento.

CONTEXTO

Esta línea de investigación se ejecuta en el contexto de un Proyecto de Grupo de Investigación de la Universidad Nacional del Sur y de un Proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de Buenos Aires (PIT-AP-BA) de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-PBA) llamado “*Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses*”.

1. INTRODUCCIÓN

Una ontología es una formalización de una parte de un dominio de aplicación dando una visión conceptual de los repositorios de datos y que se ha venido haciendo más y más popular desde el 2001 aproximadamente, en particular en las áreas de la Integración de Aplicaciones Empresariales, la Integración de Datos y la Web Semántica [Calvanese et al., 2006].

El acceso a datos basado en ontologías (OBDA) [Cali et al., 2012b; Calvanese et al., 2013] es visto como un ingrediente clave en la nueva generación de sistemas de información. En el paradigma de OBDA, una ontología define un esquema global de alto nivel de fuentes de datos (existentes) y brinda un vocabulario para consultas de usuario. Un sistema de OBDA reescribe tales consultas y ontologías en el vocabulario de las fuentes de datos y luego delega la evaluación real de la consulta a un sistema adecuado como por ejemplo una base de datos relacional o un motor Datalog [Kontchakov et al., 2013].

Como plantea [Stoilos, 2014], si bien la utilización de ontologías OWL provee un marco para la conceptualización formal y semántica de las fuentes de datos subyaciendo muchas aplicaciones modernas, el poder expresivo de OWL DL tiene un alto precio respecto a su complejidad computacional, aún luego del diseño de implementaciones con modernas optimizaciones, los razonadores OWL DL

todavía no fueron capaces de lidiar con grandes bases de datos conteniendo miles de millones de registros. Como una consecuencia de ello, en aplicaciones del mundo real, los desarrolladores a menudo emplean sistemas de respuestas de consultas escalables y eficientes que soportan solamente un perfil de OWL2 en alguno de sus perfiles, sea OWL2-EL, OWL2-QL u OWL-RL. En consecuencia, ya que OWL2 es compatible hacia atrás con OWL, tales sistemas cargan una ontología OWL pero ignoran todas sus partes que caen fuera del fragmento que soportan. En consecuencia, son *incompletas*; es decir, para algunas ontologías, consultas e instancia de base de datos fallarán en computar todas las respuestas ciertas. A pesar de que la escalabilidad es muy atractiva, la resolución de consultas incompleta no es aceptable en ciertas aplicaciones críticas como por ejemplo los dominios de la salud y la defensa militar; por lo tanto, el mejoramiento de la completitud por medio de la computación de tantas respuestas *perdidas* como sea posible sin afectar la performance sería muy beneficioso para muchas aplicaciones.

2. LÍNEAS DE INVESTIGACIÓN

Lenguajes de representación de ontologías: El estándar moderno de representación de conocimiento en la Web Semántica está dado por el lenguaje OWL2 [Hitzler et al., 2012]. OWL2 utiliza una sintaxis XML para el intercambio de datos pero su semántica está basada en las Lógicas para la Descripción (DL) [Baader et al., 2003]. Una ontología DL consiste de dos conjuntos finitos y mutuamente disjuntos: una Tbox que introduce la terminología y una Abox que contiene aserciones acerca de objetos particulares en el dominio de aplicación. En particular, OWL2 DL tiene tres *perfiles* con complejidad tratable: OWL2-EL, OWL2-QL y OWL2-RL dependiendo de la lógica para la descripción subyacente que le da significado [Hitzler et al., 2012]. Respecto de su complejidad temporal, OWL2-EL tiene tiempo polinomial para esquema y datos, OWL2-QL

permite respuestas a consultas en forma rápida (LOGSPACE) usando sistemas gestores de base de datos relacional vía SQL y OWL2-RL permite respuestas a consultas en tiempo polinomial usando bases de datos extendidas con reglas. En base a esto, cada perfil es útil para un cierto tipo de problema: OWL2-EL es útil para ontologías con una parte conceptual grande, OWL2-QL es útil para grandes datasets almacenados en RDBs, y OWL2-RL es útil para grandes conjuntos de datos almacenados como triplas RDF.

Acceso a datos basado en ontologías: En el acceso a datos basado en ontologías (OBDA) una ontología define un esquema global de alto nivel de una base de datos brindando un vocabulario para consultas, reescribiendo las consultas y ontologías en el vocabulario de las fuentes de datos y luego delegando la evaluación a un motor relacional o Datalog [Kontchakov et al., 2013]. Los sistemas de OBDA son importantes porque (i) brindan una vista conceptual del alto nivel de los datos, (ii) brindan al usuario un vocabulario conveniente para consultas, (iii) permiten al sistema enriquecer datos incompletos con conocimiento del dominio, y (iv) soportan consultas sobre múltiples fuentes de datos heterogéneas. Se pueden distinguir varios tipos de OBDA dependiendo del poder expresivo de las DL involucradas: (a) OBDA con bases de datos (por ejemplo, las lógicas de la familia de DL-Lite [Calvanese et al., 2013] y su implementación en XML OWL2-QL permiten la reducción de consultas conjuntivas sobre ontologías a consultas de primer orden sobre bases de datos relacionales estándar); (b) OBDA sobre motores Datalog (por ejemplo, las DL de la familia EL [Lutz et al., 2009] y su implementación XML OWL2 EL, Horn-*SHIQ* y Horn-*SROIQ* soportan un reducción a Datalog, y (c) OBDA con DLs expresivas (como *ALC* o *SHIQ* que requieren técnicas especiales para analizar consultas conjuntivas).

OBDA con bases de datos relacionales: Las lógicas de la familia DL-Lite [Calvanese et al., 2013] (y por ende el perfil OWL2 QL [Motik et al., 2012]) permiten una reducción de las consultas conjuntivas sobre ontologías a consultas de primer orden sobre bases de datos relacionales estándar. Así, una de las nociones fundamentales en OBDA es la de *reescritura de consultas*; en ella, dada una ontología (T,A) , el usuario formula una consulta $q(x)$ expresada en el vocabulario una terminología T de una ontología, la tarea del sistema de OBDA es la de reescribir $q(x)$ y T en una nueva consulta equivalente $q'(x)$ expresada en el vocabulario de los datos A tal que para cualquier conjunto de datos, las respuestas de $q(x)$ sobre la (T,A) son las mismas que las respuestas de $q'(x)$ sobre A . Así, el problema de consultar los datos A (cuya estructura es desconocida al usuario) en términos de una ontología T (accesible por el usuario) se reduce al problema de consultar A directamente; así cuando la A es modelada con la instancia relacional de una base de datos relacional, tal tarea se puede realizar muy eficientemente aprovechando los mecanismos de optimización de consultas mediante métodos estándar que permiten reexpresar un consulta de primerorden $q'(x)$ como un conjunto de consultas SQL. Otros métodos duales al anterior permiten realizar virtualización de Aboxes; en tal caso, la instancia relacional es reexpresada como una Abox A y se utiliza un motor DL tradicional para evaluar la consulta. La relación entre los conceptos de la ontología y los datos de la base de datos relacional se expresan con un mapeo M (obtenidos manualmente o en forma semiautomática), que es un conjunto de reglas $S(x) \sqcap \varphi(x,z)$ donde S es un nombre de concepto o rol de la ontología y $\varphi(x,z)$ es una conjunción de átomos con relaciones de bases de datos (almacenadas o en vistas) y un filtro relacional. Sin embargo y con el objetivo de mantener la complejidad temporal de los algoritmos involucrados en forma tratable, ciertas restricciones de integridad de algunos modelos conceptuales de datos (e.g. entidad-relación) no

pueden ser representadas en OWL2 QL (e.g. disyunciones en el miembro derecho de axiomas de inclusión -lo que impide modelar herencia múltiple- o restricciones numéricas en la cardinalidad de relaciones o roles). Otros problemas del enfoque de reescritura de consultas se da cuando la reescritura $q'(x)$ es muy grande para ser manejada exitosamente por el gestor relacional de bases de datos donde serán ejecutadas. El mapeo M puede contener las reglas que relacionan los términos de la ontología con el esquema de la base de datos. Entonces, dada una consulta $q(x)$ se puede obtener una reescritura $q'(x)$, la cual se puede desdoblar en una consulta SQL usando evaluación parcial. La evaluación parcial aplica resolución SLD a $q'(x)$ y al mapeo M y retorna aquellas reglas cuyos cuerpos contienen sólo átomos de la base de datos. Así, cada regla de $q'(x)$ resulta en una consulta SQL de tipo Select-Project-Join que se envía al manejador de bases de datos relacional para su ejecución. Es de notar, que a pesar de que el tema de la reescritura de consultas sobre DL-Lite (y en consecuencia sobre OWL2-QL) está bastante maduro, como DL-Lite no permite modelar todas las restricciones impuestas por ciertos modelos conceptuales (por ejemplo, al no poder usarse la disyunción en los miembros derechos de los axiomas de inclusión de DL-Lite, no es posible modelar *restricciones de cubrimiento* como que dos subclases son completas, i.e. no puede haber una tercera subclase de la superclase; otras restricciones no modelables implica decir que un atributo es funcional, y también las restricciones de cardinalidad tampoco pueden representarse en esta familia de DL requiriendo un lenguaje más expresivo [Kontchakov et al.]). El problema reside en que agregar tales restricciones destruye la propiedad de *reescritura de primer orden* con lo que al final terminaría impactando en la eficiencia del sistema.

OBDA con bases de datos no-relacionales: El paisaje del tema Bases de Datos se ha diversificado significativamente durante la última década, resultando en el surgimiento de

una variedad de bases de datos no relacionales (NoSQL), por ejemplo bases de datos XML y documentos JSON, almacenes clave-valor y bases de datos en forma de grafo [Botoeva et al., 2016]. [Harris and Seaborne, 2013] están investigando generalizaciones del marco OBDA para permitir la consulta de bases de datos arbitrarias a través de ontologías mediadoras utilizando el sistema de *Ontop* para OBDA y consultas SPARQL sobre la base de documentos *MongoDB* [Botoeva et al., 2016].

Problema de la completitud del razonamiento en el acceso a datos basado en ontologías:

Como explicamos previamente, los perfiles de OWL2 para modelar fuentes de datos sacrifican completitud para mantener un nivel de eficiencia aceptable. Sin embargo, en aplicaciones como la salud o la milicia, no es posible darse el lujo de perder respuestas a consultas. Repasamos ataques al problema encontrados en la literatura.

[Ma et al., 2006] presentan un *benchmark* para evaluar las capacidades de inferencia de sistemas ontológicos para los lenguajes OWL Lite y OWL DL. Cómo proveer aproximaciones escalables y de calidad para la resolución de consultas en DLs expresivas es un importante problema en representación de conocimiento. Debido al peor caso de complejidad para el problema de razonamiento de los perfiles expresivos de OWL2, muchas veces un cambio pequeño realizado por un ingeniero de conocimiento produce unas demoras de complejidad muy acentuadas. De la misma forma, cambiar de un razonador a otro puede producir cambios significativos en los tiempos de clasificación. [Goncalves et. al, 2012] investigan la identificación en una forma sistemática de los llamados *hot spots* (puntos calientes), que son subconjuntos difíciles de las ontologías, que, al ser removidos, producen una mejora significativa en la eficiencia. [Pan and Thomas, 2007] brindan una aproximación que garantiza sensatez y que transforma una ontología OWL expresada en una DL más expresiva en una aproximación maximal en una

DL tratable. [Pan et al., 2009] estudian el mismo problema pero tratando de mantener la completitud. [Tserendorj et al., 2008] presentan un acercamiento al razonamiento aproximado para ontologías, basado en el sistema KAON2, llamado SCREECH para razonar con ontologías OWL DL en la forma de su compilación en Datalog disyuntivo. El balance que logran se da en perder completitud en el razonamiento para ganar eficiencia.

3. RESULTADOS ESPERADOS

El *objetivo general* de este Plan de Trabajo involucra investigar mejoras en las capacidades de representación de conocimiento y razonamiento con ontologías en relación a los métodos, algoritmos y herramientas del acceso a datos en bases de datos basados en ontologías.

El conjunto de *objetivos particulares* de este Plan de Trabajo comprende la investigación de las extensiones a lenguajes de representación de conocimiento para modelar ontologías y su relación con el acceso a datos basado en ontologías. Esto integra los objetivos particulares de: (i) estudiar los lenguajes aptos para realizar acceso a datos basados en ontologías, (ii) estudiar las tecnologías asociadas (manejadores de bases de datos tradicionales y no tradicionales, razonadores DL, plataformas de implementación), (iii) estudiar y ampliar los límites del poder de representación y las clases de problemas atacables, (iv) entender las propiedades de los acercamientos planteados para (v) determinar las posibilidades de implementación computacional con miras a (vi) utilizar tal implementación en mejorar el desarrollo de las tareas mencionadas arriba en relación a la ingeniería de conocimiento de ontologías OWL2-EL, OWL2-QL y OWL2-RL en el contexto de la iniciativa de la Web Semántica.

4. FORMACIÓN DE RECURSOS HUMANOS

En relación a este plan de trabajo se está desarrollando la dirección de una tesis de maestría y una tesina de grado. Se espera ampliar el número de becarios y tesistas de grado.

5. BIBLIOGRAFÍA

- [Baader et al., 2003] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, editors. *The Description Logic Handbook Theory, Implementation and Applications*. Cambridge University Press, 2003.
- [Botoeva et al., 2016] E. Botoeva, D. Calvanese, B. Cogrel, M. Rezk, G. Xiao. OBDA beyond relational DBs: A study for MongoDB. In *29th Int. Workshop on Description Logics*, volume 1577 of CEUR Electronic Workshop Proceedings, 2016.
- [Cali et al., 2012b] A. Cali, G. Gottlob, T. Lukasiewicz. A general Datalog-based framework for tractable query answering over ontologies. *Web Semantics: Sciences, Services and Agents on the WWW*, No. 14, pp. 57-83, 2012.
- [Calvanese et al., 2013] D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Data Complexity of Query Answering in Description Logics. *Artificial Intelligence*, Volume 195, February 2013, pp. 335-360.
- [Eiter et al., 2012b] T. Eiter, M. Ortiz, M. Simkus, T. Tran, G. Xiao. Query Rewriting for Horn-SHIQ plus Rules. In *AAAI-2012*, Toronto, Canada, July 22-26, 2012.
- [Goncalves et al., 2012] R. Goncalves, B. Parsia, U. Sattler. Performance Heterogeneity and Approximate Reasoning in Description Logic Ontologies, *International Semantic Web Conference (ISWC 2012)*, pp. 82-98, 2012.
- [Harris and Seaborne, 2013] S. Harris and A. Seaborne. *SPARQL 1.1 Query Language: W3C Recommendation* 21 March 2013,
- [Hitzler et al., 2012] P. Hitzler, M. Krötzsch, B. Parsia, P. Patel-Schneider, and S. Rudolph. *OWL 2 Web Ontology Language Primer (Second Edition)*, W3C Recommendation 11 December 2012.
- [Kontchakov et al., 2013] R. Kontchakov, M. Rodriguez-Muro, M. Zakharyashev. *Ontology-Based Data Access with Databases: A Short Course. Reasoning Web: Semantic Technologies for Intelligent Data Access*, Vol. 8067, LNCS, pp. 194-229, Springer, 2013.
- [Lutz et al., 2009] C. Lutz, D. Toman, and F. Wolter. Conjunctive query answering in the description logic EL using a relational database system, (*IJCAI 2009*), pp. 2070-2075, 2009.
- [Motik et al., 2012] B. Motik, B. Cuenca Grau, I. Horrocks, Z. Wu, A. Fokoue, and C. Lutz. *OWL 2 Web Ontology Language: Profiles (Second Edition) - W3C Recommendation* 11 December 2012.
- [Pan et al., 2009] J. Pan, E. Thomas, Y. Zhao. Completeness Guaranteed Approximation for OWL DL Query Answering, *Proc. of DL*, 477, 2009.
- [Stoilos, 2014] G. Stoilos. *Ontology-Based Data Access Using Rewriting, OWL 2 RL Systems and Repairing*. In V. Presutti et al. (Eds.): *ESWC 2014*, LNCS 8465, pp. 317-332, 2014.
- [Tserendorj et al., 2008] T. Tserendorj, S. Rudolph, M. Krötzsch, P. Hitzler. *Approximate OWL-Reasoning with SCREECH*, Technical Report, Wright State University CORE Scholar, Computer Science and Engineering Faculty Publications, 2008.

Construcción de Sistemas Basados en Redes de Conocimiento para la Gestión

autor: **Gustavo Tripodi** - gtripodi@exa.unicen.edu.ar

co-autora: **Josefina Tripodi** - josefina.tripodi@alumnos.econ.unicen.edu.ar

Instituto de Investigación en Tecnología Informática Avanzada (INTIA) – Departamento de Computación y Sistemas - Facultad de Ciencias Exactas - Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA). Teléfono: +54 249-4439680. Dirección postal: Campus Universitario, Paraje Arroyo Seco, (7000) Tandil, ARGENTINA

1. RESUMEN

Las Tecnologías de la Información no constituyen un fin en sí mismo, son un medio, sin duda un medio particular que afecta nuestra manera de pensar y constituye uno de los caminos para mejorar la calidad de la Investigación. Dentro de este contexto, nos corresponde como investigadores la generación de las ideas, el diseño de las experiencias, la aplicación y la reflexión evaluativa que aporte conocimiento para el mejoramiento de la acción. En este sentido, el Análisis de Redes de Conocimiento ha pasado de ser una metáfora sugerente para constituirse en un enfoque analítico y un paradigma, con sus principios teóricos, métodos de software para el análisis y líneas de investigación.

Los Modelos propios surgidos de nuestras investigaciones anteriores y aplicados luego en cada Relevamiento, Análisis, Diseño e Implementación de las Organizaciones abordadas, nos develaron una hipótesis superadora: no solo podíamos construir la Red de Conocimiento para la Gestión, sino que se daban las condiciones para incorporar las propiedades específicas del tema abordado. Esto nos permitirá configurar los escenarios para implementar Sistemas que dan el sustento necesario para el control de las Operaciones, la gestión de trazabilidad, el acompañamiento Táctico y apoyo Estratégico a partir de los cambios de Estados y las relaciones entre las Tareas Estándares de una Red de Conocimiento para la Gestión.

Palabras claves: Tecnologías de la Información, Redes de Conocimiento, Redes de Gestión, Modelos Holísticos de las Redes, Trazabilidad, Acompañamiento Táctico, Apoyo Estratégico.

2. CONTEXTO

En nuestra Investigación anterior desarrollamos y consolidamos siguientes instrumentos:

- El sustento informático para el armado de los Modelos Holísticos de las Redes.
- El Sistema que nos permitió implementar las Redes definidas en los Modelos.
- Una Base de Datos/Conocimiento que contiene las propiedades y funciones necesarias.

Estos nos ayudaron en el contexto del abordaje de una Organización para construir las Redes de Gestión apropiadas. En el relevamiento deviene la confusión, la profusión de datos, el exceso de información y la incertidumbre de procesar, analizar e interpretar la información adecuada. Por consiguiente, los instrumentos desarrollados fueron esenciales para realizar relevamientos donde surgía el dilema de la incertidumbre de la cantidad de tiempo necesario para abarcar la cantidad y complejidad de información que podría ser recolectada sobre las Redes subyacentes. Ante esta caótica situación, sólo es posible reaccionar de forma selectiva (Rodríguez, 1999), donde nuestros Modelos se convirtieron en un filtro de comunicación de la información, para transformarla en un idioma coloquial dentro de una Organización (Guevara, 2002).

La Investigaciones anteriores y la presente se radican en el Instituto INTIA. El Proyecto que la contiene es el MERAIS V (período 2017/2019) categorizado para incentivos. Es transdisciplinar, en el que participan otros Institutos de la Facultad de Ciencias Exactas, de la Facultad de Ciencias Económicas, otras Universidades dentro del marco del PROMINF, Municipios y Empresas del Polo Informático de Tandil.

3. OBJETIVOS

• *Objetivo Principal*

Establecer las condiciones, propiedades, funcionalidades y herramientas para la construcción de Sistemas de Gestión sobre una plataforma que sustenta Redes de Conocimiento a través de la especificación de un Objeto Principal.

• *Objetivos específicos*

- Incorporar la máquina de estados y Gantt sobre las Tareas Estándar.
- Incorporar los Objetos Generales para clasificar las Propiedades del Objeto Principal.

4. INTRODUCCIÓN

El contexto en el que se procesa socialmente el conocimiento ha cambiado profundamente en los últimos años, como resultado de varias tendencias dinámicas y complejas. El nuevo contexto plantea fuertes desafíos al incorporar nuevos enfoques relativos al flujo de la información y a la difusión, así como a la apropiación de los conocimientos. En consonancia con tales tendencias, se ha consolidado un discurso que anuncia la emergencia de un nuevo tipo de sociedad a la que se denomina “sociedad de la información” o “sociedad del conocimiento” (Albornoz, 2007).

La base material de la nueva sociedad es la llamada “infraestructura global de la información”, de la cual las Redes se nutren. La rápida expansión de esta infraestructura está transformando la forma de Organización de la comunidad científica, incluyendo la Investigación en sí misma y los procesos de difusión y aplicación del conocimiento.

El sustento de estos paradigmas lo proveen las TICs, cuyas herramientas informáticas ponen de manifiesto el potencial y la efectividad de las Redes sociales online (‘software social’), operando en tres ámbitos, “las 3Cs”, de forma relacionada:

- Comunicación (nos ayudan a poner en común conocimientos).
- Comunidad (nos ayudan a formar e integrar comunidades).
- Cooperación (nos ayudan a hacer cosas

juntos).

La metodología que nos guía en el Proceso de Investigación para reconocer la Red de Conocimiento para la Gestión subyacente en una Organización está basada en dos actividades consecutivas y conjuntas: i) la modelación de la Red Comunicacional de una Organización o un Área de la misma a través de los Modelos Holísticos y ii) la generación de una nueva clasificación o la ubicación de la nueva Red dentro de la Taxonomía existente hasta ese momento.

Para aplicar la metodología transitamos el camino existente entre dos realidades isomorfas: una, la realidad en sí misma; la otra: su representación. Partimos desde Modelos Naturales que se encuentran dentro de las formas de expresión natural de los seres humanos, con alguna sofisticación de acuerdo al perfil. Luego se recorren diferentes Modelos que aportan precisión y claridad a la estructura de la Red emergente, hasta llegar al Modelo Estratégico que se pone al servicio del Plan Estratégico de la Organización.

5. LINEA DE INVESTIGACION Y DESARROLLO

El desarrollo de las interfaces gráficas permite plasmar mediante iconos y gráficos representaciones formales de la Red que se quiere implementar. Se trata de transformar una realidad dada y los Modelos Naturales que de ella existan, en distintos pero conexos y relacionados Modelos Formales Evolutivos.

Para nuestra Metodología, los Modelos más importantes y obligatorios son los Modelos de Análisis (de Escenario -MAE- y de Contenido -MAC-) sobre los cuales se construyó la Taxonomía. Los Modelos restantes nos ayudan a entender la situación con mayor precisión y lograr en el equipo una mayor cohesión con respecto a la Red de Conocimiento para la Gestión óptima a implementar.

Las etapas para Modelar la Red apropiada y óptima en una Organización o un Área de la misma son:

A. **Relevamiento:** El modelo más importante en esta etapa es el MRB (Modelo de Relevamiento Básico) que se complementa

con el MRE (Modelo de Relevamiento de Estímulo Comunicacional). La evolución de este nos lleva al MAE (Modelo de Análisis de Escenarios) y al MAC Modelo de Análisis de Contenido). El MRB se encuentra influenciado por las consideraciones y estructuras que el Usuario y la realidad nos plantea. En estos Modelos se involucran el Analista y los RRHH de la Organización con visión Táctica.

- B. **Análisis:** En esta etapa modelamos el MAE y el MAC que son esenciales para la implementación de la Red y la generación de una nueva Taxonomía o la inclusión en una existente. En esta etapa el Analista con todos los datos e información recolectados realiza el mapeo (podemos decir también la evolución y transformación) entre la realidad con sus Modelos Naturales hacia los Modelos de la Red de Conocimiento para la Gestión.
- C. **Diseño:** En esta etapa trabajamos sobre modelos que se encuentran en extremos opuestos y permiten dar un valor agregado al análisis que estamos realizando. Por un lado el MDE (Modelo de Diseño Estratégico) que aporta formalidad y claridad a los macro-procesos de la Organización, y en el cual el Analista trabaja junto a la Dirección de la Organización con una visión estratégica. En el polo opuesto la construcción del MDT (Modelo de Diseño Táctico) que despeja cualquier duda sobre los detalles de los procesos operativos; lo construye el Implementador bajo la supervisión del Analista y el asesoramiento del Usuario.
- D. **Documento final:** Este documento se denomina DRC (Deployment de la Red de Conocimiento para la Gestión) donde exponemos la Solución, más los conceptos y definiciones necesarios para la Implementación. Los apartados de este documento son cercanos a las funciones de cada uno de los integrantes del Equipo y cada uno participará, consensuará y aprobará el contenido según su perfil. El

DRC estará moderado por el Analista.

Con los resultados obtenidos en la Investigación anterior respecto a los Modelos y Etapas mencionados, probamos la siguiente Hipótesis: ***podemos construir a partir de los Modelos Holísticos propuestos la Red de Conocimiento para la Gestión que existe en cualquier Organización, Área o Proceso.*** A medida que avanzamos observando la verosimilitud de la primera Hipótesis e incorporando Organizaciones, dedicadas a muy diferentes actividades, probamos también la Hipótesis complementaria: ***los Modelos obtenidos responden a un patrón estructural y de funcionamiento, es decir, existe una Taxonomía de Redes de Conocimiento para la Gestión donde cada Organización puede mapearse.***

Los Modelos de Redes de Conocimiento para la Gestión fueron implementados en Organizaciones, Áreas y Procesos con características singulares y cubriendo un amplio espectro de situaciones.

Por otro lado las Redes se basan en Tareas y sus cambios de estado, por lo tanto los ensayos realizados nos permitieron normalizar las Propiedades de una Tarea Estándar y Universal. En el mismo sentido, definimos las Funcionalidades de esta Tarea que serán analizadas en profundidad en la presente Investigación.

El nuevo enfoque es encontrar un Objeto Principal con las Propiedades y Funcionalidades necesarias para representar objetos del mundo real y de esta manera, embebidos dentro de las Tareas Universales, conseguir aplicar sus cambios de estado. Como así también, generar su trazabilidad automática, establecer el Gantt con otras Tareas y conectarlos con el exterior y el interior de una Red. Este Objeto Principal estará compuesto por Objetos Generales que se utilizarán para clasificar las Propiedades. El Objeto General también necesitará de definiciones y funcionalidades que se reconocerán en los ensayos.

Para conectar el Objeto Principal con otros nodos y el mundo trabajaremos sobre las

variables y funcionalidades de una Central Inteligente que permita una comunicación precisa entre Usuarios en forma pública, privada y restringida.

6. RESULTADOS ESPERABLES

La aplicación de la metodología en el Trabajo de Investigación es de gran aporte para el Conocimiento y Aprendizaje sobre la construcción de Sistemas basados en Redes de Conocimiento para la Gestión.

Esta Investigación nos tendría que acercar más aún a la comprensión y solución de problemas de interacción y comunicación entre Usuarios. Para ello se analizarán y diseñarán los Objetos necesarios para la Implementación de Sistemas sobre una Plataforma de Red de Conocimiento. Estos Sistemas se enfocarán en la Trazabilidad de los objetos del mundo real a partir de las propiedades, cambios de estado y relación o Gantt entre Tareas Estándar/Universales.

El espectro de aplicación del tema es abarcativo y se ubica dentro del paradigma de innovación tecnológica, la cual ha cambiado radicalmente en estos últimos años. Los Modelos llamados "interactivos" reemplazaron el utilizado anteriormente, llamado "Modelo lineal". Los Modelos actuales subrayan el papel central que desempeña la concepción industrial sobre las relaciones entre las fases "hacia adelante" (ligadas al mercado) y las fases "hacia atrás" (ligadas a la tecnología) de la innovación, como también sobre las numerosas interacciones entre la ciencia, la tecnología y las actividades vinculadas a la innovación, tanto si se realizan en el interior de la empresa como si lo hacen en el marco de diversas cooperaciones.

Nos basamos en el Modelo del triángulo de las interacciones, elaborado por Sábato para ilustrar la estrategia más viable para alcanzar el objetivo, es el inspirado en la actual teoría de la triple hélice con la que algunos autores contemporáneos fundamentan determinadas estrategias de vinculación entre los actores políticos, académicos y empresariales como camino que conduce a la innovación (Albornoz, 2007). Desde nuestro punto de vista

el eje de esta triple hélice está compuesto por la Sociedad, las Organizaciones sin fines de lucro, las Entidades Educativas y los Emprendedores.

7. FORMACIÓN DE RECURSOS HUMANOS

Cuatro Becarios Proyecto PROMINF

Colaboradores en los siguientes Proyectos

- MERAIS IV - Coordinador Línea III Redes Sociales para la Gestión. Generación de una taxonomía de redes a partir de modelos holísticos
- MERAIS V - Construcción de sistemas basados en redes de conocimiento para la gestión.
- UTN Trenque Lauquen - Elaboración y Ensayo de Aplicación de Algoritmos de Ingeniería del Conocimiento en la Gestión de Redes Sociales

Proyectos Finales en la Facultad de Ciencias Económicas, UNICEN

- Tesis de Grado de alumno la Carrera Lic. en Administración.
 - Tesis Doctoral de Alumno de Administración
- Cinco Tesis de Grado de la Carrera de Ingeniería de Sistemas de la UNICEN*
- Aplicaciones Móviles (finalizada en 2016) vigentes:

- Seguimiento de Tesinas
- Gestión de Tutorías
- Coordinación de Prácticas Profesionales Supervisadas
- E-Learning para la Gestión Educativa

8. ANTECEDENTES

Proyectos de Investigación UNICEN-Tandil

Los Proyectos que se enumeran que contienen Investigaciones sobre Redes de Conocimiento para la Gestión

La denominación común de los Proyectos es MERAIS: Métodos de Razonamiento Aproximado en Investigación Socio-económica

- 2013-2016: MERAIS IV - Coordinador Línea III Redes Sociales para la Gestión. Generación de una taxonomía de redes a partir de modelos holísticos

Proyectos de Investigación UTN - Trenque Lauquen

- 2013-2014: Co-Director Elaboración y Ensayo de Aplicación de Algoritmos de

Ingeniería del Conocimiento en la Gestión de Redes Sociales

Créditos del Estado: Concurso Capital Semilla **Registros de Propiedad y Marca**

- Registro de Propiedad intelectual sobre los Modelos y Conceptos de Redes de Conocimiento para la Gestión
- Registro de Logo y Marca del Sistema de Red de Conocimiento para la Gestión

Proyectos implementados

- Municipio de Tandil y Vicente López
- Proyectos Koinonia, ONG para el apoyo tecnológico a entidades solidarias y Encontr@rg: portal para la búsqueda de ayuda solidaria
- Proyecto PROMINF: Gestión de Tutorías
- e-learning Asociación de Coop. Argentinas
- Portal CICE: Centro de Innovación y Creación de Empresas UNICEN

9. REFERENCIAS BIBLIOGRÁFICAS

- (Rodríguez, 1999) Rodríguez, Darío: *Diagnóstico Organizacional*. Capítulo 3: Modelos de Análisis. 1999
- (Guevara, 2002) Guevara Injoque, M. A.; Flores Nazario, C, R.: *ALLFusion: Administrando Procesos Empresariales*. 2002
- (Albornoz, 2007) Albornoz, M.; Estébanez, M. E.: *Política científica y tecnológica en Argentina*. Temas de Iberoamérica Globalización, Ciencia y Tecnología. Pp. 89. y Pp. 82. 2007

10. PUBLICACIONES PROPIAS SOBRE EL TEMA

- XVII Jornadas de Docentes Universitarios e Investigadores de Recursos Humanos de la Argentina y del Conosur - APUARH 2012 Tandil, Buenos Aires. Dos artículos sobre Redes Sociales para la Gestión. Autores: Gustavo Tripodi, Josefina Tripodi.
 - Ambiente para la Gestión del Conocimiento e Investigación del Comportamiento de Redes de Gestión - Generación de una taxonomía de Redes de Gestión Comunicacional Jerárquicas (RG CJ) a partir de modelos holísticos -
 - Construcción de la Red de Gestión Comunicacional Jerárquica (RG CJ) de un Partido Político

- Aciti, C.; Illescas G.; Tripodi G.; Tripodi J.; Xodo D.: *Redes Sociales – Redes de Gestión. Modelo holístico para la gestión de la información y el conocimiento en proyectos colaborativos interdisciplinarios*. XIV Workshop de Investigadores en Ciencias de la Computación. 1a ed. Posadas: Universidad Nacional de Misiones, 2012.
- Daniel Xodo, Ricardo Puleo Zubillaga, Gustavo Tripodi, Gustavo Illescas, Moisés Bueno - *Elaboración y Ensayo de Aplicación de Algoritmos de Ingeniería del Conocimiento en la Gestión de Redes Sociales*. XV WICC. Universidad Autónoma de Entre Ríos - Paraná

11. BIBLIOGRAFÍA

- Senge, Peter [et al.]: *La Quinta Disciplina en la práctica: estrategias y herramientas para construir la Organización abierta al aprendizaje*. 3a ed. Buenos Aires: Granica, 2009. 593 p.
- Blanchard, Ken; Randolph, Alan; Grazier, Peter: *Trabajo en equipo*. 1a ed. Buenos Aires: Deusto, 2005. 170 p. Líderes del Management: Recursos humanos.
- Freeman, Linton: *The Development of Social Network Analysis*. Vancouver: Empirical Pres, 2006; Wellman, Barry and S.D. Berkowitz, eds., 1988. *Social Structures: A Network Approach*. Cambridge: Cambridge University Press.
- Donati, Pierpaolo: *Repensar la Sociedad*, Ediciones Internacionales Universitarias, Madrid. 2006
- Ferraro, Ricardo A. Editores: Mario Albornoz; Claudio Alfaraz. *Redes de Conocimiento como nueva forma de creación colaborativa*. Editado por la Red Iberoamericana de Indicadores de Ciencia y Tecnología del Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo y la Oficina Regional de Ciencia para América Latina y el Caribe de la UNESCO. Agosto de 2006

Material de Internet: SocNet_TheoryApp.pdf. Social Network Analysis. Theory and Applications. Padiapress.com

Diseño de Sistema IoT de Monitoreo y Alarma para Personas Mayores

Daniel A. Giulianelli, Graciela De Luca, Sebastián Barillaro, Gerardo G. García, Esteban A. Carnuccio, Waldo A. Valiente, Mariano L. Volker

Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza

Dirección: Florencio Varela 1703 – CP 1754 – {sbarillaro, gdeluca, wvaliente, ecarnuccio, ggarcia, mvolker, dgiulian}@unlam.edu.ar

RESUMEN

Esta investigación intenta encontrar una forma de derribar esas barreras tomando ventaja de las tecnologías actuales como sistemas embebidos con sensores de movimiento, sensores biométricos, conexión inalámbrica, geoposicionamiento, etc. Haciendo uso de estos dispositivos y de la computación en la nube, se busca brindar a las personas adultas mayores y de su entorno mayor comodidad, autonomía, independencia, reducción de costos en los cuidados, mejores controles de la salud y mayor celeridad en la respuesta ante emergencias.

Se busca crear una solución que monitoree la salud del usuario y la reporte a los cuidadores, familiares y médicos a través de internet durante las 24hs los 7 días de la semana. Además, emita alertas en los casos que el usuario requiera atención inmediata.

En esta oportunidad, presentamos los avances realizados durante el primer año de desarrollo de este proyecto. Se analizan las diferentes líneas de investigación desarrolladas al seleccionar diversas configuraciones de sensores. Se describe algunos de los problemas no previstos relacionados con la arquitectura de hardware y software de la placa Intel Galileo y su solución.

Palabras clave: *computación en la nube, adultos mayores, geolocalización, monitoreo, IoT, Intel Galileo, Arduino, giroscopio, acelerómetro, Kallman.*

CONTEXTO

Nuestra Línea de Investigación es parte del proyecto *Sistema de monitoreo y alarma para personas adultos*

mayores ambulantes, dependiente de la Unidad Académica del *Departamento de Ingeniería e Investigaciones Tecnológicas*, perteneciente al programa de Investigaciones PROINCE de la Universidad Nacional de La Matanza, el cual es formado por docentes e investigadores de las carreras de ingeniería en informática e ingeniería en electrónica. Este proyecto es continuación de los trabajos que viene realizando el grupo de investigación, en sistemas operativos, computación de alto rendimiento, entre otros, en el área de Internet de las cosas (IoT).

1. INTRODUCCIÓN

Una caída no asistida a tiempo en una persona anciana puede provocarle su deterioro funcional, requiriendo de cuidadores o personas de apoyo. La prevención de caídas es tan importante que el ministerio de Desarrollo Social de la Nación a través la Dirección Nacional de Políticas para Adultos Mayores confeccionó una guía (1), en donde el tema de caídas es una parte principal de las recomendaciones antes situaciones de riesgo. Señalando que los efectos de las caídas pueden ser irreversibles sin una intervención adecuada e inmediata.

A nivel mundial, un informe de *World Health Organization* (2) indica que un tercio de la población de mayores de 64 años sufre algún tipo de caída cada año. Esta es una de las primordiales causas de lesiones en adultos mayores y una de las principales causas de muerte por accidente en los mayores de 75 años. Además, las fracturas de caderas provocadas por caídas representan al 90% de los casos. Una oportuna atención por parte de servicios médicos o familiares,

luego de una caída, puede brindar una recuperación y contención más rápida, mejorando la confianza, o hasta incluso salvándole la vida.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los sistemas de detección y aviso de caídas son herramientas que brindan rápida respuesta. Estos pueden ser clasificados en dos tipos: los sensibles al contexto y los portátiles.

Los sistemas sensibles al contexto requieren de la instalación de sensores diseminados por la vivienda. Estos sensores pueden estar formados por cámaras con análisis de imágenes para capturar el momento de la caída, sensores de infrarrojo, de ultra sonido, pisos con sensores de presión (3), etiquetas de identificación por radiofrecuencia (RFID), sensores inerciales y magnetómetros (4). Estas soluciones suele ser las más precisas ya que identifican el lugar de la caída. Además la persona adulta no tiene que cargar todo el día con un equipo (no intrusivo), pero es limitado por el alto costo de instalación y su rango de acción solo aplica en ámbitos cerrados o en donde esté instalado el equipamiento.

Los sistemas portátiles se basan en la incorporación de sensores en un aparato o ajustados en la ropa que supervisa las actividades de los adultos mayores en tiempo real, detectando caídas sobre la base de los cambios en algunos parámetros de movimiento. Los sensores utilizados para reconocer movimientos se componen de un acelerómetro y un giroscopio. Procesando la información sobre la orientación y la aceleración puede describir el movimiento de una persona. Este es un enfoque es de bajo costo y fácil de emplear, no requiere de instalación, funciona tanto dentro como fuera de la casa. En contra partida, al ser llevado en la ropa puede ser olvidado por la persona mayor. Si es integrado a un celular puede tener reticencia a utilizarlo. Además un golpe involuntario en el sensor puede provocar una falsa alarma.

El proyecto en el que estamos trabajando actualmente, se basa en el segundo tipo de sistema de detección de caídas.

3. RESULTADOS OBTENIDOS/ESPERADOS

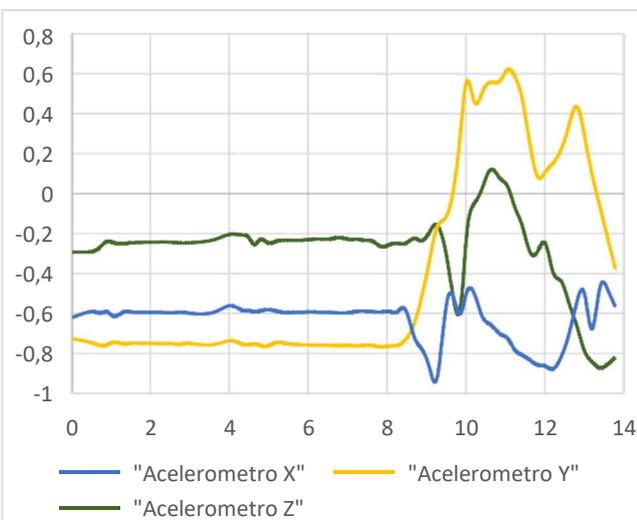
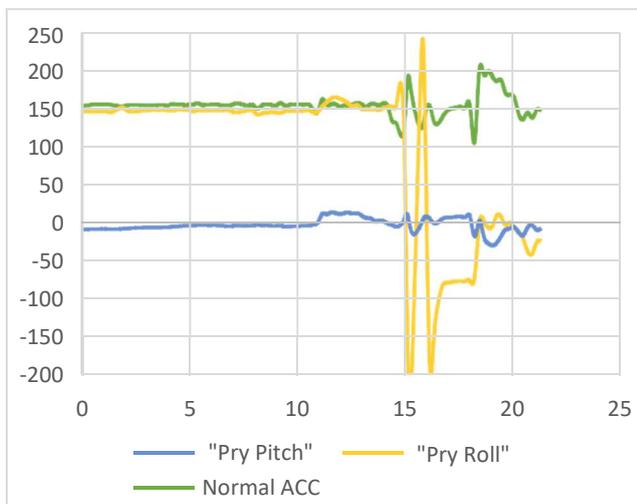
En una primera etapa en el desarrollo de esta investigación se ha podido diseñar la arquitectura del sistema. Esta arquitectura comprende los sensores biométricos y de movimiento; la placa de desarrollo de prototipos encargada de conectar los sensores, recolectar sus datos, procesarlos y enviar la información a un sistema en la nube encargado de analizarlos y comunicar las anomalías detectadas a los responsables del cuidado del usuario.

Debido a que las caídas es un riesgo que afecta a casi todos los ancianos, se comenzó con el estudio de la detección de este tipo de accidente. En ese sentido, se comenzó una línea de investigación que intentaba detectar la caída al interpretar una variación significativa en las lecturas de los datos entregados por un sensor giroscopio. Debido a la gran cantidad de falsos positivos y falsos negativos, se optó por utilizar un sensor acelerómetro. La certeza en la detección de caídas no mejoró con el cambio de sensor. Posteriormente, se comenzó una nueva línea de investigación en la que se hacía uso de ambos sensores *-giroscopio y acelerómetro-* combinando sus lecturas. Esta nueva forma de sensar las caídas resultó ser más eficiente, reduciendo la proporción de falsos negativos y falsos positivos.

Varias dificultades técnicas no previstas durante la etapa de diseño han sido resueltas durante el desarrollo del prototipo. Una de ellas fue el retardo introducido en la comunicación con el sensor giroscopio. Mientras que las lecturas funcionaban apropiadamente cuando se utilizaba un sistema Aduino, el sistema fallaba al ejecutar el mismo software en una placa de desarrollo Intel Galileo I. Luego de una investigación, se encontró que la placa de Intel introducía grandes demoras por la multiplexación de puertos GPIO. Además, el uso de un sistema operativo que no es de tiempo real aumentaba los retardos. Estas anomalías se resolvieron cuando comenzamos a utilizar la Placa de desarrollo Intel Galileo v2.0.

Posteriormente se procesan esos datos aplicando filtros Kalman para depurar la información ante lecturas erróneas. Se han realizado pruebas de campo en donde se parametrizó y validó el algoritmo de detección de caídas. Este algoritmo consiste en observar la variación

en las aceleraciones y en el giro y disparar un evento alarma cuando esa variación supera un umbral tolerable. No es relevante el sentido o dirección de la aceleración ni del giro. Lo que es interesante, es detectar un movimiento brusco característico de cualquier caída. También se han estudiado técnicas para detectar movimientos bruscos que no son producto de una caída (falsos positivos) y movimientos no bruscos por dentro del umbral de tolerancia pero que pertenecen a una caída (falsos negativos) requiriendo la atención inmediata de los cuidadores.



En las siguientes etapas del proyecto, se espera poder desarrollar un prototipo funcional para validar la estructura del sistema y la integración de sus partes. Posteriormente, se espera integrar cada vez más sensores biométricos, combinando los datos que entregan en algoritmos de fusión de sensores (*Sensor Fusion*, en inglés) para aumentar la certeza en las mediciones y detecciones de caídas.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación dentro del departamento de Ingeniería e Investigaciones Tecnológicas forma parte del trabajo que el Ingeniero Esteban Carnuccio se encuentra realizando para su maestría.

Dentro del grupo de investigación hay, además, dos investigadores en formación y se incorporó a un alumno que se encuentra finalizando su carrera, para realizar la iniciación en investigación.

5. BIBLIOGRAFÍA

- [1] M. Weiser y R. Gold, «The origins of ubiquitous computing research at PARC in the late 1980s,» 1999.
- [2] J. Wherton, R. Procter, P. Sugarhood y S. Hinder, «Co-production in practice: how people with assisted living needs can help design and evolve technologies and services,» 2015.
- [3] A. Alonso, «Siembra de Datos,» *MIT Information Technology*, pp. 66-69, Diciembre 2015.
- [4] G. Atkinson y K. Karimi, «What the Internet of Things (IoT) Needs to Become Reality,» Texas;USA, 2014.
- [5] biodatadevices, «biodatadevices,» 2014. [En línea]. Available: <http://www.biodatadevices.com/index.php?lang=es>.
- [6] J. S. & S. K. A. Aquino, «La tecnología como apoyo para alertas y ubicación de grupos de interés prioritario.,» *Revista Científica y Tecnológica UPSE*, 2(2), pp. 1-6., 2015.

Diseño y Desarrollo de Interfaces con Interacción Física Utilizando Dispositivos Móviles

Pablo M. Vera, Rocío A. Rodríguez, M. Roxana Martínez

CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

{pablomartin.vera, rocioandrea.rodriguez, roxana.martinez} @uai.edu.ar

RESUMEN

Todo cambia en búsqueda de una evolución, y los sistemas informáticos no escapan a ello. En los últimos tiempos las interfaces de usuario han evolucionado hacia formas más amigables con el auge de las pantallas táctiles. Los Smartphone son la clara prueba de ello. Pero justamente ellos son los que disponen de la tecnología para crear interfaces aún más cercanas a la gente. Gracias al hardware disponible en los teléfonos actuales es posible crear interfaces físicas que permitan controlar aplicaciones de forma más natural, moviendo girando, o simplemente apoyando el dispositivo sobre una superficie. Este proyecto tiene por objetivo diseñar y crear nuevas interfaces físicas para permitir controlar aplicaciones de forma natural más cercanas al mundo físico en el que vivimos.

Palabras clave: Interfaz, Interacción Física, Dispositivos Móviles, Pantallas Táctiles,

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Laboratorio de Algoritmos y Software del Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). El proyecto cuenta con financiamiento asignado y una duración de 2 años habiendo iniciado a mediados del 2016.

1. INTRODUCCIÓN

Los dispositivos móviles, en particular los teléfonos celulares, incorporan una gran cantidad de sensores que abren un abanico

muy grande a la hora de diseñar aplicaciones (ver figura 1).

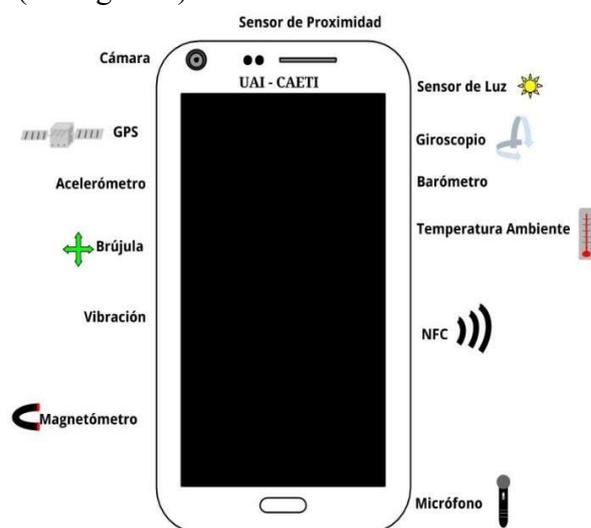


Figura 1. Sensores y Componentes de Hardware que pueden estar presentes en un Smartphone

No obstante, son pocas las aplicaciones que toman ventaja de los distintos sensores presentes. La interacción con el usuario generalmente se diseña pensando únicamente en el uso mediante la pantalla táctil, pero existen una gran cantidad de sensores que pueden contribuir a mejorar la experiencia del usuario o simplemente brindar alternativas de uso.

En el presente proyecto se trabajará en el desarrollo de una aplicación nativa en Android por ser el sistema operativo que más inserción tiene en el mercado [1]. Y por medio de diversos sensores poder desencadenar acciones que permitan al usuario interactuar con el dispositivo de una forma innovadora. HCI (Human Computer Interacción [2]) es un área en constante

crecimiento que incorpora a los móviles actualmente en el análisis de la forma en que el usuario interactúa con el mismo.

La innovación está en poder aprovechar el hardware disponible y crear una interfaz en la que un usuario que no sea experto pueda interactuar con la aplicación en forma más natural sacando provecho de distintos sensores para ejecutar acciones simples desde abrir una aplicación con sólo apoyar el dispositivo sobre una superficie, hasta realizar acciones más complejas como controlar la aplicación mediante sensores de movimiento.

Actualmente las personas llevan consigo dispositivos móviles cada vez más potentes, con múltiples procesadores y memoria RAM de 2, 3 y hasta 6GB, pero lo más importante es que incorporan una gran cantidad de sensores que permiten detectar el ambiente que lo rodea e incluso características del uso o de la persona que lo está utilizando. Estos sensores permiten ser utilizados para mejorar la usabilidad de las aplicaciones.

Uno de los primeros sensores incorporados en los dispositivos móviles fue el acelerómetro que permite que al girar el teléfono y ponerlo horizontalmente la pantalla gire y se adapte para una correcta visualización. Rápidamente surgieron distintas aplicaciones que sacan provecho de dicho sensor, por ejemplo, los juegos, donde girando el equipo es posible manejar un auto como si se estuviera girando el volante creando una interfaz más natural y relacionada con el modelo mental que una persona tiene en cuanto a cómo se conduce un automóvil.

Los seres humanos estamos acostumbrados a interactuar físicamente con los objetos, nuestro modelo mental del uso de las cosas indica que si vemos un picaporte de una puerta sabemos que debemos girarlo, si vemos un botón se puede apretar si hay una manija es porque para abrir algo hay que tirar de ella. El diseño de las interfaces digitales trae el desafío de salir de ese modelo mental ya que se debe interactuar con una pantalla

plana donde los objetos son simulados y no hay contacto físico posible más que tocar la pantalla.

“La gente forma modelos mentales mediante la experiencia, la formación y la instrucción. El modelo mental de un dispositivo se forma en gran parte mediante la interpretación de sus actos percibidos y de su estructura visible” [3].

Interactuar con los dispositivos requiere de cierta habilidad, “La habilidad manual, consistente en pulsar o mover ratones, teclas, punteros, iconos, pantallas táctiles, etc., para activar eficazmente las funciones implementadas en el Instrumento infotecnológico, modela sus redes neuronales y podría configurar una clase nueva y útil de inteligencia instrumental (basada en un lenguaje, no verbal ni relacionado con una mente lógica, sino expresable con los dedos), a la que podríamos llamar inteligencia digital...” cabe destacar que no todas las personas han nacido con ese tipo de habilidad o la han logrado adquirir. [4].

Para las nuevas generaciones que son “nativos digitales” resulta más simple generar esa habilidad. Prensky describe a los nativos digitales como aquellos que “han nacido y se han formado utilizando la particular “lengua digital” de juegos por ordenador, video e Internet” [5]. Para aquellos que nacieron con la tecnología ese nuevo modelo mental de interacción ya lo traen incorporado, pero para muchas personas mayores es extremadamente difícil adaptarse y poder controlar una interfaz gráfica.

Este proyecto busca diseñar interfaces innovadoras para romper esa barrera creando interfaces alternativas para los sistemas aprovechando los distintos sensores de los dispositivos, principalmente de los sensores de movimiento y del NFC (Near Field Communication) [6].

En cuanto a los sensores de movimiento ya no solo los equipos cuentan con acelerómetro,

sino que se complementan usando además brújula y giroscopio lo que permite saber con mayor precisión como se mueve el equipo en cada momento, también se puede complementar con la ubicación del usuario tomada por el GPS, e incluso es posible medir si la persona está caminando y cuanto camina con el contador de pasos que incorporan diversos dispositivos.

NFC es una tecnología de transmisión de información a muy corta distancia, “desarrollada en 2002 por las empresas Phillips y Sony. NFC opera a 13,56MHz, con una tasa de transmisión de 424Kbps y un alcance de lectura de aproximadamente 10cm” [7]. En pequeñas etiquetas que pueden ser autoadhesivas o llaveros que no requieren de batería para funcionar, se puede almacenar información y fácilmente leer esa información al apoyar el teléfono sobre las mismas. Estas etiquetas pueden ser configuradas para iniciar automáticamente una aplicación, cambiar la configuración del celular o realizar un pago, entre otras posibilidades. NFC puede utilizarse para disparar aplicaciones en forma automática sin tener que entrar a la pantalla del celular y buscar un ícono determinado y también puede utilizarse para realizar aplicaciones más complejas combinando su uso con otros sensores.

[1] LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Los ejes principales del trabajo son:

- Análisis de los sensores disponibles en los equipos móviles comprendiendo su funcionamiento. De esta forma se podrá determinar las posibilidades y limitaciones al momento de acceder a la información de los sensores mediante una aplicación
- Diseño de interfaces innovadoras que mediante el uso de los sensores permitan manejar distintas aplicaciones.

[2] RESULTADOS OBTENIDOS/ESPERADOS

Se ha comenzado a investigar sobre el manejo de los sensores en el sistema operativo Android analizando sus posibilidades y limitaciones de forma de determinar la mejor forma de uso y como pueden aplicarse para el control de aplicaciones.

Se ha comenzado a desarrollar una primera aplicación de prueba a fin de familiarizarse con el uso de los sensores.

El objetivo es desarrollar una aplicación que permita configurar rápidamente un temporizador girando el dispositivo para establecer el tiempo deseado. Se asume el teléfono va estar apoyado sobre una superficie plana para poder utilizarlo. Girando el dispositivo a la derecha se debe aumentar el tiempo del temporizador y girando a la izquierda se disminuye el máximo tiempo configurable será de 1 hora.



Figura 2. Esquema de uso de una aplicación para controlar un temporizador mediante sensores de movimiento

Esta aplicación debe poder controlarse de dos formas:

- De forma tradicional utilizando la pantalla táctil arrastrando la aguja al tiempo deseado.
- Sin necesidad de tocar la pantalla (por ejemplo, para cuando la persona está cocinando y tiene las manos sucias). Este segundo modo solo se habilitará

cuando el dispositivo esté apoyado sobre una superficie plana y se configurará girándolo. Una vez que se detecte la finalización del movimiento se le dará un tiempo al usuario (por ejemplo 5 segundos) y se comenzará la cuenta atrás automáticamente.

Adicionalmente se ha comenzado a investigar sobre la tecnología NFC para poder realizar paneles de control que con solo apoyar el dispositivo sobre un indicador dispare determinada aplicación o acción.

Por ejemplo, suponiendo que en la cocina hay un tag NFC sobre la mesada, se podría configurar para que al apoyar el celular sobre el mismo se habrá el temporizador mencionado anteriormente sin necesidad de buscar la aplicación e iniciarla manualmente.

También utilizando NFC se puede realizar un panel de control para adultos mayores, donde poniendo varios tags juntos se le puedan configurar acciones predeterminadas. Por ejemplo tener tags puestos en fotos de sus familiares y que con solo apoyar el dispositivo sobre el mismo se inicie automáticamente la llamada, y otro tag que le recuerde la fecha en que debe ir al médico o comprar medicamentos (ver figura 3).



Figura 3. Modelo de Panel de NFC en donde cada etiqueta puede desencadenar una acción diferente

Este mismo concepto podría aplicarse para realizar un panel de control para gente no vidente con indicaciones en braille para iniciar distintas aplicaciones que la persona pueda necesitar.

También mediante el uso de NFC se planea realizar una aplicación para que niños con TEA (Trastorno del Espectro Autista) puedan reconocer distintos objetos y trabajar el tema de secuencias. Esta última aplicación se realizará en conjunto con otro grupo de investigación del CAETI que trabaja con interfaces para niños con TEA.

[3] FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 7 personas.

- 3 Docentes (2 de Postgrado y 1 de Grado).
- 4 (2 Alumnos de Posgrado, 2 Alumnos de Grado).

En el área de dispositivos móviles se encuentran en realización 2 tesis de maestría en la UAI (Universidad Abierta Interamericana) y una de doctorado en la UNLP (Universidad Nacional de La Plata).

5. BIBLIOGRAFIA

- [1] Gartnet. "Ventas Globales de Smartphone según sistema operativo" (2015).
- [2] Dunlop, Mark, and Stephen Brewster. "The challenge of mobile devices for human computer interaction." *Personal and ubiquitous computing* 6.4 (2002): 235-236.
- [3] Norman, Donald A. *La psicología de los objetos cotidianos*. Vol. 6. Editorial Nerea, 1998.
- [4] Sáez Vacas, Fernando. "Nativos digitales, inteligencia digital ¿Homo digitalis?." *Telos* 86 (2011): 6-8.
- [5] Prensky, Marc. *Nativos e inmigrantes digitales*. Distribuidora SEK, 2010.
- [6] NFC, NearFieldCommunication.org <http://nearfieldcommunication.org/>

- [7] Tapia, Dante I., et al. "Identificación por radiofrecuencia: fundamentos y aplicaciones." Proceedings de las primeras Jornadas Científicas sobre RFID. Ciudad Real, Spain (2007): 1-5.

Educción Emocional de un Individuo en Contextos Multimodales en Computación Afectiva

Federico M. Rossi[†], Enrique P. Calot[†], Jorge S. Ierache[†]

[†] Laboratorio de Sistemas de Información Avanzados,
Departamento Computación,
Facultad de Ingeniería, Universidad de Buenos Aires
Av. Paseo Colón 850 - C1063ACV - Buenos Aires - Argentina
Tel +54 (11) 4343-0893 / 4343-0092.
{ecalot, frossi, jierache}@lsia.fi.uba.ar

Resumen

El presente artículo describe la investigación en desarrollo en el contexto de computación afectiva, específicamente en la educación de emociones de un individuo, mediante la estimulación de los participantes con imágenes y videos de alto contenido emocional. Se describen las interfaces biométricas de lectura del pulso cardíaco, EEG empleando un BMI, la cadencia de tecleo, registro de voz y filmación del rostro del participante. Se presenta el ambiente de educación denominado cámara de inmersión emocional, y el ciclo simplificado de recolección de datos en un contexto multimodal.

Palabras clave: *Dispositivos de bioseñales, Educación de emociones, Adquisición multimodal.*

Contexto

El Proyecto articula líneas de trabajo de *Keystroke Dynamics* y *Brain Machine Interface* (BMI) orientados tanto a la seguridad como a la extracción de patrones emocionales en el ser humano. El proyecto es financiado por UBACyT 2014-2017 GEF, con radicación en el Laboratorio de Sistemas de Información Avanzados de la Facultad de Ingeniería de la Universidad de Buenos Aires. El mencionado proyecto se realiza con la colaboración del ISIER de FICCTE UM, apoyando específicamente en el área de aplicación

de BMI, en el marco del PID 01-001/12/14 [Calot Ierache, *et al.*, 2015; 2016].

Introducción

La computación afectiva es el estudio y el desarrollo de sistemas y dispositivos que pueden reconocer, interpretar, procesar y estimular las emociones humanas. Rosalind Picard define que la computación afectiva es "la informática que se relaciona con las emociones, no sólo con las consideradas más importantes, como la alegría o la tristeza, sino también con el interés, el aburrimiento o la frustración, que son las que se dan en relación con los ordenadores". [Picard, 2000]. La educación emocional de un individuo en un contexto multimodal permite la registración simultánea de diversa información biométrica que permite educir el estado emocional del individuo.

La computación afectiva y el desarrollo de aplicaciones en esta área requieren el registro y aprendizaje en los futuros sistemas que interactúen con los humanos en forma empática. En el contexto multimodal de registros biométricos como así también las técnicas de estimulación, son diversas, solo realizaremos una breve descripción en esta sección de las empleadas en nuestro proyecto.

Estimulación

Para realizar la estimulación hemos solicitado un conjunto de datos recopilado por el *Center for*

the Study of Emotion & Attention de la University of Florida en Gainesville, FL, Estados Unidos. El conjunto contiene una base de datos de imágenes (IAPS [Lang *et al.*, 2008]). La base de datos es utilizada en diversos estudios científicos que analizan emociones y es considerada un estándar de facto.

También hemos descargado cuatro videos con alto contenido emocional.

El experimento barre los cuatro cuadrantes del modelo excitación-valencia (*valence-arousal*) [Bradley y Lang 1994] (ver figura 1) considerando cuatro fases en las que se aplica una misma secuencia estimuladora.

Primeramente, se muestra un subconjunto de imágenes tomadas de IAPS (100 imágenes, 25 por fase) y luego un video o audio, cuya temática se encuentra relacionada a la fase afectiva transitada.



Figura 1. Modelo de excitación-valencia

Finalizada la estimulación mediante imágenes, se solicita al voluntario que complete una encuesta SAM (por sus siglas en inglés *Self-Assessment Manikin*) propuesta por Lang [1980]. (véase apartado VI de la sección Interfaces de Entrada).

A su vez, cumplida la estimulación mediante video o audio, se vuelve a solicitar al voluntario que complete una encuesta SAM. Acto seguido, se procede a la captura de voz y dinámica de tecleo.

La figura 3 resume el proceso completo.

Interfaces de Entrada

I. RITMO CARDÍACO

Se adquirió el ritmo cardíaco mediante una pulsera colocada al participante (Xiaomi Mi Band 1s) en su muñeca, que registra el ritmo cardíaco de manera continua en el tiempo.

II. DINÁMICA DE TECLEO

La dinámica de tecleo se adquiere luego de la estimulación al solicitar al participante que ingrese un texto de carga emocional neutral, este texto consta de varias frases mezcladas de manera aleatoria y las mismas serán utilizadas en cada uno de los cuadrantes. Es guardado tanto el *flight time* como el *hold time* para así poder reconstruir todos los eventos del teclado. En este aspecto se realizaron avances importantes dentro del laboratorio publicándose un nuevo método para clasificación de patrones y una réplica del método más utilizando hasta el momento [González *et al.*, 2016].

III. AUDIO

El experimento guardó todas las pistas de audio en estéreo para su posterior análisis. El formato es OGA. Se le solicita al participante leer un texto de carga emocional neutral con una metodología análoga a la de la dinámica de tecleo.

IV. VIDEO

La captura de video se realiza en todo momento de manera continua. Para analizar gestos emocionales en el video del usuario es posible utilizar la herramienta *clmtrackr* [Øygard, 2015], que detecta caras y luego interpreta emociones a partir de un set de datos de puntos mapeados de una base de datos. Su implementación utiliza modelos locales restringidos ajustados por puntos de referencia regularizados (*constrained local models fitted by regularized landmark mean-shift*) desarrollado por Audun Mathias Øygard y basado en el artículo [Saragih *et al.*, 2011]. Para detectar los estados emocionales se utilizó la base de datos MUCT [Milborrow *et al.*, 2010] agregando algunas imágenes propias de Øygard. El modelo

emocional lo provee con 18 coeficientes [Calot *et al.*, 2016].

El formato de video es de 640 x 480, con *codec* VP9 y *framerate* de 30 cuadros por segundo. Puede, además, ser analizado con cualquier otra herramienta.

V. EEG

Una Interfaz Cerebro-Computadora (Brain-Machine Interface, BMI) facilita la comunicación entre las funciones mentales o cognitivas creadas a partir del cerebro de una persona, captando las señales eléctricas, para ser procesadas, clasificadas y comunicadas con aplicaciones o dispositivos específicos [Ierache *et al.* 2012, 2013a, 2013b; Pereira *et al.*, 2015; Calot *et al.*, 2015]. En el contexto de emociones, el trabajo de [Ierache *et al.*, 2014; 2015] presenta los primeros resultados del Estado Emocional Centrado en Estímulos, Aplicando Interfaz Cerebro-Máquina. Este experimento cuenta con el uso del *headset* Emotiv EPOC para la adquisición de señales encefalográficas mediante 14 electrodos, de los cuales se consideraron inicialmente los ubicados en la corteza prefrontal (F3, F4, AF3, AF4, según el sistema internacional 10-20 [Jasper, 1958]).

OpenViBE es utilizado para recibir y almacenar las muestras crudas para luego aplicar filtros sobre los electrodos y realizar el posprocesamiento. Esto puede ser utilizado para calcular la posición del estado emocional en el modelo bidimensional de excitación-valencia [Feldman, 1995; Oude Bos, 2006]

VI. ENCUESTA (SAM)

Se realiza una encuesta a último momento preguntando al participante en qué estado emocional se encuentra. La misma hace uso de la metodología ampliamente utilizada para capturar emociones: SAM (por sus siglas en inglés *Self-Assessment Manikin*) propuesta por Lang [1980].

Tal como lo describen sus autores en Bradley y Lang [1994], SAM es un método no verbal fácil de administrar para evaluar rápidamente el placer, la excitación y la dominancia asociadas con la reacción emocional de una persona ante un evento.

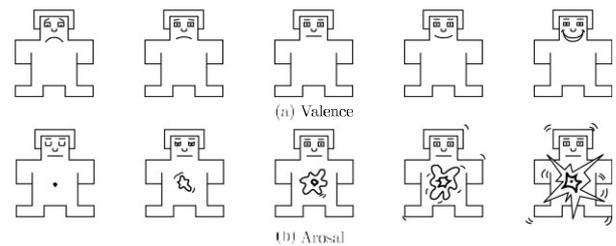


Figura 2. Dibujos SAM para (a) valencia/*valence* y (b) excitación/*arousal*.

Consiste de 5 dibujos que caracterizan si la persona se encuentra contenta o triste (*valence*) y cuán excitada está (*arousal*). Para cada dominio existen 9 puntos, un punto por cada figura, y un punto intermedio entre las mismas (ver figura 2).

Resultados, Objetivos

Se ha creado una base de datos a partir de un experimento que permitirá un posterior análisis y el avance en líneas de investigación relacionadas con tesis de grado y doctorado (ver sección Resultados y Objetivos). Para ello doce personas han participado de una experimentación en la Cámara de Inmersión Emocional diseñada a tal efecto en las instalaciones del Laboratorio de Sistemas de Información Avanzados en la Facultad de Ingeniería de la UBA. Para la realización de los experimentos se implementó un ámbito que denominamos “Cámara de Inmersión Emocional” (EIC, ver figura 3) es un espacio de aproximadamente 4m² con una sala de control adyacente donde los investigadores controlan permanentemente la conectividad y las capturas del experimento. Cuenta con una sala de espera donde los participantes aguardan a que los investigadores preparen los equipos y comience el experimento. La sala se encuentra aislada tanto en sonido como en visibilidad, encontrándose las ventanas cubiertas con paneles de color verde para facilitar la extracción de las imágenes del video.

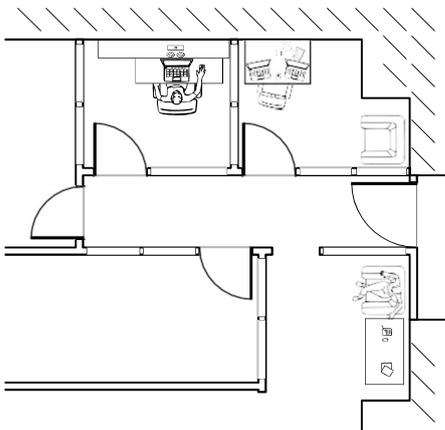


Figura 3. Diseño de la Cámara de Inmersión Emocional

Con el experimento realizado se ha creado un *dataset* que registra las bioseñales enunciadas en la sección introducción. El data set generado es aplicado en el marco de patrones de tecleo y su variación en función del estado emocional del individuo. La arquitectura de ECI (cámara de inmersión emocional) permite el que los datos registrados contribuyan en futuras líneas de investigación en el contexto de computación emocional. La figura 4 resume el proceso de experimentación desarrollado en la ECI.

Formación de Recursos Humanos

Una tesis de grado en ingeniería cuyo objetivo se encuadra en el desarrollo de un *framework* que para la generación de eventos y captura de señales cerebrales, su comportamiento ante estímulos afectivos y el establecimiento de un marco comparativo entre estas y otros tipos de bioseñales, profundizando variadas perspectivas que mejoren la interpretación y el entendimiento de las emociones humanas en el contexto de la computación afectiva.

Una tesis doctoral cuyo objetivo se enmarca en encontrar variaciones emocionales en la cadencia de tecleo de los individuos.

Líneas de investigación encuadradas bajo el objetivo de combinar múltiples bioseñales que se complementan entre sí para reforzar la educación el estado emocional de los humanos.

Referencias

BRADLEY, M. M.; & LANG, P. J. (1999). International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings (Tech. Rep. No. B-2). Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

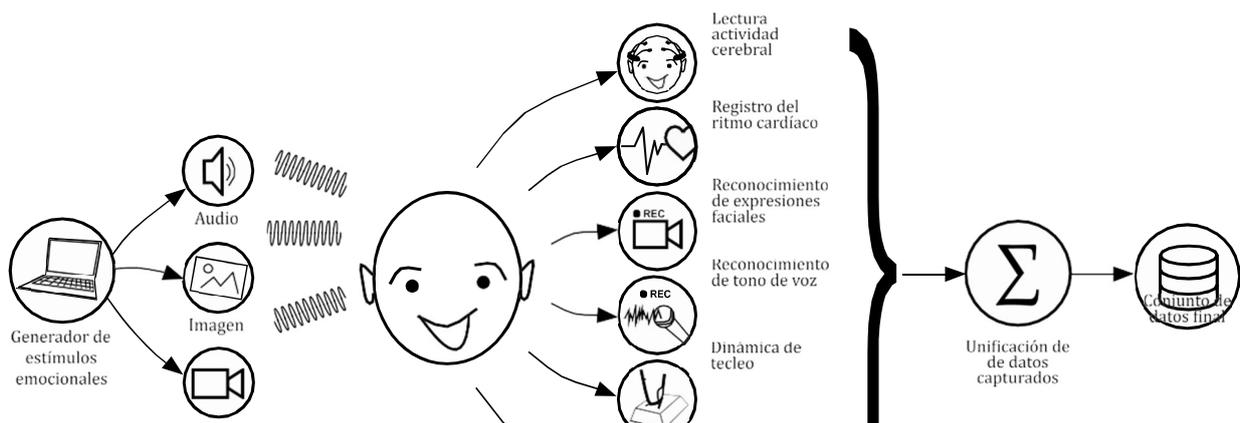


Figura 4. Ciclo simplificado de recolección de datos.

- BRADLEY, M. M.; & LANG, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.
- CALOT, E. P.; ACETO, E. L.; RODRÍGUEZ, J. M.; LIGUORI, A.; OCHOA, M. A.; MERLINO, H.; FERNÁNDEZ, E.; GONZÁLEZ, N.; PIRRA, F.; & IERACHE, J. S. (2015). Líneas de investigación del Laboratorio de Sistemas de Información Avanzados. In *Proceedings del XVII Workshop de Investigadores en Ciencias de la (pp. 4 p.)*.
- CALOT, E. P.; RODRÍGUEZ, J. M.; & IERACHE, J. S. (2014). Improving versatility in keystroke dynamic systems. In *Computer Science & Technology Series. XIX Argentine Congress of Computer Science, Selected papers*. ISBN 978-987-1985-49-4, Editorial de la Universidad Nacional de La Plata (EDULP).
- CALOT, E., ROSSI, F. M., GONZALEZ, N. F., HASPERUÉ, W., & IERACHE, J. S. (2016). Avances en educación de dinámica de tecleo y el contexto emocional de un individuo aplicando interfaz cerebro computadora. In *XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina)*.
- FELDMAN, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of personality and social psychology*, 69(1), 153.
- GONZALEZ, N., CALOT, E. P., & IERACHE, J. S. (2016). A Replication of Two Free Text Keystroke Dynamics Experiments under Harsher Conditions. In *Biometrics Special Interest Group (BIOSIG), 2016 International Conference of the (pp. 1-6)*. IEEE.
- IERACHE, J.; PEREIRA, G.; IRIBARREN, J.; & SATTOLO, I. (2013^a). Robot Control on the Basis of Bio- electrical Signals. *International Conference on Robot Intelligence Technology and Applications (RiTA)*, 2012. Gwangju, Korea on December 16— 18, 2012. *Series Advances in Intelligent and Soft Computing of Springer*.
- IERACHE, J.; PEREIRA, G.; CALOT, E.; & IRIBARREN, J. (2013^b). Framework for Brain Computer Interface implemented to control devices in the context of home automation. *XIX Congreso Argentino de Ciencias de la Computación*, ISBN 978-897-23963-1-2.
- IERACHE, J.; PEREIRA, G.; & IRIBARREN, J. (2012). Demostración de los resultados en la integración de Interfases Lectoras de Bioseñales aplicadas al Control de un Robot. *VII Congreso Educación en Tecnología y Tecnología en Educación Universidad Nacional del Noroeste de la Provincia de Buenos Aires. UNNOBA, demos educativas*. ISBN 978-987-28186-3-0.
- IERACHE, J.; PEREIRA, G.; & IRIBARREN, J. (2014). Navigation Control of a Robot from a remote location via the Internet using Brain-Machine Interface; *Robot Intelligence Technology and Applications 2.: Springer*. 2014. p 297-310. ISBN 978-3-319-05581
- IERACHE, J.; NERVO, F.; PEREIRA, G.; & IRIBARREN, J. (2015). Emotional Status Focused on Stimuli by Applying Brain-Machine Interface, *Computer Science and Technology Series Selected Papers XX Congreso Argentino de Ciencias de la Computación*, pp 245-256, Edulp. Edición 2015, ISBN 978-987-1985-71-5
- JASPER, H. H. (1958). "Report of the committee on methods of clinical examination in electroencephalography: 1957". *Electroencephalography and Clinical Neurophysiology*. 10 (2): 370–375. doi:10.1016/0013-4694(58)90053-1.
- LANG, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. Páginas 119–137. Ablex, Norwood, NJ, 1980. ISBN 978-0-89391-023-5.
- LANG, P. J.; BRADLEY, M. M.; & CUTHBERT, B.N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.
- MILBORROW, S.; MORKEL, J.; & NICOLLS, F. (2010). The MUCT Landmarked Face Database. *Pattern Recognition Association of South Africa*.
- OUDE BOS, D. (2006). EEG-based emotion recognition- The Influence of Visual and Auditory Stimuli. *Capita Selecta (MSc course)*.
- ØYGDARD, A. M. (2015). Emotion detection example. *clmtrackr*. URL: http://auduno.github.io/clmtrackr/examples/clm_emotiondetection.html. Vigente al 2 de febrero de 2016.
- PEREIRA, G.; NERVO, F.; IRIBARREN, J.; CALOT, E. P.; IERACHE, J. S.; & MAZZA, N. (2015). Líneas de investigación aplicada al empleo de interfases cerebro-maquina. In *XVII Workshop de Investigadores en Ciencias de la Computación, Salta*
- PICARD, R. (2000). *Affective Computing*. USA: MIT Press, ISBN: 978-0-262-66115-7.
- SARAGIH, J. M.; LUCEY, S.; & COHN, J. F. (2011). Deformable Model Fitting by Regularized Landmark Mean-Shift. *Int. J. Comput. Vision*, 91(2), 200–215. <http://dx.doi.org/10.1007/s11263-010-0380-4>

Elaboración de Modelos Cuantitativos para la Evaluación Institucional

Ana Funes, Mario Berón, Aristides Dasso

SEG / Departamento de Informática / Facultad de Ciencias Físico-Matemáticas y Naturales
/ Universidad Nacional de San Luis

Ejército de los Andes 950, D5700HHW San Luis, Argentina

+54 (0) 266 4520300, ext. 2126

{afunes, arisdas, mberon}@unsl.edu.ar

Resumen

Dentro del contexto de desarrollo de modelos de evaluación de sistemas complejos, esta investigación tiene como objetivo concretar la creación de un modelo que permita la evaluación cualitativa y cuantitativa de instituciones y de sus servicios, en especial instituciones educativas de tercer nivel. A tal efecto, proponemos un modelo en el cual comenzamos por establecer un conjunto de características (los requisitos que deben cumplir las instituciones) en un formato jerárquico. Esta tarea la llevamos a cabo basándonos en los estándares propuestos por el Consejo Nacional de Evaluación y Acreditación Universitaria (CONEAU) establecidos a través de decretos, resoluciones y guías, y aplicando el método de evaluación Logic Scoring of Preference (LSP), lo que nos lleva a construir un modelo adecuado que permita obtener un resultado numérico final entre 0 y 100. Dicho número es un indicador del porcentaje de adaptación a la norma del sistema bajo evaluación.

Palabras clave: Educación tercer nivel. Evaluación de instituciones educativas. Logic Scoring of Preference (LSP). Modelo de Evaluación. Métodos de Evaluación. Planes de estudio. Universidad.

Contexto

Este trabajo de investigación se viene

llevando a cabo dentro del SEG (Software Engineering Group), en el ámbito de la Universidad Nacional de San Luis y se encuentra enmarcado dentro de una línea de investigación sobre la construcción de modelos de evaluación de sistemas complejos, la cual viene desarrollándose desde hace unos años en el ámbito del Proyecto de Ciencia y Técnica P-031516 “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad” (Director: Daniel Riesco, Co-Director: Roberto Uzal; Acreditado con evaluación externa; Financiamiento: Universidad Nacional de San Luis), en donde se han obtenido resultados plasmados en varias publicaciones nacionales e internacionales (ver por ejemplo [1], [3], [4], [20], [21], [22]).

Introducción

Existen diversos organismos nacionales e internacionales de evaluación los que cuentan con guías para evaluar no sólo instituciones académicas sino también los servicios prestados por estas (cursos, carreras, títulos, investigación, transferencia, etc.) además de cómo certificar las instituciones. A veces estos organismos también ofrecen un ranking de las instituciones evaluadas.

Muchos de estos organismos que establecen rankings de universidades siguen diversos criterios [2], [27], [11], [12]. Por ejemplo, the Association of American Universities consideró

la aplicación de Star Metrics [23] para la evaluación de aspectos de las universidades en distintos niveles [5]. En [24], [25] se pueden hallar varios ejemplos de métodos de evaluación en distintos países. En la literatura pueden encontrarse también diferentes criterios, por ejemplo [6], [10], [12], [28].

En la República Argentina, CONEAU es el organismo encargado de la tarea de acreditación y evaluación de instituciones universitarias. CONEAU tiene como misión institucional "...asegurar y mejorar la calidad de las carreras e instituciones universitarias que operan en el sistema universitario argentino por medio de actividades de evaluación y acreditación de la calidad de la educación universitaria." [8].

Emplear y construir modelos de evaluación de sistemas complejos, entre los que se encuentran las instituciones educativas y los sistemas de enseñanza, constituye una necesidad primordial para garantizar que los mismos cumplan con requisitos establecidos por aquellos organismos encargados de su evaluación y seguimiento [11]. En este sentido, es importante garantizar que los evaluadores y los evaluados puedan considerar que tanto las medidas, como las herramientas y los métodos empleados para dicha tarea sean las más adecuadas, objetivas y transparentes posible.

En este sentido, el método LSP (Logic Score of Preferences) [16], [13], [14], [15] brinda la posibilidad de contar con un modelo cuantificable, que reduzca la subjetividad inherente a la evaluación de muchos de los aspectos involucrados.

LSP es un método que se basa en el empleo de una lógica continua, que permite la creación de funciones complejas de evaluación y su aplicación en la evaluación y comparación de sistemas de índole general, permitiendo la creación de modelos precisos y fácilmente adaptables a las necesidades del usuario, en este caso los estándares de calidad de una institución o carrera universitaria.

Líneas de Investigación, Desarrollo e Innovación

El proceso general propuesto por el método LSP es mostrado en la Figura 1. El desarrollo e identificación de la lista de características principales a tener en cuenta (requisitos) corresponde al primer nivel del *árbol de requerimientos* que el método prescribe construir en una de sus etapas.

Cada una de estas características del primer nivel comprende varios ítems o sub-categorías más específicas de acuerdo con propiedades similares, tales que las mismas puedan ser razonablemente agrupadas, y así sucesivamente, llegando hasta ítems que no se descomponen más, es decir, las 'hojas' del árbol de requerimientos, que el método llama *variables de performance*.

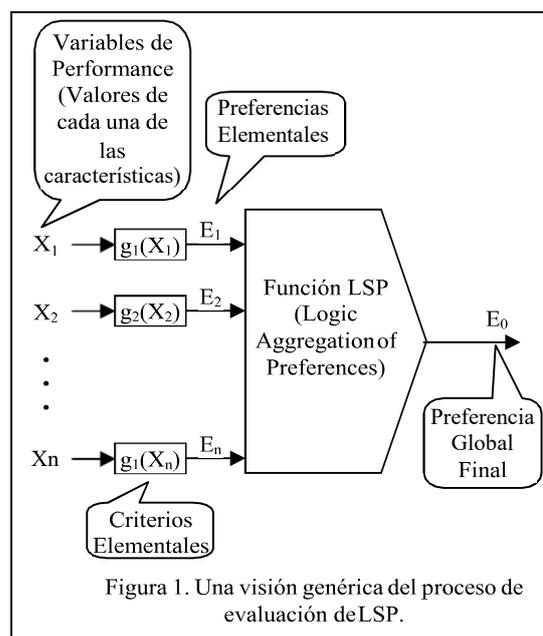


Figura 1. Una visión genérica del proceso de evaluación de LSP.

Las hojas del árbol de preferencias son empleadas para construir una *estructura de agregación* junto con los operadores de Lógica Continua provistos por el método. Estos operadores o funciones GCD (Generalized Conjunction Disjunction) nos permiten agregar los valores observados de cada una de las variables de performance, previamente

mapeados a valores en el intervalo $[0, 100]$ llamados *preferencias elementales*, por medio de funciones llamadas *criterios elementales*. Las preferencias elementales representan el grado de satisfacción o cumplimiento de un requisito del sistema bajo evaluación.

La estructura de agregación tiene como objetivo entregar como resultado final un único valor entre 0 y 100 (preferencia global final E_0 en la Figura 1) que representa el grado de adecuación o de satisfacción de todas las características consideradas para el sistema bajo evaluación.

Es importante destacar que el árbol de preferencias y las variables de performance consideradas son elección de quién o quienes construyen el modelo sobre la base de las necesidades del usuario. A partir del árbol de requerimientos, se pueden generar diversas estructuras de agregación, usando operadores de la lógica de LSP, como modelos de evaluación, previa clasificación de los distintos aspectos que el usuario considere mandatorios, opcionales y deseables. En este caso, se pretende consensuar un modelo que sirva de base para unificar criterios para los pares evaluadores.

Así, por ejemplo, si se considera alguna de las características que una institución debería poseer, el valor asignado a la correspondiente variable de performance corresponderá a la valoración que se haga del mismo por medio de una métrica directa o indirecta adecuada; dicho valor se transformará, por medio del correspondiente criterio elemental, en un valor del intervalo $[0,100]$. La relación entre el valor asignado a la variable y el intervalo forma parte también de los criterios de los evaluadores lo que se verá reflejado en la definición de la función de criterio elemental.

Resultados y Objetivos

En esta primera etapa nos encontramos desarrollando un modelo de evaluación que sigue de manera general las directivas establecidas en las normas y guías que emplea

CONEAU.

CONEAU abarca un campo muy extenso en el tipo de evaluaciones y acreditaciones que realiza. Dichas actividades incluyen no sólo planes de estudio de muy diversas disciplinas, sino que, como consecuencia de esto, debe evaluar también instituciones educativas en las que dichos planes se implementan. En esta primera etapa, hemos decidido abordar el problema construyendo un modelo para la evaluación de la carrera de grado Licenciatura en Ciencias de la Computación, y como paso siguiente, se planifica abarcar otras carreras afines trabajando sobre la base de este primer modelo. Para ello, nos basamos en los Contenidos Curriculares Básicos, Cargas Horarias, Criterios de Intensidad de la Formación Práctica y Estándares para la Acreditación establecidos en la Resolución 786/2009 del Ministerio de Educación [19] para la carrera en cuestión y otras carreras afines. Asimismo, para la confección del modelo y las métricas, se tiene en cuenta la información que las instituciones universitarias deben entregar a CONEAU al momento de la presentación a la convocatoria de acreditación. Esta información, al momento de la evaluación, se encuentra disponible para los pares evaluadores a través del sistema

CONEAU Global (www.coneau.gob.ar/global), lo cual permite la instanciación del modelo de evaluación aquí propuesto.

Como parte del trabajo futuro, esperamos, en una etapa siguiente, ampliar y calibrar el modelo producido así como extenderlo y adaptarlo a otras carreras afines.

También consideramos explorar en la generación de modelos que incorporen otros ítems para la evaluación, teniendo en cuenta los aportes existentes en otros países y en la literatura en general, algunos de los cuales han sido citados más arriba, con el objeto de ampliar y/o mejorar nuestra propuesta.

Formación de Recursos Humanos

Dentro del SEG (Software Engineering

Group), en el ámbito de la Universidad Nacional de San Luis, en el que se realiza el Proyecto de Incentivos código 22/F222 “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad”, se han llevado a cabo numerosas tesis de grado y de posgrado.

Entre otros, nos hemos concentrado en la evaluación de sitios de gobierno electrónico lo que ha dado como resultado una tesis de maestría en 2010 [7]; mientras que hay otras dos en preparación, una de ellas sobre la propuesta de un modelo integral para la evaluación de la calidad del atributo Accesibilidad al Contenido Web, y la otra sobre creación y evaluación de modelos LSP en un contexto MDA.

La propuesta aquí presentada, también, tiene como objetivo ser motivo de tesis, como lo han sido la construcción de otras herramientas en el ámbito del proyecto.

Referencias

- [1] Aristides Dasso, Ana Funes, Mario Peralta, Carlos Salgado, “Una Herramienta para la Evaluación de Sistemas”, Workshop de Investigadores en Ciencias de la Computación, WICC 2001, Universidad Nacional de San Luis, San Luis, Argentina, May 2001.
- [2] American Evaluation Association, “Guiding Principles For Evaluators”. file:///C:/Users/arisdas/AppData/Local/Temp/GuidingPrinciplesPrintable.html. (Recuperado febrero 2017).
- [3] Ana Funes, Aristides Dasso, “Web Application Frameworks Evaluation”, CONAIIISI 2014, 13 y 14 de noviembre de 2014, San Luis, Argentina. pp. 1063-1070. ISSN: 2346-9927.
- [4] Ana Funes, Aristides Dasso, Carlos Salgado, Mario Peralta, “UML Tool Evaluation Requirements”. Argentine Symposium on Information Systems ASIS 2005. Rosario, Argentina. September 29-30, 2005.
- [5] Association of American Universities, “STAR METRICS Workshop”. <http://www.aau.edu/members/article.aspx?id=12784&terms=star+metrics> (Recuperado febrero 2017).
- [6] Bruce G Charlton and Peter Andras. “Evaluating universities using simple scientometric research-output metrics: total citation counts per university for a retrospective seven-year rolling sample Science and Public Policy, 34(8), October 2007, pages 555–563
<https://www.staff.ncl.ac.uk/peter.andras/SPPoet07Charlton.pdf> (Recuperado febrero 2017).
- [7] Castro, Marcelo. “Análisis de las propiedades y atributos propios de sitios de gobierno electrónico”. Tesis de Posgrado, Maestría en Ingeniería de Software, Universidad Nacional de San Luis, Argentina, 2010.
- [8] CONEAU. “¿Qué es la CONEAU?”. http://www.coneau.gob.ar/CONEAU/?page_id=7 Recuperado marzo 2017.
- [9] Consortium of Universities for Evaluation Education (CUEE)
<http://www.evaluationeducation.ca/index.html> (Recuperado febrero 2017).
- [10] ELLEN NOLTE, CAROLINE VIOLA FRY, ELEANOR WINPENNY, LAURA BRERETON. “Use of outcome metrics to measure quality in education and training of healthcare professionals”. RAND Europe. February 2011. https://www.rand.org/content/dam/rand/pubs/workings_papers/2012/RAND_WR883.pdf (Recuperado febrero 2017).
- [11] FINDING MEANINGFUL PERFORMANCE MEASURES FOR HIGHER EDUCATION <http://researchanalytics.thomsonreuters.com/m/pdfs/higher-ed-exec-report.pdf> (Recuperado febrero 2017).
- [12] Innovation Policy Platform (IPP). “Metrics and evaluation for universities and PRIs”. <https://www.innovationpolicyplatform.org/content/metrics-and-evaluation-universities-and-pris> (Recuperado febrero 2017).
- [13] J. J. Dujmovic and A. Bayucan, “Evaluation and Comparison of Windowed environments”, Proceedings of the IASTED International Conference Software Engineering (SE'97), pp 102-105, 1997.
- [14] J. J. Dujmovic, “A Method for Evaluation and Selection of Complex Hardware and Software Systems”, The 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG96 Proceedings, vol. 1, pp.368-378, 1996.
- [15] J. J. Dujmovic, “Quantitative Evaluation of Software”, Proceedings of the IASTED International Conference on Software Engineering, edited by M.H. Hamza, pp. 3-7, IASTED/Acta Press, 1997.
- [16] Jozo J. Dujmovic, “Continuous Preference Logic for System Evaluation”, IEEE Transactions on Fuzzy Systems, Vol. 15, N° 6, December 2007
- [17] M. Castro, A. Dasso, A. Funes. “Modelo de Evaluación para Sitios de Gobierno Electrónico”. 38 JAIHO/SIE 2009, Simposio de Informática en el Estado 2009, Mar del Plata, Argentina, August 26- 28, 2009.

- [18] Metrics for the Evaluation of Knowledge Transfer Activities at Universities http://ec.europa.eu/invest-in-research/pdf/download_en/library_house_2008_unico.pdf (Recuperado febrero 2017).
- [19] Ministerio de Educación de la República Argentina. Resolución 786/2009. Boletín Oficial, N° 31.667, 4 de junio de 2009, pg. 91
- [20] N. Debnath, A. Dasso, A. Funes, G. Montejano, D. Riesco, R. Uzal, "The LSP Method Applied to Human Resources Evaluation and Selection", Journal of Computer Science and Information Management, Publication of the Association of Management/International Association of Management, Volume 3, Number 2, 2000, ISBN 1525-4372, pp.1-12.
- [21] N. Debnath, A. Dasso, A. Funes, G. Montejano, D. Riesco, R. Uzal, "The LSP Method Applied to Human Resources Evaluation and Selection", Journal of Computer Science and Information Management, Publication of the Association of Management/International Association of Management, Volume 3, Number 2, 2000, ISBN 1525-4372, pp.1-12.
- [22] Narayan Debnath, Aristides Dasso, Ana Funes, Roberto Uzal, José Paganini. "E-government Services Offerings Evaluation Using Continuous Logic". 2007 ACS/IEEE International Conference on Computer Systems and Applications, AICCSA '2007, Amman, Jordan. Sponsored by IEEE Computer Society, Arab Computer Society, and Philadelphia University, Jordan. May 13-16, 2007.
- [23] National Institutes of Health (NIH) and the National Science Foundation (NSF). "STAR METRICS". <https://www.starmetrics.nih.gov/> (Recuperado febrero 2017).
- [24] Organisation for Economic Co-operation and Development (OECD) <http://www.oecd.org/>
- [25] ShanghaiRanking Consultancy "Ranking Resources". <http://www.shanghairanking.com/resources.html> (Recuperado febrero 2017).
- [26] ShanghaiRanking Consultancy. "Academic Ranking of World Universities". <http://www.shanghairanking.com/aboutarwu.html> (Recuperado febrero 2017).
- [27] Times Higher Education (THE) <https://www.timeshighereducation.com/> (Recuperado febrero 2017).
- [28] UNESCO. "Research Evaluation Metrics". Published in 2015 by the United Nations Educational, Scientific and Cultural Organization, 7, place de Fontenoy, 75352 Paris 07 SP, France. <http://unesdoc.unesco.org/images/0023/002322/232210E.pdf> (Recuperado febrero 2017).

Estrategias de IoT para Lograr Ciudades Digitales Seguras, más Inclusivas y Sustentables

Javier Díaz, Paula Venosa, Laura Fava, Néstor Castro, Diego Vilches, Fernando López
LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata

Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

{jdiaz, pvenosa, lfava, ncastro, dvilches}@info.unlp.edu.ar, flopez@mail.linti.unlp.edu.ar

RESUMEN

El término "smart cities" -ciudades inteligentes-, es más que una frase de moda: es una de las tendencias más desafiantes vinculada a otro concepto actual "Internet de las cosas" o simplemente IoT. El término IoT se refiere generalmente a escenarios donde la capacidad de cómputo y la conectividad de las redes se extienden a objetos, sensores y elementos cotidianos - no computadoras personales-, permitiendo que estos dispositivos generen, intercambien y consuman datos.

En todo el mundo, más y más ciudades se están convirtiendo en "inteligentes" mediante el uso de tecnología avanzada para mejorar sus servicios y la calidad de vida de sus ciudadanos, y esta revolución liderada por IoT está ayudando tanto a las grandes metrópolis, como a las pequeñas ciudades a transformarse en ciudades inteligentes y verdaderamente conectadas. En este artículo se describen líneas de investigación y desarrollo que forman parte de un proyecto integral destinado a proveer soluciones que aporten a la transformación de ciudades en entornos digitales y más inteligentes, teniendo en cuenta la seguridad de la información desde su concepción.

Palabras claves: smart cities, Internet de las Cosas, IoT, sensores, seguridad en IoT, wearables, medio ambiente.

CONTEXTO

El Laboratorio de Investigación de Nuevas Tecnologías Informáticas LINTI de la Facultad de Informática, viene trabajando en proyectos relacionados con cuidado del medio ambiente y el uso racional de los recursos energéticos.

Asimismo, desde el año 2002 el LINTI también trabaja de manera integral en temas relacionados con la seguridad y privacidad, tanto en la investigación como en formación académica de grado y postgrado, así como en campo a través de diversas experiencias con organismos nacionales e internacionales. Es por ello que a la hora de abordar nuevos desafíos como los proyectos a continuación descriptos, se analizan los aspectos de seguridad inherentes y las posibles mejoras a implementar desde su concepción.

Las líneas de trabajo que se describen en este artículo se desarrollan en el LINTI y están enmarcadas en el proyecto *Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro*, acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

1. INTRODUCCIÓN

En la actualidad, se escucha con frecuencia el término "smart cities", sin embargo, no existe una definición

absoluta y estandarizada sobre este concepto. Algunos de los aspectos que son necesarios para considerar que un determinado servicio es provisto por una ciudad como parte de un acercamiento a la idea de smart cities, tiene que ver con la recolección, análisis y puesta en disponibilidad de cualquier tipo de información que sea útil a todos los ciudadanos. Por otro lado, el conocimiento y uso de la información, permite de alguna manera, mejorar la calidad de vida y el vínculo del ciudadano con los servicios que la ciudad provee [1].

Los aspectos claves para el desarrollo de ciudades inteligentes son:

- La existencia de una infraestructura digital moderna que permita el desarrollo de los servicios inteligentes (Internet, sensores, aplicaciones y redes sociales, aplicaciones móviles, etc.) combinada con una política de acceso abierto al consumo y la generación de la información pública en todo momento.
- Pensar y desarrollar los servicios centrados en el ciudadano y sus necesidades finales.
- Transparencia en la comunicación a los ciudadanos de los resultados y la performance alcanzada por los servicios inteligentes.
- Seguridad en la comunicación de los componentes de la infraestructura y en la transmisión de datos.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Como se ha mencionado, el concepto de Smart City y de IoT están muy relacionados ya que no consisten solamente en la conexión de cada vez más personas, sino en el planteamiento de un mundo digital en el que “todo” podrá estar conectado entre sí. En este contexto,

en el LINTI se está trabajando en varias líneas de investigación y desarrollo como movilidad urbana, eficiencia energética, producción florifrutihortícola, salud y deporte.

Si bien los servicios a desarrollar a partir de las líneas anteriormente citadas podrían resolverse de manera individual, alcanzan mayor potencial cuando se integran en una visión conjunta con el resto de los servicios. Por ejemplo, una aplicación que ayude a gestionar el tráfico en tiempo real de una ciudad podría ayudar a identificar las zonas con mayor concentración de contaminación ambiental, información que a su vez puede utilizarse para analizar si las condiciones ambientales en determinadas zonas producen una mayor incidencia de enfermedades respiratorias.

Lo anteriormente dicho nos lleva a pensar en soluciones donde se hace necesario aproximaciones del tipo interdisciplinarias.

Por otro lado, y en los casos que la información además de estar disponible sea íntegra y confiable, la misma resulta fundamental no solo para quien gestiona las soluciones sino también para los ciudadanos que las usan, ya que los mismos tienen la posibilidad de ejercer un rol de contralor y de esta manera, poder exigir a los responsables el cumplimiento de alguna normativa relacionada con dicha información.

La Figura 1 ilustra una infraestructura digital moderna que permite el desarrollo de los servicios inteligentes usando Internet, sensores, aplicaciones, etc., combinada con una política de acceso abierto al consumo y la generación de la información pública, de manera segura.

en tiempo real, el consumo de energía y que permitirá a los responsables gerenciar eficientemente el uso de la misma. Si bien el prototipo está dirigido a nuestra Facultad, el análisis de las tecnologías y los resultados podrían aplicarse a otros escenarios.

En esta misma línea se está trabajando en otro proyecto denominado *WENU*, que plantea abordar la problemática del consumo energético a partir de la toma de mediciones de los dispositivos de mayor consumo, los cuales tienen el alto impacto en épocas de temperaturas extremas: los sistemas de climatización [7]. Es común observar el uso irresponsable de aires acondicionados y estufas eléctricas en comercios y edificios de oficinas, por ejemplo, lugares vacíos por horas o hasta días donde por accidente queda encendido alguno de estos dispositivos u oficinas con un aire acondicionado configurado en una temperatura extremadamente baja. Por esto se propone una solución que promueve la concientización sobre el uso responsable de los sistemas de climatización, alertando a los usuarios cuando la temperatura de un ambiente es muy diferente de una medición de control tomada en el mismo edificio en un lugar sin climatización artificial y dando la posibilidad de apagar aires acondicionados de forma remota o automática cuando haya indicios de que un lugar está vacío pero estos dispositivos siguen encendidos. Como parte de este proyecto se están estudiando protocolos y mecanismos para proveer una comunicación segura entre los distintos componentes [8][9][10], analizando ventajas y desventajas para luego aplicarlos al resto de los proyectos de IoT.

Wearables

Las nuevas tecnologías informáticas

también están alcanzando la industria del deporte, sin embargo en Argentina aún no se encuentran ofertas interesantes destinadas a optimizar el entrenamiento de profesionales del deporte. Los entrenadores y deportistas siguen utilizando hojas de papel y pizarras y en el mejor de los casos planillas de cálculo o software de alguna plataforma específica para realizar los seguimientos del desarrollo atlético. Esto tiene como consecuencia la generación de compartimientos de información estancos que dificultan un abordaje integral y continuo en el tiempo.

En esta línea, se están analizando tecnologías de vanguardia basada en sensores y GPS, con el objetivo de desarrollar dispositivos y software que permitan a los preparadores físicos de equipos de fútbol, rugby, hockey, etc. administrar el entrenamiento de los deportistas para mejorar su eficiencia cognitiva y neuromuscular y que a su vez, sirvan como complemento motivador a las prácticas tradicionales [5][6].

3. RESULTADOS Y OBJETIVOS

De los proyectos presentados, r-Parking es un prototipo que está operativo.

Por su parte, WENU se encuentra en fase de prueba en dos oficinas de la Facultad de Informática de la UNLP, tomando mediciones de referencia de la estación meteorológica instalada en esta institución y proveyendo una funcionalidad parcial. Como próximo paso se prevee instalar un sistema piloto completamente funcional con cifrado en capa de enlace y capa de aplicación, a fin de disponer de un sistema fiable y seguro en distintos ambientes de esta institución.

En cuanto a las otras líneas se espera:

- Incorporar nuevas funcionalidades

para promover el uso eficiente de energía eléctrica en la Facultad de Informática de la UNLP.

- Implementar un prototipo de hardware y software que permita a los preparadores físicos configurar rutinas, monitorear la performance de los jugadores y motivar las prácticas deportivas.
- Analizar los requerimientos de seguridad en todos los proyectos de IOT abordados, y evaluar la factibilidad de incorporar las soluciones implementadas en WENU, a los demás proyectos.

En síntesis, este artículo presenta varias líneas de trabajo, todas destinadas a proveer a los ciudadanos una serie de servicios basados en tecnologías modernas y digitales, que les permite un uso más eficiente y una mayor interacción con los mismos, generando una retroalimentación al sistema global que permita mejorar su calidad de vida, teniendo en cuenta la sensibilidad de la información que los mismos manejan y los requerimientos de seguridad asociados.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de la línea de I+D+i presentada en este artículo se encuentra formado por docentes investigadores categorizados del LINTI y alumnos avanzados de las carreras de Lic. en Informática y Lic. en Sistemas de la Facultad de Informática.

Basada en las líneas de investigación de este proyecto se ha finalizado una tesina en diciembre de 2016 y se está desarrollando otra cuya propuesta fue presentada hacia fines del 2016. Asimismo se están dirigiendo tres

proyectos con estudiantes vinculados a los temas detallados en las líneas de Investigación, Desarrollo e Innovación.

5. REFERENCIAS

- [1] Department for Business, Innovation and Skills, UK Government, Smart Cities: Background paper, octubre 2013.
- [2] Fundación Telefónica, "Smart Cities: un primer paso hacia la internet de las cosas", 2011.
- [3] Sitio oficial de Urbiótica, <http://www.urbiotica.com/>
- [4] Presente y Futuro del Gas en la Argentina, Alieto Aldo Guadagni, http://www.iae.org.ar/Econometrica_FEB_2012.pdf
- [5] SMART High Intensity Cognitive Training, The Science Behind SMARTfit's Gamified Approach to Sports Performance Training, Multisensory Fitness Inc, 2016.
- [6] Our Brain Awareness Q&A Session on Movement, Exercise, and the Brain, K. Merzenish, 2010.
- [7] Modelo Estacional de Demanda de Energía Eléctrica, Guillermo Beyrne, Facundo Malvicino y Luis Alberto Trajtenberg, Secretaría de política económica y planificación del desarrollo, 2015
- [8] Practical security overview of IEEE 802.15.4, Adam Reziouk, Enzo Laurent, Jonathan, Christofer Demay, 2016.
- [9] A Comprehensive Taxonomy and Analysis of IEEE 802.15.4 Attacks, Yasmin M. Amin and Amr T. Abdel-Hamid, 2016.
- [10] Countering Three Denial-of-Sleep Attacks on ContikiMAC, Konrad-Felix Krentz, Christoph Meinel and Hendrik Graupner, 2017.

Evaluación de Software para el Desarrollo Industrial

Alicia Mon; Horacio Del Giorgio, Matías Quere!
 Departamento de Ingeniería e Investigaciones Tecnológicas. Escuela de Posgrado
 Universidad Nacional de La Matanza.
 Florencio Varela 1903 - San Justo (CP 1754)
 Tel: 4480-8952

alicialmon@gmail.com; hdelgiorgio@unlam.edu.ar; matias.quere!@gmail.com

Resumen

El desarrollo de las Tecnologías de la Información y la Comunicación (TICs) en general y del software en particular cumplen un rol central en la estimulación de la productividad industrial y en la posibilidad de potenciar el crecimiento en diversas ramas y sectores industriales. La línea que se expone en el presente artículo, se enmarca en una línea de investigación que se propone estudiar los diferentes tipos de productos software y de TICs que incorporan los sectores productivos, de modo tal de poder detectar los tipos de tecnologías instaladas en la actualidad, analizar cuánto valor agrega el uso de estas tecnologías en los diferentes sectores industriales y divisar las necesidades de desarrollo e implementación de software en las cadenas de valor.

En esta línea, se ha realizado un análisis de la conformación industrial del Partido de La Matanza y una tipificación de las TICs, agrupándolas por Área de Negocio dentro de las empresas como así también de forma transversal en función a los equipos, la infraestructura y el software que éstas utilizan. También se han analizado diferentes criterios y experiencias previas para la generación de indicadores.

Palabras clave: TICs, industria del software, desarrollo industrial, valor agregado.

Contexto

En el año 2012, el Departamento de Ingeniería e Investigaciones Tecnológicas

(DIIT) de la UNLaM creó un Centro de Desarrollo para la promoción y radicación de empresas de la industria de software, de modo tal de generar un alto grado de inserción laboral local de los estudiantes de Ingeniería Informática. El Centro se encuentra funcionando, aunque su infraestructura edilicia está en el proceso final de construcción dentro del predio de la UNLaM.

En este contexto, la línea de investigación que se está desarrollando se propone como objetivo la generación de información que permita vincular a las empresas de software radicadas en la UNLaM con la Secretaría de la Producción del Municipio de La Matanza y con las entidades locales que nuclean a las diversas empresas, como son la Cámara de Industria y Comercio de La Matanza y la Unión Industrial del Partido de La Matanza, de modo tal de facilitar la detección de necesidades de desarrollo e implementación de productos software.

Esta línea, incluye un proyecto de investigación del DIIT dentro del programa de incentivos, un Proyecto PICTO aprobado por el Ministerio de Ciencia y Tecnología, una red de Misiones de la Secretaría de Políticas Universitarias sobre “Cooperación para la difusión de la informática como soporte de la innovación productiva vinculada al desarrollo regional”, en colaboración con la Universidad Nacional del Noroeste de la Provincia de Buenos Aires y la Universidad Nacional de San Juan y el inicio de un proyecto de vinculación tecnológica entre la UNLaM y la UNLZ.

Sobre esta línea, el DIIT se propone estudiar los niveles de innovación en el

desarrollo de software y establecer un canal de vinculación con el medio productivo local, además de elaborar información pertinente para el desarrollo tecnológico, la inserción laboral de alumnos y graduados y la formación/actualización de los docentes de la Universidad.

Introducción

El uso de las TICs impacta directamente en la estimulación de la productividad industrial y ofrece un considerable potencial para el crecimiento de las industrias, tal como se indica en el Libro Blanco de la Prospectiva TIC - Proyecto 2020 [MinCyT, 2009]. Sin embargo, poner en acción dicho potencial depende crucialmente de la realización de profundos cambios en la estructura productiva, reorganización de los negocios, desarrollo de capital humano y las estrategias de promoción de las políticas públicas. Estas consideraciones son válidas no solamente para los países en desarrollo sino también para los países más avanzados.

Las TICs aportan valor a la producción y en la competitividad, al tiempo que constituyen uno de los factores intangibles que plantean mayor dificultad en su gestión. Cada actividad industrial generadora de valor, contiene algún tipo o nivel de tecnología.

Las TICs en general y la implantación de software en particular tienden a facilitar un reordenamiento de los procesos productivos, de logística y distribución, así como el control sobre las cadenas de comercialización, generando un mayor valor agregado sobre el producto final.

Sin embargo, la incorporación de tecnologías requiere de la definición de estrategias basadas en el conocimiento de un conjunto de instrumentos que permitan la gestión de los recursos tecnológicos y la incorporación de nuevos desarrollos que agreguen valor y formen recursos, mejorando los niveles de empleo y valorización del capital.

Respecto del desarrollo de la industria del software en Argentina, si bien ha generado un crecimiento sustancial en sus volúmenes de producción, los desarrollos tecnológicos han sido focalizados hacia los servicios financieros, destinando más de un 50% de la producción de software, en tanto que la industria solo demanda un 9% del total del software desarrollado [CESSI].

La industria local pareciera no tener definidas estrategias de actualización tecnológica en la cual basar la mejora de la competitividad, dado que no resulta ser un sector demandante de productos software ni de la incorporación de TICs en sus procesos productivos.

La incorporación de nuevas tecnologías en los sectores industriales requiere de un profundo conocimiento sobre la capacidad existente, es decir que, sin información relativa a las TICs instaladas y utilizadas en los diferentes procesos, no es posible definir necesidades de incorporación tecnológica para generar una reconversión en las cadenas de valor.

En lo que respecta a la Industria, el uso de las TICs puede referirse a tareas específicas implicadas en la creación de un producto (*tecnologías de producto*), a tareas involucradas en el desarrollo de un proceso productivo (*tecnologías de proceso*), a las prácticas implicadas para la operación de las distintas funciones de una unidad productiva (*tecnologías de gestión*), o bien a las prácticas realizadas para garantizar la correcta apropiación de las competencias por parte de consumidores y usuarios (*tecnologías de uso*) [Zubieta, 2013].

Si bien existe diversa bibliografía sobre el desarrollo productivo y los desarrollos tecnológicos, no se ha encontrado aún una forma específica de medir los diferentes niveles de TICs y el impacto que generan en los niveles de productividad y en las estrategias de innovación requeridas por la industria.

Líneas de Investigación, Desarrollo e Innovación ²

La posibilidad de conocer las diferentes tecnologías, los tipos de productos software instalados así como la agregación de valor que aportan en la productividad, resulta una información clave para la toma de decisiones estratégicas tanto en la industria del software como en los diferentes sectores industriales.

En este sentido, los principales ejes de investigación del presente proyecto son:

- Analizar los tipos de productos software y de TICs que se implementan en la industria de modo de diferenciarlos según el valor agregado que aportan en los niveles de productividad e innovación.
- Estudiar la conformación industrial del Partido de La Matanza diferenciado por rama, sector y tipo de empresa.
- Vincular los diferentes tipos de productos software y de TICs que cada rama o sector de la industria local tiene implementado.
- Elaborar indicadores de tipos de software y de TICs en la Industria.
- Detectar las necesidades de desarrollo de software y de implantación de TICs que pueden ser incorporadas en las diferentes áreas productivas.

El estudio de los “Tipos” de productos software y de las TICs permite ordenar, sistematizar y jerarquizar la combinación de diferentes software instalados, la infraestructura disponible y las comunicaciones según su desarrollo tecnológico, permitiendo conocer, según la rama de actividad, qué áreas de proceso agregan mayor valor en cada sector industrial y qué tecnología específica requiere ser incorporada para mejorar el desarrollo productivo.

Resultados y Objetivos

Como resultado en esta línea de investigación, se han elaborado un conjunto

de Tipologías ordenadas en base a diferentes taxonomías que permiten analizar las áreas al interior de las industrias y las tecnologías insertas en cada área.

La estructura básica de la tipificación permite inicialmente distinguir dos Taxonomías. La primera las diferencia en productos software, equipos o hardware y comunicaciones o infraestructura, tal como se expone en las siguientes figuras:

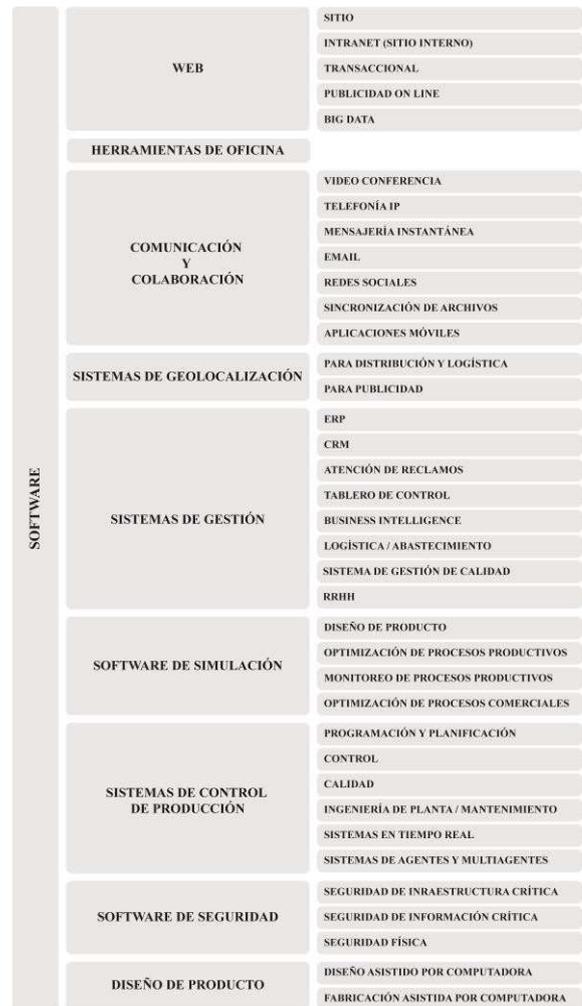


Figura 1. Elaboración propia



Figura 2. Elaboración propia

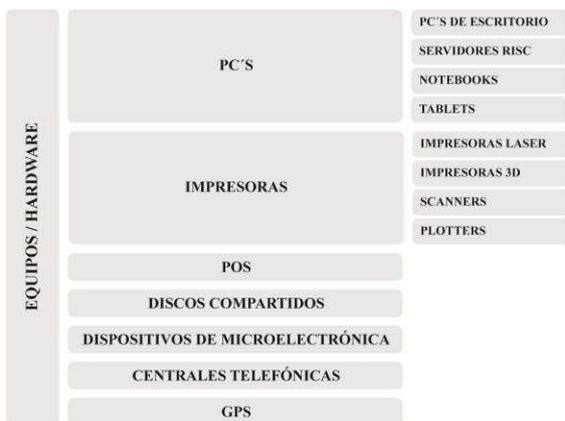


Figura 3. Elaboración propia

El análisis que se está elaborando en base a dicha taxonomía de TICs diferencia a cada una de ellas por los tipos de productos que tienen implementados, evaluando su aporte en base al mayor nivel de desarrollo tecnológico que contenga y al valor que agregue a la productividad.

La segunda taxonomía analiza las áreas de negocios al interior de una empresa, independientemente de la rama a la que pertenezca y del tamaño de la misma. Esta taxonomía permite detectar los sistemas involucrados en una organización, que incluyen diversos tipos de TICs, tal como se expone en el siguiente cuadro basado en un modelo adaptado de Cadena de Valor de Porter [Porter, 1985].

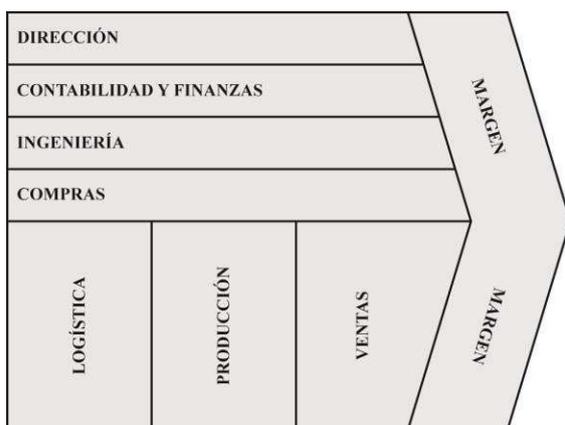


Figura 4. Elaboración propia

Cada uno de los tipos de software y TICs aporta valor según el área de negocio en la que esté implementado. Esta lista de

opciones es sólo indicativa, no pretende ser un listado exhaustivo y se encuentra en etapa de discusión.

Posteriormente se han analizado aspectos de mejores prácticas para la generación de Indicadores en general, y luego apuntándolos hacia el desarrollo tecnológico de las Empresas. A título de ejemplo, se han analizado los términos de Relación con los Clientes y Capacidad de Innovación. Dos términos no sólo muy importantes, sino también íntimamente relacionados entre sí, ya que la Innovación Empresarial (más precisamente, actividades de I+D+i) es un elemento crucial para la mejora de la Relación con los Clientes.

Finalmente, se ha trabajado sobre algunas formas pre-existentes que algunos autores consideraron para clasificar a las empresas según el grado de innovación y apropiación de las TICs, observándose que en algunos casos se habla de grados generales de desarrollo de TICs (aplicables a toda la empresa) y en otros casos se hace hincapié en el modo en que las TICs dan soporte a determinadas secciones o estrategias de la misma (por ejemplo, Recursos Humanos, Procesos Internos, la visión hacia el afuera, entre otros aspectos).

Actualmente se está trabajando en la creación de los instrumentos metodológicos de relevamiento y análisis para poder validar las tipologías con la medición de las mismas en la industria del Partido de La Matanza. Se espera que al finalizar el proyecto se pueda evaluar el nivel de desarrollo tecnológico de los sectores industriales en La Matanza y definir las necesidades de investigación, desarrollo e innovación de TICs que tiene el distrito.

Formación de Recursos Humanos

El grupo de investigación GIS se ha conformado para este proyecto como un grupo interdisciplinario e interuniversitario, integrado por Ingenieros Informáticos, Industriales y Electrónicos, todos docentes-investigadores.

Un integrante del grupo GIS se encuentra desarrollando una tesis de la Maestría en Dirección Estratégica y Tecnológica del Instituto Tecnológico Buenos Aires, en tanto que otro de los investigadores se encuentra desarrollando su tesis del Doctorado en Ciencias Económicas de la UNLaM sobre el tema abordado en este Proyecto.

Se prevé la incorporación de 2 alumnos de grado para que realicen su proyecto final de carrera en el marco del proyecto.

Bibliografía

- Ca' Zorzi, Antonio (2011) - Las TIC en el desarrollo de la PyME: Algunas experiencias de América Latina. Disponible en http://www.oitcenterfor.org/sites/default/files/file_publicacion/tic_pyme.pdf
- Llano Naranjo, Nicolás (2009) - Política para la promoción en el acceso y uso de TIC en micro, pequeñas y medianas empresas colombianas. Disponible https://spi.dnp.gov.co/App_Themes/SeguimientoProyectos/ResumenEjecutivo/090518%20-%20Politica%20Uso%20de%20TIC%20en%20mipymes%20-%20MinComunicaciones.pdf
- Ministerio de Ciencia, Tecnología e Innovación Productiva (2009) - Libro Blanco de la Prospectiva TIC - Proyecto 2020. Disponible en <http://cdi.mecon.gov.ar/bases/docelec/va1028.pdf>
- Novick, Marta & Ritondo, Sebastián (2013) - El desafío de las TIC en Argentina. Crear capacidades para la generación de empleo. CEPAL, Ministerio de Trabajo, Empleo y Seguridad Social. Disponible en http://repositorio.cepal.org/bitstream/handle/11362/3009/1/S2013168_es.pdf
- Porter, Michael (1985) - Competitive Advantage: Creating and Sustaining Superior Performance, Editorial The Free Press, Nueva York, Estados Unidos.
- Saavedra García, María L. & Tapia Sánchez, Blanca (2013) - El uso de las

tecnologías de información y comunicación TIC en las micro, pequeñas y medianas empresas (MIPyME) industriales mexicanas. Disponible en <http://www.redalyc.org/pdf/823/82326270007.pdf>

- Yoguel, Gabriel & Novick, Marta & Milesi, Darío & Roitter, Sonia & Borello, José (2004)

- Información y conocimiento: la difusión de las tecnologías de información y comunicación en la industria manufacturera argentina. Disponible en <http://www.cepal.org/publicaciones/xml/2/19412/lcg2220e-yoguel.pdf>

- Zubieta, Roberto & Villadeamigo, José & Cianci, Luciano (2013): Los Índices de Nivel Tecnológico – Su papel en una Estrategia de Desarrollo. http://www.uba.ar/archivos_secyt/image/SI-MPOSIO%20VIII%20Documento.pdf

Generación de Metamodelos Ontológicos que Representen al Artefacto Sistema de Información (SI)

María Claudia Gómez, Sebastián José García, María Romagnano, Alejandra Orellana, María del Carmen Becerra, Alicia Aballay, Adriana Sarmiento.

Proyecto “Representación Genérica de Modelos Conceptuales en el campo de los Sistemas de Información”

Departamento de Informática / Facultad de Ciencias Exactas, Físicas y Naturales / Universidad Nacional de San Juan

cgomez@unsj-cuim.edu.ar; cacugomez@yahoo.com.ar; sgarcia@unsj.iinfo.edu.ar; maritaroma@iinfo.unsj.edu.ar; ale_ore@yahoo.com.ar; mcbecerra2008@gmail.com; profe.alicia@yahoo.com.ar; adriva2005@yahoo.com.ar;

Resumen

Hacen ya varios años que los Sistemas de Información (SI) impactan profundamente en la sociedad y vienen evolucionando de manera vertiginosa. Pero se puede detectar que el pensamiento sobre su formalización no lo hace en la misma medida que sus aplicaciones. Esto ha generado planteos académicos que no llegan a lograr los objetivos pretendidos cuando se crean carreras universitarias que los contemplen, como así también los grupos de investigación que buscan su avance. Existe un debate sobre la crisis en la disciplina SI, que se atribuye principalmente a que es indefinida su identidad. Por tanto, es esencial contar con un marco teórico que explique los fundamentos de los SI. Las ontologías cumplen un rol importante en la resolución de la interoperabilidad semántica entre los SI y su uso. Por lo tanto, se propone emplearlas para representar modelos de sistemas de información de diferentes dominios y así lograr analogías que permitan encontrar representaciones genéricas relacionando los mismos, con el fin de estimular la utilización de modelos conceptuales que mejoren los mecanismos de abstracción de la realidad.

Palabras clave:

Sistemas de Información. Modelos. Metamodelos. Ontologías.

Contexto

El proyecto de Investigación está desarrollado por Profesores del Departamento e Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNSJ. Está financiado por la Universidad, en la convocatoria de Proyectos que se inició el 1 de enero de 2016 y culmina el 31 de diciembre de 2017. El grupo de investigación ha evolucionado su pensamiento y aplicaciones comenzando con el proyecto “Apoyo a Organizaciones Locales en la Caracterización de sus Sistemas de Información”. Luego, en función de las conclusiones obtenidas, se generó el proyecto “Análisis e Integración de los Sistemas de Información en la Organización”, el que a través de su trabajo de campo verificó la necesidad de integrar la información en las organizaciones y la pertinencia de la utilización de sistemas centralizados. Posteriormente se analizó la hipótesis de conflicto que generaban los sistemas en la

organización mediante el proyecto “Aplicación de las Técnicas de Resolución de Conflictos a la Gestión de los Sistemas de Información”, así se identificó la importancia del recurso humano en la gestión de cualquier SI. Mediante el proyecto “Impacto de Confundir los Sistemas de Información y las Tecnologías de Información en la Organización” se concluyó que el gran inconveniente, en la actualidad, por el que pasan los SI se debe a la carencia de fundamentos conceptuales, siendo éste una de las principales causantes de esta confusión. La falta de claridad en la conceptualización de los SI generó en el grupo de investigación el interrogante de la importancia que tienen los modelos conceptuales para los sistemas realizando el proyecto “Identificación de Modelos Conceptuales en el Campo de los Sistemas de Información”. Las conclusiones obtenidas del proyecto llevó al análisis de situaciones comunes que conduzcan a generalizar el proceso de modelización.

1. Introducción

El surgimiento y evolución de plataformas y tecnologías de desarrollo de software, la evolución del hardware, las nuevas formas de interacción humano-computadora, los avances en el cómputo móvil, el incremento de los usuarios de los sistemas y de los datos que deben gestionar y el surgimiento de la Web de datos, son algunas de las razones por las cuales los sistemas de software son cada vez más complejos.

Un medio para manejar esta complejidad y continuar proporcionando sistemas que soporten las nuevas y crecientes necesidades del mercado, es el llevar a cabo el desarrollo de software utilizando

métodos apropiados de abstracción (Anda).

En este contexto, un paradigma de desarrollo de software que se ha vuelto popular en la academia y en la industria, en los últimos años, es el Desarrollo Dirigido por Modelos (Model Driven Development – MDD). MDD basa el desarrollo de software en la construcción de modelos conceptuales que describen el sistema a desarrollar; es decir su estructura, comportamiento, estructura de datos, aspectos de presentación, etc.). Se parte de un modelo que conceptualiza el dominio del problema. De manera sistemática este modelo se va transformando en otros modelos, de niveles de abstracción más bajos, hasta generar implementaciones concretas.

Asimismo, se ha planteado que los paradigmas que han sustentado el desarrollo de los SI se han basado en diferentes núcleos de interés, tales como generación y mantenimiento automático de registros, flujos de trabajo, reingeniería y gestión de datos/información. Actualmente, estos paradigmas no son suficientes para abordar los fenómenos y las situaciones problemáticas que surgen frente a los avances de las tecnologías de la información y la comunicación (Álvarez, Palliotto, y Barchini).

Por su parte, Nicola Guarino (2006) plantea que una ontología es un artefacto de la ingeniería, constituido por un vocabulario específico usado para describir una cierta realidad, más un conjunto de asunciones explícitas con respecto al significado intencional de las palabras del vocabulario.

Ya en el año 1996 Studer, Benhamins y Fensel y en 1998 Uschold y Gruninger plantearon que resulta de especial interés el uso de las ontologías como una especificación formal y explícita de una conceptualización compartida en la Gestión del Conocimiento, debido a que

éstas pueden funcionar como un marco para la unificación de diferentes puntos de vista del conocimiento y servir como base para:

- a) La comunicación entre personas con diferentes necesidades, que compartan un área de conocimiento.
- b) Facilitar la interoperabilidad entre sistemas.
- c) La reutilización de software, la realización de chequeos de consistencia, la adquisición de conocimiento y para la especificación de requerimientos.

Con las ontologías se intenta expresar un esquema conceptual exhaustivo y riguroso de un dominio particular, con la finalidad de facilitar la comunicación y reutilización de conocimiento e información entre los miembros de las organizaciones, computadoras y humanos.

Una ontología define un vocabulario común que incluye, además, la interpretación de los conceptos básicos del dominio y sus relaciones. Las ontologías, como productos de software, deben ser desarrolladas siguiendo los estándares establecidos para esto, por supuesto, adaptados a las características de las ontologías.

Por lo tanto, nuestra propuesta consiste en usar ontologías para representar modelos de sistemas de información de diferentes dominios y así lograr analogías que permitan encontrar representaciones genéricas mediante relaciones entre modelos, con el fin de estimular la utilización de modelos conceptuales que mejoren los mecanismos de abstracción de la realidad.

2. Líneas de Investigación, Desarrollo e Innovación

Teniendo en cuenta las disciplinas bajo las cuales puede abordarse la problemática planteada como lo son las Ciencias Naturales, Ingeniería de Software y Ciencias del Comportamiento (Figura 1), no siendo área de competencia las Ciencias Naturales se está trabajando en tres líneas de investigación, en función de la formación que poseen los integrantes del grupo de trabajo:

1. Fundamentos conceptuales de los SI.
2. Representaciones que generen modelos genéricos a través de la Ingeniería de Software y Ontologías.
3. Recursos humanos y seguridad informática.

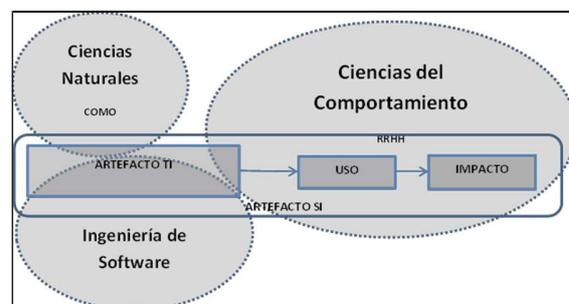


Figura 1. Visión interdisciplinaria de SI

Estas líneas de investigación permiten abrir un amplio espectro que ayuda a la formalización de los Sistemas de Información.

A través de la lectura de trabajos relacionados y de la observación se pudieron determinar las siguientes dimensiones bajo estudio:

- *Dimensión Académica.* La crisis detectada como problema ha disminuido significativamente el interés de su estudio, reduciendo la cantidad de alumnos. No obstante existe, a nivel internacional, significativos esfuerzos en dar un marco teórico que identifique a la disciplina y de formular un modelo global de

currícula en SI. Además, en nuestro país, y como lo establece la Resolución N° 786/2009 del Ministerio de Educación, difiere del modelo global. Por lo tanto las Universidades Nacionales están en proceso de adaptación de sus currículas. Indubitablemente, se avizoran cambios radicales en la disciplina.

- Otra dimensión en la que se ha visto reflejada esta falta de claridad está dada por los usuarios de los Sistemas de Información en todos los ámbitos, ya que su utilización se ha generalizado tanto que hoy la información se ha vuelto el recurso más importante, a nivel individual, organizacional, estatal, etc.
- Una tercera dimensión en la que se manifiesta la dificultad es en el ámbito de los desarrolladores de Sistemas de Información, quienes tienen una gran responsabilidad en el entendimiento del dominio del problema a solucionar.

Teniendo en cuenta las situaciones problemáticas planteadas en diferentes dimensiones surge la necesidad de buscar algún mecanismo que permita generar conceptualizaciones ordenadas bajo un criterio formal y con un respaldo teórico asociado.

3. Resultados obtenidos/esperados

El objetivo general propuesto consiste en generar representaciones genéricas que relacionen los modelos conceptuales encontrados en el campo de los SI que permitan conceptualizar generalizaciones que tiendan a definir las características esenciales de la disciplina.

Los resultados del proyecto que se pretenden obtener son:

- Estudio del estado del arte mediante la búsqueda de antecedentes en la materia. En este aspecto se ha realizado un análisis exhaustivo de publicaciones y bibliografías asociadas al tema bajo estudio.
- Caracterización de los sistemas de información planteados en las cátedras asociadas (Sistemas de Información I y II, Ingeniería de Software II y III).
- Utilización de herramientas de la Ontología, de la Ingeniería de Software, de la Inteligencia Artificial, y de la Teoría General de Sistemas para detectar relaciones entre modelos.
- Descripción de las representaciones conceptuales identificadas, que permitan realizar generalizaciones utilizando metamodelos.

El modelo del que se pretende cotejar como genérico se ha denominado **Modelo de las 5 vocales** (Figura 2). Éste toma como referencia la teoría del Éxito de William DeLone y Ephraim McLean. En este se enfatiza el uso y el impacto de los SI, como así también pone principal esmero en la satisfacción del usuario y sus beneficios netos, considerando que la calidad del sistema está en relación directa con estas variables (Delone, McLean, 2003).

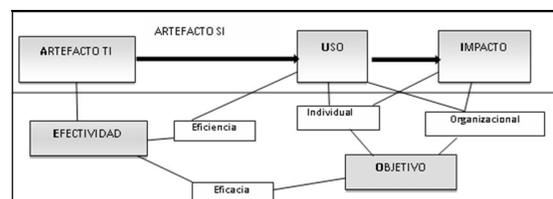


Figura 2. Modelo de las 5 vocales

La incorporación del Objetivo individual y organizacional y conceptos como el de

efectividad que se basa en la eficacia (cumplimiento de objetivos) y la eficiencia (relación insumo/producto) pretende realizar una representación de los SI desde la perspectiva de su análisis cuantitativo y cualitativo.

De esta forma todos los artefactos de Tecnologías de Información (TI), en la mayoría de los casos representados por proyectos de software, normalmente evaluados mediante métricas específicas, requieren de una perspectiva desde las diferentes ciencias. Esto conduce al equipo de trabajo a pensar que los SI necesitan de una visión y una evaluación interdisciplinaria. En este sentido, aparece una nueva perspectiva definida como *Artefacto SI* que tiene como componentes: el Artefacto TI, su Uso y su Impacto.

Las presentaciones en congresos realizadas hasta la fecha son:

- “Un Modelo de Artefacto” SI. CONAISI 2016.
- “Teorizando en SI”. SABTIC 2016.
- “Modelo Genérico para Representar Sistemas de Información”. SABTIC 2016.
- “Herramientas para generar contratos electrónicos en entornos de comercio electrónico basado en ontologías” JAIIO 2016.
- “Modelo de Integración de Estándares para la Gestión de la Identidad y Privacidad” CACIC 2016.
- “La Modelización Conceptual como Herramienta para Formalizar Sistemas de Información” JATIC 2016.

4. Formación de Recursos Humanos

El grupo de investigación está conformado por profesores de la Carrera Licenciatura en Sistemas de Información,

entre los cuales hay 4 magister, dos especialistas. En la actualidad se están formando un doctorando en Ingeniería y un maestrando en Informática.

Se está dirigiendo una tesis de maestría. Se han dirigido tres tesinas de grado.

5. Referencias

Álvarez, M., Palliotto, D. & Barchini, G. (2006). “Qué son los Sistemas de Información basados en Ontologías?”. Encuentro Informático Riojano. Argentina, 2006.

Anda, et al. (2006). "Experiences from Introducing UML-based Development in a Large Safety-Critical Project". *Empirical Software Engineering*. Vol. 11, Pp. 555-581.

DeLone, W, McLean, E.(2003). “Model of Information Systems Success: a ten years update”. *Journal of Management Information Systems / Spring 2003*, Vol. 19, No. 4, pp. 9–30. © 2003 M.E. Sharpe, Inc. 0742–1222 / 2003.

A Ten-Year Update Guarino, N. (2006). “Formal Ontology and Information Systems”. *Proceedings of FOIS '98*.

Studer R., Benjamins V. R. & Fensel D. (1998). “Knowledge Engineering: Principles and Methods”. *Data and Knowledge Engineering*. Elsevier. DOI: 10.1016/s0169-023x(97)00056-6.

Volumen 25. Issue (1-2). Pp. 161-197. Uschold M. & Gruninger M. “Ontologies: Principles, Methods and Applications”. *Knowledge Engineering Review*. 1996. 11(02): 93-136.

Herramienta de Modelado y Análisis Estocástico de Sistemas Biológicos

Vilallonga, Gabriel^{1,2}; Riesco, Daniel¹, Matías Truglio¹, Antônio-Carlos G. Almeida³, Antônio M. Rodrigues³

¹Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis. Ejercito de los Andes 950. San Luis
Tel: +54 (0266) 4520300. Int. 2127

²Departamento Sistemas de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca Maximio Victoria N° 55 - C.P: 4700 - San Fernando del Valle de Catamarca Teléfono:
03834- 435112 – int 168

³ Department of Biosystems Engineering, Federal University of São João del Rei, MG, Brazil.

gvilallo@tecno.unca.edu.ar/unsl.edu.ar, driesco@unsl.edu.ar, mtrugli@unsl.edu.ar, acga@ufsj.edu.br, amr@ufsj.edu.br.

Resumen

Los sistemas híbridos cuentan con la atención de gran parte de la comunidad científica por lo atractivo del estudio de sus dinámicas, las continuas y las discretas, y la comprensión de sus interacciones. Entre las más comunes se encuentran aquellas que cambian de estado entre diferentes procesos continuos. También se pueden encontrar formas de interacción que incluyen transiciones discretas superpuestas a evoluciones continuas, u otras como resultados de una decisión, o por la ocurrencia de determinados eventos.

El estudio de la distribución probabilísticas de la dinámica discreta y la continua se lleva a cabo por medio del análisis del sistema híbrido estocástico (SHE) que lo modela.

Debido a que el modelo numérico, de un SH, es afectado por el problema de la explosión de estados y de ser sumamente exigente en lo que a recursos se refiere, aparece como una opción aceptable la formulación del mismo sistema por medio de un SHP.

Estos tipos de sistemas pueden ser analizados por medio de herramientas informáticas con solida base matemática como son los Model Checkers Probabilísticos.

Los sistemas biológicos encuadran perfectamente en la clasificación de SHE.

En el ámbito de estudio de SB no se cuenta con herramientas que permitan una traducción

directa de un SHE, como por ejemplo un sistema de reacciones, a un modelo estocástico factible de ser analizado por herramientas informáticas disponibles en la actualidad como son los model checkers antes mencionados.

En nuestra línea de investigación proponemos el estudio de factibilidad y de propuesta de desarrollo de una herramienta de análisis de SB basado en su formulación estocástica.

Tenemos como hipótesis de trabajo que el desarrollo de este prototipo de herramienta que permite la obtención de un modelo estocástico a partir de su formulación por medio de reacciones que permitirá analizar el sistema e incrementar la productividad en el estudio de SBs habilitando a su verificación y validación con herramientas novedosas en el área biológica. En una primera etapa se centrará en la obtención de modelos estocásticos de un sistema de reacciones para luego habilitar el análisis basado en probabilidades y en simulaciones probabilísticas basadas en su semántica estocástica.

Palabras claves: Sistemas Híbridos (SH), Sistemas Híbridos Estocásticos (SHE), Model Checking (MC), Model Checking Probabilístico (MCP), Sistemas Biológicos (SB), Ingeniería de Software (IS).

Contexto

El proyecto de investigación “*Fortalecimiento de la Calidad y la Productividad en Ingeniería de Software*”, posee una línea de investigación dedicada a la verificación y validación de sistemas y específicamente al estudio y análisis de SB. Dentro de esta línea se prevé desarrollar una herramienta que permita acercar el área de análisis de SHE, como son los SB, con las herramientas informáticas con base matemática como los son los MCP.

El mencionado proyecto es evaluado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca y se desarrolla y ejecuta en ámbitos del Departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCa y en el Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis donde los integrantes se desempeñan como docentes de la carrera de Ingeniería en Informática (UNCa), e Ingeniería en Informática y Licenciatura en Ciencias de la Computación (UNSL) respectivamente. También se cuenta con el convenio establecidos entre la UNSL y la Universidad Federal de Minas Gerais (UFMG), en el marco del programa Centros Asociados para el Fortalecimiento de Posgrados Brasil/Argentina (CAFP-BA) tanto para maestrías como para doctorados. Por medio de esta se ha logrado generar un nexo con investigadores responsables del LABORATÓRIO DE NEUROCIÊNCIA EXPERIMENTAL E COMPUTACIONAL DR. ARISTIDES AZEVEDO PACHECO LEÃO (LANEC) de la Universidad Federal de Sao João del-Rei, (UFSJ) en lo que respecta a los sistemas biológicos, los cuales se propone estudiar.

Esta línea de investigación, inserta en el proyecto, establece como prioridad la formación científico-técnico de los integrantes con la premisa de proceder a la transferencia de resultados a los medios informáticos como así también al ámbito biológico.

Introducción

Los SB están enmarcados dentro de la categoría de SH, y en muchos casos como

SHE ya que los procesos que los caracterizan cuentan con magnitudes aleatorias que varían con el tiempo.

Estos tipos de sistemas presentan un gran desafío en el área computacional no solo por su estudio, sino que también por las herramientas que se poseen para su modelado y análisis. La coexistencia de la dinámica continua y discreta plantean una mejora a los modelos matemáticos hasta ahora vigentes, como así también de las herramientas de análisis disponibles en la actualidad para paliar las carencias de los acercamientos numérico de estos.

Los SB cuentan con una complejidad que desafía a las herramientas informáticas disponibles.

Muchas veces la manera de reducir la complejidad en el estudio de estos tipos de sistemas ha sido empleando modelos probabilísticos. La aleatoriedad se ha convertido en un método estándar en el modelado y análisis de estos tipos de modelos complejos. El hecho es que los SHE [1] en la actualidad son un tópico de gran interés. El estudio de este tipo de sistemas permite un acercamiento a familias de modelos que resultan de varios tipos de aleatoriedad de sistemas híbridos. En la forma más general, un SHE considera la distribución de probabilidad para ambos, transiciones discretas y continuas, y, además, estas distribuciones pueden depender de otras.

Los modelos estocásticos y las lógicas temporales, como las lógicas estocásticas continuas y las lógicas arborescentes computacionales probabilísticas son usados para el modelado y análisis de sistemas con el fin de conocer acerca de su rendimiento y confiabilidad. En nuestro trabajo estamos interesados en el estudio del comportamiento estocástico de los modelos a ser analizados y en la disponibilidad de herramientas informáticas con que se cuenta para trabajar sobre estos.

En el área informática se cuenta con Model Checkers con distintos formalismos que subyacen y que permiten el análisis de distintas características. En la actualidad se destacan, entre otros, model checkers con

capacidades de análisis estadístico como son UPPAAL [2–4] y PRISM [5]. Ambos han sido desarrollados en ámbitos académicos con desempeños destacados. En estos se ha priorizado su sustento matemático y la eficiencia en los algoritmos de análisis.

La línea de investigación presentada propone el estudio de SB por medio de herramientas informáticas que hasta el momento no ha sido utilizada en el área como son los Model Checkers Probabilísticos.

En el estudio de los SB, la naturaleza estocástica de los procesos celulares ha motivado el uso de SHE para el modelado de estos tipos de fenómenos, ya que combinan las generalidades de SH con procesos probabilísticos [6].

Los SHE han mostrado proveer un marco de trabajo adecuado para el modelado de la evolución temporal de poblaciones de especies químicas involucradas en conjuntos de reacciones químicas. Las reacciones químicas que se producen dentro de la células son frecuentemente modeladas usando formulaciones probabilísticas, las cuales toman en cuenta la aleatoriedad inherente de movimiento molecular térmico [7]. En estos sistemas las reacciones son tratadas como eventos probabilísticos que cambian la población de las especies individuales basadas en la estequiometría de las reacciones. Desde el punto de vista de la IS, la complejidad de estos tipos de sistemas desafía la capacidad de las herramientas actuales, donde, de las experiencias realizadas por medio de experimentos sobre las herramientas, pueden surgir sugerencias de mejoras y uso de las experiencias de usuarios resultantes. Este último ítem se basa en que muchas veces los sistemas estudiados pertenecen a áreas muy diferentes a la informática, como por ejemplo la biológica, y el lenguaje de comunicación debe ser revisado para una mejor comprensión entre los investigadores de las distintas disciplinas.

Una de las mayores dificultades que se detectan en el estudio de los SB está dado en la dificultad manifiesta con que se encuentran los investigadores del área biológica al momento de modelar con las herramientas informáticas

disponible.

Atentos a esta brecha es que se propone un acercamiento entre ambas disciplinas. Este está dado por el desarrollo de una herramienta que permita a biólogos poder expresar sus modelos en su propio lenguaje, y por medio de la herramienta traducir el modelo a uno estocástico factible de analizar por medio de un Model Checker Estocástico.

Los resultados en esta línea de investigación pueden mejorar la producción en el área biológica como así también el proceso de formación académica, de docentes y alumnos en el desarrollo de aptitudes de carreras informáticas elevando de esta manera la calidad de la educación que favorezca la formación de ingenieros en la Universidad.

Línea de Investigación, Desarrollo, e Innovación

Inserto en el marco del proyecto se encuentra la línea de investigación referida a la validación y verificación de sistemas. En este caso puntual se hace fuerte hincapié en el área de SHE, donde la mayoría de los casos extra informáticos, han sido llevados a cabo por investigadores de la misma área, sin intervención de investigadores de nuestra disciplina. Esto es sumamente notorio en los primeros contactos interdisciplinarios.

Esta línea de investigación permite abrir una perspectiva multidisciplinaria que tiene como efecto sinérgico el trabajo con sistemas reales, de otras áreas, que permiten la revisión de las herramientas informáticas utilizadas en el área de validación y verificación de sistemas complejos, como así también el desarrollo de nuevas herramientas que permitan aumentar la productividad del área biológica.

Estas actividades han exigido realización de cursos acerca de la temática, como así también el estudio de material bibliográfico y de publicaciones científicas.

Es de hacer notar que estos grupos de trabajos están coordinados por docentes investigadores de las distintas universidades.

Resultados y Objetivos

El objetivo principal de esta línea de

investigación es el de contribuir de manera efectiva y medible en la concreción de aportes concretos al área de verificación y validación insertas en IS aplicada, lo que implica revisión, o nuevas propuestas, de técnicas, metodologías, y el desarrollo de herramientas que asistan al desarrollo de software específico para el trabajo con SHE de una disciplina determinada, en nuestro caso en biología.

El efecto deseado, también, es incidir significativamente en las actividades académicas de grado y posgrado, la formación de recursos humanos, y la transferencia al medio. Estos objetivos están siendo alcanzados gracias al trabajo conjunto entre el equipo de la UNSL-UNCa y UFSJ. El trabajo interdisciplinar está generando aportes significativos gracias a la interacción con científicos del LANECS, referido al área biológica, específicamente en el trabajo con Bombas de Sodio-Potasio afectadas con Palitoxina [8,9].

Formación de Recursos Humanos

Esta línea de investigación está en una etapa intermedia, donde integrantes del proyecto desarrollan sus tesis de doctorado y maestría en el marco de la carrera de doctorado en Ingeniería de Software, como así también de la maestría en Ingeniería del Software en temas específicamente relacionados a la línea de verificación y validación. Además los participantes pertenecen a distintas cátedras de las carreras de ingeniería y licenciatura de las universidades que participan.

Se ha procedido a la incorporación de alumnos de los últimos años con la finalidad de incluirlos en actividades de investigación y desarrollo en las áreas específicas del proyecto. El proyecto prevé un programa de capacitación y formación de recursos humanos, que contempla las siguientes actividades de dirección de tesis de grado de la carrera de Ingeniería en Informática y en actividades de actualización y posgrado en el área de estudio. Se propone la participación de los integrantes en eventos nacionales e internacionales de la especialidad, como congresos, simposios,

seminarios y cursos.

Bibliografía

- [1] G. Pola, M. L. Bujorianu, J. Lygeros, and M. D. D. Benedetto. Stochastic hybrid models: an overview. Proc. of the IFAC Conference on Analysis and Design of Hybrid Systems, pages 45-50, 2003.
- [2] David, A., Larsen, K.G., Mikucionis, M., Poulsen, D.B.. Statistical Model Checking for Stochastic Hybrid Systems. In Proc. HSB 2012. Computational Engineering, Finance, and Science (cs.CE); Software Engineering (cs.SE). arXiv:1208.3856. 2012.
- [3] Bulychev, P., David, A., Larsen, K.G., Mikucionis, M., Legay, A., Wang, Z.. UPPAAL-SMC: Statistical Model Checking for Priced Timed Automata. Logic in Computer Science, Formal Languages and Automata Theory. In Proceedings QAPL 2012. DOI:10.4204/EPTCS.85.1
- [4] Uppaal in a nutshell. Kim G. Larsen, Paul Pettersson, Wang Yi. International Journal on Software Tools for Technology Transfer. December 1997, Volume 1, Issue 1-2, pp 134-152.
- [5] Kwiatkowska, M., Norman, G., Parker, D.: PRISM 4.0: Verification of probabilistic real-time systems. In: Proc. CAV. Springer (2011).
- [6] Singh, Abhyudai, and Joo P. Hespanha. "Stochastic hybrid systems for studying biochemical processes." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368.1930 (2010): 4995-5011.
- [7] Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. of Comp. Physics, 22, 403-434, 1976.
- [8] Rodrigues, A. M., Almeida, A. C. G., Infantes, A. F., Teixeira, H. Z., & Duarte, M. A. (2008). Model and simulation of Na⁺/K⁺ pump phosphorylation in the presence of palytoxin. Computational biology and chemistry, 32(1), 5-16.
- [9] Artigas, P., & Gadsby, D. C. (2003). Na⁺/K⁺-pump ligands modulate gating of palytoxin-induced ion channels. Proceedings of the National Academy of Sciences, 100(2), 501-505.

Informática y Tecnologías Emergentes

Russo C.¹, Sarobe M., Ado M., Ahmad H., Alonso N., Alvarez E., Bendati N., Charne J., Cicerchia B.², De Vito M., Di Cicco A., Esnaola L., Fernandez D., Guasch M., Jaszczyszyn A., Jatip N., Lencina P., Llanos E., Luengo P., Moran M., Moretti N., Muscia A., Occhipinti S., Osella Massa G., Picco T., Piergallini R., Pompei S., Ramon H.³, Rodriguez M., Serrano E., Schiavoni A., Serafino S., Smail A., Terzano J., Tessore J.⁴, Yamel L.

Instituto de Investigación y Transferencia de Tecnología (ITT) - CIC⁵

Escuela de Tecnología (ET)

Universidad Nacional del Noroeste de la Provincia de Buenos Aires

Sarmiento y Newbery, Junín (B) – TE: (0236) 4477050

{claudia.russo, monica.sarobe, tamara.ahmad, nicolas.alonso, eduardo.alvarez, natalia.bendati, javier.charne, lucas.cicerchia, cecilia.devito, carlos.dicicco, leonardo.esnaola, adrian.jaszczyszyn, nicolas.jatip, paula.lencina, emanuel.llanos, pablo.luengo, alex.muscia, pocchipinti, german.osella, trinidadpicco, sabrina.pompei, hugo.ramon, antonela.schiavoni, sandra.serafino, ana.smail, julieta.terzano, juanpablo.tessore}@itt.unnoba.edu.ar

Resumen

Las tecnologías emergentes son innovaciones en desarrollo que como su nombre lo dice en un futuro cambiarán la forma de vivir del ser humano brindándole mayor facilidad a la hora de realizar sus actividades. Estas tecnologías innegablemente están modelando nuestra sociedad, nuestras costumbres, la forma de relacionarnos y comunicarnos, la forma en la que las empresas producen, la forma en la que se educa.

Hemos pasado de un modelo de sociedad industrial a un modelo de producción del conocimiento donde las demandas de las tanto de las empresas como de la sociedad, han cambiado. Incluso el modo de interactuar con esta tecnología está cambiando. Hoy en día esa interacción es mucho más dinámica dando al usuario un rol activo, convirtiendo al mismo usuario en parte de la tecnología.

El proyecto descrito en este documento tiene como propósito identificar, contextualizar,

evaluar, desarrollar y aplicar diversas herramientas informáticas en tecnologías emergentes, las cuales tendrán un impacto en forma directa en áreas tales como: telecomunicaciones, salud, seguridad, gobierno, educación, industria, entre otras. El trabajo se cimentará en cuatro ejes fundamentales: tecnologías exponenciales (IT), tratamiento masivo de datos (big data), tecnología en educación (e-tecnología), y robótica e interacción hombre-máquina (HCI).

Palabras clave: tecnologías emergentes, tecnologías exponenciales, big data, HCI

¹ Investigador Asociado CIC (Adjunto sin Director)

² Becario CIC

³ Investigador Asociado CIC (Adjunto sin Director)

⁴ Becario CIC

⁵ Instituto Asociado CIC

Contexto

Las líneas descriptas en el presente documento están insertas en el proyecto de investigación “Informática y tecnologías emergentes”, acreditado y financiado por la Universidad Nacional del Noroeste de Buenos Aires (UNNOBA) en la convocatoria “Subsidios de investigación bianuales” 2017. Las actividades de investigación de dicho proyecto tienen lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) de la UNNOBA.

1- Introducción

En las últimas décadas del siglo XX e inicio del siglo XXI se ha producido una aceleración inédita en las vinculaciones entre las personas y la tecnología. Si bien esta es una característica propia del siglo XX, en los últimos cincuenta años esta aceleración ha crecido de forma geométrica marcando un nuevo paradigma. La sociedad actual está inmersa en un mundo de productos tecnológicos, este hecho está directamente relacionado con la revolución científico-técnica iniciada en el siglo XX y que se profundiza en el siglo XXI.

Las tecnologías se presentan cada vez más como una necesidad en el contexto de la sociedad donde los rápidos cambios, el aumento de la demanda y actualización de la información y de nuevos conocimientos se convierten en una exigencia permanente.

Es necesario considerar los cambios producidos en nuestra sociedad desde las últimas décadas, muchos de los cuales persisten hoy en día. Una de estas transformaciones, que continúa motivando cambios en nuestros hábitos y costumbres, ha tenido lugar gracias a la incorporación de nuevas y más sofisticadas Tecnologías de la Información y la Comunicación (TIC).

Las TIC son herramientas necesarias en la sociedad de la información en general y en particular son útiles en nuestra vida personal,

social y laboral. Ofrecen muchas oportunidades: nuevos empleos, participación más amplia en los procesos políticos, participación ciudadana en la toma de decisiones, acceso fácil a información y comunicación con el mundo más allá del hogar y la comunidad, desarrollo de redes, entre otras; aunque también presentan riesgos. Mantenerse al margen de las TIC hoy significa desaprovechar todas las oportunidades que ofrecen.

Con la aparición de Internet, en la década de los sesenta, se inauguró un modelo de intercambio de información, ideas y contenidos, radicalmente distinto al paradigma de los medios de comunicación convencionales. A lo largo de la evolución que han experimentado las aplicaciones de Internet, se ha puesto de manifiesto las virtudes del uso de las TIC en la sociedad actual.

En la Declaración de Principios de Ginebra se enunciaron los principios fundamentales de la construcción de una sociedad de la información para todos. Allí se reafirmó la decisión de proseguir la búsqueda para garantizar que todos se beneficien de las oportunidades que puedan brindar las TIC, recordando que los gobiernos y también el sector privado, la sociedad civil, las Naciones Unidas y otras organizaciones internacionales, deben colaborar para acrecentar el acceso a la infraestructura y las tecnologías de la información y la comunicación.

Persiguiendo estos principios fundamentales, Internet ha evolucionado en los últimos años a una velocidad de vértigo. Se ha pasado del concepto de Web 1.0 definida por Sir Timothy Berners-Lee en el año 1989 al incipiente concepto de la Web 4.0. Han sido muchas las tecnologías que se han aportado en las dos últimas décadas, empezando por Web 1.0 como una red de información y conocimiento, la Web 2.0 como una red de la comunicación, la Web 3.0 como una red de cooperación y la Web 4.0 como una red de

integración (Aghaei S, Nematbakhsh M.A. and Khosravi H. 2012).

La Web 1.0 es una red de sólo lectura, estática y unidireccional. En ella, las empresas, organizaciones y usuarios en general, podían ofrecer información y presentar sus productos y servicios. En el año 2004, Dale Dougherty (Vicepresidente de O'Reilly Media) acuñó el término de la Web 2.0 en una sesión de la conferencia de intercambio de ideas. Una de las características sobresalientes de la Web 2.0 es la colaboración y la inteligencia colectiva. Las principales tecnologías y servicios incluidos en la Web 2.0 son los blogs, las wikis, la folcsonomía, los mashups y las nubes de etiquetas (Berners-Lee T. 1998).

John Markoff (escritor del New York Times) sugirió la idea de la Web 3.0 como la tercera generación de la web en el año 2006. La Web 3.0, también conocida como Web Semántica, trata de vincular, integrar y analizar datos de los distintos conjuntos para obtener el flujo de información nueva, y es capaz de mejorar la accesibilidad, movilidad, simulación, creatividad y la innovación, fomentando el fenómeno de la globalización. La Web Semántica es una web que permite mostrar las cosas con un enfoque que la computadora puede entender (Palmer S.B., 2001).

La reciente Web 4.0 surge como un concepto, y no hay una definición formal y exacta de cómo será. Esta nueva generación de la web es conocida como la Web Simbiótica. El futuro de la web es la interacción simbiótica entre humanos y máquinas. Con ello, será posible la construcción de interfaces más potentes, tales como las interfaces controlado con la mente. En definitiva, la Web 4.0 será una web inteligente, abierta y adaptativa en la lectura de los contenidos digitales, y podrá reaccionar en la forma de ejecutar y decidir lo que se ejecute personalizando para cada usuario la información presentada (Aghaei S, Nematbakhsh M.A. and Khosravi H. 2012).

El desarrollo tecnológico actual, pone a nuestra disposición gran variedad de sensores, altas capacidades de procesamiento, conectividad alámbrica e inalámbrica, dispositivos de observación tripulados y no tripulados, y todo tipo de plataformas móviles que, debidamente automatizadas, podrán brindar una colaboración indispensable en la obtención de los objetivos buscados. Dentro del área de sensores, se destaca la importancia de la captura de información por reconocimiento de imagen, la que, por su relevancia, dividiremos en dos partes. Por un lado, la toma de imagen en sí misma y por otro, el tratamiento que de ella se realice.

En lo referido a plataformas móviles aptas para la toma de información, se cuenta con mecanismos de desplazamiento terrestres como ser mini tractores y todo tipo de plataformas de arrastre. En el caso de requerir tomas aéreas, se recurre a todo tipo de dispositivos como ser, vehículos no tripulados (Drones) que van desde aviones y helicópteros radio controlados, multicópteros, e incluso torres de observación elevadas y hasta globos aerostáticos.

En estos contextos las nociones de ubicuidad implican un cambio en un orden epistémico que aún no ha alcanzado un siglo de desarrollo, si tomamos como parámetro de cambio la entrada en la videosfera a partir de 1969 con la transmisión a escala global de los Juegos Olímpicos, parámetro propuesto por Regis Debray se considera fundamental aproximarse a una mirada sobre estas nuevas herramientas que modifican y extienden nuestras formas de percepción.

Para entender lo que está pasando en el mundo de la ciencia y la tecnología y su vertiginoso avance, nos debemos remontar al año 1969 cuando el científico Gordon Moore (cofundador de *Intel Corporation*), definió un fenómeno que impacta fuertemente a la sociedad hasta el día de hoy: “el costo de la capacidad de cómputo disminuirá a la mitad cada 18 meses”. Desde el año 1970 hasta el día de hoy, este fenómeno, conocido como la

ley de Moore, se ha cumplido cabalmente y es responsable de la penetración casi imperceptible de la computación en todos los aspectos de la vida humana.

De acuerdo a un estudio desarrollado por el Richard Smalley Institute, los diez principales problemas que afectan al desarrollo de la humanidad son: energía, agua, alimento, pobreza, medio ambiente, enfermedades, educación, guerra, democracia, población. Cada uno de estos problemas está afectando fuertemente a nuestra sociedad en la actualidad y la perspectiva futura, mirada desde el paradigma actual, se ve pesimista. Sin embargo, aplicando el rápido avance de la ciencia y tecnología a la resolución de estos problemas, nos podemos dar cuenta que el mundo cuenta o contará con las soluciones adecuadas, en la medida que la humanidad logre asimilar y utilizar los avances adecuadamente a través de sus estructuras políticas, legales y finalmente sociales.

Este desarrollo exponencial de la computación está penetrando otras áreas de la ciencia y tecnología, debido a la utilización cada día más intensa de la informática para el avance científico.

Así se reconocen un grupo de tecnologías llamadas emergentes que son víctimas de este fenómeno: la informática y las redes, la biotecnología, la robótica, la inteligencia artificial, la nanotecnología y la neurociencia.

La disminución de costos en la computación y las tecnologías de las redes de comunicaciones están produciendo un rápido crecimiento de la cobertura y velocidad de Internet. También han sufrido una gran disminución de costos los sensores y micro dispositivos inteligentes que conectados a Internet permiten la conectividad inteligente de miles de millones de objetos, la llamada Internet de las Cosas, dando lugar a múltiples servicios asociados a la localización de dispositivos (Servicios de

Localización) y a la Inteligencia de Negocios basada en esta gran cantidad de información disponible (*Big Data*).

Durante 150.000 años el desarrollo humano ha sido local y lineal. Se vive una era en la cual todo el desarrollo, particularmente el de la ciencia y la tecnología, se dan en forma global y exponencial. Este cambio de paradigmas produce una gran oportunidad de generación de valor económico y social que es difícil de predecir y que finalmente se instala en la sociedad provocando un gran impacto.

En la historia, han habido tres grandes auges del conocimiento y la innovación, la Antigua Grecia, el Renacimiento y la Sociedad de la Información. Estos tres momentos se han producido porque la humanidad ha dado un gran paso en conectividad. Hoy día, existen más de 2500 millones de personas conectadas a Internet. En los próximos 5-6 años, hasta el año 2020, se agregarán otros 3000 millones de personas a Internet, lo cual asegura un crecimiento aún mayor en la generación de conocimiento y un gran auge de la innovación.

El potencial de las tecnologías emergentes se amplifica cuando interactúan y se combinan de forma innovadora. El impacto se amplifica aún más cuando las tecnologías se unen en plataformas abiertas y ecosistemas. Estos reducen la inversión y plazo de ejecución necesario para impulsar la próxima ola de innovación en los mercados al permitir que las personas y las tecnologías creen rápidamente mayor innovación sobre las olas anteriores. Esta capacidad de combinar y recombinar las tecnologías y aprovechar las innovaciones existentes es el núcleo de las tecnologías emergentes.

2- Líneas de Investigación y Desarrollo

El objetivo general del proyecto es investigar cómo la informática impacta en el desarrollo de tecnologías emergentes de manera de

analizar, definir y desarrollar herramientas y estrategias innovadoras que impacten en el desarrollo de la sociedad.

Dicho objetivo será abordado a través de cuatro ejes: tecnologías exponenciales, tratamiento masivo de datos, tecnología en educación y robótica e interacción hombre-máquina (HCI).

Por un lado, se espera diseñar, documentar e implementar diferentes soluciones basadas en tecnologías exponenciales, analizando los requisitos de hardware y software y los desafíos e impacto que implican su implementación en diferentes contextos.

Por otro parte se busca identificar, analizar, diseñar, desarrollar y documentar distintas técnicas que pueden aplicarse al tratamiento masivo de los datos con el propósito de detectar patrones de comportamiento que contribuyan a resolver distintas problemáticas de la sociedad.

También se pretende diseñar, desarrollar y documentar diferentes soluciones innovadoras de hardware y software basadas en e-Tecnología que permitan la innovación en el campo educativo.

Por último, se busca además identificar, evaluar, analizar las capacidades y posibilidades del desarrollo de soluciones HCI aplicadas a las tecnologías emergentes como estrategias para resolver los problemas en contextos heterogéneos.

3- Resultados Obtenidos/Esperados

Con el desarrollo del proyecto se pretende difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes congresos, jornadas y workshops de carácter nacional e internacional relacionados con el uso de las tecnologías emergentes.

4- Formación de Recursos Humanos

El equipo de trabajo está compuesto por 12 investigadores formados (11 del área de Sistemas y 1 del área de Ciencias de la Educación), 12 investigadores en formación, 3 becarios de posgrado y 13 becarios de grado.

En relación a las líneas de investigación presentadas se espera para los próximos dos años contribuir al inicio y concreción de 2 (dos) Tesinas de Licenciatura en Sistemas, 3 (tres) Prácticas Profesionales Supervisadas de la Ingeniería en Informática. Actualmente se encuentra en desarrollo 1 (un) trabajo de Especialista, 4 (cuatro) Tesis de Magister y 4 (cuatro) Tesis Doctorales.

5- Bibliografía

- [1] Debray Régis, Vida y Muerte de la Imagen, Paidós, Barcelona, 1994 (1992).
- [2] Moore, Gordon E. (1965), "Cramming more components into integrated circuits". Revista Electronics N° 38
- [3] The Richard E. Smalley Institute for Nanoscale Science and Techology: <http://cnst.rice.edu/>
- [4] Patricia Halaban. "La comunicación virtual en educación a distancia, un estudio en interacciones comunicacionales y procesos pedagógicos en internet". CICCUS Febrero 2010.
- [5] Cumbre Mundial sobre la Sociedad de la Información (Ginebra 2003). Disponible en: http://www.itu.int/wsis/documents/doc_multi.asp?lang=es&id=1161|1160
- [6] Internet Usage and World Population Statistics are for June 30, 2014. <http://www.internetworldstats.com/stats.htm>

Ingeniería de Software para Desarrollar una Plataforma Académica Prototípica de eHealth

Marcelo Fabio Roldán¹, Germán Montejano², Ana Funes²

¹Departamento de Ciencias Exactas, Físicas y Naturales –
Universidad Nacional de La Rioja
La Rioja – Argentina
marcelofabio01@yahoo.com.ar

²Universidad Nacional de San Luis, Ejército de los Andes 950
5700 San Luis, Argentina
{gmonte, afunes}@unsl.edu.ar

Resumen

En la presente línea de investigación nos proponemos utilizar la Ingeniería de Software, sus métodos, técnicas y procesos como herramienta para desarrollar un sistema que implemente algunas subespecialidades informáticas en medicina con competencias específicas orientadas a la Telemedicina y a la salud electrónica (eSalud o eHealth), aplicando en su construcción artefactos de programación computacional y embebida para la adquisición, procesamiento, almacenamiento y transmisión de datos obtenidos desde diferentes elementos sensores y hacia los especialistas que interpretan tal información en terminales remotas.

Esto es posible, en concordancia con la mayor disponibilidad de elementos necesarios para la construcción de prototipos que permiten la creación de interfaces entre los parámetros físicos y el medio de comunicación que facilita el acceso desde la computadora y su software de procesamiento.

Esta construcción, que brinde los componentes principales de una plataforma de esta complejidad, se realizará a partir de la cátedra de Programación II y de la cátedra de Ingeniería de Software de la carrera de Ingeniería en Sistemas de

Información, de la Universidad Nacional de La Rioja (UNLaR).

A partir de esto, se espera generar vínculos inter organizativos, que de manera transversal, permitan utilizar y evaluar el sistema desarrollado y determinar su factibilidad de éxito en los medios donde pudiera aplicarse.

Palabras clave: eHealth, telemedicina, salud electrónica, bioinformática.

Contexto

El presente trabajo de investigación se encuentra enmarcado en una colaboración entre investigadores del Proyecto de Incentivos “Ingeniería de Software para clasificar patrones cognitivo conductuales. Clasificación taxonómica predictiva y su impacto en la graduación de los aspirantes a la carrera de suboficiales de policías.” Código de Proyecto 27/A554, de la Universidad Nacional de La Rioja, investigadores del Proyecto de Ciencia y Técnica P-031516 “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad” de la Universidad Nacional de San Luis e investigadores médicos de la Provincia de La Rioja, Fundación Vínculos, Asociación de Trabajadores de la Sanidad (ATSA) y

aprovechando los convenios existentes con Microsoft Argentina.

Introducción

La Telemedicina, uno de los servicios de eHealth, ha sido abordada desde hace más de tres décadas por diferentes investigadores que han visto reflejada en ella la posibilidad de eliminar algunas restricciones relativas a la lejanía, la escasez de medios o recursos y la aplicación de las nuevas tecnologías [AMA01].

Desde un punto de vista docente, el abordaje de un proyecto de Telemedicina brinda a los alumnos diferentes situaciones problemáticas que deberán encarar en su realidad profesional, convirtiéndose en una suerte de práctica profesional supervisada durante la cursada de las asignaturas Ingeniería de Software y Programación II de la carrera Ingeniería en Sistemas de Información de la UNLaR.

De igual manera, para los estudiantes de carreras de Medicina, Enfermería y afines, la Telemedicina proveerá un instrumento que los acercará a las tecnologías actuales de eHealth en el mundo.

El conocimiento obtenido en esta primera etapa de investigación se convertirá en base para nuevas investigaciones, pudiendo responder preguntas de difícil resolución o que demandan excesivo tiempo, como por ejemplo las dificultades en la comunicación y la necesidad de protocolos o métodos de compresión especiales.

Con esta tecnología se pretende adquirir datos desde una terminal remota asistida por personal idóneo y conectada a un paciente. Estos datos serán transmitidos por internet a uno o varios expertos en salud, quienes podrán interpretar y/o brindar un primer diagnóstico de forma remota, utilizando dichos parámetros de salud, los cuales serán visibles en el sistema local. [BBL02]

Teniendo en cuenta lo dicho anteriormente, creemos que este uso se traducirá en nuevas aplicaciones de la medicina para brindar cobertura y asistencia a lugares remotos de manera significativa o a personas con dificultad de acceso.

Resultados y Objetivos

Este trabajo de investigación tiene como objetivo principal, por un lado, que los alumnos de Licenciatura e Ingeniería en Sistemas de Información de la UNLaR apliquen una metodología de desarrollo de Ingeniería de Software, basada en las estructuras conceptuales de la Programación Orientada a Objetos para la creación de un modelo prototipo con los componentes principales de un sistema de Telemedicina. Por otro lado, se espera que dicho desarrollo sirva de referencia y base para su aplicación posterior a otros ámbitos de las subespecialidades informáticas de la medicina.

Se espera, además, que mediante el uso de equipamiento de Telemedicina y las aplicaciones de la Telemedicina en contextos remotos, resulte posible desarrollar los artefactos principales que permitan la adquisición de datos primordiales para el experto de la salud con un grado de error aceptable, proveyendo la tecnología necesaria para la comunicación efectiva entre los distintos actores del sistema.

Líneas de Investigación y Desarrollo

El trabajo se llevará a cabo a partir de casos de estudio, y de bibliografía de referencia que permitan facilitar el uso de técnicas de Telemedicina.

Durante el proyecto se prevé utilizar las herramientas de Microsoft Argentina, en especial las tecnologías .Net para la construcción del software siguiendo los lineamientos establecidos por acuerdos

vigentes en cuanto al uso de software con licencias gratuitas por alumnos de la UNLaR.

En su contexto de aplicación, este desarrollo de sistema de información, proveerá soluciones para aquellas problemáticas de salud cuya dificultad principal viene dada por la distancia y la escasez de especialistas en áreas remotas, pondrá a disposición de estudiantes y docentes de las carreras de Medicina, Odontología y Enfermería, de un recurso actual cuya tendencia en el mundo es inminente [AMA01].

Esto es posible, además, debido a que factores restrictivos de esta disciplina tienden a desaparecer, tal como las tecnologías de la comunicación, internet y la electrónica, además de los procesos de reforma y modernización de la salud para una mayor equidad, accesibilidad y prestación de servicios.

Siendo que la UNLR cuenta con un Departamento de Ciencias de la Salud y la Educación, el escenario de pruebas se encuentra al alcance, promoviendo estrategias conjuntas de trabajo orientadas a la investigación en este campo de desarrollo, no solo científico, sino laboral y disciplinar.

Ante este escenario, se espera que la aplicación de la tecnología a la Telemedicina, facilite la producción de conocimiento a partir del aprendizaje y del descubrimiento.

Se desarrollarán, durante el proyecto, las actividades necesarias para crear los módulos de software para normalizar los datos de parámetros vitales y otros orientativos para diagnóstico del médico y transmitirlos a su ubicación. A través de la aplicación de las tecnológicas informáticas se facilitará el almacenamiento y consulta posterior de los datos conformando una historia clínica unificada para Telemedicina.

Para el desarrollo de la aplicación prototípica, en este caso, se trabajará

consultando a profesionales locales que poseen alguna experiencia, en especial con AMD-Intel, principal proveedor de equipamientos de eHealth.

En una etapa posterior, se correlacionarán los resultados arrojados por los instrumentos desarrollados con los equipos que pudieran estar disponibles o sus equivalentes de uso actual. Estos resultados nos servirán como retroalimentación al analizar la calidad de medición de los modelos generados.

Formación de Recursos Humanos

El presente trabajo surge como una de las tantas aplicaciones posibles de la Ingeniería de Software y de la Programación Orientada a Objetos siendo sus resultados relevantes y formativos para los alumnos de las carreras de Licenciatura e Ingeniería de Sistemas de la Universidad Nacional de La Rioja, de la Universidad Nacional de San Luis y de manera complementaria para los alumnos y docentes de las carreras de Medicina, Odontología y Enfermería de la misma Universidad.

Fomenta, además, el trabajo en equipo e interdisciplinar, fundamental en el área del desarrollo informático, y la participación de diferentes estamentos sociales, como sindicatos, universidades y gobierno.

Referencias

- [AMA01] Andrés Martínez. Bases metodológicas para evaluar la viabilidad y el impacto de proyectos de Telemedicina. Organización Panamericana de la Salud. Washington 2001.
- [BBL02] B. Blobel. Analysis, Design And Implementation Secure And Interoperable Distributed Health Information Systems, Amsterdam, Netherlands. Ed., IOS PRESS 2002
- [CUE93] Cueto, E. G., Introducción a la psicometría. Editores: Madrid: España, 1993
- [IEE07] Watson, H.J.; Wixom, B.H. The Current state of the Business Intelligence. Univ. of Georgia,

Athens. 2007. Último acceso 2011
<http://ieeexplore.ieee.org/>

[OFE01] O.Ferrer-Roca Ed. Telemedicina, Edición Medica Panamericana 2001, Madrid, España

[RMF12] Roldán, M. F., Montejano, G., Funes, A, Una Metodología para el Desarrollo de Aplicaciones Autoadaptativas basada en Business Intelligence. Aplicación en Medicina. Tesis de Maestría. 2012

[RDF13] Roldán, M. F., Debnath, N., Funes, A, Montejano, G., Riesco, D. A Methodology Based on Business Intelligence for the Development of Predictive Applications in Self-Adapting Environments, a aparecer en proceedings de ICCSEE 2013, Hangzhou, China.

TELEMEDICINA: UN DESAFÍO PARA AMÉRICA LATINA

[MAA04] Martínez A, Villarroel V, Seoane J, Del Pozo F. A study of a rural telemedicine system in the Amazon region of Peru. Journal of Telemedicine and Telecare. 2004;10(4):219–226. [PubMed]

[MAA05] Martínez A, Villarroel V, Seoane J, Del Pozo F. Analysis of information and communication needs in rural primary healthcare in developing countries. IEEE Transactions on Information Technology in Biomedicine. 2005; 9(1):66–72. [PubMed]

[SIR03] Silverman RD. Current legal and ethical concerns in telemedicine and e-medicine. Journal of Telemedicine and Telecare. 2003;9(1):67–69. [PubMed]

[SLI05] Sergio Litewka. Telemedicina: Un Desafío Para América Latina. Acta Bioeth. 2005; 11(2): 127–132. [PubMed]

Interoperabilidad entre Lenguajes de Modelado Conceptual en *crowd*

Christian Gimenez¹ Germán BraunError!Laura CecchiError!
Pablo FillottraniError!

email: {christian.gimenez, german.braun, lcecchi}@fi.uncoma.edu.ar,
prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general del trabajo de investigación es desarrollar una nueva versión de la herramienta Web *crowd*, que permita la interoperabilidad de los lenguajes de modelado conceptual UML, EER y ORM, a partir de un metamodelo de unificación. Se pretende trabajar también en el soporte gráfico para sendos lenguajes, favoreciendo la comunicación entre usuarios mediante el uso de lenguajes comunes entre ellos. Como trabajo futuro, planeamos extender la codificación actual de la herramienta, basada en Lógicas Descriptivas, para una validación intra- e inter-lenguajes.

Palabras Clave: Ingeniería de Software basada en Conocimiento, Lógicas Descriptivas, Ontologías, Interoperabilidad de Lenguajes de Modelado Conceptual.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el marco del proyecto de investigación *Agentes Inteligentes y Web Semántica (04/F014)*, por la Universidad Nacional del Sur a través del proyecto de investigación *Integración de Información y Servicios en la Web (24/N027)*, por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral, y por el Consejo Interuniversitario Nacional (CIN), a través de una beca estímulo a las vocaciones científicas. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019. La beca CIN tiene una duración de un año.

1. Introducción

crowd, es una herramienta cliente-servidor creada en respuesta a la complejidad inherente al modelado conceptual y ontológico. En los trabajos [1, 2], se presentó la arquitectura Web de *crowd* y un prototipo implementado que permite, en primer instancia, determinar la

consistencia de un modelo gráfico representando una ontología. El *front-end* permite al usuario modelar de forma gráfica usando diagramas de clases UML, mientras que el *back-end* trabaja del lado del servidor con un razonador capaz de inferir posibles restricciones implícitas a los modelos. Los módulos en el servidor traducen el modelo inicial en un modelo lógico basado en Lógicas Descriptivas (DL) [3], como propone [4]. La comunicación entre el cliente y el servidor es a través del protocolo OWLlink [5].

Si bien, UML [6] es un lenguaje de modelado muy utilizado en la actualidad, existen otros de amplio uso, como EER [7] para bases de datos y ORM [8] para aquellos expertos en dominios e interacción entre el análisis de requerimientos y datos. Así, cada modelador podría utilizar el lenguaje que considere familiar o más adecuado para sus modelos. Por esta razón, en los sistemas de software complejos, la interoperabilidad de los lenguajes de modelado conceptual se ha vuelto una necesidad, dado que los modeladores podrían requerir vincular entidades entre modelos representados en diferentes lenguajes de modelado conceptual y comunicarse con otros *skateholders* mediante lenguajes comunes.

En el ámbito de este trabajo, proponemos extender la arquitectura de *crowd* para soportar el metamodelo de unificación de lenguajes UML, EER y ORM, presentado en [9, 10, 11], creado como una representación lógica intermedia que permitirá el mapeo de un lenguaje gráfico a otro bajo la interpretación definida en [9]. Ésta representación es ampliamente diferente a la que actualmente posee *crowd* puesto que debe ser compatible con los diferentes lenguajes que soportaremos a futuro, y por ello debe considerar la compatibilidad entre varios lenguajes de modelado conceptual y la determinación de la satisficibilidad de los diagramas. Además, se hace necesaria una modificación de la interfáz gráfica para proveer al usuario de las herramientas para visualizar y editar los otros lenguajes gráficos, como así también de las opciones para la interoperabilidad entre ellos.

Por otra parte, se prevé desarrollar una traducción a DL de los diagramas UML basada en la de ICOM [12], ampliando la expresividad

y la capacidad de representar conocimiento implícito en forma gráfica y la interacción con los sistemas de razonamiento.

La robustez de la arquitectura desarrollada para *crowd* permite su adaptación y extensión a fin de contemplar nuevas y varias traducciones e, incluso, para incrementar el conjunto de servicios de razonamientos, ampliando la posibilidad de experimentación con diversas traducciones con un esfuerzo reducido gracias al soporte base del sistema.

La estructura del presente trabajo es la siguiente. En la sección **Error! Reference source not found.** presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección **Error! Reference source not found.** indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

El proyecto de investigación *Agentes Inteligentes y Web Semántica*, UNCo, tiene varios objetivos generales. Uno de ellos es el de *generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica*. En este sentido, se estudian técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos. Fundamentalmente, se busca aplicar estos conceptos como soporte para comunidades de desarrollo de ontologías.

Por otro lado, en el proyecto de investigación *Integración de Información y Servicios en la Web* se propone investigar y desarrollar metodologías y herramientas que favorezcan la interoperabilidad semántica de información y de servicios en la Web, fundamentados en los últimos avances en el área de lenguajes de representación del conocimiento, ontologías y modelado

conceptual.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la representación del conocimiento, las Lógicas Descriptivas [13], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento. Particularmente, se ha escogido experimentar sobre metodologías que integren razonamiento con un *front-end* gráfico para dar soporte a la ingeniería de ontologías.

En esta línea de investigación se propone, como principal objetivo, continuar con el desarrollo de nuestra herramienta Web denominada *crowd*, la cual contribuye con el modelador en el diseño y la visualización de ontologías durante su evolución, actualmente por medio del lenguaje UML, asistidos por técnicas de razonamiento automáticas. La extensión de la arquitectura consistirá en permitir que los modeladores puedan utilizar diferentes lenguajes de modelado conceptual en forma conjunta y simultánea.

JointJS¹ debido a su conjunto de funcionalidades y a la utilización de Backbone².

Actualmente, del lado del servidor el prototipo soporta un módulo traductor para OWL 2 y OWLlink basado en las primitivas dadas en [14] y [5] respectivamente, y además, posee en funcionamiento un módulo generador de consultas. La salida de dicho traductor es enviada al razonador con el fin de verificar la satisfacibilidad de la base de conocimiento generada y, como consecuencia, la consistencia del diagrama de clases representado. El razonador utilizado es Racer [15] debido a que también implementa una interfaz DIG [16], permitiendo compatibilidad con las versiones previas del protocolo OWLlink. Con respecto a la implementación del metamodelo, se está trabajando sobre el mapeo de las entidades estructurales estáticas, clasificadas en [9, 17]. Por otro lado, para brindar soporte gráfico al metamodelado conceptual, nos encontramos ampliando la biblioteca JointJS mediante un plugin para ORM e incorporando el ya

crowd proporcionará soporte de visualización y edición en el lenguaje que le resulte más adecuado al modelador y permitirá la interoperabilidad entre los lenguajes, sea UML [6], ORM [8] o EER [7].

Si bien, la bibliografía relevada presenta herramientas gráficas para modelado conceptual usando los diferentes lenguajes ya mencionados, no se han encontrado implementaciones tales como la propuesta en este trabajo.

3. Resultados Obtenidos y Trabajo Futuro

El desarrollo de *crowd* provee de soporte a las primitivas gráficas de los diagramas de clases UML, con una interfaz Web que permite al usuario visualizar y modelar en forma gráfica, sin recurrir a elementos textuales. Para ello fue necesario el relevamiento de una serie de bibliotecas gráficas realizadas en JavaScript, optando por

disponible ERD plugin³ para modelos entidad-relación. Además, se implementarán las funcionalidades necesarias desde el *front-end* de *crowd* para la interacción y visualización de los lenguajes.

Finalmente, como trabajo futuro a este desarrollo [18, 2], se espera la implementación de otros algoritmos de codificación basada en DL, de los lenguajes de modelado, como los definidos en [19, 20].

4. Formación de Recursos Humanos

Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas CIN para estimular la vocación científica. Una de esas becas fue otorgada a uno de los autores de este trabajo, que está desarrollando su tesis

¹ <http://www.jointjs.com/>

² <http://backbonejs.org>

³ Plugin ERD para JoinJS disponible en <http://resources.jointjs.com/demos/erd>

de grado de la Licenciatura en Ciencias de la Computación en esta temática.

5. Referencias

- [1] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una Arquitectura Cliente-Servidor para Modelado Conceptual Asistido por Razonamiento Automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [2] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *the 2nd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '16 JAIIO '16 - to appear*, 2016.
- [3] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003.
- [4] Daniela Berardi, Diego Calvanese, and Giuseppe De Giacomo. Reasoning on UML class diagrams. *Artif. Intell.*, 168(1-2):70–118, 2005.
- [5] Thorsten Liebig, Marko Luther, Olaf Noppens, and Michael Wessel. OwlLink. *Semantic Web*, 2(1):23–32, 2011.
- [6] Grady Booch, James Rumbaugh, and Ivar Jacobson. *Unified Modeling Language User Guide*. Addison-Wesley Professional, 2005.
- [7] Martin Gogolla. *Extended Entity-Relationship Model: Fundamentals and Pragmatics*. Springer-Verlag, 1994.
- [8] Terry Halpin and Tony Morgan. *Information Modeling and Relational Databases*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2 edition, 2008.
- [9] C. Maria Keet and Pablo Rubén Fillottrani. An ontology-driven unifying metamodel of UML Class Diagrams, EER, and ORM2. *Data & Knowledge Engineering*, 2015.
- [10] Pablo Rubén Fillottrani and C. Maria Keet. Conceptual model interoperability: A metamodel-driven approach. In *Rules on the Web. From Theory to Applications - 8th International Symposium, RuleML 2014, Co-located with the 21st European Conference on Artificial Intelligence, ECAI 2014, Prague, Czech Republic, August 18-20, 2014. Proceedings*, 2014.
- [11] C. Maria Keet and Pablo Rubén Fillottrani. Toward an ontology-driven unifying metamodel for UML class diagrams, EER, and ORM2. In *Conceptual Modeling - 32th International Conference, ER 2013, Hong-Kong, China, November 11-13, 2013. Proceedings*, 2013.
- [12] P. Fillottrani, E. Franconi, and S. Tessaris. The new ICOM Ontology Editor. In *Description Logics*, CEUR Workshop Proceedings. CEUR-WS.org, 2006.
- [13] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [14] W3C OWL Working Group. *OWL 2 Web Ontology Language: Document Overview*. W3C Recommendation, 27 October 2009. Available at <http://www.w3.org/TR/owl2-overview/>.
- [15] V. Haarslev and R. Möller. Racer system description. In R. Goré, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJCAR'2001, June 18-23, Siena, Italy*, pages 701–705. Springer-Verlag, 2001.
- [16] S. Bechhofer, R. Moller, and P. Crowther. The DIG Description Logic Interface. In *In Proc. of International Workshop on Description Logics (DL2003)*, 2003.
- [17] Pablo R. Fillottrani and C. Maria Keet. KF metamodel formalization. *CoRR*,

abs/1412.6545, 2014.

[18] Germán Braun, Christian Gimenez, Laura Cecchi, and Pablo Fillottrani. Towards a Visualisation Process for Ontology-Based Conceptual Modelling. 2016.

[19] Enrico Franconi, Alessandro Mosca, and Dmitry Solomakhin. ORM2: formalisation and encoding in OWL2. In *On the Move to Meaningful Internet Systems: OTM 2012*

Workshops, Confederated International Workshops: OTM Academy, Industry Case Studies Program, EI2N, INBAST, META4eS, OnToContent, ORM, SeDeS, SINCOM, and SOMOCO 2012, 2012.

[20] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Unifying class-based representation formalisms. *J. Artif. Intell. Res. (JAIR)*, 11, 1999.

Linked Data y Ontologías en una Herramienta Gráfica Web

Gaston Michelan¹ Germán Braun^{1;2;3} Laura Cecchi¹
Pablo Fillottrani^{2;4}

email: {gaston.michelan,german.braun,lcecchig}@fi.uncoma.edu.ar, prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)

Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general del trabajo de investigación es permitir la interacción entre fuentes de datos enlazados disponibles en la Web y la herramienta cliente-servidor para el modelado conceptual gráfico con soporte de razonamiento: *crowd*. De este modo, se espera poder navegar cualquier ontología asociada a los datos y observar sus relaciones de una manera gráfica, esto último, con el fin de facilitar su interpretación al usuario convencional.

Palabras Clave: Linked Open Data, Web Semántica, Ingeniería de Software.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Integración de fuentes de datos, su representación, la extracción de información a partir de ellos y para establecer sus relaciones y correctitud. Para los primeros casos tenemos, como área de estudio, lo que se llama Linked

Comahue, en el marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014) y a través de una beca de Iniciación a la Investigación para alumnos; por la Universidad Nacional del Sur a través del proyecto de investigación Integración de Información y Servicios en la Web (24/N027), y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019. La beca de Iniciación a la Investigación tiene una duración de un año.

1. Introducción

En la actualidad, el rol de la Tecnología de la Información (TI) es indispensable en el éxito de los procesos intensivos de datos tales como la inteligencia de negocios y la toma de decisiones [1]. Sin embargo, existe una importante brecha entre las personas que interpretan y usan los datos, y los sistemas, en los cuales dichos datos son almacenados y procesados. Esto redundando en la necesidad de nuevas tecnologías y herramientas para la Data [2, 3, 4], movimiento propuesto por Tim-Berners Lee, cuyo principal objetivo es publicar y conectar datos estructurados en la Web [5, 6], a partir de un conjunto de prácticas para tal fin. Para el segundo caso podemos encontrar

herramientas de modelado conceptual, las cuales nos permiten, a través de relaciones entre conceptos u ontologías, y un razonador[7, 8, 9], evaluar la correctitud de esos datos en cuanto a sus relaciones y a la vez, proponer nuevas interacciones implícitas entre ellos[10].

En este sentido y como un primer enfoque para la solución de estos problemas, hemos desarrollado en nuestros grupos de investigación, dos herramientas: *crowd* [11, 12, 13] y LinkedGILIA [14, 15]. *Crowd* es una herramienta cliente-servidor para el modelado conceptual gráfico y diseño de ontologías con soporte de razonamiento para asistir a usuarios en esos procesos. Además, permite ejecutar chequeos de satisfacibilidad e inferencia de restricciones implícitas, cuyos resultados se muestran utilizando el mismo lenguaje gráfico. Por otro lado, LinkedGILIA es una herramienta que nos permite convertir bases de datos relacionales a RDF [16], consultarlas a través de SPARQL [17] y enlazar esos datos con otras fuentes RDF existentes en la Web, aplicando el concepto de LinkedData y haciendo que estos sean de público acceso. Esta herramienta está orientada particularmente al dominio académico y de información científica.

En el ámbito de este trabajo proponemos, como primera instancia, agregar datos enlazados a *crowd* para permitir navegar cualquier ontología asociada a ellos y poder observar sus relaciones gráficamente, esto también nos habilitará a realizar el proceso inverso, es decir, a partir de una ontología obtener sus datos relacionados. A su vez, se planea colaborar con el modelador para mejorar las ontologías diseñadas, aplicando herramientas como DL-learner [18, 19], continuando con el desarrollo del ciclo de vida de aplicaciones Linked Data [20] comenzado en [15].

Finalmente, se propone extender *crowd* para visualizar los datos en RDF y en algún formato más amigable utilizando alguna biblioteca PHP para consultar las tripletas producidas.

La estructura del presente trabajo es la siguiente. En la sección 2, presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación, el problema que se

estudia y los objetivos. En la sección 3, indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

Dentro de los objetivos del nuevo proyecto de investigación Agentes inteligentes y Web Semántica, UNCo, se planteó el de generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en sistemas computacionales basados en la web, es decir lo que se ha llamado la Web Semántica. Para ello, es necesario profundizar el estudio de técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos. Se pretende aplicar estos conceptos como soporte para comunidades de desarrollo de ontologías, entre otros.

El proyecto de investigación Integración de Información y Servicios en la Web, UNS, se propone la integración y búsqueda, en forma semántica, de la información y de servicios en la Web utilizando lenguajes y tecnologías que están innovando en la actualidad.

Las investigaciones de ambos proyectos pueden relacionarse y ser aplicadas en el estudio de tecnologías y formalismos para solventar las necesidades de compartir, intercambiar y asegurar la correctitud de la información, que se encuentra disponible en sistemas existentes, de una forma precisa y consistente.

Particularmente, en esta línea de investigación se experimentará sobre la integración de tecnologías y lenguajes ya existentes para la generación de datos abiertos, enlazados y correctos y la publicación de los mismos en la Web. Del mismo modo, se espera, identificar falencias, proponer mejoras y evaluar el impacto en los usuarios, de esta nueva infraestructura.

Como primer paso se trabajará en el análisis de la arquitectura y del funcionamiento in-terno

de la herramienta *crowd*, para determinar las formas de inserción de datos y el tipo de los mismos que admite. Luego, se adaptará la arquitectura de la herramienta, de modo de crear una conexión entre diferentes fuentes de datos RDF y *crowd*. Esto mejorará la asistencia al modelador ya que permitirá visualizar gráficamente, además de las relaciones entre los datos, las ontologías subyacentes a dichas fuentes enlazadas.

En una segunda instancia, se pretende extraer nuevo conocimiento a partir de los datos, usando herramientas basadas en machine learning y generando axiomas y conceptos que se correspondan con los datos subyacentes. De esta manera, soportaremos el proceso de integración de los datos [21] con las ontologías de alto nivel.

Nuestra intención con esto, es establecer relaciones entre datos pertenecientes a bases de datos linkeadas en la Web, como así también exponer toda esta información a través de front-ends visuales para habilitar búsquedas semánticas a usuarios convencionales.

3. Resultados Obtenidos y Trabajo Futuro

De anteriores trabajos surgió LinkedGILIA [15], que es una aplicación Web específica para el dominio académico, que enlaza información científica desde diferentes fuentes como, por ejemplo, DBLP, entre otras. Dicha aplicación permite visualizar la información actualizada asociada a los datos disponibles y realizar consultas en el lenguaje SPARQL [17]. Como caso de estudio se utilizó, para la implementación del prototipo, al grupo de investigación GILIA del Departamento de Teoría de la Computación, de la Facultad de Informática, de la Universidad Nacional del Comahue.

Actualmente, nos encontramos trabajando en el análisis de la arquitectura y funcionalidades de *crowd* [12, 13] y en la evaluación de las diferentes formas de interacción que esta posee con elementos externos. A partir de los resultados de estos estudios, se extenderá la arquitectura *crowd* y se desarrollará una interfaz, que nos permita navegar por los datos

y la ontología modelada, accediendo a través de sus enlaces y visualizando información relacionada con otras ontologías existentes en la Web.

Se espera implementar un prototipo de la arquitectura extendida y evaluarla sobre distintos casos de estudio.

4. Formación de Recursos Humanos

Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

Otro de los autores, alumno avanzado de la carrera Licenciatura en Ciencias de la Computación, ha obtenido una Beca de Iniciación en la Investigación para Alumnos Universitarios de la Universidad Nacional del Comahue. Dicho becario realizará su tesis de grado en la temática de la línea de investigación presentada en el marco del GILIA.

5. Referencias

- [1] Bou-Wen Lin. *Information technology capability and value creation: Evidence from the us banking industry*. Technology in Society, 2007.
- [2] Tim Berners-Lee. Linked Data, 2006. <https://www.w3.org/DesignIssues/Linkedata.html>, accedida en septiembre de 2016.
- [3] Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked Data - The Story So Far. *Int. J. Semantic Web Inf. Syst.*, 2009.
- [4] Tom Heath and Christian Bizer. *Linked Data: Evolving the Web into a Global Data Space*. Morgan & Claypool, 1st edition, 2011.
- [5] T. Berners-Lee, J. Hendler, and O. Lassila. *The semantic web*. *Scientific American*, May 2001.

- [6] Liyang Yu. *A Developer's Guide to the Semantic Web*, Second Edition. Springer, 2014.
- [7] V. Haarslev and R. Moller. Racer system description. In R. Gore, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJ-CAR'2001*, June 18-23, Siena, Italy, pages 701–705. Springer-Verlag, 2001.
- [8] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. Pellet: A practical owl-dl reasoner. *Web Semant.*, 5(2):51–53, June 2007.
- [9] Dmitry Tsarkov and Ian Horrocks. Fact++ description logic reasoner: System description. In *In Proc. of the Int. Joint Conf. on Automated Reasoning (IJCAR 2006)*, pages 292–297. Springer, 2006.
- [10] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003.
- [11] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una arquitectura cliente-servidor para modelado conceptual asistido por razonamiento automático. In *Proceedings of XVIII Workshop de Investigadores en Ciencias de la Computación*, Argentina, 2016. Universidad Nacional de Entre Ríos.
- [12] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. crowd: A tool for conceptual modelling assisted by automated reasoning - preliminary report. In *Proceedings of Simposio Argentino de Ontologías y sus Aplicaciones*. 45 JAIIO, Buenos Aires, 2016. Universidad de Tres de Febrero.
- [13] Germán Braun, Christian Gimenez, Laura Cecchi, and Pablo Fillottrani. Towards a visualisation process for ontology-based conceptual modelling. In *Proceedings of VIII Brazilian Seminar Ontology. ONTOBRAS*, Curitiba-PR, Brasil.
- [14] Gaston Michelan, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Linked open data para la integración de información científica. In *Proceedings of XVIII Workshop de Investigadores en Ciencias de la Computación*, Argentina, 2016. Universidad Nacional de Entre Ríos.
- [15] Gaston Michelan, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Integration of scientific information through linked data - preliminary report. In *Proceedings of Simposio Argentino de Ontologías y sus Aplicaciones*. 45 JAIIO, Buenos Aires, 2016. Universidad de Tres de Febrero.
- [16] Marcelo Arenas, Claudio Gutierrez, and Jorge Pérez. Foundations of rdf databases. In *Reasoning Web*, pages 158–204, 2009.
- [17] World Wide Web Consortium. SPARQL 1.1 Query Language. Available at <http://www.w3.org/TR/sparql11-query/>, last accessed May 2016
- [18] DL-Learner Homepage. <http://dl-learner.org/>.
- [19] Lorenz Buhmann, Jens Lehmann, and Patrick Westphal. DL-learner—a framework for inductive learning on the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 39:15–24, 2016.

- [20] Anja Jentzsch, Max Schmachtenberg, Christian Bizer and Richard Cyganiak. Linking Open Data cloud diagram, 2014. <http://lod-cloud.net/>, accedida en septiembre de 2016.
- [21] AnHai Doan, Alon Halevy, and Zachary Ives. *Principles of Data Integration*. Elsevier, 2012.

Métodos y Técnicas para Desarrollos de Aplicaciones Ubicuas

Elena Durán¹, Silvina Unzaga¹, Margarita M. Álvarez¹, Nevelin I. Salazar¹, Gabriela Gonzalez¹, Beatriz Fernández Reuter¹, Patricia P. Zachman²

¹ Instituto de Investigaciones en Informática y Sistemas de Información (IISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)
e-mail: {edurán, alvarez, sunzaga, nsalazar, ggonzalez, bfreuter}@unse.edu.ar

²Dpto de Ciencias Básicas y Tec.Aplicadas, Universidad Nacional del Chaco Austral
e-mail: {ppzsp1640@gmail.com}

CONTEXTO

En este trabajo se presenta la línea de investigación del proyecto “Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas”, correspondiente a la convocatoria 2016 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SICYT - UNSE). El proyecto de reciente aprobación tiene un período de ejecución 2017-2019. Lo allí propuesto es una continuación de la línea de investigación iniciada en 2012, en el proyecto "Sistemas de información web personalizados, basados en ontologías, para soporte al aprendizaje ubicuo", aprobado y financiado por SICYT – UNSE. La línea de investigación presentada en este trabajo, propone favorecer el desarrollo de conocimiento científico- tecnológico de relevancia sobre Computación Ubicua, realizando propuestas de técnicas, métodos y estrategias para el diseño y construcción de aplicaciones ubicuas.

RESUMEN

Los nuevos desarrollos de las Ciencias de la Computación/Informática se orientan hacia la Computación Ubicua, en la que los ordenadores están incluidos en nuestros movimientos naturales y en las interacciones con nuestro entorno, tanto físico como social. Este nuevo paradigma de la computación involucra la movilidad; es decir, la capacidad de mover los servicios informáticos con nosotros; y la pervasividad, es decir, la capacidad de obtener información del entorno en el que está inmersa una aplicación informática y utilizarla para construir

dinámicamente modelos de computación. En consecuencia, los principales desafíos en la Computación Ubicua se originan en la integración de la movilidad con la pervasividad, en la búsqueda de que todo dispositivo computacional, mientras se mueve con nosotros, pueda ir construyendo de forma incremental modelos dinámicos de sus diversos entornos y configurar sus servicios en consecuencia.

Nuevos retos emergen en términos de cómo se debe diseñar y desarrollar aplicaciones ubicuas. Esto requiere el replanteo de las arquitecturas de software posibles, el diseño de ontologías y de modelos de dominio, nuevos escenarios de interacciones y el análisis de nuevas familias de requisitos no funcionales (como configurabilidad y la adaptabilidad).

En este proyecto se propone favorecer el desarrollo de conocimiento científico-tecnológico de relevancia sobre Computación Ubicua, realizando propuestas de técnicas, métodos y estrategias para el diseño y construcción de aplicaciones ubicuas.

Para esta investigación, es necesario un estudio de tipo exploratorio, descriptivo, analítico y, en algunos casos, correlacional, e investigación aplicada, en la que se usarán métodos y técnicas cuantitativas y cualitativas para alcanzar los objetivos propuestos. La investigación está orientada a analizar, proponer y aplicar técnicas, métodos y/o estrategias para conocer el contexto, razonar en base a él y adaptar los servicios de las aplicaciones ubicuas, en miras de mejorar el

nivel de satisfacción del usuario y el desempeño de estas aplicaciones.

Palabras clave: *Computación Ubicua, Técnicas de Diseño de Software, Razonamiento Automático, Sistemas Adaptativos, Ontologías, Evaluación de Sistemas Ubicuos.*

1. INTRODUCCIÓN

La rápida expansión de las comunicaciones inalámbricas y la masificación de los dispositivos móviles de computación, han impactado fuertemente en las actividades cotidianas, especialmente en las formas de acceso al conocimiento, dando lugar al surgimiento de una nueva forma de computación denominada computación ubicua (u-computing).

El término "computación ubicua" se atribuye a Mark Weiser (Weiser, 1991) quien manifestaba que las tecnologías más profundas son aquellas que desaparecen, quedando inmersas en la vida cotidiana, de forma tal que no se pueden distinguir de ella. En la misma, se combinan tecnologías de información y comunicación utilizando un gran número de pequeñas computadoras embebidas, equipadas con sensores y actuadores que interactúan con el medio ambiente para intercambiar datos (Sakamura & Koshizuka, 2005). Su principal propósito es facilitar las tareas cotidianas, de modo que puedan realizarse de modo más rápido y efectivo.

Esta nueva forma de computación posee las siguientes características (Ding et al., 2010):

- Invisible: interviene en nuestra vida diaria, pero es tan insignificante, que las personas se olvidan de su gran impacto en lo cotidiano.
- Ubicua: permite la integración entre la computadora y el ambiente en el que se desenvuelven las personas por medio de dispositivos embebidos (PADs, tabletas, celulares, tarjetas y lectores de identificación por radiofrecuencia (RFID), entre otros) en el ambiente de manera que permitan el acceso y comunicación en "cualquier lugar" y en "cualquier momento".

- Dinámica: por un lado, se encuentran los usuarios que están continuamente en movimiento dentro de un ambiente de computación ubicua; por el otro, los dispositivos que entran y salen de un ambiente de computación, lo que lleva a que la estructura de los sistemas cambie dinámicamente.
- Autoadaptativa: un sistema ubicuo es capaz de detectar y/o inferir las necesidades de sus usuarios, y de proveer voluntariamente a los mismos los servicios de información requeridos.

Para que un sistema de computación ubicua, sea capaz de asistir al usuario proactivamente, es necesario que posea la habilidad de sensor el contexto del mismo, razonar frente a cambios producidos en dicho contexto y reaccionar de manera acorde, de ser necesario (Ding et al., 2010).

El contexto en un sistema ubicuo se define como cualquier información que puede ser usada para caracterizar la situación de una entidad. Una entidad es una persona, lugar u objeto que es considerado relevante para la interacción entre el usuario y la aplicación, eventualmente incluyendo al usuario y a la aplicación misma (Dey, 2001). Dependiendo de la entidad y del tipo de soporte personalizado que se pretenda brindar, será necesario considerar diferentes aspectos (físicos, ambientales, personales, sociales, del sistema, de la aplicación, etc.) al modelar el contexto en una aplicación ubicua.

Dada su característica de sensibilidad al contexto, puede decirse que un sistema ubicuo realiza tres funciones principales (Loke, 2006):

- Sensor: los sensores proporcionan un medio de adquirir datos o información acerca del mundo físico o de algún aspecto del mismo (contexto).
- Pensar: los datos capturados se procesan y, en combinación con otra información poseída por el sistema, se infiere nueva información acerca de la situación de las diferentes entidades que están utilizando el sistema.
- Actuar: una vez que el sistema ha caracterizado las situaciones de las diferentes entidades, es decir, ha

comprendido su contexto, está en condiciones de ejecutar las acciones más convenientes para el logro de sus objetivos.

Teniendo en cuenta la funcionalidad que deben presentar esta clase de sistemas, se puede observar que su diseño presenta interrogantes que no se manifiestan en otro tipo de aplicaciones, como por ejemplo: cómo modelar contexto heterogéneo y dinámico de forma eficiente, cómo proporcionar adaptación contextual adecuada y oportuna a cada usuario particular, cómo integrar eficazmente los diferentes dispositivos que conforman el sistema, cómo identificar contenido relevante en repositorios digitales con miras a la adaptación, y cómo combinarlo con la información sensada por el sistema. Es decir, el desarrollo de un sistema ubicuo requiere de la utilización de técnicas especializadas de adquisición y modelado del contexto, así como de razonamiento y adaptación.

Esta necesidad se ve reforzada por el gran potencial que exhibe esta clase de aplicaciones, tanto porque facilitan la interacción hombre-computadora ofreciendo asistencia personalizada y oportuna al usuario, como por los diversos contextos de uso en los que es posible implementarlas.

2. DESCRIPCIÓN DE LA LINEA DE INVESTIGACION

Atendiendo a las problemáticas planteadas en el apartado anterior, en Durán et al. (2014) hemos propuesto una arquitectura para desarrollar aplicaciones de aprendizaje ubicuo; la que puede ser adaptada sin mayores cambios para aplicaciones de computación ubicua en general. Sobre la base de esta arquitectura es posible el desarrollo de aplicaciones que operen en entornos dinámicos de computación ubicua y se adapten a los cambios de contexto. Esta arquitectura ofrece, además, un enfoque de desarrollo de aplicaciones basado en modelos ontológicos que facilita la adaptación dinámica y automática de los servicios disponibles para el usuario.

Considerando los diversos componentes de esta arquitectura (módulos de software que

gestionan servicios, modelos, ontologías y repositorios) y que existen distintas técnicas y métodos que se pueden aplicar para el modelado y creación de estos componentes, se propone en este proyecto investigar las técnicas y métodos para el modelado y construcción de aplicaciones ubicuas con el fin de favorecer el desarrollo de conocimiento científico-tecnológico de relevancia sobre Computación Ubicua, realizando propuestas de técnicas, métodos y estrategias para el diseño y construcción de aplicaciones ubicuas.

Para alcanzar lo expuesto anteriormente, se plantean los siguientes objetivos específicos:

1. Revisar y analizar las técnicas y métodos existentes para el modelado y la construcción de aplicaciones ubicuas.
2. Construir y evaluar el modelo de contexto ubicuo.
3. Evaluar y proponer técnicas y métodos para el razonamiento y la adaptación en aplicaciones ubicuas.
4. Aplicar las técnicas y métodos propuestos en la construcción de aplicaciones ubicuas de impacto local y regional.
5. Evaluar, en contextos reales, el nivel de satisfacción del usuario y el desempeño de las aplicaciones ubicuas construidas.

3. METODOLOGÍA

Trabajos previos, de los últimos años, del grupo de investigación: (Figuroa et al, 2014; Durán, et al., 2014a; Durán, et al., 2014b; Salazar y Durán, 2014; Gonzalez y Durán, 2014; Alvarez et al., 2015 a; Alvarez et al. 2015b; Unzaga et al., 2015; Durán et al.2016a; Loto y Durán, 2015, Fernández Reuter et al. 2014; Durán 2015.

), entre otros, constituirán el punto de partida para esta investigación.

Luego, y con el propósito de dar cumplimiento al objetivo específico 1, se realizarán las siguientes actividades:

- a) Exploración e investigación bibliográfica sobre técnicas y métodos utilizados actualmente para el modelado y construcción de aplicaciones ubicuas.
- b) Análisis de los antecedentes encontrados

c) Síntesis de los antecedentes encontrados.

Con el fin de obtener el objetivo específico 2, se realizará:

- a) Definición de categorías de información a incluir en el modelo de contexto.
- b) Determinación de las estrategias de representación de la información del modelo de contexto y diseño de su estructura.
- c) Determinación de las estrategias de adquisición de la información del modelo de contexto.
- d) Determinación de las estrategias de composición de la información y/o razonamiento de nuevo conocimiento del modelo de contexto
- e) Evaluación del modelo de contexto construido.

Para el objetivo específico 3 se plantean las siguientes actividades:

- a) Profundización del estudio de las técnicas y métodos usados para el razonamiento y adaptación de aplicaciones ubicuas.
- b) Evaluación de las técnicas y métodos estudiados.
- c) Diseño de nuevas técnicas y métodos para mejorar las falencias encontradas.

Con el propósito de dar cumplimiento al objetivo específico 4, se realizarán las siguientes actividades:

- a) Identificación de problemáticas locales y/o regionales que justifiquen el desarrollo de aplicaciones ubicuas.
- b) Análisis y Diseño de repositorios de contenido para aplicaciones ubicuas
- c) Diseño de las aplicaciones ubicuas.

Por último, para dar cumplimiento con el objetivo específico 5 se realizará:

- a) Operacionalización de las variables nivel de satisfacción del usuario y desempeño de aplicaciones ubicuas.
- b) Ejecución de las aplicaciones en contextos reales.
- c) Análisis de resultados.
- d) Elaboración de conclusiones.

4. RESULTADOS ESPERADOS

Los resultados que se esperan obtener de esta investigación son:

- Estado del arte sobre técnicas y métodos usados actualmente para adquirir el contexto, razonar en base a él, adaptar los servicios y contenidos, y crear y gestionar repositorios para aplicaciones ubicuas.
- Modelo de contexto para aplicaciones ubicuas.
- Técnicas y métodos para el razonamiento y la adaptación de aplicaciones ubicuas.
- Prototipos de aplicaciones ubicuas que atiendan problemáticas locales y/o regionales con licencia y/o patente.
- Diseño de repositorios para ser gestionados desde aplicaciones ubicuas.

El proyecto tendrá también un importante impacto a nivel local, ya que los desarrollos concretados en el marco del proyecto serán transferibles de modo directo a organizaciones del medio cuyas problemáticas se atienden en estos desarrollos.

5. FORMACIÓN DE RECURSOS HUMANOS

El desarrollo de la línea de investigación presentada, a través de la ejecución del proyecto facilitará la formación de recursos humanos de la Facultad de Ciencias Exactas y Tecnologías de la UNSE. En este sentido, dos integrantes son Becarias CONICET y están desarrollando su Plan de Beca y su Tesis de Doctorado en Ciencias de la Computación.

Otra integrante del proyecto, elaborará su Trabajo Final Integrador de la carrera de postgrado Especialización en Enseñanza de la Tecnología de la UNSE. Además de los resultados esperados indicados en el apartado 4, se considera que el desarrollo de este proyecto impulsará el afianzamiento en líneas de investigación ya existentes sobre computación ubicua, personalización, y ontologías, lo que contribuirá a una mejora en el fondo de conocimiento disciplinar disponible no sólo a nivel local sino también regional y nacional.

6. BIBLIOGRAFÍA

- Álvarez, M., Únzaga S. y Durán E. (2015a). "Modelo de dominio en sistemas de aprendizaje ubicuo". 10° Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. 21 y 22 de mayo de 2015. Universidad Nacional de Salta. ISBN N°: 978-987-633-133-3. Pág. 116.
- Álvarez, M., Únzaga, S. y Durán, E. (2015b). "Recomendaciones Personalizadas para Aplicaciones de Soporte al Aprendizaje Ubicuo". XXI Congreso Argentino de Ciencias de la Computación. CACIC 2015. Junín.
- Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, 5(1), 4-7.
- Ding, G. J., Li, H., & Tingting, Z. (2010). A preliminary study of personal learning environment based on Ubiquitous Computing Model. In 3rd IEEE International Conference on Ubimedia Computing UMedia 2010 (pp. 350–354). Ieee. Retrieved from <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5544429>
- Durán, E.; Álvarez, M. y Únzaga, S. (2014a). "Ontological model-driven architecture for ubiquitous learning applications". EATIS 2014 - 7th Euro American Association on Telematics and Information Systems. Valparaíso (Chile). 2 al 4 de Abril de 2014.
- Durán, E.; Alvarez, M. y Unzaga, S. (2014b). "Design of a Personalization Module for U-learning application". Fifth International Conference on Advances in New Technologies, Interactive Interfaces and Communicability (ADNTIIC 2014), ALAIPO, Huerta Grande, Córdoba, Argentina, ISBN 978.88.96.471.37.1.
- Durán, E., Álvarez, M., Únzaga, S. y González, G. (2016a). "Personalization Module for U-learning Applications". Handbook of Research on Human Interaction and the Impact of Information Technologies/IGI Global. En edición.
- Durán, E., Únzaga S. y Álvarez, M. (2015). "Instanciación del Modelo de Servicios para una aplicación de apoyo al Aprendizaje Ubicuo en un curso de Redes de Computadoras". Revista TEKNO de la Facultad de Ingeniería de la Fundación Universitaria Tecnológico Comfenalco. ISSN 1900–7388. Vol.15; N°1; pag. 31-42.
- Fernandez Reuter, B. y Durán, E. (2014). "Framework de recomendación automática de contenidos en foros de discusión para entornos de e-learning". 7th Euro American Association on Telematic and Information Systems (EATIS 2014). Valparaíso, Chile. Proceedings published by ACM Digital Library within its International Conference Proceedings Series. ISBN 978-1-4503-2435-9. Article N°: 38.
- Figueroa, S; Cordero, R. Leiva, V; Aoad, V. (2014) "Modelo de entorno de aprendizaje ubicuo: su aplicación en el proceso de capacitación docente". IV Jornadas del NOA y II Jornadas nacionales de educación a distancia y tecnologías educativas. Argentina.
- González, G. y Durán E. (2014). "Modelo del estudiante para sistemas de aprendizaje ubicuo: representación por medio de ontologías". IX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2014), Universidad Nacional de Chilecito, La Rioja, ISBN 978-987-24611-1-9, pp. 298-305.
- Loke, S. (2006). *Context Aware Pervasive Systems* (1st ed.). Auerbach Publications. <http://doi.org/QA76.5915.L65.2006>.
- Loto M. y Durán E. (2015). "Diseño de una aplicación móvil personalizada de apoyo al aprendizaje de Redes de Computadoras". X Congreso de Tecnología en Educación & Educación en Tecnología. Corrientes.
- Sakamura, K., & Koshizuka, N. (2005). Ubiquitous computing technologies for ubiquitous learning. In IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'05), pp. 11- 20, IEEE.
- Salazar, N. y Durán, E. (2014). "Objeto de Aprendizaje para la enseñanza de la Simulación". IX Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2014).
- Únzaga S., Álvarez M., Durán E. (2015). "Modelo de Requerimientos de una Aplicación de Apoyo al Aprendizaje Ubicuo para el Ingreso Universitario". TE&ET'15: X Congreso de Tecnología en Educación y Educación en Tecnología 2015. Argentina.
- Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, 265(3), 66–75. Recuperado de http://www.syssoft.univ-trier.de/systemsoftware/Download/Fruehere_Veranstaltungen/Ubiquitous_Computing/2004/02/ParadigmPrint.pdf

Misión CubeSat FS2017: Desarrollo de Software para una Misión Satelital Universitaria

Ezequiel González¹, Pablo Soligo², Eduardo Sufán, Emmanuel Arias, Ricardo Barbieri, Pablo Estrada, Alfonso Montilla, José Robin, Javier Uranga, M. Cecilia Valenti y Elbio Zapata

Comisión Nacional de Actividades Espaciales (CONAE)
Universidad Nacional de La Matanza (UNLaM)

⁽¹⁾ ezequielg@alumno.unlam.edu.ar

⁽²⁾ psoligo@unlam.edu.ar

RESUMEN

Siguiendo los lineamientos establecidos en el Plan Espacial Nacional [1], la Comisión Nacional de Actividades Espaciales (CONAE) ha creado la Unidad de Formación Superior (UFS) para capacitar recursos humanos de alto nivel en ciencia y tecnologías espaciales, a fin de satisfacer las necesidades de las instituciones/empresas científico-tecnológicas argentinas. En asociación con varias universidades nacionales, la UFS ha implementado nuevos programas de maestría en diferentes áreas tecnológicas que cubren las especialidades de un sistema espacial. En ese contexto, el objetivo primario de la misión FS2017 (*Formador Satelital 2017*) es proveer a los maestrandos de la UFS entrenamiento práctico en los aspectos técnicos y programáticos de una misión espacial. Este trabajo presenta un resumen de la implementación del software (SW) de la misión, dividido en 2 segmentos: SW de tierra y SW de vuelo. Se abordan conceptos generales de la misión, el enfoque de management, la arquitectura y las características técnicas de la implementación.

Palabras Clave: Ingeniería de SW, Desarrollo de SW, CubeSats, SW para misiones espaciales, Operaciones de misiones espaciales, Proyectos espaciales universitarios.

CONTEXTO

CONAE, en asociación con tres universidades nacionales -Universidad Nacional de Córdoba, Universidad Tecnológica Nacional (Facultades Regionales Córdoba y Mendoza) y Universidad Nacional de La Matanza-, ha implementado 4 nuevos programas de Maestría: en Aplicaciones

de Información Espacial (MAIE), en Tecnología Satelital (MTS), en Instrumentos Satelitales (MIS) y en Desarrollos Informáticos de Aplicación Espacial (MDIAE). La misión FS2017 es un desarrollo conjunto de las 4 maestrías de la UFS, dependiente de la CONAE, en asociación con sus socios académicos. Dicha misión se compone de dos proyectos [2]: el que está siendo realizado por las cohortes 2015-2017, incluyendo únicamente las fases A y B, y la continuación del proyecto FS2017 (desde la Fase C hasta el decomisionado del satélite) a ser realizado por futuras cohortes (figura 1).

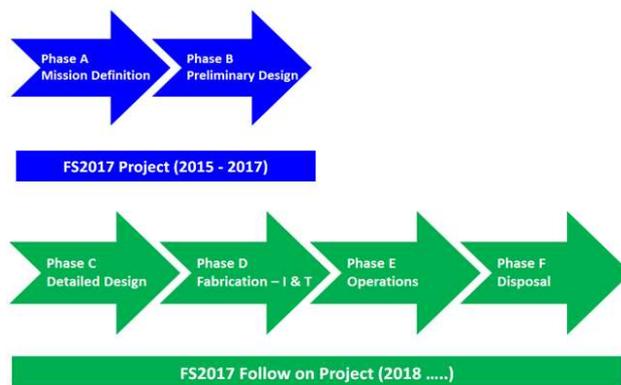


Figura 1: Proyecto y Misión 2017

El objetivo primario de la misión FS2017 es académico, esto es, la formación de los alumnos de la UFS en las áreas técnicas y programáticas de una misión satelital, según su especialidad [2]. Para satisfacer este objetivo, los estudiantes deben participar en todas las fases correspondientes al ciclo de vida de una misión espacial [3][4], desde la Fase A (estudio de factibilidad) hasta la Fase F (decomisionado). Para verificar este objetivo, los estudiantes deben aprobar todas las revisiones formales previstas en este tipo de misión/proyecto.

Los criterios de éxito de la misión [2], propuestos para satisfacer el objetivo primario, son (en orden de prioridad): (1) Familiarización de los maestrands con el ciclo de vida de un proyecto espacial; (2) Diseño, fabricación, integración y ensayos del satélite; (3) Campaña previa al lanzamiento; (4) Envío del satélite al espacio y actividades de comisionado; (5) Proceso, análisis, interpretación y difusión de los datos de la misión; (6) Entrenamiento de futuras cohortes durante la fase operativa de la misión.

1. INTRODUCCIÓN

A. Concepto de Misión

De acuerdo a estudios previamente realizados [2], la única opción que permite la realización de misiones del tipo de la FS2017 en un plazo de dos años y con las limitaciones de presupuesto propias del ambiente académico es la categoría de Pico o Nanosatélite (ver recuadro en línea punteada en la Figura 2).

Femtosatellite	Picosatellite	Nanosatellite	Microsatellite	Minisatellite
1 to 100 [gr]	0.1 to 1 [kg]	1 to 10 [kg]	10 to 100 [kg]	100 to 500 [kg]
1 to 10 [KUSD]	10 to 300 [KUSD]	0.1 to 10 [MUSD]	0.5 to 30 [MUSD]	5 to 300 [MUSD]

Figura 2: Categorías de Satélites

Dentro de esta categorización, se ha decidido usar específicamente el concepto CubeSat [5] (sombreado en rojo). Las ventajas de este concepto son: propósito educacional, múltiples oportunidades de lanzamiento con mínimo riesgo, amplia disponibilidad de proveedores de componentes, corto tiempo de desarrollo y baja relación costo-beneficio. Para poder cumplir con los objetivos de misión, se han definido dos instrumentos o cargas útiles: (1) una cámara que calificará el detector usado en las cámaras NIR-SWIR (Near/Short-Wavelength InfraRed) [6] de los satélites SABIA-Mar [7] de CONAE y (2) un sistema de colecta de datos -DCS [8]- que podrá adquirir datos provenientes de estaciones en tierra, permitiendo proveer servicios a instituciones que necesitan recabar información desde áreas remotas del país.

B. Organización del Proyecto

El proyecto FS2017 presenta cuatro ramas: plataforma, instrumentos, aplicaciones y soft-

ware. Cada una de ellas define su propio organigrama y en su conjunto reportan al Project Manager de la misión [2]. Con el objetivo de tratar temas transversales a las distintas áreas/maestrías se han definido grupos específicos de trabajo a nivel misión para resolver cuestiones de gestión de proyectos, calidad, ingeniería de sistemas e integración y ensayos.

C. Metodología de Desarrollo de Software

El modelo de desarrollo utilizado para el proyecto FS2017 es el *iterativo incremental* [9]. La metodología seleccionada para el desarrollo de los incrementos es *Scrum* [10][11]. La razón principal de dichas elecciones radica en la distribución horaria de las actividades del proyecto FS2017 en el calendario de la MDIAE [12]. Los períodos pre-asignados al proyecto son alternados y breves, motivo más que suficiente para pensar en *sprints* cortos, con objetivos claros, reuniones breves y eficientes, como la mejor forma de trabajo para el equipo de SW.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo trata sobre el análisis, diseño, desarrollo, testing y operación del SW requerido para una misión espacial [13][14][15][16], la FS2017. El producto SW se ha dividido en 2 grandes áreas [17]: SW del segmento Terreno (inciso A) y SW del segmento de vuelo (inciso B).

A. Software del Segmento Terreno

El SW del segmento terreno tiene como objetivo permitir la operación y el mantenimiento del estado de salud del satélite FS2017, pero con los atributos de generalidad para que sea aplicable a futuras misiones. Dentro de las funcionalidades principales se encuentra la bajada y análisis de telemetría (TM), la subida de telecomandos (TC), procesamiento y publicación de los datos de ciencia. A tal fin se desarrolló un SW multimisión [17] que sirve de soporte a las tareas de operación del satélite con un alto nivel de flexibilidad y robustez.

La arquitectura del segmento terreno se presenta en la figura 3. Las principales funciones que debe cumplir son:

- **Mantenimiento del estado de salud:**
 - Visualización de TM en tiempo real e histórica
 - Análisis de tendencia.
 - Alarmas, Proced. y acciones de contingencia.
- **Planificación y ejecución de adquisiciones:**
 - Creación de scripts de comandos.
 - Calendarización de actividades de pasadas.
 - Subida de comandos y revisión de ejecución.
 - Generación de productos orbitales.
 - Pasadas planificadas.
 - Eclipses.
- **Descarga, procesamiento y publicación de datos ciencia:**
 - Bajada de datos, publicación vía portales y servicios web.

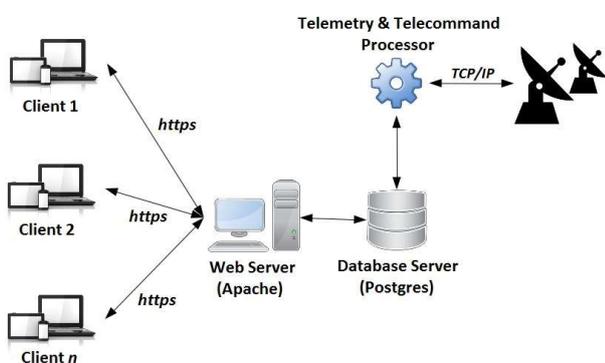


Figura 3: Arquitectura del SW del segmento Terreno

El sistema está montado sobre una arquitectura cliente-servidor clásica y fue desarrollado en Python sobre el framework Django [18]. El lenguaje es de tipo dinámico, el framework fuerza el uso del paradigma orientado a objetos y un diseño del tipo Model-Driven Architecture (MDA). El motor de la base de datos es PostgreSQL 9 y el acceso a los datos por parte de las aplicaciones se realizó en su totalidad mediante el ORM (Object-Relational Mapper) disponible en el framework. Los servicios 7x24 (Decodificación de TM y codificación de comandos) hacen uso parcial del framework utilizando únicamente el ORM integrado. La aplicación es hosteada sobre apache en sistema operativo Linux Ubuntu 15.10.

El sistema es una alternativa a otros desarrollos ad-hoc ofreciendo una variante multiplataforma y multimisión que maximiza los atributos de:

- **Interoperabilidad:** Acceso estándar, común y centralizado para aplicaciones del dominio.
- **Eficiencia:** Optimización del espacio de almacenamiento, optimización de las capacidades de búsqueda y recuperación.

- **Accesibilidad/Disponibilidad:** Acceso a los datos desde cualquier punto de la red en cualquier momento. Acceso desde diferentes contextos, aplicaciones y medios, volviéndolos útiles a un mayor número de usuarios.
- **Integridad:** Modelo de datos referencial, control de integridad, trazabilidad y coherencia de datos.
- **Seguridad:** Accesos controlados por el SGDB.
- **Recuperación:** Copias de seguridad diferenciales automatizadas. Recuperación de datos.

Dentro del esquema cliente/servidor el SW del segmento terreno aplica sólo dos capas, donde las aplicaciones se conectan directamente al RDBMS con la única excepción de los servicios de descarga de datos ciencia, los cuales pueden ser utilizados por otras aplicaciones derivando en un modelo de 3 capas. Esto último excede el alcance del proyecto en su fase académica.

B. Software del segmento de Vuelo

El SW de segmento de vuelo [17] se encarga de la recolección de HouseKeeping (HK) de los distintos subsistemas que componen al segmento de vuelo y su transmisión al segmento terreno para el monitoreo del estado de salud del satélite. Por otro lado, se encarga de la recepción, procesamiento y ejecución de tele-comandos provenientes del segmento terreno, que tienen la finalidad de operar el satélite. Este SW se ejecuta sobre un sistema de tiempo real (FreeRTOS, ANSI C), debido al estricto control y ejecución en tiempo y forma de sus rutinas.

La figura 4 presenta el diagrama de arquitectura del SW del segmento de vuelo. En ella pueden apreciarse sus principales sistemas y bloques internos: el Multi-Mission Platform (MMP) y el Payload, detallados a continuación.

- **MMP:** Attitude Determination Control System (ADCS), Command and Data Handling (C&DH), Telemetry, Telecommand and Control (TT&C) y Electric Power Subsystem (EPS).
- **Payload:** Cámara, Payload Onboard Computer (POBC) y Data Collector System (DCS).

Se considera que los módulos de SW que tendrán modificaciones sobre su SW por defecto se corresponden a los subsistemas C&DH y POBC. Dentro del SW de C&DH se realizan la mayoría de las actividades de control y operación del satélite. Las principales funciones son:

- **Procesamiento y ejecución de telecomandos:** de TC provenientes del segmento terreno.
- **Procesamiento de HK:** recolección de HK de los diferentes subsistemas.
- **Procesamiento de Telemetría:** construcción de frames de TM para enviar al segmento terreno.
- **FDIR:** detección, aislamiento y recuperación ante fallas.

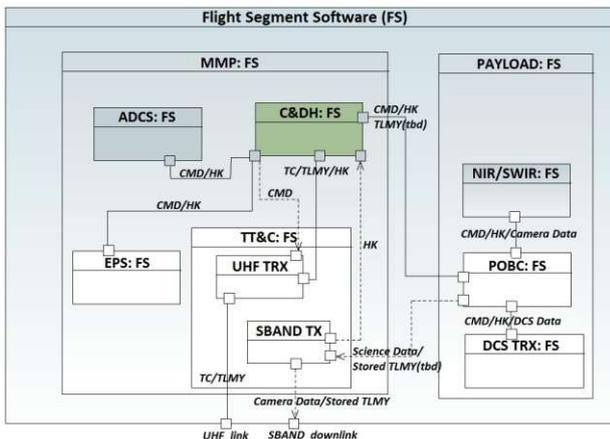


Figura 4: Arquitectura del SW del segmento de Vuelo

El SW de Payload es el responsable de la gestión de instrumentos y de la adquisición de datos de ciencia. Sus principales funciones son:

- **Gestor de comandos:** recibe y ejecuta los comandos provenientes de C&DH.
- **Gestor de Cámara NIR-SWIR:** gestiona las adquisiciones de imágenes en las áreas de interés.
- **Gestor de DCS:** gestiona las adquisiciones del DCS desde las DCP (Data Collection Platform).

3. RESULTADOS OBTENIDOS

La misión FS2017 ha permitido a los alumnos de las cuatro maestrías de la UFS abordar el desarrollo de un sistema satelital complejo con componentes de vuelo y de tierra. Esto plantea un desafío a nivel de desarrollo de SW cuya ejecución ha requerido una organización, una metodología y una aproximación técnica específica para el área espacial [2][3][4]. En relación a los hitos programáticos, se ha superado el PMSR (Preliminary Mission & System Review) y se están desarrollando las actividades en miras al hito de fin de la fase B: el PDR (Preliminary Design Review).

Respecto a la ingeniería de SW del proyecto, al día de hoy se ha implementado exitosamente

una versión inicial del SW del Segmento Terreno multimisión incluyendo módulos de bajada, decodificación, almacenamiento y visualización de telemetría (histórica y tiempo real). Se han desarrollado también productos soporte de dinámica orbital y publicación de datos ciencia. En todos los casos, se han utilizado técnicas y herramientas de desarrollo de uso extendido y propósito general. En relación al SW del segmento de vuelo, se ha realizado un diseño detallado de la arquitectura del sistema, contemplando sus funciones y requerimientos, y se han realizado pruebas iniciales y actividades de preparación de entorno. La figura 5 muestra el CubeSat en el laboratorio de la UFS.



Figura 5: Satélite FS2017 en el laboratorio de la UFS

Hasta el momento, el proyecto ha cumplido los objetivos, lo que permite concluir que una misión como la FS2017 es adecuada para la capacitación en ciencia y tecnología espacial.

4. FORMACIÓN DE RECURSOS HUMANOS

A. Organización del Equipo de Software

La organización del equipo de SW [12] se detalla en la figura 6. En ella se observan las dos áreas principales: (1) SW del segmento de vuelo, (2) SW del segmento terreno.

El equipo de segmento de vuelo está subdividido en 3 áreas de trabajo: (i) CD&H SW, (ii) ADCS SW, (iii) POBC SW. El equipo de segmento terreno está subdividido en 4 áreas de trabajo: (i) Mission Unit SW, (ii) Control Unit SW, (iii) Orbit Dynamics SW, (iv) Ground Station SW. Ambos equipos tienen bajo su responsabilidad el análisis, diseño, desarrollo, testing unitarios y despliegue en producción del SW del segmento correspondiente. Además, se cuenta con tres áreas transversales: (a) Integration & Testing, (b) Models & Simulations, y (c) Assurance & Configuration Management.

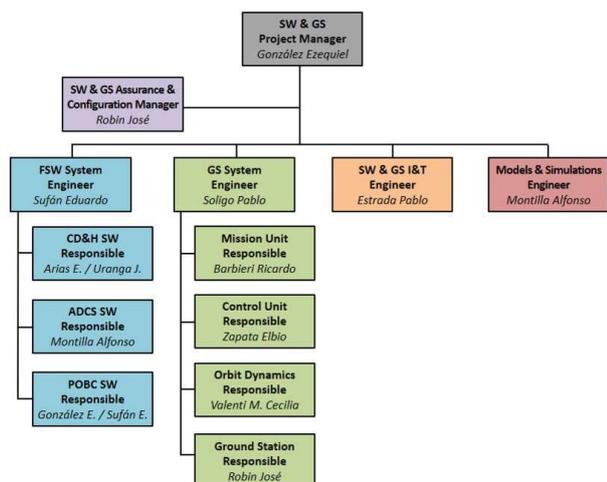


Figura 6: Organigrama del Equipo de Software

La figura 7 presenta al equipo de SW del FS2017 en las instalaciones de la UFS, ubicada en el Centro Espacial Teófilo Tabanera, Córdoba.



Figura 7: Equipo de SW del FS2017 en el CETT

B. Tesis de Maestría del Equipo (en curso)

Las tesis de los maestrandos de la MDIAE son:

- Arias, E. “Diseño de una Arquitectura de Aviónica tolerante a fallas basada en componentes COTS para vehículos satelitales de nueva generación”.
- Barbieri, R. “Sistema de Soporte de decisión espacial de Leishmaniasis operacional en Argentina basado en tecnología geoespacial”.
- Estrada, P. “Sistema Automatizado para la Colocalización de datos intersatelitales y mediciones de campo”.
- González, E. “Diseño de una Arquitectura Satelital Segmentada basada en Sistemas Multi-agente para la Gestión de Emergencias”.
- Montilla, A. “Uso de técnicas de Computer Vision para la detección en vuelo de focos de calor en imágenes satelitales”.
- Robin, J. “Diseño de una Arquitectura de Software orientada al desarrollo de un sistema informático para el control de una Antena Satelital”.
- Soligo, P. “Análisis y Simulación de Redes DTN aplicadas a Constelaciones Satelitales”.
- Sufán, E. “Diseño, Implementación y Verificación de un Procesador L1 Time Domain Back Projection para Imágenes de SAR Aerotransportado”.
- Uranga, J. “Implementación Paralela en GP-GPU Portable del Algoritmo Omega-K para enfoque SAR”.
- Valenti, M. C. “ARxCODE, Prototipo de Software para el análisis de riesgo por Colisión con Desechos Espaciales”.
- Zapata, E. “Detección de Patrones en Micro-Nano Satélites utilizando metodologías de Visión Artificial”.

5. BIBLIOGRAFÍA

- [1] CONAE (2010). *Plan Espacial Nacional Argentina en el espacio 2004-2015*.
- [2] CONAE (2016). *FS2017 Project Implementation Plan (FS-410000-PL-00100-A)*.
- [3] ECSS (2009). *Project Planning and Implementation (ECSS-M-ST-10C Rev.1)*.
- [4] ECSS (2017). *System Engineering General Requirements (ECSS-E-ST-10C Rev.1)*.
- [5] The CubeSat Program, Cal Poly SLO (2014). *CubeSat Design Specification (CDS) Rev. 13*.
- [6] CONAE (2016). *FS2017 Parámetros Relevantes de la Cámara NIR-SWIR (SA-150100-IA-00100-A)*.
- [7] CONAE (27 de marzo de 2017). Misión SABIA-Mar. Recuperado de <http://www.conae.gov.ar/index.php/espanol/misiones-satelitales/sabiamar/objetivos>.
- [8] CONAE (2016). *FS2017 DCS - Selección de Dispositivos*.
- [9] Alexander, L. & Davis, A. (1991). Criteria for Selecting Software Process Models. En *Proc. 15th COMPSAC*, pp. 521-528.
- [10] Schwaber, K. (2004). *Agile Project Management with Scrum*. Redmond, WA, U.S.: Microsoft Press.
- [11] Trimble, J. & Webster, C. (2012). Agile development methods for space operations. En *SpaceOps 2012 Conf.*
- [12] CONAE (2016). *FS2017 Software Project Management Plan (FS-210100-GS-00100-A)*.
- [13] ECSS (2009). *Software (ECSS-E-ST-40C)*.
- [14] ECSS (2013). *Software Engineering Handbook (ECSS-E-HB-40A)*.
- [15] ECSS (2008). *Ground Systems and Operations (ECSS-E-ST-70C)*.
- [16] ECSS (2009). *Software Product Assurance (ECSS-Q-ST-80C)*.
- [17] CONAE (2016). *FS2017 Software Concept Document (FS-230000-SP-00100-A)*.
- [18] Django Software Foundation (2017). *Django Documentation (Release 1.10.7.dev.20170317114851)*. Recuperado de <https://docs.djangoproject.com/en/1.10>.

Procesamiento y Análisis de Datos Espaciales y Temporales Relativos a Espacios Urbanos

Romina Stickar¹³, Damián Barry¹³, Rodrigo René Cura¹³, Leonardo Ordinez¹, Carlos Buckle¹³, Claudio Delrieux¹²³

¹Depto. de Informática, Fac. de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB).

Puerto Madryn, Argentina.
+54 280-4472885 – Int. 117

²Depto. de Ingeniería Eléctrica y Computadoras - Universidad Nacional del Sur (UNS)

Bahia Blanca, Argentina
+54 291-4595000 - Int. 3371.

³LINVI – Laboratorio de Investigación en Informática – Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB)

romistickar@gmail.com, damián_barry@unpata.edu.ar, rodrigo.renecura@gmail.com, leo.ordinez@gmail.com, carlos.buckle@gmail.com, cad@uns.edu.ar

Resumen

Actualmente, muchos de los gobiernos locales no cuentan con información de calidad, ni mecanismos de adquisición continua de datos que permitan un monitoreo de la dinámica global de la ciudad, como herramienta de apoyo para la toma de decisiones y la definición de políticas públicas. Este proyecto apunta a estudiar, experimentar y optimizar infraestructuras de Big Data enfocadas en información heterogénea fuertemente ligada a datos espaciales y temporales. Nuestro estudio, se concentra en diseñar un entorno que permita gestionar la recolección de datos urbanos desde diferentes fuentes (registros de actividad de teléfonos móviles, redes sociales, aplicaciones móviles de turismo, tarjeta SUBE, sensores urbanos, etc) para realizar análisis de datos geo- temporales, construir modelos predictivos y generar servicios para la visualización y síntesis de la información resultante. Los resultados serán de directa aplicación tanto en organismos

públicos como privados, para objetivos como por ejemplo: planificación del transporte público, cartografía social, acceso a centros de salud y educación, turismo, etc.

Palabras clave: big data, datos espaciales, datos temporales.

Contexto

Este proyecto se desarrollará en el LINVI (Laboratorio de Investigación en Informática) de la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB) y está integrado por docentes investigadores de la UNPSJB Sede Puerto Madryn y de la Universidad Nacional del Sur (UNS). El proyecto fue avalado por el Consejo Directivo de la Facultad de Ingeniería de la UNPSJB. Será financiado por la Secretaría de Ciencia y Técnica de la UNPSJB para llevar a cabo durante el período 2017 - 2019.

1. Introducción

La Región Patagónica, en general, y la Provincia del Chubut, en particular, tienen la característica de tener una baja densidad poblacional. La mayor parte de ésta se concentra en aislados centros urbanos, dejando extensas áreas geográficas con escasos habitantes (en algunos casos, menor a uno) por km². Asimismo, la región tiene otras características geográficas interesantes, como son grandes extensiones de bosque nativo y de mar, reservas naturales y áreas protegidas con fauna y flora única, vías de comunicación terrestre de cientos de km, centros urbanos en crecimiento de manera no planificada, actividad económica principalmente primaria y extensiva, entre otras.

En este contexto, la necesidad de gestionar información impone que los datos que se manipulen posean además una referencia espacial y temporal respecto a su generación. De esta manera, cada dato manipulado, no solo modela parcialmente un dominio de aplicación particular, sino que lo hace en un tiempo y en un espacio determinado. Esto complejiza el modelado de los sistemas en tanto agrega nuevas dimensiones al problema.

Este nuevo dominio de aplicación que ahora incorpora las dimensiones temporal y espacial, requiere un tratamiento especial. Aquí se involucran diferentes áreas de conocimiento en variadas etapas de madurez. Desde el punto de vista de las comunicaciones y las redes móviles, se presenta una problemática concreta, la cual es la baja (e incluso inexistente) conectividad en grandes extensiones de terreno. Sobre esta temática se han desarrollado diferentes propuestas que tienen que ver con redes oportunistas [1, 2, 8] y esquemas de manejo de micromensajes orientados a redes de sensores [3]. Por el lado de la captura de datos específicamente se pueden mencionar los aportes de [4], referido a combate de fuego en

bosques, [5, 6], referidos a manejo de situaciones de emergencia, entre otros.

En el plano de las tecnologías de la información y la comunicación (TIC), un sistema de información efectivo es el que provee a sus usuarios con la información necesaria para responder a los requerimientos medioambientales [7]. En sentido amplio, el campo de la computación colaborativa, abarca el uso de computadoras para apoyar la coordinación y la cooperación de dos o más personas que tratan de realizar una tarea o resolver un problema juntos [9, 10, 11]. En particular, la problemática de relevar información, observar y analizar datos e información del ambiente físico, mediante sistemas colaborativos, que involucran participación ciudadana, es relativamente nueva [12]. A partir de las últimas dos décadas, los avances en las TIC han creado capacidades accesibles para integrar y coordinar información de distintos actores y construir conocimiento a partir de ella [13]. En este punto es fundamental contar con lenguajes y modelos que permitan describir los dominios de aplicación en un contexto espacio-temporal [14, 16, 17, 18]. Asimismo, resulta determinante que dichos modelos permitan elaborar inferencias no triviales mediante técnicas de aprendizaje automatizado.

Por otro lado, un aspecto necesario en el análisis del dominio de aplicación planteado tiene que ver con la realización de simulaciones. Para ello, se explorarán en particular las técnicas basadas en Autómatas Celulares [19, 20, 21, 22] y Agentes [23, 24, 25].

Otro aspecto importante en evaluación del trabajo y su análisis, está relacionado con la visualización de la información. En particular, la visualización de grandes volúmenes de datos en el contexto espacio-temporal [15].

La motivación principal de la proyecto se asienta en el área de aplicación de los sistemas a desarrollar, los cuales tienen que ver con

situaciones complejas, multicausales, posiblemente extraordinarias y que involucran la interacción de diferentes actores, donde la dinámica e interpretación de sucesos ocurridos en espacios geográfico-temporales, hacen que tanto la recolección de los datos como su transformación, análisis, interpretación y explotación de los mismos sea un desafío.

El objetivo general de la presente investigación es el estudio y desarrollo de una infraestructura común a nivel municipal/provincial para la coexistencia y la colaboración de múltiples sistemas orientados al procesamiento y análisis de datos espaciales y temporales relativos a entornos urbanos.

2. Líneas de Investigación, Desarrollo e Innovación

Las líneas de investigación estarán orientadas a:

Estudiar, experimentar y optimizar infraestructuras de Big Data enfocada en información heterogénea, pero fuertemente ligada a datos espaciales y temporales.

Aplicar técnicas de adquisición, almacenamiento y distribución de información en ambientes heterogéneos. Además de poder contar con una infraestructura capaz de procesar grandes volúmenes de datos.

Diseñar y adaptar técnicas de normalización y homogeneización de datos, mediante un vocabulario común, que permita la manipulación de manera agnóstica de los datos propios del dominio de aplicación contextualizados en espacio y tiempo.

3. Resultados y Objetivos

Actualmente el proyecto se encuentra en una etapa inicial de revisión y estudio de los antecedentes.

Los integrantes del equipo ya nos encontramos realizando las capacitaciones necesarias para el entendimiento profundo de

los estudios y el avance hacia las etapas posteriores.

Los objetivos de la investigación y desarrollo de este trabajo se pueden resumir en:

- Investigar, diseñar, montar y desplegar un entorno de Software y Hardware que permita tanto gestionar la recolección de datos urbanos espaciales y temporales, como su posterior explotación mediante herramientas de visualización, procesos de análisis y correlación de datos, servicios de publicación, etc.
- Estudiar, proponer, definir e implementar protocolos y mecanismos estandarizados de recolección y explotación de datos para facilitar el acceso a la infraestructura y fomentar un ambiente open-data alrededor de la misma.
- Promover la adopción de la tecnología desarrollada, para su implementación en la gestión pública y privada, mediante convenios de transferencia tecnológica.
- Dar continuidad a las actividades realizadas por el Laboratorio de Investigación en Informática (LINVI) perteneciente al Departamento de Informática de la UNPSJB, así como incrementar el número de proyectos acreditados y de trabajos publicados por este.

Los resultados en esta etapa incipiente del proyecto se encuentran orientados al dominio de aplicación del transporte y la movilidad urbana. En particular, se está trabajando para modelar y optimizar el flujo de las líneas de colectivo en ciudades tamaño medio-bajo (menores a 300.000 habitantes). A la vez, se está desarrollando un trabajo para la implementación de ciclovías en dichas ciudades, que involucra la recolección de datos de recorridos de ciclistas y la integración de los mismos con otras fuentes de datos para su análisis.

4. Formación de Recursos Humanos

Del proyecto forman parte docentes del Departamento de Informática de la Facultad de Ingeniería de la UNPSJB Sede Puerto Madryn y de la UNS. Dos de ellos están realizando sus doctorados y otros dos comenzarán con su carrera de posgrado. Uno de los doctorandos es beneficiario de beca CONICET.

También forman parte del proyecto graduados de la carrera de Licenciatura en Informática y alumnos del ciclo superior. Su participación en el mismo les permitirá introducirse en temas de investigación en una disciplina que mundialmente se encuentra en auge y que se utiliza diariamente. Además, podrán participar y exponer sus ideas en un grupo de trabajo, realizar investigaciones, diseñar sus propias soluciones, contrastarlas con las de sus pares y las puedan validar con la realidad.

Es intención del proyecto, que los alumnos próximos a graduarse puedan elaborar sus tesis sobre la disciplina y sus ramas, logrando un aporte extra a la investigación.

5. Bibliografía

- [1] Vinícius F.S. Mota, Felipe D. Cunha, Daniel F. Macedo, José M.S. Nogueira, Antonio A.F. Loureiro, Protocols, mobility models and tools in opportunistic networks: A survey, *Computer Communications*, Volume 48, 15 July 2014, Pages 5-19, ISSN 0140-3664.
- [2] Chilipirea, C., Petre, A.-C., Dobre, C., Pop, F., and Suciú, G. (2017) A simulator for opportunistic networks. *Concurrency Computat.: Pract. Exper.*, 29: e3814
- [3] Message Queue Telemetry Transport, <http://mqtt.org/>
- [4] "Emerging technologies: Smarter ways to fight wildlife crime" - *Emerging technologies / Environmental Development* 12 (2014) 62 – 72
- [5] Rajiv Ranjan, Samee Khan, Joanna Kolodziej, Albert Y. Zomaya "Guest Editors' Introduction: Cloud-Based Smart Evacuation Systems for Emergency Management", *IEEE Cloud Computing*, Volume 1, Number 4, November 2014
- [6] Karan Mitra, Saguna, Christer Åhlund, "A Mobile Cloud Computing System for Emergency Management", *IEEE Cloud Computing*, Volume 1, Number 4, November 2014
- [7] El-Bibany, H. E. D., Katz, G., & Vij, S. (1991). *Collaborative Information Systems* (No. 48). Stanford University, Center for Integrated Facility Engineering.
- [8] Sergio F. Ochoa; Rodrigo Santos, Human-centric Wireless Sensor Networks to Improve Information Availability During Urban Search and Rescue Activities, *INFORMATION FUSION*; Lugar: Amsterdam; Año: 2015 vol. 22 p. 71 – 84.
- [9] Borenstein, N. S. (1992, December). Computational mail as network infrastructure for computer-supported cooperative work. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work* (pp. 67-74). ACM.
- [10] Schooler, E. M. (1996). Conferencing and collaborative computing. *Multimedia Systems*, 4(5), 210-225.
- [11] Bafoutsou, G., & Mentzas, G. (2002). Review and functional classification of collaborative systems. *International journal of information management*, 22(4), 281-305.
- [12] Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J., & Parrish, J. K. (2014). Next steps for citizen science. *Science*, 343(6178), 1436-1437.
- [13] Newman, J., & Thomas-Alyea, K. E. (2012). *Electrochemical systems*. John Wiley & Sons.
- [14] Haining, R. P. (2012), *Statistics for Spatio-Temporal Data* by Noel Cressie and Christopher K. Wikle. Hoboken, NJ: Wiley, 2011, 624 pp.. *Geogr Anal*, 44: 411–412.
- [15] A. Malik, R. Maciejewski, E. Hodgess and D. S. Ebert, "Describing Temporal Correlation Spatially in a Visual Analytics Environment," *2011 44th Hawaii International Conference on System Sciences*, Kauai, HI, 2011, pp. 1-8.
- [16] Martin Erwig, Ralf Hartmut Güting, Markus Schneider, Michalis Vazirgiannis."Spatio-Temporal Data Types: An Approach to Modeling and

Querying Moving Objects in Databases”
GeoInformatica (1999) 3: 269.

[17] Brodeur, J., Y. Bédard & M.-J. Proulx, 2000, Modelling Geospatial Application Database using UML-based Repositories Aligned with International Standards in Geomatics, ACMGIS 2000, November 10-11, Washington DC, United-States

[18] Mehrdad Salehi, Yvan Bédard, Mir Abolfazl Mostafavi, Jean Brodeur, Formal classification of integrity constraints in spatiotemporal database applications, *Journal of Visual Languages & Computing*, Volume 22, Issue 5, October 2011, Pages 323-339, ISSN 1045-926X

[19] Sven Maerivoet, Bart De Moor. Cellular Automata Models of Road Traffic. *Physics Reports*, vol. 419, nr. 1, pages 1-64, november 2005

[20] K. Nagel, M. Schreckenberg. A cellular automaton model for freeway traffic. 1992. *Journal of Physics I* 2, 2221.

[21] Inés Santé, Andrés M. García, David Miranda, Rafael Crecente. Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning* 96 (2010) 108–122.

[22] Wolfram, Stephen. Twenty Problems in the Theory of Cellular Automata. *Physica Scripta*, Volume 9, Issue, pp. 170-183 (1985).

[23] Keith, Clarke. Cellular Automata and Agent-Based Models. M.M. Fischer, P. Nijkamp (eds.), *Handbook of Regional Science* (2014).

[24] Daniela Fecht, Linda Beale, David Briggs. A GIS-based urban simulation model for environmental health analysis. *Environmental Modelling & Software* 58 (2014) 1-11.

[25] Harvey J. Miller. Collaborative mobility: using geographic information science to cultivate cooperative transportation systems. *Procedia Social and Behavioral Sciences* 21 (2011) 24–28

Raspberry Pi como Servidor Portátil de Contenidos para Ser Consumidos desde Dispositivos Móviles

Daniel A Giulianelli, Rocío A Rodríguez, Pablo M Vera, Víctor M Fernández, Claudia G Alderete

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
 Departamento de Ingeniería e Investigaciones Tecnológicas
 Universidad Nacional de La Matanza
 Florencio Varela 1903, San Justo, Provincia de Buenos Aires
 {dgiulian, rocio.rodriguez, pvera, vfernandez, calderete} @unlam.edu.ar

RESUMEN

Los alumnos “nativos digitales” suelen traer al aula teléfonos celulares y otros dispositivos como tablets o lectores de libros. Estos dispositivos en su mayoría cuentan con conectividad wifi y acceso a internet por lo que pueden utilizarse para acceder a recursos para trabajar en clase. Sin embargo, no siempre se dispone de una conexión a internet o en algunos casos la señal no es estable y no permite acceder a recursos con la fluidez suficiente. Una posible solución es crear una red dentro del aula donde el docente pueda llevar consigo un servidor de contenidos portátil y los alumnos puedan acceder al mismo utilizando sus propios dispositivos.

En este proyecto se generará una solución de bajo costo, implementando como servidor una Raspberry Pi, sobre la cual se instalarán programas específicos planificados para ser utilizados en el aula. Este servidor permitirá consumir los contenidos accediendo por medio de dispositivos móviles utilizando conexión wifi a una red generada desde el propio servidor. En el prototipo a realizar, se conectará una pantalla táctil de tamaño reducido lo que permitirá la portabilidad de la solución, en donde el docente podrá visualizar una síntesis de los resultados propuestos por los alumnos a los problemas planteados, así como su respuesta a preguntas específicas.

Palabras clave: Solución Portable, Raspberry, Servidor de Contenidos, Materiales Didácticos

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al programa PROINCE y es realizado por el grupo de investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza (UNLaM). El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D alumnos de la universidad de las carreras de Ingeniería.

Este proyecto es continuación de un proyecto previo en el que se ha trabajado con Raspberrys para utilizarlas como nodos de una red para Video-Conferencias.

1. INTRODUCCIÓN

La Raspberry Pi (RPi) es un SBC (Single Board Computer) que tiene el tamaño de una tarjeta de crédito, desarrollada en Reino Unido por la fundación Raspberry Pi.

La tabla 1, muestra las prestaciones en cuanto a hardware de los distintos modelos.

Tabla 1. Características de las Raspberry Pi

	Raspberry Pi Modelos				
	1 A	1 B	1 B+	2 B	3 B
RAM	256 MB	512 MB		1 GB	
Almacenamiento	SD			Micro SD	
Procesador	ARM11			ARM Cortex - A7	ARM v8
Velocidad	700 MHz			900 MHz	1,2 GHz
Conectividad de Red	Ninguna	Ethernet			Ethernet, WiFi, Bluetooth
Nº puertos USB	1	2	4		
Cantidad de GPIO	26			40	
Alimentación	5v				
Tamaño	85 x 56 x 17 mm				

El modelo más reciente RPi 3 incorpora wifi, a los modelos previos era necesario agregarle por USB una pequeña placa externa.

Por medio de una tarjeta de memoria se le puede instalar un sistema operativo. Existe una distribución de Linux particular para la RPi denominada Raspbian, pero también se pueden instalar otros sistemas operativos existiendo por ejemplo una versión específica de Windows 10 para este dispositivo. En cuanto al uso que se le puede dar, existen diversos trabajos realizados en los últimos años, entre ellos se pueden mencionar aplicaciones dedicadas a:

- Domótica [1], [2], [3], [4];
- Monitoreo para Seguridad [5], [6], [7], [8];
- Simulación de consolas de juegos [9], [10];
- Estación meteorológica [11], [12], [13];
- Agro (Riego Automático [14], Secadora de Granos de Café [15]);

Estos trabajos mencionados previamente muestran el gran abanico de acción utilizando la Raspberry Pi. En el presente proyecto se plantea utilizar la RPi como servidor de contenidos didácticos, siendo los alumnos a través de sus dispositivos móviles los clientes de esa red.

“El auge en las telecomunicaciones inalámbricas de las últimas décadas unidos a la explotación de la movilidad que ofrecen los dispositivos portátiles y que requieren los

usuarios de hoy día, ha permitido que se presente un nuevo escenario de operación” [16].

Al generar un punto de acceso portátil mediante la raspberry pi será posible desplegar contenidos aun cuando no se cuente con una infraestructura de red existente. En la figura 1 se presenta gráficamente la solución a construir.

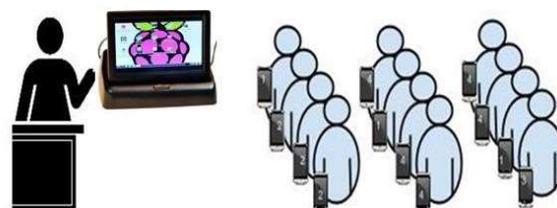


Figura 1. Solución a Construir –RPi (Servidor) y Dispositivos Móviles (Clientes)

La solución a generar podrá ser utilizada en ámbitos donde no se cuenta con conectividad wifi, sacando provecho de los dispositivos móviles con los que se cuentan para consumir recursos provistos en un servidor de contenidos implementado en una Raspberry pi. Para esto, será necesario analizar algunas cuestiones en cuanto a Hardware que permitan asegurar la portabilidad de la solución, incluso la posibilidad de conectar una batería que alimente la solución pensada para entornos donde no hay enchufes disponibles en forma cercana y cómoda para que el docente pueda utilizar la solución. En cuanto a software es necesario estudiar algunas cuestiones de seguridad, por ejemplo: que permitan que desde un mismo dispositivo

no se lance más de una respuesta a un problema planteado. Si el acceso a la red será libre o si desde cada dispositivo a conectarse deberá ingresarse una clave específica para tener acceso a la red. Por supuesto deberá tomarse en cuenta algunas cuestiones de accesibilidad y compatibilidad de la solución con los dispositivos móviles.

Por otra parte, para que la solución pueda implementarse de forma exitosa será necesario pensar que contenidos, problemas, ejercicios, preguntas; pueden ser presentados y cuál es la mejor manera para que los mismos puedan ser exhibidos. En el equipo de trabajo hay docentes vinculados con diversas cátedras que consideran que podrían hacer uso de este tipo de aplicaciones. Incluso estando en el equipo de investigación jefes de cátedras entusiasmados y dispuestos a utilizar esta futura solución. Lo expuesto motiva a encarar su implementación real en el ámbito universitario.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Dispositivos Móviles
- Redes AdHoc
- Software a Medida
- Diseño de Interfaz táctil
- Pruebas de acceso desde Dispositivos Móviles
- Presentación de Resultados

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto ha iniciado en el presente año por lo cual no hay resultados aún para reportar. Se espera poder tener una solución de hardware con software específico que permita ofrecer contenidos que puedan ser accedidos desde dispositivos móviles, aún en zonas donde no se cuente con conectividad.

Siendo la Raspberry Pi una alternativa portátil para conseguir este objetivo.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por:

- 10 Docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 4 alumnos de ingeniería quienes se están formando en actividades de Investigación y Desarrollo los cuales cuentan con becas asignadas

Vinculado con esta temática se encuentra en realización una tesis de maestría.

5. BIBLIOGRAFIA

- [1] Jain, S., Vaibhav, A. and Goyal, L., 2014, February. Raspberry Pi based interactive home automation system through E-mail. In Optimization, Reliability, and Information Technology (ICROIT), 2014 International Conference on (pp. 277-280). IEEE.
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6798330&isnumber=6798279>
- [2] Haro, L.F.D., Cordoba, R., Rivero, J.I.R., de la Fuente, J.D., Peces, D.A. and Mera, J.M.B., 2014. Low-Cost Speaker and Language Recognition Systems Running on a Raspberry Pi. IEEE Latin America Transactions, 12(4), pp.755-763.
- [3] Öztürk, T., Albayrak, Y. and Polat, Ö., 2015, May. Object tracking by PI control and image processing on embedded systems. In 2015 23rd Signal Processing and Communications Applications Conference (SIU) (pp. 2178-2181). IEEE.
http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7130305
- [4] Zhao, C.W., Jegatheesan, J. and Loon, S.C., 2015. Exploring IOT Application Using Raspberry Pi. International Journal of Computer Networks and Applications, 2(1), pp.27-34.

<http://www.ijcna.org/Manuscripts%5CVolume-2%5CIssue-1%5CVol-2-issue-1-M-04.pdf>

- [5] Ibrahim, M., Elgamri, A., Babiker, S. and Mohamed, A., 2015, October. Internet of things based smart environmental monitoring using the Raspberry-Pi computer. In Digital Information Processing and Communications (ICDIPC), 2015 Fifth International Conference on (pp. 159-164). IEEE.
- [6] Ahmad, T., Studiawan, H. and Ramadhan, T.T., 2014. Developing a Raspberry Pi-based Monitoring System for Detecting and Securing an Object. <http://kbj.if.its.ac.id/wp-content/uploads/2015/03/2014-IES-PENS.pdf>
- [7] Snyder, R.M., 2014. Power monitoring using the Raspberry Pi. Association Supporting Computer Users in Education "Our Second Quarter Century of Resource Sharing", p.82.
- [8] Yaldaie, A., 2016. Home automation and security system with the Raspberry Pi. <http://www.theseus.fi/handle/10024/106677>
- [9] Richardson, M. and Wallace, S., 2012. Getting started with raspberry PI. "O'Reilly Media, Inc."
- [10] Hussain, S.R., Naidu, K.R., Lokesh, C.R., Vamsikrishna, P. and Rohan, G., 2016, February. 2D-game development using Raspberry Pi. In 2016 International Conference on Information Communication and Embedded Systems (ICICES) (pp. 1-8). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7518858
- [11] Salcedo Tovar, M.L. and Cendrós, J., 2016. Uso del minicomputador de bajo costo "Raspberry Pi" en estaciones meteorológicas. *Télématique*, 15(1), pp.62-84.
- [12] Anchundia, R. and Michael, J., 2015. Telemetría de Estación Meteorológica. <http://www.dspace.espol.edu.ec/xmlui/handle/123456789/31545>
- [13] Malfatti, G.M., Pavan, W. and Cunha, J.M., 2014. Rede de estações meteorológicas automáticas sem fio. <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/133214/1/ID-43372-2014-TCC-MiniColetorDados.pdf>
- [14] Escalas Rodríguez, G., 2015. Diseño y desarrollo de un prototipo de riego automático controlado con Raspberry Pi y Arduino. <http://upcommons.upc.edu/handle/2099.1/25074>
- [15] Vaccaro Acosta, F.E., Soriano, A. and Alberto, J., 2015. Telecontrol de secadora de granos de cacao con energía solar y biogas. <http://www.dspace.espol.edu.ec/xmlui/handle/123456789/30149>
- [16] Calderón, O. J., & QUINTERO, V. M. (2004). Un nuevo aspecto de la movilidad: Redes Ad Hoc-Conceptos. *Revista Colombiana de Tecnologías de Avanzada*, 1(3), 59-64. http://201.245.175.152/unipamplona/portaIG/home_40/recursos/01_general/revista_3/13102011/08.pdf

Realidad Virtual y Adiestramiento en Sistemas Críticos

Oscar Martín Bianchi ^(a,b)

oscarmartinbianchi@gmail.com

Javier García Polak ^(a,b)

javi.polak@gmail.com

Ignacio Arrascaeta ^(a)

ignacioarrascaeta@gmail.com

German Luis Vila Krause ^(a)

g.vilakrause@gmail.com

^(a)CIDESO⁰, DIGID¹- Ejército Argentino

^(b)EST², IESE³ - Ejército Argentino

RESUMEN

La complejidad de los ejercicios de adiestramiento en un ámbito tan exigente como el de la defensa – tanto en medidas de seguridad, como en los costos asociados a una operación de instrucción – convierte a estas prácticas en una actividad crítica.

Los costos asociados a dichas actividades, aumentan a medida que se asciende en el grado de exposición, del personal al que se desea dar instrucción, a procedimientos cercanos al uso de material bélico o sistemas de armas.

A través de la inclusión de tecnologías de simulación y virtualización se busca reducir dichos factores. Utilizándolas para dar contexto y aportar valor agregado, sumando una cuota de realismo a bajo costo y riesgo controlado – tanto humano como material –, para dar apoyo a las operaciones de adiestramiento.

Por eso, desde el proyecto SATAC se propuso la utilización de tecnologías de Realidad Virtual, considerando sus ventajas. El fin de esta integración será aportar realismo y complejidad al proyecto sin la necesidad de realizar despliegues adicionales de material y con riesgos mínimos para los participantes, potenciando a las operaciones convencionales con simulaciones vivas, virtuales y constructivas (Live, virtual, constructive – LVC–) (1).

Palabras Clave: *Realidad Virtual, Adiestramiento, Simulación, Simulación Constructiva, Entidades Sintéticas, Ingeniería del Software, LVC.*

CONTEXTO

El Ejército Argentino, a través del CIDESO, lleva adelante el desarrollo de Sistemas de Comando y Control (C2) para el apoyo a la toma de decisiones y la dirección de las operaciones, y posee amplia experiencia en lo que respecta a sistemas de simulación para aplicaciones militares, tanto para problemas militares operativos – Batalla Virtual (BV) – como para operaciones militares de paz – Simupaz –.

Conceptualmente, BV muestra de manera simulada una visión análoga a lo que debería ser un sistema de C2, donde los distintos niveles decisores poseen un tablero de control para visualizar la situación de los elementos que comandan de manera gráfica.

Por otro lado, el Ejército Argentino lleva adelante el desarrollo de un Sistema Automatizado para el Tiro de Artillería de Campaña (SATAC), el cual tiene por objetivo gestionar la red de Apoyo de Fuego de una Unidad de Combate, interactuando con los sistemas de mayor nivel para su gestión a nivel Brigada^a.

En este contexto, la posibilidad de sumar elementos del mundo virtual a un sistema de C2 es invaluable, ya que no solo complementa

^a Conjunto de dos o más Unidades de Combate.

¹ CIDESO: Centro de Investigación y Desarrollo de Software

² DIGID: Dirección General de Investigación y Desarrollo

³ EST: Escuela Superior Técnica - Facultad de Ingeniería del Ejército Argentino

⁴ IESE: Instituto de Enseñanza Superior del Ejército - Universidad del Ejército Argentino

el adiestramiento volviéndolo una experiencia mucho más realista y constructiva, sino que además, puede utilizarse para evaluar las posibles reacciones de los elementos reales, implementando modelos de simulación concebidos para emular el comportamiento de entidades virtuales o sintéticas.

Teniendo en mente la existencia de una gran cantidad de sistemas legados (legacy) y la flexibilidad de un sistema de C2 como SATAC, se propone la integración a partir de un modelo de tres componentes complementarios: 1) Batalla Virtual, que aporta el contexto de simulación sobre el que se llevaría a cabo la instrucción, 2) Una herramienta de simulación (actualmente en desarrollo en la plataforma Unity), a cargo de la generación de entidades sintéticas (2), basada en la información aportada por los modelos de simulación de BV y la aportada como información real al tercer sistema 3) SATAC, que utilizará, a efectos prácticos, las entidades generadas por dicha herramienta asumiendo los datos de la misma como provenientes de agentes reales del sistema.

Así mismo, se investigarán y desarrollarán las herramientas necesarias para indicar el modo en que SATAC utiliza las entidades del mundo real para poder generar en el mundo virtual análogas sintéticas (agencias de inteligencia, operadores de radares, entre otros) y así poder integrar simulación y realidad de manera simultánea en el mismo escenario.

1. INTRODUCCIÓN

Ambos sistemas por separado aportan gran valor y una capacidad fundamental para cualquier Ejército del siglo XXI. Estas capacidades, a su vez, pueden ser potenciadas de manera considerable si se logra la sinergia entre ambas.

En general, cuando se desarrolla un ejercicio de adiestramiento en el ámbito de la defensa, se plantea una situación ficticia con cierto grado de realismo, decidido y establecido por el director de la actividad, el cual pone en contexto a todos los participantes explicando y

dando sentido a las actividades que cada uno debe desempeñar. En este caso, el foco está en la tarea que cada uno debe realizar, dejando el contexto físico en segundo plano. Las ventajas e influencias aportadas por el contexto físico son muchas veces desatendidas, dándole mayor interés a las decisiones de carácter operativo. Es decir, la situación debe ser, en la mayoría de los casos, “imaginada” por los participantes según las orientaciones del director del ejercicio.

Desde el punto de vista de la organización tanto la creación como la puesta en marcha de un ejercicio real, acarrea un muy alto costo considerando el material y los recursos humanos involucrados.

En un nivel superior de adiestramiento, se suelen realizar ejercicios virtuales utilizando sistemas de simulación constructiva, como BV. El contexto virtual creado toma una relevancia elevada, ya que el foco del adiestramiento se encuentra en las decisiones a nivel táctico-estratégico dejando de lado el proceso de decisiones a bajo nivel.

Esto se vuelve especialmente relevante en sistemas del tipo de SATAC (sistemas de propósito crítico), donde el adiestramiento con sistemas de armas acarrea el costo intrínseco de la operación de los mismos, que vuelve dicha actividad altamente gravosa.

La utilización de tecnologías de la información que permitan integrar, con esfuerzo controlado, las ventajas de ambos tipos de adiestramiento, en conjunto con sistemas reales de C2, ofrece una nueva y valiosa capacidad a la organización. Estas tecnologías involucran sistemas de simulación LVC (3), visualización, generación de entidades sintéticas, componentes de realidad aumentada, entre otras.

Desde la presentación del primer artículo donde se definió el concepto de realidad virtual (4) los avances realizados en el campo de la interacción entre las realidades o mundos virtuales y la realidad física se han sucedido uno a otro de forma constante, generando una infinidad de posibilidades de aplicación en las tecnologías orientadas al adiestramiento en el campo de la defensa.

Las líneas de investigación del presente trabajo evaluarán las ventajas de la integración de sistemas de simulación para complementar ejercicios de adiestramiento reales, explotando la posibilidad de la creación de mundos virtuales para dar contexto a actividades de instrucción realizadas en el mundo real, y como dichas realidades podrían coexistir.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Dentro del marco del proyecto SATAC, se llevaron adelante dos líneas de investigación en la materia de adiestramiento:

1- Estudio de posibles técnicas alternativas de simulación que redujeran de forma considerable los costos y la complejidad de los procedimientos relacionados al adiestramiento, tanto en el uso del sistema, como de los procedimientos que el sistema pretende automatizar.

2- Estudio del Impacto de las técnicas de Realidad Virtual y Realidad Aumentada (5) en el adiestramiento, así como posibles aplicaciones operacionales a través del monitoreo de la evolución del prototipo operacional. La implementación de esta propuesta permite los siguientes avances:

- Enriquecer el ejercicio del personal a cargo otorgándoles un contexto claro, utilizando tecnologías de Realidad Virtual y generación de entidades sintéticas a través de sistemas de simulación.
- Generar situaciones dinámicas con problemáticas más cercanas a la realidad, permitiendo adaptaciones ante las acciones del enemigo simulado.
- Dar la posibilidad de asociar ejercicios de toma de decisiones de alto nivel (generalmente virtuales) con ejercicios operativos (idealmente reales) de manera integrada. (6)
- Aportar desde los simuladores y a través de la interface, distintos grados de inteligencia artificial como valor agregado al ejercicio.

Así, pues, las líneas de investigación que abre SATAC dentro del CIDESO se resumen en las siguientes:

- Realidad Virtual: incorporación de elementos de realidad virtual para dispositivos de visualización y móviles.
- Simulación Viva: para la interacción de entidades reales dentro de los sistemas de simulación (1).
- Generación de entidades sintéticas: para la incorporación de entidades virtuales a ejercicios reales.

3. RESULTADOS OBTENIDOS/ESPERADOS

A lo largo del año 2016 se analizó y diseñó una solución que permita insertar el concepto planteado, en el esfuerzo coordinado de los desarrollos principales del CIDESO, específicamente dentro del mencionado proyecto, SATAC. La misma se está realizando a través de la implementación de tipo “prototipo evolutivo” que permite validar el producto y que pretende que sea escalable.

Cabe mencionar que los requerimientos de la propuesta surgen de las observaciones realizadas en actividades desarrolladas previamente por el CIDESO en conjunto con unidades operacionales de las Fuerzas Armadas Argentinas.

Se considera a la implementación del concepto una extensión de las capacidades operativas de SATAC. El sistema deja de ser únicamente un sistema de Comando y Control (C2), para convertirse también en una herramienta de adiestramiento de nivel operacional, táctico y estratégico, mediante la implementación de la interface previamente mencionada (7).

De la presente línea de investigación se espera obtener los siguientes beneficios para el proyecto:

- Reducción de costos operativos al momento de realizar tanto los ejercicios de capacitación como de adiestramiento.
- Simplificación y agilización de los procedimientos para la preparación y despliegue de ejercicios.
- Aumento de la portabilidad del sistema, ya que el mismo podrá ser desplegado en cualquier ambiente o instalación en el marco de un ejercicio de adiestramiento.

- Aumento de la seguridad a través de la reducción en la exposición de los educandos a situaciones de riesgo.

Uno de los beneficios más significativos que se espera obtener, es un modelo de bajo costo y fácil despliegue (8) (9) (5). Esto es de suma importancia para el éxito de la implementación del sistema, ya que aprovisionar a los distintos elementos de la organización con el material apropiado para la utilización del mismo en otras instancias, ha probado ser, según la experiencia adquirida en otros desarrollos, una enorme dificultad y generalmente uno de los factores de riesgo de mayor impacto.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo principal que desarrollará las líneas de investigación propuestas se encuentra contenido en su totalidad dentro del Proyecto SATAC. Dicho equipo consta de ingenieros en informática y electrónica, analistas de sistemas y estudiantes/becarios con distintas capacidades específicas, como ser modelado y construcción de Sistemas de Información Geográfica (GIS), sistemas distribuidos, aplicaciones móviles, sistemas de simulación de comportamiento organizacional, tecnologías de objetos en diferentes lenguajes y, por supuesto, tecnologías de Realidad Virtual (Unity), entre otros.

Los trabajos desarrollados dentro del equipo del proyecto SATAC si bien tienen como prioridad el desarrollo del mencionado sistema, también contribuyen a la difusión de nuevo conocimiento dentro de la organización y la interacción con grupos de investigación externos en busca de potenciar las capacidades globales.

Adicionalmente, se impulsa en toda la organización la realización de prácticas profesionales supervisadas, tesis y tesinas de grado y posgrado asociadas a los desarrollos internos del laboratorio, así como la realización de cursos de especialización por parte de los integrantes.

En el caso particular del Proyecto SATAC, se dispone de un equipo integrado por alumnos

de grado de la Escuela Superior Técnica, de la Universidad de Palermo y de la UTN Regional Buenos Aires.

Si bien el CIDESO dispone de investigadores aptos para seguir las líneas propuestas, la colaboración entre laboratorios de informática busca el beneficio de la “sinergia” en I+D y constituye la materialización del aporte al desarrollo científico – tecnológico que el Ejército pretende en bien de la comunidad. Sobre las líneas de investigación propuestas, se buscará colaboración especialmente en la integración de sistemas. Dada la amplitud de sistemas a integrar, desde interfaces para sensores hasta sistemas legados desarrollados en distintos lenguajes de programación y con arquitecturas dispares, se buscarán contactos que posean conocimientos previos en el tema, con el fin de acortar los tiempos de investigación.

5. BIBLIOGRAFÍA

1. **Joseph J. Testa, Mike Aldinger, Kris N. Wilson, Christopher J. Caruana.** *Live-Virtual-Constructive Systems, Interactions, And Test And Training Benefits.* Fort Walton Beach, FL : DRS Training & Control Systems, Inc, 2006.
2. **Moriello, Sergio A.** *Inteligencias Sintéticas.* Buenos Aires, Argentina : Alsina, 2001.
3. **Mr. Phil Harvey, Mr. Steven Hatter, Maj Michael Davis Revision.** *Joint Training: Live, Virtual, and Constructive (L-V-C).* Kirtland AFB, NM : Air Force Distributed Mission Operations Center, 2008.
4. **Sutherland, Ivan E.** *The Ultimate Display.* s.l. : Information Processing Techniques, Office, ARPA, OSD, 1965.
5. **Jonathan J. Hull, Berna Erol, Jamey Graham, Qifa Ke, Hidenobu Kishi, Jorge Moraleda, Daniel G. Van Olst.** *Paper-Based Augmented Reality.* Menlo Park : Ricoh Innovations, Inc., California Research Center, 2007.
6. *Interoperability-Ready, Training-Focused Architecture for Command and Control Systems.* **Repetto, Alejandro Juan**

Manuel. Orlando, FL : s.n., 2011. Interservice/Industry Training, Simulation and Education Conference.

7. *Framework de Interoperabilidad para Sistemas de Comando y Control.* **Repetto, Alejandro Juan Manuel.** El Calafate, Santa Cruz : WICC'10, 2010. ISBN 978-950-34-0652-6.

8. **Institute for Defense Analyses, Science and Technology Div.** *Cost-Effectiveness of Computer-Based Instruction in Military Training.* Alexandria : Institute for Defense Analyses.

9. **Jesse Orlandy and Joseph String.** *The Cost-Effectiveness of Military Training.* [ed.] Office of the Secretary of Defense. Alexandria : Institute for Defense Analyses.

10. **Stanley, Major Bruce E.** *Wargames, Training, and Decision-Making.* [ed.] School of Advanced Military Studies. Fort Leavenworth : United States Army, Command and General Staff College.

11. **Schroeder, Ralph.** *The Usability of Collaborative Virtual Environments and Methods for the Analysis of Interaction.* [ed.] University of Oxford Oxford Internet Institute. Oxford : Massachusetts Institute of Technology, 2006.

12. **Enrico Costanza, Andreas Kunz, and Morten Fjeld.** *Mixed Reality: A Survey.* 2009.

13. **Craig, Donald.** *Advantages of Simulation.* St. John's : Memorial University, Faculty of Science.

14. **Bianchi, Oscar Martin and Repetto, Alejandro.** *Real-Virtual World Interaction for Training Simulations (WInter Training).* Posadas : WICC 2012, 2012. 978-950-766-082-5.

15. **David S. Alberts, Richard E. Hayes.** *Understanding Command And Control.* [ed.] Command and Control Research Program. s.l. : Office of the Secretary of Defense.

16. **Kresimir Cosic, Miroslav Slamic and Drazen Penzar.** *Combat and Security Related Modeling and Simulation.* s.l. : Information and Security, 2003.

17. **Gregory Harrison, Eric Worden, Jason Smith, Jonathan Brant, Dave**

Maynard, Tom Wonneberger. *Adaptive Artificial Enemy for Embedded Simulation.* Orlando : Lockheed Martin Global Training and Logistics.

Realidad Virtual y Aumentada, Big Data y Dispositivos Móviles: Aplicaciones en Turismo

Feierherd Guillermo, Depetris Beatriz, Huertas Francisco, González Federico, Romano Lucas, Viera Leonel, Horas Fabiola, Delia Lisandro

Instituto de Desarrollo Económico e Innovación
 Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
 Hipolito Irigoyen 880 - Ushuaia - Tierra del Fuego
 {gfeierherd, bdepetris, fhuertas, fgonzalez, lromano, lviera}@untdf.edu.ar
 fabiolahoras@gmail.com, lisandro.delia@gmail.com

RESUMEN

Las Tecnologías de la Información y la Comunicación (TICs) son “tecnologías disruptivas”, pues permiten resolver antiguos problemas mediante soluciones impensables antes de que estas tecnologías existieran.

La mayoría de las actividades humanas las han incorporado, produciendo la aparición de soluciones novedosas y un importante aumento de la productividad.

La actividad turística no es ajena a estas transformaciones. Basta ver cómo se realizan hoy las reservas y contrataciones de los servicios turísticos para advertir los cambios introducidos para productores y consumidores.

Por otra parte, la introducción de teléfonos inteligentes (smartphones), contribuyó a la ubicuidad de la computación y con ello, a la generación de grandes volúmenes de datos (big data) y nuevos paradigmas sobre las formas en las que se utilizan estos recursos. A su vez, dispositivos cada vez más potentes y económicos han puesto tecnologías como las de realidad virtual y realidad

aumentada al alcance de todos.

El proyecto busca relevar los usos que la industria turística está haciendo de estas tecnologías en forma individual o combinada, para proponer alternativas de aplicación en el ámbito de nuestra provincia. A fin de demostrar la factibilidad de las propuestas se propone desarrollar algunas aplicaciones experimentales.

Palabras clave: Realidad virtual; Realidad Aumentada; Big Data; Dispositivos Móviles; Turismo; Smart Destinations

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral

en trámite).

El financiamiento requerido es suministrado por la UNTDF.

El proyecto se desarrolla formalmente desde el 01/03/2017 hasta el 28/02/2019.

INTRODUCCIÓN

El rápido desarrollo de las Tecnologías de la Información y la Comunicación (TICs) produce cambios en todos los niveles organizacionales, tanto en el ámbito público como privado.

Entre las transformaciones que interesan al proyecto están las vinculadas al concepto de Destinos Turísticos Inteligentes (STD por Smart Tourism Destinations), el que constituye una derivación directa del concepto de Ciudades Inteligentes (Smart Cities).

Si bien es difícil encontrar definiciones que estén universalmente aceptadas, podemos decir que una ciudad inteligente es aquella en la que las TICs se introducen estratégicamente buscando mejorar la competitividad de la ciudad y, al mismo tiempo, la calidad de vida de sus ciudadanos. [1]

Por su parte, un Destino Turístico Inteligente puede ser definido de distintas maneras. Como lo señala Alfonso Vargas-Sánchez [2] en su revisión de la literatura sobre el tema, un STD es definido por López de Ávila como “un destino turístico innovador, construido sobre la infraestructura de tecnología actualizada, garantizando el desarrollo sostenible de las áreas turísticas, accesible para todos, facilitando la interacción de los visitantes y su integración con el entorno, incrementando la calidad de la experiencia en el destino y mejorando la calidad de vida de los residentes”. A su vez señala Vargas-Sánchez que Gretzel, Sigala, Xiang & Koo consideran que el

turismo inteligente es el “turismo apoyado por esfuerzos integrados en un destino para recopilar y aprovechar los datos derivados de la infraestructura física, las conexiones sociales, las fuentes gubernamentales y organizativas en combinación con el uso de tecnologías avanzadas para transformar esos datos en experiencias in situ y propuestas de valor comercial con un enfoque claro en la eficiencia, la sostenibilidad y el enriquecimiento de la experiencia.”

Como surge de las definiciones anteriores es evidente que las TICs son facilitadoras. Como bien señalan Boes, Buhalis e Inversini, “los destinos pueden desarrollar su inteligencia alineando las dimensiones clave de liderazgo, capital social, innovación y capital humano, utilizando las TICs como la “infraestructura” que facilita la co-creación de valor / experiencias para sus visitantes y competitividad para su industria. [3] Al mejorar la inteligencia de las dimensiones de Ciudad Inteligente (personas, vida, movilidad, medio ambiente, economía y gobierno), los destinos crean las condiciones para apoyar el desarrollo de Destinos de Turismo Inteligente donde se prioriza la interconexión, co-creación y la creación de valor, a través de la implementación de aplicaciones tecnológicas e infraestructuras TIC como Cloud Computing e Internet de las Cosas [4]. Las sinergias entre el interés y la preferencia garantizan que todas las partes interesadas se beneficien del proceso y que se desarrollen mejores experiencias y calidad de vida para todas las partes interesadas que participan en el destino turístico.

Con la tecnología inmersa en prácticamente cualquier organización o entidad, los destinos turísticos van a

potenciar las sinergias entre la tecnología ubicua y distintos componentes sociales para proveer experiencias enriquecedoras a los turistas, ya sea antes, durante o luego de su viaje. Como ha ocurrido en todos los ámbitos en los que criteriosamente se introduce tecnología, los destinos que hagan un buen uso de éstas podrán incrementar sus niveles de competitividad.

Las TICs hacen que las ciudades sean más accesibles y disfrutables, tanto para residentes como para turistas, gracias a servicios interactivos que interconectan distintos niveles de gobierno con empresas y proveen información en tiempo real de utilidad para todos ellos. Además, los datos resultantes pueden ser analizados por los organismos involucrados para el desarrollo de mejores políticas.

Desde una perspectiva turística, las TICs pueden contribuir generando valor agregado a las experiencias de los turistas y a la vez mejorar la eficiencia de las organizaciones relacionadas, facilitando la automatización de algunos procesos o la obtención de información valiosa. Es un hecho que las tecnologías de la información y la comunicación (TIC) han significado una innovación disruptiva en las relaciones entre oferta y demanda a nivel de turismo.

El entorno digital permite promocionar destinos, productos y servicios y, a la vez, conocer a los turistas como nunca antes se había hecho, ofreciéndoles experiencias personalizadas y de mayor calidad: segmentación e hipersegmentación del mercado, reducción de costos, mayor eficiencia y competitividad, etc.

Desde este punto de vista, las TIC han abierto un nuevo horizonte en el sector turístico, que plantea retos y

oportunidades y que requiere un gran esfuerzo de adaptación, tanto por parte de las empresas que prestan estos servicios, como por la gestión del destino liderada por el sector público.

La democratización de la tecnología ha sido la causa de que el sector haya pasado de estar controlado por la oferta (de las empresas) a estar dominado por la demanda (de la sociedad civil).

Por otra parte ha cambiado la forma de acceder a internet: se ha pasado de la computadora de escritorio y la notebook a los dispositivos móviles (tablets y teléfonos principalmente). A su vez, estos nuevos dispositivos ya no solo se usan para comunicarse entre personas, sino que poseen una serie de sensores y accesorios que los convierten en aparatos “todo en uno” (cámara de fotos, filmadora, gps, medio de pago, etc.) En consecuencia, tanto el sector público como el privado tienen que estar preparados para cubrir los requerimientos y necesidades del turista en el siglo XXI: más informado (a través de webs, redes sociales, etc.), que organiza personalmente su viaje y que requiere una serie de servicios digitales en el destino, al igual que los que tiene en su residencia habitual (por ejemplo, conectividad inalámbrica para todos sus dispositivos).

La llamada “Internet de las cosas” (IoT por Internet of Things) está ganando terreno rápidamente entre las TICs. El término fue propuesto en 1999 por Kevin Ashton -investigador del MIT- quien definió IoT como una red que conecta cualquier cosa en cualquier momento y lugar, para identificar, localizar, administrar y monitorear objetos inteligentes.

La idea detrás de IoT es generar interacciones automáticas en tiempo real entre distintos objetos conectados a

internet. Así, IoT facilita que distintas plataformas puedan transmitir datos obtenidos mediante diferentes sensores y que esos datos puedan ser procesados e interpretados en tiempo real.

En un contexto turístico, los turistas pueden usar sus teléfonos móviles para explorar el destino y sus eventos de interés, utilizando información provista por los gobiernos, agencias privadas e incluso otros turistas o ciudadanos.

Todas estas actividades llevadas a una escala masiva producen una cantidad de información multidimensional conocida como Big Data.

Haciendo uso del Big Data, las organizaciones turísticas pueden extraer información valiosa y mejorar la experiencia de los usuarios, proponiendo una nueva forma de relación con los turistas. Los destinos turísticos que primero aprovechen estas condiciones, harán una importante diferencia comparada con otras ciudades.

Por su parte, la Realidad Virtual (VR) permite a los turistas tener una experiencia inmersiva en los destinos, aún a distancia, mostrándole a la persona “cómo es y cómo se siente el lugar propuesto”. La VR facilita a los turistas visualizar y recorrer distintos atractivos, obteniendo información complementaria en formatos multimediales (texto, audio, imagen, video) en un entorno 3D o de 360 grados.

A su vez, la Realidad Aumentada (AR) puede utilizarse para mostrar cómo se veía un mismo atractivo en distintas épocas (por ejemplo, cómo se veía Ushuaia hace 100 años) o agregar en tiempo real, información de interés a un elemento en cuestión.

Con Realidad Virtual y Realidad Aumentada la promoción de los destinos turísticos se vuelve una atracción en sí

misma, incrementando el interés de las personas y aumentando así el mercado potencial de turistas.

Es deseable entonces contar con espacios turísticos consolidados sobre la base de una infraestructura tecnológica de vanguardia, incluyendo sistemas inteligentes que obtengan la información de forma automática, la analicen y comprendan los acontecimientos en tiempo real, facilitando la toma de decisiones y la interacción del visitante con el entorno turístico.

La innovación, el aprovechamiento de las nuevas tecnologías, la sostenibilidad ambiental y económica, como así también la cooperación público-pública y público-privada son condiciones sine qua non para el desarrollo turístico actual, donde el gobierno, los ciudadanos y las empresas locales deben asociarse con los turistas para comprenderlos, ofrecerles lo mejor y aprender de ellos.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Como ha sido mencionado previamente los temas de investigación / desarrollo / innovación para el proyecto con los siguientes:

- Realidad Virtual
- Realidad Aumentada
- IoT
- Big Data

especialmente en su vinculación con dispositivos móviles.

RESULTADOS Y OBJETIVOS

El proyecto ha sido formalmente iniciado el 01/03/2017, razón por la cual a la fecha no se han obtenido resultados.

En cuanto a los resultados esperados, cabe mencionar que, en líneas generales, el objetivo del proyecto es analizar el estado del arte en las tecnologías de (Realidad Virtual [VR], Realidad Aumentada [AR] y Big Data [BD]), poniendo especial atención al uso que se hace de las mismas en el dominio del turismo, con el fin último de proponer y desarrollar prototipos de aplicaciones enfocadas al turismo local.

En ese marco se busca:

- 1.- Identificar oportunidades de aplicación de VR, AR y BD al turismo de Tierra del Fuego.
- 2.- Desarrollar, utilizando las tecnologías mencionadas, aplicaciones experimentales para uso turístico.
- 3.- Vincular la información turística oficial existente y la recolectada con sensores de infraestructura, o la que pueda inferirse a partir del uso de smartphones y headsets (visores de realidad virtual), para su posterior uso en la gestión turística.
- 4.- Evaluar la posibilidad de aplicar las tecnologías analizadas para desarrollar un centro de interpretación Antártica interactivo-virtual y multidimensional, con base en la ciudad de Ushuaia.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por tres Profesores Titulares (dos vinculados a la Licenciatura en Sistemas y uno a la Licenciatura en Turismo), dos Asistentes, dos Alumnos (uno de ellos también Asistente Alumno) y un Profesor Adjunto externo (UNLP), también vinculado a las TICs.

Uno de los Asistentes está finalizando su Máster en Ciudades Inteligentes en la Universidad de Girona,

España, con prácticas en Schiedam, Holanda. Su tesis “Boosting attractiveness of heritage areas in Schiedam, Netherlands, with new technologies” está íntimamente ligada al proyecto esperándose una retroalimentación constante en su trabajo.

El Profesor Adjunto externo está finalizando el trabajo final para obtener el grado de Especialista en Ingeniería de Software. Su trabajo titulado "Desarrollo de Aplicaciones Móviles Multiplataforma" también tiene aplicación directa al tema del presente proyecto.

Uno de los alumnos ha finalizado, en el tiempo transcurrido desde la presentación del proyecto hasta la fecha, su tesina de grado en la Universidad Nacional de la Patagonia San Juan Bosco.

La misma, consistente en un “Marco de trabajo para el desarrollo de aplicaciones móviles híbridas”, es de gran interés para el proyecto debido a su aplicación directa.

Por su parte, la alumna que integra el proyecto tiene previsto realizar su tesina de Licenciatura en el “Análisis de Experiencia de Usuario en Aplicaciones Móviles”.

Cabe mencionar también que el integrante vinculado a la Licenciatura en Turismo dirige el proyecto “Relevamiento y puesta en valor del Patrimonio Artístico (bustos, composiciones, monolitos, placas, estatuas y mástiles) en espacios públicos de la ciudad de Ushuaia”, que fuera presentado y aprobado en la misma convocatoria que el que aquí se presenta.

REFERENCIAS

- [1] Andrea Caragliu, Chiara Del Bo, and Peter Nijkamp. (2017, 03). Smart Cities in Europe. *Journal of Urban Technology*. [Online]. 18(2). Disponible en

https://www.researchgate.net/publication/46433693_Smart_Cities_in_Europe

[2] Alfonso Vargas-Sánchez. (2017, 03). Exploring the concept of Smart Tourist Destination. Enlightening Tourism. A Pathmaking Journal. [Online] 6(2). Disponible en https://www.academia.edu/30399967/Vol_6_No_2_2016_July-December

[3] Kim Boes, Dimitrios Buhalis and Alessandro Inversini. (2017, 03). Conceptualising Smart Tourist Destination Dimensions. Disponible en https://www.researchgate.net/profile/Dimitrios_Buhalis/publication/272576525_Conceptualising_Smart_Tourism_Destination_Dimensions/links/54e9d4390cf25ba91c7ff25c.pdf

[4] Taewo Nam y Theresa Pardo, (2017, 03), “Conceptualising Smart City with Dimensions of Technology, People and Institutions”, presentado en 12th Annual International Conference on Digital Government Research, College Park, MD, USA, Junio 12 - 15, 2011. Disponible en https://inta-aivn.org/images/cc/Urbanism/background%20documents/dgo_2011_smartcity.pdf

Servicios de Cloud Computing e Informática Contextualizada Aplicada a una App Turística para la Ciudad de Puerto Iguazú Misiones

Cristian A. Kornuta, Diego I. Motta

Departamento de Informática, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones. Félix de Azara 1552. (3300) Posadas, Misiones. Argentina.

{cristian.kornuta24, diegomotta18}@gmail.com

Resumen

La plataforma tecnológica de Cloud Computing, es la innovación tecnológica por excelencia de la década, la presente tecnología ha permitido y permite la creación servicios por parte de los gigantes de Internet como ser Mapas en líneas, Base de datos en tiempo real o la creación de sitios como plataformas sociales o plataforma de videos, plataformas que nos permiten transmitir y comunicarnos libremente. Lo anterior permitió que en la actualidad sea tan importante un bloguero en línea como un corresponsal de una cadena de televisión.

Hoy en día surgen innovaciones de nuevos modelos de negocios que se basan en esta tecnología, modelos que hace unos años atrás no se nos hubiese imaginado implementarlo. La tecnología de Cloud Computing conjuntamente con el uso masivo de dispositivos móviles y el concepto de computación contextualizada han propiciado un nuevo camino para todo lo referente a la asistencia en línea de los usuarios. Dentro del sin fin de usuarios de estas tecnologías encontramos a los turistas. Este nuevo turista llamado 2.0 o digital, el cual se convirtió en hiperconectado y multicanal, acostumbrado a usar sus dispositivos móviles se vio beneficiado con la aparición de app móvil que lo hacen sentir como en su casa en el destino donde elige visitar.

Palabras Clave: Cloud Computing – Informática sensible al contexto – Sistemas de archivo de Google.

Contexto

El presente trabajo expone un estudio de tres temas como ser Cloud Computing, Informática contextualizada y el Sistemas de archivo de Google tratados en el curso de Sistemas Distribuidos perteneciente a la Maestría en Tecnología de la Información entre la Universidad Nacional de Misiones y la Universidad Nacional del Nordeste a partir de la revisión bibliografía y material del curso utilizado y referencias externas. Estos temas fueron seleccionados y se desarrollan en el marco del actual proyecto que nos encontramos desarrollando en la ciudad de Puerto Iguazú, Misiones “*Puerto Iguazú móvil*”.

Introducción

La plataforma tecnológica de Cloud Computing, es la innovación tecnológica por excelencia de la década, esta tecnología ha permitido y permite la creación servicios por parte de los gigantes de Internet en su mayoría. Servicios que antes solo unos pocos podrían imaginar a un futuro muy lejano. Podemos decir que nos encontramos frente a un nuevo paradigma tecnológico que se impondrá de a poco como lo hizo el internet y la web en su momento afirman. Tanto es el impacto que ha producido esta innovación tecnológica que el concepto de Cloud Computing o “la nube” se encuentra en boca de los medios de comunicación y de la sociedad en común. La mayoría de las empresas de una u otra forma ya se encuentran utilizando estos servicios y migrando a este nuevo modelo de negocio.

Además, son cada vez más las empresas que en mayor o menor medida ofrecen los mismos servicios que ofrecen los gigantes de internet.

Este nuevo paradigma tecnológica que nos encontramos viviendo vino propiciado por el uso masivo de dispositivos móviles; Dispositivos que permitieron y permiten que el usuario común que antes se encontraba frente a una computadora “de forma fija” o teniendo el problema de tener que desplazar sus documentos de un lugar a otro para poder trabajar en ellos desde otra computadora. A pasar a ser un usuario que la mayoría de las veces de manera gratuita puede desplazarse y tiene acceso a todos los recursos que posee sean estos archivos en su conjunto y de las aplicaciones con las cuales trabaja; con una simple conexión a internet.

No solo esto ha permitido esta tecnología, sino que los prestadores de estos servicios pueden realizar un seguimiento realmente profundo del usuario final con respecto a sus hábitos, consumos y tareas diarias o en que ocupa su tiempo. Lo que llevó a una nueva forma de ofrecer productos y servicios, como ser el caso de Google Maps, o Google Calendar, con el hecho de registrar un evento en algún lugar, como ser la casa por ejemplo y el evento de sacar la basura, él te lo recordará cuando te encuentres en ese lugar porque sabe dónde te encuentras y se lo dijiste.

Todos los días aumenta el número de usuarios en Internet, usuarios que de apoco van formando parte de uno u otro grupo de usuarios que utilizan algún servicios de Internet. Entre estos usuarios beneficiados se encuentran los turistas, los cuales actualmente son llamados turistas 2.0 o llamado digital, el cual se convirtió gracias a estas tecnologías en hiperconectado y multicanal, acostumbrado a usar sus dispositivos móviles se vio beneficiado con la aparición de app móvil que lo hacen sentir como en su casa en el destino donde elige visitar el uso de estas APP móviles

no sólo se da durante el periodo de visita del turista sino antes, durante y después.

Es por eso que este nuevo paradigma de la nube ha impulsado al sector turístico a desarrollar aplicaciones referentes al transporte, gastronomía y cultura por nombrar algunas.

Una de las características que poseen estas aplicaciones, y que será desarrollado en este trabajo, es referente a la informática contextualizada cuyas aplicaciones en el sector es muy amplio.

Otro gran sector beneficiado por los servicios de la nube, y que en mayor o menor medida, son los impulsores de estas nuevas tendencias son los desarrolladores en este sector podemos encontrar Servicios como ser Firebase que es un sistema de Base de datos en tiempo real y con todo el soporte Backend para el desarrollo de aplicaciones.

Línea de investigación y desarrollo

El presente trabajo orienta su línea de investigación en dos grupos; el primer grupo referente al Geo referenciamiento, Geo posicionamiento y al Geo marketing y una segundo grupo a la informática contextualizada como ser la identificación de usuario y el ofrecimiento de servicios o productos orientados a personas específicas.

Objetivos y resultados

Puerto Iguazú es una ciudad 100% turística visitada anualmente por miles de turistas locales, regionales y del resto del mundo, si bien existe mucha promoción a través de los medios de comunicación, folletería e internet; aún no se ha solucionado el problema de la falta de información disponible organizada, concentrada y principalmente actualizada en un solo lugar, que permita al turista mejorar la experiencia en el destino y que el mismo incremente su promedio de estadía en la ciudad y así consuma más; hasta el momento dicho promedio es de dos días y medio (según IPEC), Iguazú no cuenta con una app para dispositivos móviles que concentre toda la información, basada en la geo localización que

permite detectar dónde está el turista y mostrarle todos los puntos turísticos, hoteles, restaurantes, empresas de turismo, entretenimiento nocturno, estaciones de servicio, cajeros automáticos, servicios de salud y todo lo que pueda ser de interés para un turista. La geo localización juntamente con las redes sociales es una nueva forma de hacer marketing, tienen un gran potencial en el turismo ya que el viajero está constantemente conectado y le gusta compartir sus experiencias en las redes sociales (fotos, el lugar donde se encuentran). Existen aplicaciones como:

- **Foursquare:** servicio basado en geo localización web aplicada a las redes sociales.
- **Tripadvisor:** es una aplicación móvil que permite conocer puntos turísticos, restaurantes y hoteles en un determinado lugar a través de la geo localización y el geo marketing.

Ninguna de las aplicaciones mencionadas anteriormente soluciona el 100% del problema expuesto, y la mayor desventaja que presentan es que no proporcionan información actualizada y oportuna para el turista.

“Iguazú en tu móvil” pretende por un lado ser una aplicación que permita al turista vivir una experiencia satisfactoria en su estadía en la ciudad y hacerlo sentir como un local y por otro lado tiene como objetivo principal contribuir al desarrollo económico de la ciudad, a través, del trabajo en conjunto de los actores principales: la municipalidad de Iguazú, la Asociación Civil Atractivos turísticos de Iguazú, los diferentes alojamientos, restaurantes, comerciantes, transporte y otros. Ellos proveerán periódicamente y constantemente la información para la aplicación (promociones, espectáculos etc.). Los actores antes mencionados serán los encargados de promocionar la aplicación a través de publicidad utilizando cartelera, folletería y las TICs.

Con respecto al desarrollo de la aplicación y mantenimiento estará a cargo del equipo responsable del presente proyecto.

Objetivo general

Lograr a través del desarrollo de una APP con geo localización (en varios idiomas) concentrar, organizar y mantener actualizada toda la información turística y de servicios disponible de la ciudad de Puerto Iguazú para aumentar la oferta turística y así incrementar el promedio de estadía y el consumo del turista en el lugar, beneficiando y contribuyendo de esta manera el desarrollo económico de los comerciantes, empresarios locales y todos los actores involucrados en el proyecto.

Objetivos Específicos

- Promocionar los otros puntos turísticos de Iguazú.
- Hacer sentir al turista como un local.
- Proporcionar al turista no solo una guía turística sino también una guía de servicios e información actualizada sobre promociones, espectáculos y otros.
- Ofrecer recomendaciones personalizadas sobre restaurantes, alojamientos y actividades a realizar.
- Contribuir al desarrollo económico de las empresas y comercios locales relacionados con el turismo.
- Brindar una solución al problema urbanístico de la ciudad. (Posee muchas calles diagonales).
- Comprometer a la Municipalidad de Puerto Iguazú, ACATI, Empresas hoteleras, turísticas, restaurantes, comerciantes, empresas de servicios, expertos en idiomas y otros a trabajar en conjunto.
- A continuación se presentara las pantallas que fueron desarrolladas en carácter de prototipo de la aplicación con el propósito de ser presentado frentes a la municipalidad de la ciudad

y las principales empresas turísticas de la localidad para su financiamiento.



Imagen 1. Pantalla principal de la app

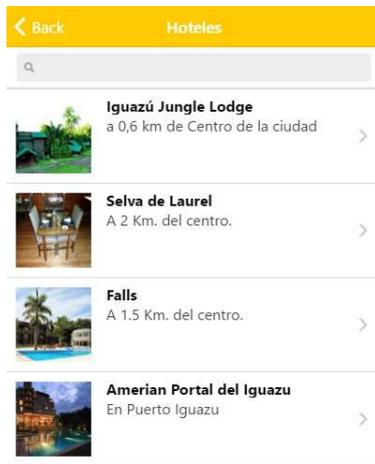


Imagen 2. Pantalla listado de hoteles.

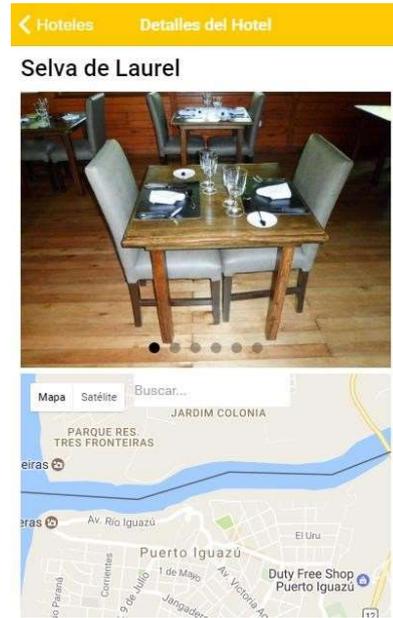


Imagen 3. Información de uno de los restaurantes seleccionado

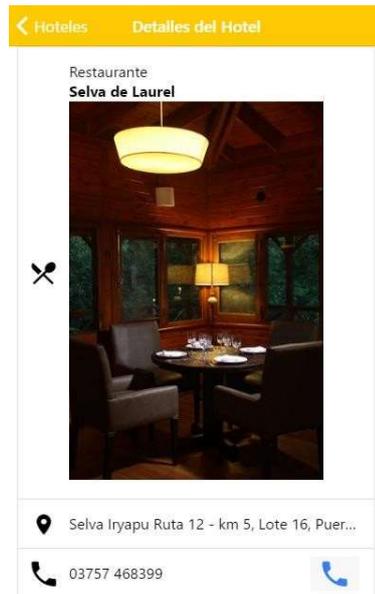


Imagen 4. Información de uno de los restaurantes seleccionado

Conclusión

El presente trabajo deja en evidencia el actual estado que se encuentra los servicios de informática, y las condiciones que se encuentran dadas para que cualquier persona pueda desplegar una idea que llegue muy lejos.

Con respecto a la tecnología presente hoy en día en cuanto al cloud computing es una tecnología que planteó un nuevo paradigma y que vino para quedarse, y será una verdadera revolución que se dará como se dio en su momento con el Internet y la Web.

El presente proyecto representa una gran innovación para la Ciudad de Puerto Iguazú, para ciudades aledañas y localidades en un futuro, debido al gran número de teléfonos inteligentes y su uso masivo por parte de los turistas.

Los turistas son grandes consumidores de aplicaciones móviles que les permitan mejorar su experiencia en el lugar que se encuentran visitando. La aplicación que se desarrolla con este proyecto juntamente con la incorporación de tecnologías como ser GeoMarketing, Geolocalización y el concepto de SoLoMo servirán para estudiar los hábitos de los turistas y mejorar la oferta turística actual y así ampliarla. Además, dará origen a nuevos emprendimiento a partir del análisis detallado de los datos que proporcione el uso de la aplicación y el estudio del consumo por parte de los turista; como por ejemplo a partir del análisis de flujo de personas por ciertos lugares, el recorrido que hacen los turistas dentro de la ciudad y caminos por donde se desplazan se pueden armar nuevos circuitos turísticos.

Todo lo anterior llevará a aumentar la tasa de permanencia del turista en la ciudad de Iguazú, propiciando así el aumento del consumo y el fomento de la región por parte del sector turístico.

Se debe aclarar que el hecho de ofrecer y promocionar puntos turísticos que no pertenecen a la ciudad de Puerto Iguazú, no representa una amenaza al cumplimiento del objetivo general del presente proyecto.

Referencias

COMPUTACIÓN EN LA NUBE (CLOUD COMPUTING). Diapositiva del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad

Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). Consultado el 1 de diciembre de 2016.

L. Joyanes Aguilar. Cloud Computing. Estrategias digitales para organizaciones y empresas. AlfaOmega, México, 2011. Consultado el 1 de diciembre de 2016.

Cómputo Ubicuo e Interacción Humano-Computadora. Jesus Favela. CICESE. Material del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). Consultado el 1 de diciembre de 2016.

Novatica. Computación Ubicua. Material del curso de Sistemas Distribuidos de la Maestría en Tecnologías de la Información. Universidad Nacional del Nordeste (UNNE) y la Universidad Nacional de Misiones (U.Na.M.). de la cátedra. Consultado el 1 de diciembre de 2016.

Presentación. Desarrollo de Aplicaciones Móviles Sensibles al Contexto. Lic. en Cs. de la Comp. e Ingeniería en Computación. Depto. de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur. 1er. Cuatrimestre de 2016. Disponible en <http://cs.uns.edu.ar/~mvm/CAIA-2016/index.php?accion=download&dir=downloads/Clases> Consultado el 1 de diciembre de 2016.

Sistema Automático para el Monitoreo Operacional de los Desmontes Usando Imágenes Satelitales

Francisco Darío Maldonado¹, Walter Fabián Sione^{1,3}, Francisco Mesías Viva^{1,2}

¹ CEREGEO - Centro Regional de Geomática, Facultad de Ciencia y Tecnología. Universidad Autónoma de Entre Ríos. Ruta 11, km 10,5, Oro Verde. Entre Ríos.

² CICyTTP - Centro de Investigaciones Científicas y Transferencia de Tecnología a la Producción, CONICET. Matteri y España s/n. (3105). Diamante. Entre Ríos.

³ PRODITEL- Laboratorio de Teledetección, Departamento de Ciencias Básicas. Universidad Nacional de Luján – UNLU. Ruta 5 (6700), Lujan. Buenos Aires. Argentina.

e-mail: {francisco.dario.maldonado; wsione; fmvm483}@gmail.com

RESUMEN

En este trabajo se presenta la línea de Investigación en desarrollo en el Centro Regional de Geomática - CEREGEO, Facultad de Ciencia y Tecnología de la Universidad Autónoma de Entre Ríos - UADER. Esta tiene como objetivo el desarrollo de un sistema de software automático para el monitoreo operacional de los desmontes en la cobertura vegetal del bosque nativo del centro-norte de Entre Ríos, utilizando imágenes satelitales Landsat. La metodología informatizada para la detección de los desmontes, articula técnicas de procesamiento de imágenes desarrolladas en Python y Matlab. Esta posibilitará el monitoreo automático operacional a partir de imágenes satelitales adquiridas mensualmente. Las técnicas informatizadas permitirán la transformación no supervisada de datos continuos en datos temáticos agrupados en Monte, Desmonte y No monte. Los resultados del proyecto permitirán establecer en el futuro un sistema automático operacional de monitoreo de la cobertura de bosques, generando alertas automáticas de desmontes y el mapeo mensual de los bosques.

Palabras clave: Imágenes satelitales, teledetección, monitoreo, sistema informatizado.

CONTEXTO

Esta línea de investigación y desarrollo “Detección automática de los desmontes con imágenes satelitales” se enmarca en los Proyectos “Desarrollo de un Sistema operacional para detección de cambios y monitoreo del bosque nativo usando imágenes satelitarias. Norte de la Provincia de Entre Ríos.”, “Monitoreo histórico de los desmontes y recuperación” y el proyecto de Tesis que inicia este año “Desarrollo y Validación de un Sistema Informatizado de Monitoreo Automático de los desmontes para un área del Espinal del Centro-Norte de Entre Ríos, en base a Grandes Volúmenes de Datos de Teleobservación”.

Esta línea de desarrollo se orienta a proporcionar, a las autoridades provinciales de medio ambiente, herramientas para la aplicación de la Ley de Presupuestos Mínimos de Protección Ambiental de los Bosques Nativos de la Ley Nacional de Bosques 26331. Estos proyectos se iniciaron en Marzo de 2013 en el CEREGEO - Centro Regional de

Geomática la Facultad de Ciencia y Tecnología - FCyT/UADER - Universidad Autónoma de Entre Ríos y se orientan a suplir necesidades apuntadas en el Plan Argentina 2020 (Núcleo Socio - Productivo Estratégico, 2012).

1. INTRODUCCIÓN

Los organismos gubernamentales de control y gestión de los recursos naturales necesitan herramientas ágiles para la aplicación de la Ley de Presupuestos Mínimos para el Bosque Nativo (Ley N° 26331) y para el control del ordenamiento territorial.

Los sistema que utilizan imágenes digitales obtenidas periódicamente por satélites orbitales en el espectro óptico, han demostrado la agilidad necesaria para el monitoreo del territorio. Actualmente, el aumento de la resolución espacial y radiométrica de las imágenes plantea la necesidad de la informatización para el manejo de un volumen creciente de datos adquiridos por estas imágenes.

La adquisición digital de datos ambientales se inició hace 40 años con la Serie de satélites Landsat y esta garantizada en el futuro por la LDCM - Landsat Data Continuity Mission, por lo que el desarrollo de un sistema de monitoreo basado en imágenes Landsat es adecuado para el estudio histórico del territorio y de su evolución futura.

Las imágenes son adecuadas para el monitoreo del ambiente a escala regional por su bajo costo y características espectrales y radiométricas que permiten analizar el uso y cobertura de las tierras, obteniendo buenos resultados para el mapeo y detección de cambios, como señalado por Coppin et al. (2004).

En la Provincia de Entre Ríos, el bosque nativo fue mapeado por Sabbatini et al. (2009), con imágenes Landsat, obteniendo la

“Zonificación del bosque natural” solicitada por el Gobierno Provincial y según estos resultados la provincia contaba, en 2008, con más de un millón de hectáreas de monte nativo.

Los proyectos de investigación y desarrollo iniciados en 2013 en el CEREGEO, comenzaron generando mapas manualmente a intervalos de 5 años, para el análisis histórico de los desmontes desde 1980 a 2010, Maldonado et al. (2012). Esta experiencia mostró la necesidad de informatizar las técnicas y articularlas en un Sistema operacional para el monitoreo continuo de los procesos ambientales a lo largo del tiempo.

Las técnicas automáticas de detección y monitoreo de cambios en la cobertura vegetal, actualmente no obtienen buenos resultados como mencionado por Coppin et al. (2004), y en general introducen al sistema el fenómeno llamado “propagación de los errores”. Este amplifica los errores en los mapas finales por la multiplicación de los errores de los mapas antecedentes, y es en general inaceptable para fines cuantitativos, según muestran Lunetta et al. (1981), Congalton y Green (1999).

Una técnica automática capaz de la detección digital que evita la propagación de errores por la necesidad de mapas intermedios, es la “ROCAM- Rotación Radiométrica Controlada por la Moda”, fue desarrollada y presentada en Maldonado et al. (2007). Esta se basa en la “RCEN - Rotación radiométrica controlada por Eje de No-Cambio”, Maldonado et al. (2002) y Maldonado et al. (2007). Esta técnica de detección digital de cambios, permite el uso simple de imágenes de diferentes sensores y ya ha mostrado buenos resultados en regiones semiáridas de la catinga de Brasil y chaco árido en Maldonado (2007) entre otros ambientes. La característica principal de esta técnica es la rotación controlada del espacio bi-temporal de la radiometría de las imágenes

digitales, produciendo buenos resultados con imágenes calibradas o sin calibrar o de diferentes sensores con diferente sensibilidad y radiometría, siendo adecuada para incorporar nuevas imágenes y desarrollos futuros.

Posteriormente a la detección de cambios y transformación de datos continuos a temáticos, necesita la definición automática de umbrales en algunos casos de transformación, (Zhu et al. 2015; Cattaneo et al. 2011).

La informatización de estas rutinas de procesamiento y transformación de datos continuos a temáticos se efectuara con Python (Python, 2017), Matlab (Mathwork, 2017), La Librería GDAL - Geospatial Data Abstraction Library (GDAL, 2017), Rutina de Extracción de nubes Fmask (Zhu & Woodcock, 2015) y Quantum Gis (QGIS, 2017).

El objetivos de la investigación y desarrollo es un Sistema automático informatizado que articule técnicas de detección de cambios con técnicas de caracterización temática de los cambios. El sistema informatizado propuesto se basa en la obtención automática de una leyenda numérica relacionada con la frecuencia de ocurrencia en el tiempo del suelo desnudo y posterior caracterización ambiental de esta frecuencia y mapeo temático en tres clases Monte, Desmonte y No monte.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas son ejecutado por un equipo multidisciplinario con capacidad para el procesamiento de imágenes, la adaptación y aplicación de técnicas de levantamiento de campo y comprender la dinámica ambiental en una región donde predomina la actividad agropecuaria en áreas de bosque natural.

Una de las líneas en desarrollo se relacionan con la “Disminución del tamaño de los subsets de datos” para mejorar la transformación automática de datos continuos a temáticos,

usando técnicas no supervisadas Kmean o Isodata. Otra línea se orienta al cálculo automático de un Índice de vegetación ajustado al suelo, SAVI -Soil Adjusted Vegetation Index (Huete & Glenn, 2011), que permita la detección de cambios. Y la línea de Aplicación de una Leyenda temática con base numérica igual al promedio del valor de los pixeles de la misma clase Kmean en una banda Índice SAVI. El mapeo de los cambios y trabajos de levantamiento de la verdad de campo, es una de las etapas necesarias en los proyectos que utilizan imágenes para la obtención de la precisión del mapeo (Lunetta, 1981; Congalton & Green, 1999). En esta línea se aplican los levantamientos fisonómico-estructurales de la vegetación para caracterizar los cambios, según Maldonado et al. (2005), para el ajuste y test de la metodología.

3. RESULTADOS ESPERADOS

El principal resultado es una metodología automática operacional para detección y monitoreo mensual de los desmontes en el bosque nativo. Esta metodología operacional minimiza el efecto de las fuentes de errores que son propagados a los largo del tiempo. La principal fuente de errores está asociada con la subjetividad de la intervención humana, sobre todo cuando el personal en operación del sistema cambia durante el tiempo de funcionamiento de sistema. El sistema producirá periódicamente: alertas de desmontes, mapas de intensidad de los cambios del bosque natural, mapas de cambios de uso producidos en áreas de bosque natural y tablas cuantitativas de los desmontes en el centro-norte de Entre Ríos.

4. FORMACIÓN DE RECURSOS HUMANOS

Durante la ejecución de los proyectos fueron ofrecidas una becas de iniciación por año para alumnos de las licenciaturas en Biología y para los proyectos futuros se ofrece una beca por año también para Licenciatura en Sistemas de Información de la Facultad de Ciencia y Tecnología, para programación de rutinas Python en QGIS.

Los resultados del proyecto que se inicia para la informatización de la etapa de detección, serán producto de la Tesis de Doctorado “Desarrollo y Validación de un Sistema Informatizado de Monitoreo Automático de los desmontes para un área del Espinal del Centro-Norte de Entre Ríos, en base a Grandes Volúmenes de Datos de Teleobservación”, Doctorado en Ciencias Aplicadas de la Universidad de Lujan, financiado por el CONICET.

4. BIBLIOGRAFÍA

- Congalton, R.G.; Green, K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practice. Mapping sciences series. New York: CRC Press Inc., 1999. 98p.
- Coppin, P.; Jonckheere, I. ; Nackaerts, K.; Muys, B.;Lambin, E. 2004. Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9): 1565-1596.
- Cattaneo. C.A.; Larcher, L.I.; Ruggieri, A.I.; Herrera, A.C.; BIASONI, E.M. 2011. Método de umbralización de imágenes digitales basado en entropía de Shannon y otros. En: Möller, O.; Signorelli, J. W.; Storti, M.A. (Eds). *Mecánica Computacional*, 30: 2785-2005.
- Huete, A. R.; Glenn, E. P. 2011. Remote Sensing of Ecosystem and Function. En: Weng, Q. *Advances in Environmental Remote Sensing*. London: CRC Press. cap.12, p.291-320. 600p.
- GDAL. 2017. Geospatial Data Abstraction Library. Open Source Geospatial Foundation. <<http://www.gdal.org>>
- Lunetta, R.; Congalton, R.; Frenstermaker, L.; Jensen, J.; McGwire, K.; Tinney, L. 1981. Remote Sensing and geographic information system data integration: error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, 57(6): 677-687.
- Maldonado, F.D.; Santos, J.R.; Carvalho, V.C. 2002. Land use dynamics in the semiarid region of Brazil (Quixabá-PE): characterization by principal components analysis. *International Journal of Remote Sensing*, 23(23): 5005- 5013.
- Maldonado, F.D.; Carvalho, V.C.; Souza, C.L.; Martinelli, M.; Pinheiro, O.J. ; Santos, F.F. 2005a. Determinación de la longitud de transecta para el relevamiento fisonómico-estructural de la vegetación del semiárido para suministrar datos a las técnicas de percepción remota orbital. *Multequina*, 13: 1-14, 2005.
- Maldonado, F. D.; Graça, P. M.; Santos, J. R. Detecção automática de mudanças na cobertura vegetal utilizando a técnica de Rotação radiométrica controlada pela evolução da moda das imagens componentes.. In: XIII Simposio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis. Anais 2007 Proceedings. São José dos Campos : INPE, 2007. p. 5911-5918.
- Maldonado, F. D. ; Santos, J.R.; Graça, P.M. 2007. Change detection technique based on the radiometric rotation controlled by no-change axis, applied on a semi-arid landscape. *International Journal of Remote Sensing* , 28(8): 1789-1804.
- Maldonado, F.D.; Sione, W.F.; Aceñolaza, P.G. 2012. Mapeo de desmontes en áreas de bosque nativo de la Provincia de Entre Ríos. *Ambiência*, 8(esp): 532-532.
- Matlab. 2017. Multi-paradigm programming language environment. The Mathworks Inc. <<http://www.mathworks.com>>
- Núcleo Socio-Productivo Estratégico. 2012. Sistemas de Captura, Almacenamiento y puesta en disposinibilidad de Datos

Ambientales. Argentina Innovadora 2020. Sec. de Planeamiento y Políticas. Min. de Ciencia y Tecnología e Innovación Productiva de la Nación. Julio 2012. 15p. <[http: www. argentinainnovadora2020.mincyt. gob.ar](http://www.argentinainnovadora2020.mincyt.gob.ar)>

Python. 2017. The Python Software Foundation. <[http: www. python. org](http://www.python.org)>

QGIS. 2017. Sistema de Información Geográfica. Quantum GIS Development Team. Open Source Geospatial Foundation Project. <[http: qgis. osgeo. org](http://qgis.osgeo.org)>

Sabattini, R.A.; Ledesma, S.; Brizuela, A.; Sabattini, J. 2009. Zonificación de los bosques nativos en el Departamento La Paz (Entre Ríos) según las categorías de conservación. Informes FCA UNER y la Dirección General

de Recursos Naturales, Sec. de la Producción Gob.de Entre Ríos. FCA UNER: Oro Verde, Octubre 2009. 30p.

Zhu, Z. y C. E. 2014. Woodcock. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 138: 1143-1157.

Zhu, Z.; Wang, S.; Woodcock, C. E. 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 144: 152-171.

Subsistema de Recolección de Elementos de Hardware Integrado a un Sistema de Gestión, Inventario y Monitoreo de Hardware con Alertas Automáticas (S.R.E.H.)

Silvia Edith Arias, Fabián Gibellini, Analía L. Ruhl, M. Alejandra Di Gionantonio, Nora V. Flores, Mónica M. Serna, Daniel F. Arch, Milagros Zea Cárdenas, Germán Parisi, Diego Barrionuevo.

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas / Universidad Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja S/N, 5016

s_autn@hotmail.com, fgibellini@bbs.frc.utn.edu.ar, analialorenaruhl, ing.alejandravg, ingnoraflores, sernamonicam@gmail.com, daniel.arch@pjn.gov.ar, milyzc, germanparisi, santosdiegob@gmail.com

Resumen

S.R.E.H. permite recolectar información de distintos elementos de hardware en una estación de trabajo. Trabaja tanto en sistemas operativos GNU/Linux como en Windows. Estos datos recolectados son base para un monitoreo de los equipos.

Esta herramienta se ha diseñado para interactuar con un sistema de Gestión, Inventariado y Monitoreo de Hardware con Alertas Automáticas, desarrollado durante el período 2014-2016 en el marco del Proyecto “Inventario Seguro en Ambiente Informáticos con Alertas Automáticas y Registro de historial para activos de TI” acreditado y financiado por la Secretaría de Ciencia y Técnica de Código: EIUTNCO0002226. El cual se lleva a cabo en el Laboratorio de Ingeniería en Sistemas de Información de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional (Argentina).

Palabras clave: Recolección, Alertas, Inventario, Hardware, Monitoreo, Historial, Gestión.

Contexto

El presente trabajo está inserto dentro del proyecto homologado por la Secretaría de Ciencia y Técnica de la UTN - FRC “Inventario Seguro en Ambiente Informáticos con Alertas Automáticas y Registro de historial para activos de TI” (I.S.A.I.) - Código: EIUTNCO0002226, radicado en el Laboratorio de Ingeniería en Sistemas de Información (LabSis) de la Universidad Tecnológica Nacional – Facultad Regional Córdoba (UTN - FRC), quien es a su vez es el organismo que contribuye a su financiamiento.

En el marco de la gestión de incidentes, el Laboratorio de Sistema (LabSis) tiene como principal objetivo ofrecer disponibilidad y un correcto funcionamiento de los equipos (PCs) a alumnos y docentes.

Actualmente cuando un alumno, docente o encargado reporta un funcionamiento no esperado, relacionado al hardware, se notifica a los integrantes del Área Técnica, quienes realizan pruebas sobre el equipo para detectar los componentes afectados. En caso de ser necesario, el equipo queda fuera de cualquier uso hasta su reparación. Es decir, que el personal del LabSis debe, dependiendo de la cantidad de máquinas afectadas, redistribuir los recursos o las aulas. Este proceso de

detección de inconvenientes relacionados al hardware es manual.

De lo antes planteado, derivó la necesidad de que el personal de LabSis conozca en cualquier momento: PCs habilitadas y el funcionamiento de sus respectivos componentes, para determinar con anterioridad a sus posibles usos, si surgió o no, alguna incidencia en las mismas a través de una Gestión, Inventariado y Monitoreo de Hardware que a través de alertas automáticas notifiquen ante cualquier comportamiento no esperado a los usuarios involucrados.

Introducción

La primera fase del proyecto I.S.A.I. consistió en el desarrollo y pruebas de un sistema de Gestión, Inventariado y Monitoreo de Hardware con Alertas Automáticas (G.I.M.H.A.A.) [1], cuya fuente de datos provenían desde un sistema de inventariado, llamado OCS Inventory [2]. Durante esta etapa se detectaron inconsistencias en los datos que reportaba el OCS Inventory, cuyo patrón no fue identificado. Debido a esto surgió la necesidad de identificar un procedimiento más confiable de obtención de datos de cada estación de trabajo.

Las inconsistencias encontradas se listan a continuación:

Duplicaba equipos. Durante las pruebas realizadas desde agosto hasta diciembre del 2016 surgió que en algunos casos el agente OCS (cliente instalado en las estaciones de trabajo a ser inventariadas) duplicaba las máquinas al momento de reportar su inventario al módulo OCS servidor, esto ocurrió en un aula del LabSis y en todos los casos bajo el sistema operativo Linux. En el resto de las aulas con equipos idénticos reportaba con normalidad.

En equipos con sistema operativo Windows el software reportaba en distintos tiempos, no seguía la configuración establecida.

No hay compatibilidad entre versiones. Cuando era necesario actualizar los agentes o el servidor a una versión más reciente no había compatibilidad con la versión anterior,

por lo que ante estas situaciones era obligatorio actualizar tanto el agente OCS de cada estación de trabajo como el módulo servidor que brinda servicios a los agentes OCS.

Los puntos mencionados indujeron a la decisión estratégica de desarrollar un subsistema Recolector de Elementos de Hardware (S.R.E.H.) que alimenta al subsistema G.I.M.H.A.A. La integración de ambos subsistemas conforman el sistema de nuestro proyecto I.S.A.I. Se realizó un análisis para determinar cuántos y cuáles serían los elementos de hardware a ser monitoreados, además de definir las características de cada elemento a ser recolectados.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se inscribe dentro de esta línea de investigación, enfocado en la gestión de activos de T.I. y que afecta a los ámbitos académico, gubernamental y empresarial. Favorece ampliamente al crecimiento de la gestión de activos de T.I. principalmente en las instituciones públicas del país, donde se sufren permanentemente sustracciones indetectables.

Al tratarse de un software con una licencia libre [3], todo aquel que desee implementar el sistema podrá acceder a la aplicación y su código, como así también adaptarlo para la estructura del ambiente informático sobre el cual lo desee trabajar.

Resultados y Objetivos

El objetivo de la fase II fue desarrollar el sistema S.R.E.H., que sea de libre uso y de fácil acceso para el LabSis y para cualquier organismo que lo requiera, haciendo hincapié en que este constituye al sistema I.S.A.I. brindando la base para dar soporte a una Gestión de Configuración y Activos TI del Servicio dentro de estas subredes, permitiendo monitorear, controlar, prevenir, proteger, notificar (a través del G.I.M.H.A.A) y tomar decisiones en tiempo y forma sobre

los equipos con los que cuenta el laboratorio y lograr una mejora continua de los servicios brindados a alumnos y docentes.

El sistema completo I.S.A.I. está conformado el subsistema G.I.M.H.A.A. y el subsistema S.R.E.H., este último fue desarrollado recientemente y ambos se encuentran actualmente en pruebas beta.

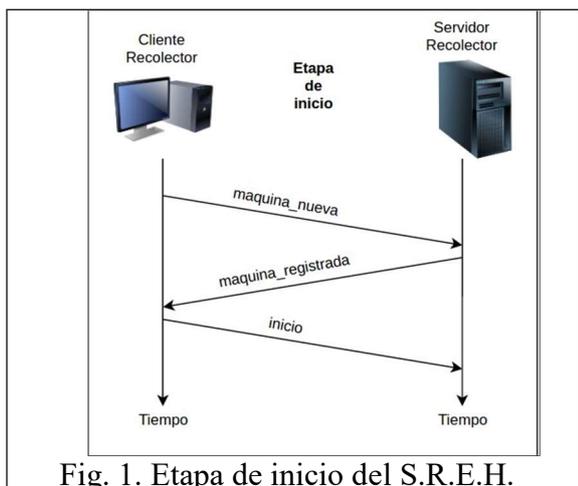


Fig. 1. Etapa de inicio del S.R.E.H.

Los elementos seleccionados a inventariar en cada estación de trabajo son:

- Almacenamiento: Disco duro (unidades físicas)
- Procesamiento: CPU (velocidad, caché, generación, fabricante, etc.).
- Memorias RAM: Memoria (slot que usa, tipo, fabricante, entre otros).

Por otro lado, también se obtienen datos del equipo, como nombre de la máquina, arquitectura (32 o 64 bits), nombre y versión del SO.

Estos datos y sus características, fueron seleccionados a partir de una priorización de componentes de hardware según las necesidades del LabSis en base a un estudio y análisis de factibilidad y confiabilidad de la obtención de los datos según cada sistema operativo.

El subsistema S.R.E.H utiliza una arquitectura cliente - servidor. El servidor se desarrolló en python 2.7 [4]. Para el desarrollo del cliente se tuvo en cuenta la heterogeneidad de sistemas operativos utilizados en las aulas del LabSis, por lo que se desarrollaron clientes tanto para sistemas

operativos Windows como para GNU/Linux, específicamente Debian [5].

Según el sistema operativo objetivo, los clientes utilizan distintas herramientas que permiten la obtención de datos de los componentes seleccionados, para el caso de Windows existe "Windows Management Instrumentation" (WMI) [6] que provee una interfaz para acceder a datos de administración del sistema y su red, lo que nos permite obtener datos del hardware y software entre otros. Esta información es simple de entender ya que se muestra de manera muy similar a un motor de base de datos relacional, pues se basa en SQL.

En el caso de GNU/Linux se usó inicialmente la herramienta dmidecode [7] con permisos de superusuario para los tres componentes (almacenamiento, procesador y memoria).

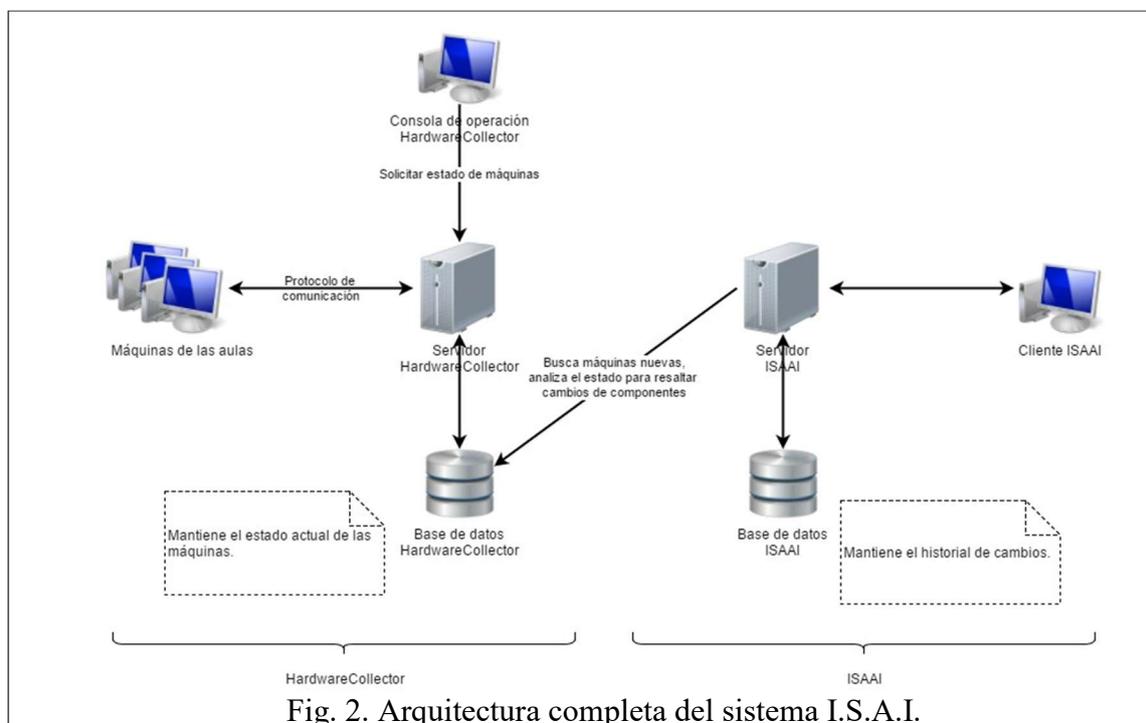
Actualmente, debido a la necesidad de minimizar el uso de permisos de usuario y otorgar al subsistema mayor flexibilidad, se usan distintos archivos y herramientas: módulo de python: platform [8], /proc/cpuinfo [9], /sys/devices/system/cpu/cpu/{id_procesador_fisico}/cpufreq/scaling_max_freq [10], dmidecode, lsblk [11], /sbin/udevadm, /sys/class/block/{nombre_dispositivo}/device /vendedor.

Para la comunicación entre cliente y servidor se ha creado un protocolo que se ejecuta sobre el protocolo TCP [12], utilizando sockets [13] como canal de comunicación.

El protocolo creado tiene tres etapas: La etapa de inicio, la cual se lleva a cabo cuando se instala un cliente en el equipo a ser inventariado, el servidor ejecuta un proceso que configura al cliente, la Fig. 1. muestra la dinámica de esta etapa. Esta configuración permite explicitar si se van a reportar los tres componentes o alguno en particular. Por último en la etapa de funcionamiento, existen dos modos de actuar, el modo pasivo y el activo. En el modo pasivo el cliente, según su configuración, reporta al servidor, y en el modo activo el servidor solicita un reporte, al cual el cliente responde.

Todos estos elementos constituyen el del subsistema Recolector para integrarlo al sistema I.S.A.I. logrando de esta manera el primer sistema autóctono de inventario

incorporarse a la carrera de investigador y un estudiante investigador de la carrera de Ingeniería en Sistemas de Información y becarios que forman parte del equipo.



automatizado con alertas automáticas ante un cambio en el hardware, permitiendo que el LabSis tenga conocimiento del estado de sus equipos en uso, cuando sea requerido y al menos una vez por cada uso de cualquier equipo. Las alertas son disparadas ante cualquier cambio de hardware en un equipo respecto a la última vez que este fue usado. A través de esto se logra visibilidad de los activos en uso en las aulas de LabSis.

La arquitectura final del sistema ISAAI se la muestra en la Fig. 2.

Si bien esta segunda fase priorizo tres elementos de hardware, posteriormente se puede ampliar a otros componentes e inclusive se puede extender a un inventario del software de cada estación de trabajo.

Formación de Recursos Humanos

El grupo está compuesto por un Director, un Co-Director, dos profesores investigadores de apoyo, tres profesores aspirantes a

Este proyecto contribuirá a la formación y crecimiento de la carrera de investigador de los integrantes del mismo.

Además existe una colaboración interproyecto “Generación de modelos descriptivos para la prevención de incidentes en equipos informáticos en el contexto del Laboratorio de Sistemas (Fase II)” [14].

Referencias

- [1] “Gestión, Inventario y Monitoreo Hardware con Alertas Automáticas”. Workshop de investigadores en Ciencias de la Computación. Concordia. Entre Ríos. 2016.
- [2] OCS Inventory, <http://wiki.ocsinventoryng.org/index.php/Documentation:Administration> Última Visita: 18-08-2016.
- [3] Yingkui, Z., Jing, Z., Liye, W. “Justification of Free Software and its Enlightenment”. Econ. & Manage, Beijing Univ. of Chem. Technol., Beijing, China. Software Engineering (WCSE). Volume: 2. Page(s):171 – 173. ISBN: 978-1-4244-9287-9 DOI: 10.1109/WCSE.2010.47. IEEE. 19-20 Dec. 2010.

- [4] Python. Página oficial.
<https://www.python.org/>. Última visita: 13-03-2017.
- [5] Debian. Página oficial
<https://www.debian.org/intro/about>. Última visita 13-03-2017.
- [6] Windows Management Instrumentation
[https://msdn.microsoft.com/en-us/library/aa394582\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/aa394582(v=vs.85).aspx). Última visita: 13-03-2017
- [7] Dmidecode.
<https://linux.die.net/man/8/dmidecode>. Última visita: 13-03-2017
- [8] Plattform. Página oficial
<https://docs.python.org/3/library/platform.html>.
Última visita 13-03-2017.
- [9] “The /proc/cpuinfo file”.
http://www.linfo.org/proc_cpuinfo.html. Última visita 13-03-2017.
- [10] “CPU frequency and voltage scaling code in the Linux(TM) kernel”.
<https://www.kernel.org/doc/Documentation/cpu-freq/user-guide.txt>. Última visita: 13-03-2017.
- [11] lsbblk. <https://linux.die.net/man/8/lsblk>.
Última visita: 13-03-2017
- [12] “Transmission Control Protocol”. RFC 793.
- [13] “The Definition of a Socket”. RFC 147. [14] “Generación de Modelo Descriptivo para la prevención de incidentes en equipos informáticos en el contexto del laboratorio de sistemas (Fase II)”. Corso, C., Maldonado, C., Gibellini, F., Ciceri, L., Martinez, G., Pereyra, F., Donnet, M. Departamento de Ingeniería en Sistemas Universidad Tecnológica Nacional, Facultad Regional Córdoba.

**Procesamiento de Señales y
Sistemas de Tiempo Real**

Alternativas Eficientes para Procesamiento y Comunicaciones en Sistemas de Tiempo Real

Fernando Romero, Mariano Mendez, Diego Encinas,
Armando De Giusti¹, Fernando G. Tinetti, Santiago Medina, Martín Pi Puig, Juan
Manuel Paniego, Matías Dell'Oso.

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

¹ CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

{fromero, mmendez, dencinas, degiusti, fernando, smedina, mpipuig, jmpaniego,
mdelloso}@lidi.info.unlp.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio y desarrollo de Sistemas de Tiempo Real (STR), en particular, la planificación de tareas y la comunicación en Sistemas Distribuidos de Tiempo Real (SDTR), experimentando con microcontroladores y simulaciones para la adquisición y control de las variables del sistema.

Contexto

Esta línea de Investigación forma parte del proyecto "Arquitecturas multiprocesador en HPC: software de base, métricas y aplicaciones" (11/F018) del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Sistemas Embebidos, Comunicaciones, Sensores, Robots, Microcontroladores.

1. Introducción

Los Sistemas de Tiempo Real (STR) interaccionan con su entorno físico y deben satisfacer restricciones de tiempo de respuesta [14] [7] [15] [8] [23], dependiendo del medio físico en el que está inmerso el sistema. El cumplimiento de dichos plazos pasa a tener radical importancia, considerándose un fallo su incumplimiento. La interacción con el sistema físico en el que están embebidos se lleva a cabo a través de sensores y detectores que proporcionan información sobre los valores actuales de las variables del sistema, permitiendo la elaboración de respuestas transmitidas a través de actuadores. Dentro de estos sistemas es posible distinguir dos tipos:

- 1) Los que elaboran la respuesta en forma puramente computacional, empleando algoritmos más o menos sofisticados, que pueden ir

¹ Investigador CONICET

de un simple cálculo hasta el empleo de complejas técnicas de inteligencia artificial.

- 2) Los que elaboran la respuesta a través de la intervención humana [34].

Dentro de los SDTR existen nodos (móviles y fijos) alimentados a batería, por lo cual supervisar el consumo es sumamente necesario. Esto requiere el uso de placas de bajo consumo energético, tales como los microcontroladores que, a pesar de tener una baja capacidad de cómputo, se ajustan a los requerimientos de la mayoría de los STR.

Otra característica a estudiar, es la latencia en la elaboración de las respuestas frente a señales externas, por lo general a través de interrupciones de hardware. Para ello se debe optimizar el software que interviene en el proceso. Generalmente, dichas latencias provienen de la utilización de SOTR (Sistemas Operativos de Tiempo Real), principalmente debido a la intervención del planificador de tareas.

Se experimenta sobre diferentes SOTR y sobre simulaciones de los mismos [20]. Además, se disponen de diversas plataformas de hardware y software para la construcción de sistemas de diferente grado de complejidad. Estos pueden estar basados en microcontroladores que soporten o no un SOTR. Entre las placas de desarrollo utilizadas se encuentran Arduino, Raspberry PI y CIAA [27], entre otras.

Una característica fundamental de los SOTR es tener un alto grado de fiabilidad, por lo que el estudio de la detección y control de condiciones de falla es un aspecto de gran importancia [2] [17].

Luego, con respecto a los microrobots, se utilizan algunos de diseño y fabricación propia como también los sofisticados robots Khepera [24] [25] [26].

Estos sistemas permiten abordar el estudio del paradigma llamado “Internet of Things (IoT)” [22] [21] [3].

En las diferentes implementaciones que se llevan a cabo en esta línea de investigación, se trabaja de manera experimental con mini-robots y sobre diferentes SOTR (Linux RT-Preempt, FreeRTOS, MQX, OSEK-OS, etc.) [20] [10].

Se estudian sistemas de comunicación basados en los protocolos RS485, SPI, I2C, CANBUS [29] [30] [31] [32] [33] y MODBUS [28]. Los sistemas basados en microcontroladores [11] pueden conectarse de manera cableada o bien de forma inalámbrica (Radio Frecuencia, Bluetooth, WiFi).

Se puede disminuir la complejidad y la probabilidad de errores en el diseño de sistemas, desarrollando una simulación específica a través de la utilización de diferentes frameworks [37]. En particular, con CANBUS se generan diferentes modelos de comportamiento y se realizan implementaciones con la herramienta de simulación Proteus [38].

2. Líneas de Investigación y Desarrollo

Se plantean como temas de estudio:

- Estudiar los sistemas robóticos con intervención humana [5] [14] en el lazo de control a través de acciones e interfaces no convencionales. Son componentes de estos sistemas: el modelo de reconocimiento de comandos gestuales [35] [36], las interfaces, el modelo de observación, los modelos dinámicos y de

realimentación de la máquina, el modelo de planificación y, eventualmente, los modelos de realimentación para el operador humano y el modelo de actuación humana (para interpretar acciones y distracciones).

- Verificación y validación del hardware por medio de simulaciones que permitan predecir posibles comportamientos y en consecuencia generar mayor eficiencia de diseño [9] [10].
- Construcción de redes de sensores sobre una determinada plataforma de microcontroladores, interactuando con un computador. Se utiliza MODBUS, RS485 o CANBUS como protocolo de comunicaciones, tanto a nivel físico como de simulación [1] [18].
- Odometría con navegadores, a través de robots Khepera [6] [16] [4] y otros de producción propia. Se enfoca principalmente en obtener vehículos autónomos [12] que puedan circular en un entorno de autopistas inteligentes y con capacidad de estacionamiento.

3. Resultados y Objetivos

De acuerdo con las tareas desarrolladas y a desarrollar, los resultados se enmarcarán en varias direcciones, todas relacionadas con los sistemas de tiempo real:

- Construcción y uso de robots tipo vehículo autónomo para el estudio de Odometría.
- Medición de consumo de diferentes tipos de microcontroladores bajo distintas condiciones de uso.
- Estudio de distintas plataformas de hardware: Arduino, Intel Galileo, CIAA, Freescale Kinetis, Raspberry Pi, entre otras.

- Construcción y estudio de dos redes de sensores, empleando Radio Frecuencia, CANBUS, MODBUS y RS485.
- Construcción de modelos de simulación de protocolos de comunicaciones en sistemas de tiempo real. Realización de pruebas de concepto utilizando herramientas de simulación.

4. Formación de Recursos Humanos

En base a estos temas se están desarrollando trabajos de varios alumnos encuadrados en el marco de la Convocatoria a Proyectos de Desarrollo e Innovación de la Facultad de Informática de la UNLP. Además, se encuentran en desarrollo y se han concluido varias tesinas de grado de alumnos de la Licenciaturas de Informática y Sistemas, como así también Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Computación.

5. Referencias

- [1] Jordi Bartolomé "El protocolo MODBUS", 2011. En <http://www.tolaemon.com/docs/modbus.htm>
- [2] Andersen, B. L. "Method of detecting systemic fault conditions in an intelligent electronic device." U.S. Patent 6,434,715, issued August 13, 2002.
- [3] Atzori, L., A. Iera, G. Morabito. "The internet of things: A survey." *Computer networks* 54, no. 15 (2010): 2787-2805.
- [4] Azizi, F., N. Houshangi. "Mobile robot position determination using data from gyro and odometry." In *Electrical and Computer Engineering*, 2004. Canadian Conference on, vol. 2, pp. 719- 722. IEEE, 2004.

- [5] Bekey, George A. Robotics: state of the art and future challenges. Imperial College Press, 2008.
- [6] Borenstein, Johann, Liqiang Feng, "Gyrodometry: A new method for combining data from gyros and odometry in mobile robots." In Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on, vol. 1, pp. 423-428. IEEE, 1996.
- [7] Burns, A, A, Wellings. "Real-Time Systems and Programming Languages: Ada, Real-Time Java and C/Real-Time POSIX", Addison-Wesley Educational Publishers Inc., 2009.
- [8] Buttazzo, G. C., "Hard RealTime Computing Systems", Third edition, Springer, 2011.
- [9] Eickhoff, J., Simulating Spacecraft Systems, Springer, 2009.
- [10] "FreeRTOS - market leading RTOS (real time operating system) for embedded systems supporting 34 microcontroller architectures". [http:// www.freertos.org/](http://www.freertos.org/).
- [11] Jenkins, T., I. Bogost. "Designing for the internet of things: prototyping material interactions." In CHI'14 Extended Abstracts on Human Factors in Computing Systems, pp. 731-740. ACM, 2014.
- [12] Jones, J. L., A. M. Flynn, Bruce A. Seiger. Mobile robots: inspiration to implementation. Vol. 2. Wellesley MA: AK peters, 1999.
- [13] Kopetz. H., "Real-Time Systems, Design Principles for Distributed Embedded Applications". Second Edition. Springer. 2011.
- [14] Krishna, C. Mani. Real-Time Systems. John Wiley & Sons, Inc., 1999.
- [15] Liu, J. W. S. Liu, "Real Time Systems", Integre Technical Publishing Co., Inc., 2000
- [16] Rekleitis, I. M., G. Dudek, E. E. Miliotis. "Multi-robot exploration of an unknown environment, efficiently reducing the odometry error." In International Joint Conference on Artificial Intelligence, vol. 15, pp. 1340-1345. Lawrence Erlbaum Associates Ltd, 1997.
- [17] Rohani, A., H. R. Zarandi. "An analysis of fault effects and propagations in AVR microcontroller ATmega103 (L)." In Availability, Reliability and Security, 2009. ARES'09. International Conference on, pp. 166-172. IEEE, 2009.
- [18] Introduction to the Controller Area Network (CAN) Texas Instrument Application Report SLOA101A–August 2002– Revised July 2008.
- [19] Silberschatz, A., P. B. Galvin, G. Gagne, Operating System Concepts, 8th Edition, ISBN : 978-0-470-12872-5, Wiley, 2009.
- [20] The Cheddar project: A free real time scheduling analyzer, <http://beru.univbrest.fr/~singhoff/cheddar/>
- [21] Weber, Rolf H., and Romana Weber. Internet of Things. New York: Springer, 2010.
- [22] Xia, Feng, L. T. Yang, L. Wang, and Alexey Vinel. "Internet of things." International Journal of Communication Systems 25, no. 9 (2012): 1101.
- [23] PHILLIP A. LAPLANTE, SEPPO J. OVASKA. REAL-TIME SYSTEMS DESIGN AND ANALYSIS Tools for the Practitioner Fourth Edition. A JOHN WILEY & SONS, INC., PUBLICATION. IEEE PRESS. 2012.
- [24] <http://ftp.k-team.com/khepera/documentation/Kh2UserManual.pdf>

- [25] <http://ftp.k-team.com/khepera/documentation/Kh2ProgrammingManual.pdf>
- [26] <http://ftp.k-team.com/khepera/documentation/Kh2IRAN.pdf>
- [27] <http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp>
- [28] <http://www.tolaemon.com/docs/modbus.htm>
- [29] http://www.bosch-semiconductors.de/en/ubk_semiconductors/ip_modules_3/produkttable_ip_modules/can_literature_1/can_literature.html
- [30] <http://www.can-cia.de/can-knowledge/can/can-fd/>
- [31] http://www.iso.org/iso/catalogue_detail.htm?csnumber=59165
- [32] http://www.freescale.com/files/microcontrollers/doc/app_note/AN1798.pdf
- [33] http://www.can-cia.org/W_Scanner_Auto_OBD2_OBD1_2x2_E_OBD_20_PIN.html
- [34] Chi-Pang Lam and Shankar Sastry, A POMDP Framework for Human-in-the-Loop System, University of California at Berkeley.
- [35] Mitra S., Acharya T.” Gesture recognition: A survey”. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol 37(3). pp 311–324. 2007.
- [36] Murthy G. R. S., Jadon R. S. “A Review of Vision Based Hand Gestures Recognition”. International Journal of Information Technology and Knowledge Management. Vol.2-2. pp 405-410. 2009.
- [37] D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
- [38] Proteus. <https://www.labcenter.com>. 2017

Detección en Tiempo Real de Malezas a través de Técnicas de Visión Artificial

Costamagna Marcelo¹, Panero Javier², Peretti Gastón³, Felissia Sergio Francisco³,
Lurgo Gerardo³, Cicioli Mauricio¹

- 1) Departamento de Mecánica de la Facultad Regional Villa María de la Universidad Tecnológica Nacional.
- 2) Departamento de Electrónica de la Facultad Regional Villa María de la Universidad Tecnológica Nacional.
- 3) Departamento de Electrónica de la Facultad Regional San Francisco de la Universidad Tecnológica Nacional

Av. De la Universidad 50- San Francisco (Pcia. de Córdoba) – CP (2400)

Tel.: 03564-421147 / e-mails: gastonperetti@gmail.com

costamagna_m@frvm.utn.edu.ar

Resumen

En este proyecto se propone desarrollar un prototipo de detección de las malezas que aparecen durante el procedimiento denominado barbecho en los cultivos agrícolas típicos de nuestra región, en particular soja, maíz y trigo. El aporte de este proyecto, se centra en desarrollar un sistema de visión artificial que permita la detección de malezas en tiempo real, aplicando técnicas de procesamiento de imágenes, que sea susceptible de incorporar en un equipo pulverizador de herbicidas. Adicionalmente se pretende también analizar comparativamente, los resultados experimentales del prototipo propuesto, con los de los sistemas de detección más utilizados actualmente.

Palabras clave: Detección malezas - Agricultura de precisión - Visión artificial.

Contexto

La investigación está inserta dentro de la línea de control y redes de información. El proyecto de investigación detección en tiempo real de malezas a través de técnicas de visión artificial se lleva a cabo en el ámbito del Departamento de Electrónica de la Facultad Regional San Francisco y en el ámbito del

Departamento de Electrónica de la Facultad Regional Villa María.

La Institución que acredita el proyecto de Investigación y desarrollo (PID) es Rectorado de la Universidad Tecnológica Nacional como proyecto de investigación y desarrollo inter-facultad mediante la financiación del mismo con fecha de aprobación en el mes de Abril del 2016.

Introducción

El control químico de malezas es, y seguirá siendo, de gran importancia en la agricultura de alta eficiencia. Sin embargo, las consideraciones tanto económicas y ecológicas dictan que nuevas reducciones en el uso de herbicidas serán necesarios en el futuro. En general, en nuestro país, la aplicación de herbicidas se realiza en forma uniforme en la totalidad del lote, no obstante el grado de infestación de malezas puede no ser homogéneo sino que aparece en forma de “manchas” y es en estos casos en donde el uso de herbicidas podría reducirse aplicándolo sólo a las áreas infestadas de maleza. Es por este motivo que el proyecto plantea la detección de malezas en barbecho en tiempo real, con el objetivo de reducir la utilización de herbicidas mediante procesamiento digital de imágenes. Para la detección de malezas en la actualidad, se utilizan métodos de tecnología óptica y

electrónica. Dentro de estas tecnologías los podemos clasificar en dos grupos:

- ✓ Sensores del espectro de luz reflejada
- ✓ Sensores basados en toma de imágenes y algoritmos de procesamiento de imágenes.

Dentro de los sensores del espectro de luz reflejada, tenemos:

- Sensores de color verde.
- Arreglos de sensores de color de distintas longitudes de onda y ancho de banda estrecho.
- Detección del color rojo y cercano al infrarrojo (RR/NIR).

Por otra parte dentro de los sensores basados en toma de imágenes tenemos:

- Detección por análisis de color.
- Detección por análisis de forma.
- Detección por análisis de textura.

Dentro de los sensores de luz reflejada, comenzaremos a detallar los sensores de color verde. Con esta estrategia, estaríamos en condiciones de distinguir las malezas verdes del suelo y, de esta manera, poder pulverizar solamente los sectores que poseen malezas [1]. Desde el punto de vista tecnológico, resulta un procedimiento sencillo. Existen en el mercado, sensores de color comerciales que se adaptarían perfectamente a este método. Cada sensor realiza una exploración del suelo y cuando detecta la presencia del color verde envía la orden a la pastilla pulverizadora para producir el rociado. Desde el punto de vista del acondicionamiento de señal, el sensor suministra una salida lógica que se activará cuando se detecte la presencia del color verde. Un microcontrolador de bajas prestaciones puede ser utilizado sin problemas, con la sola consigna de producir el retardo correspondiente para producir el disparo de la electroválvula que maneja el pico pulverizador, teniendo en cuenta la velocidad de avance. El sensor es un detector de color, el cual puede adquirirse o fabricarse mediante el uso de fototransistores y filtros ópticos.

Con respecto a los arreglos de sensores de distintas longitudes de onda, tiene como principio el estudio de las características de reflectancia espectral de los cultivos, de las malezas y del suelo. Estas

características se miden usando un espectrómetro. Para cada una de las especies de malezas o cultivos, se suele medir por separado las características espectrales de los tallos y las hojas como dos clases de objetos individuales. El suelo es tratado como una clase adicional. Se diseña el sensor seleccionando un arreglo de fotodetectores para esas longitudes de onda características y contrastantes entre las malezas y el cultivo o el suelo. La electrónica asociada a este tipo de sensores, requiere de circuitos de acondicionamiento de señal analógicos, conversores analógico digital de varios canales y microcontroladores con cierta capacidad de cálculo y velocidad de procesamiento para poder tomar decisiones en forma rápida y actuar en tiempo real.

En el caso de detección de color rojo y cercano al infrarrojo (RR/NIR) se utilizan sensores que están disponibles en el mercado actualmente. Se basa en la detección de malezas mediante la cantidad de luz reflejada por las mismas al ser iluminadas[2]. Fundamentalmente se han realizado estudios para determinar la cantidad de luz reflejada cerca del infrarrojo (NIR) y en el rango del rojo (RR) [3]. Dado que el suelo y las plantas reflejan cantidades distintas de ambos espectros, es posible utilizar esta característica para realizar una clasificación. Las plantas debido a su clorofila, absorben gran cantidad de luz roja, en el rango de longitudes de onda que van de 630 a 660 nm (RR), mientras que la luz cercana al infrarrojo, que se encuentra entre 750 y 1200 nm (NIR) es ampliamente reflejada. La relación de NIR/RR se encuentra entre 1.1 y 1.5 para el suelo y entre 6 y 15 para las plantas con hojas verdes, dependiendo del tipo de suelo y de maleza [4].

Los circuitos electrónicos para implementar un sistema de detección RR/NIR se componen de dos fotodetectores, uno dedicado a detectar la cantidad de luz roja reflejada por la planta y otro dedicado a detectar la cantidad de luz cercana al infrarrojo reflejada. Esto es posible mediante el uso de filtros ópticos de la longitud de onda deseada (RR o NIR) para cada sensor. El conjunto de sensores realiza una exploración del suelo y capta la luz reflejada de ambos espectros y realiza el procesamiento de estos datos mediante un microcontrolador, que corre un algoritmo de discriminación relativamente sencillo. Un punto a considerar es el método de iluminación del sector a explorar que puede ser a través de la luz natural o a través de iluminación activa. En este último caso

resulta conveniente ya que permitiría su utilización en el trabajo nocturno, no obstante, que esta situación requiere de una cuidadosa selección de la fuente de luz.

Dentro de las estrategias que utilizan sensores basados en la toma de imágenes, las mismas se basan en la utilización de cámaras de video o fotográficas.

En el caso en que utilicemos una cámara digital color y realicemos una clasificación de zonas verdes contra suelo, estaríamos en condiciones de realizar una pulverización selectiva. Sin embargo, existen algunos puntos que deben ser tenidos en cuenta. La clasificación por detección de color mediante visión artificial ha probado ser efectiva en gran cantidad de trabajos, pero éstos fueron realizados en laboratorios o invernaderos con condiciones de luz controladas. Desde el punto de vista tecnológico, como hemos mencionado antes, se necesitaría cámaras digitales color, una CPU que se encarga de correr un algoritmo de procesamiento de imágenes para determinar la existencia o no de malezas. Necesariamente se necesita un microcontrolador de altas prestaciones, preferentemente con capacidades de Procesamiento Digital de Señales (DSP). El algoritmo de detección, debe ser capaz de determinar las condiciones de iluminación para así poder detectar adecuadamente la maleza. Además, se debe tener en cuenta que la iluminación no se mantiene constante a lo largo del día, y varía con la nubosidad. Es necesario considerar la definición óptica de la cámara, para poder lograr una detección lo más precisa posible en el menor tiempo posible. Cabe destacar que, cuanto mayor definición tenga la cámara, mayor procesamiento es necesario y, por lo tanto, mayor es el tiempo involucrado. También hay que tener en cuenta que la cámara posee una gran parte óptica y que el polvo en el ambiente puede obstruir las lentes.

En el caso de la detección por análisis de formas, la gran diferencia radica en el software de detección que contiene la CPU. En este caso, el algoritmo es capaz de diferenciar la forma de las hojas del cultivo respecto de la forma de las hojas de la maleza. Existen trabajos que avalan este procedimiento y en donde se han logrado eficiencias en la discriminación de hasta el 73.1% [5]. No obstante existen trabajos que combinan el análisis de forma con el análisis del color,

conformando sistemas mixtos, en donde los resultados obtenidos fueron similares [6]. Sin embargo la gran mayoría de estas investigaciones fueron realizadas en laboratorios o invernaderos y requieren de imágenes de alta calidad gráfica y de un gran nivel de procesamiento [6]. Además, se debe considerar que en condiciones de campo las hojas se superponen y su orientación no es pareja. Esto genera una gran cantidad de formas distintas a reconocer. También debemos decir que a medida que las plantas se desarrollan cambian sustancialmente su forma, lo que complica aún más su detección [2]. Dada la complejidad del algoritmo de detección es preciso hacer uso de una CPU con gran capacidad de procesamiento DSP. Este tipo de sistema es más complejo pero presenta ventajas significativas en lo que respecta a la detección del color, ya que es capaz de discriminar entre cultivo, maleza y suelo. Pese a su complejidad, se han desarrollado sistemas de detección a campo en tiempo real [5]. La complejidad del algoritmo, determina que se requiera de un tiempo importante para tomar la decisión, lo que conspira como dijimos antes con la detección en tiempo real. Esta situación de demora en la actuación podría mitigarse ajustando la velocidad de avance y la distancia del detector al pico pulverizador.

Para el caso de detección por análisis de textura, Este método analiza la información sobre la textura de las imágenes, la que se encuentra relacionada con la distribución en el espacio del objeto fotografiado. Las partes brillantes y las partes oscuras se deben a variaciones de color o a sombras locales producidas por la geometría de las superficie del objeto fotografiado [7]. Dichas características pueden ser tomadas como referencia para lograr una diferenciación entre el suelo, el cultivo y las malezas. Este método se utiliza para determinar no solamente la existencia de malezas sino también la variedad de las mismas. Este tipo de clasificación requiere de un procesamiento intensivo mediante el uso de algoritmos altamente complejos. Para resolver este tipo de problemática se han utilizado gran cantidad de métodos. Existen trabajos que se basan en el uso de redes neuronales, en donde se alcanzan precisiones del orden de 95% en la detección [8] o sistemas mixtos que hacen uso de un filtro Gabor Wavelet en la primera etapa y luego, en una segunda instancia, trabajan con redes neuronales artificiales para realizar la clasificación. De lo anterior se infiere fácilmente que con el uso de un procesador más veloz, el tiempo

de procesamiento disminuye y la velocidad de avance aumenta. También se desarrollaron líneas basadas en un concepto muy diferente. En vez de utilizar imágenes de alta resolución y realizar un reconocimiento individual de cada planta; se utilizan imágenes de baja resolución que abarcan grandes áreas y se detecta la densidad de malezas en tiempo real [9]. Al igual que el método anterior este requiere de una CPU capaz de realizar una gran cantidad de cálculos en forma intensiva debido a la complejidad de los algoritmos aplicados al filtrado y detección. Esto implica la utilización de, por ejemplo, un DSP muy potente para poder realizar estos cálculos en tiempo real. El inconveniente que presenta está dado por la alta complejidad de los algoritmos que redundan en el uso de procesadores de alto rendimiento y elevado costo. Entendemos que este sistema seguirá siendo perfeccionado y constituirá en el futuro la base de la detección de malezas mediante visión artificial, probablemente formando parte de un sistema mixto.

El proyecto plantea la utilización de una cámara digital color y luego mediante el procesamiento digital de imágenes detectar la existencia o no de malezas en barbecho. Las técnicas de procesamiento de imágenes han sido utilizado con éxito para muchas aplicaciones en la agricultura y muestran un enorme potencial para la detección de malezas, ya que utiliza no sólo información espectral, sino también la información espacial y de textura.

La metodología utilizada y los resultados obtenidos serán susceptibles de ser publicados y presentados en eventos científicos tecnológicos.

El aporte del proyecto en cuanto al impacto del mismo sobre la sociedad en general, es que el desarrollo de un sistema de este tipo, permitiría reducir fuertemente la utilización de herbicidas como así también en lograr un pequeño aporte para hacer sustentable la tarea diaria en el ámbito agropecuario respetando nuestra naturaleza. Es decir el beneficio al medio ambiente, resultará en una vía para mitigar los problemas de la agricultura a gran escala y hacer más sustentable esta actividad económica.

Línea de Investigación y Desarrollo

Nuestro proyecto consta de los siguientes ejes:

1. Determinación de las necesidades de hardware y software que se requieren para la detección de malezas sobre barbecho en tiempo real.
2. Estudio de los dispositivos que pueden ser integrados en el sistema, en base a la disponibilidad y oferta en Argentina.
3. Definición de la estrategia de procesamiento digital de imágenes a utilizarse en función del objetivo planteado (detección de malezas en barbecho).
4. Desarrollo de software de control y hardware para el manejo de las válvulas de accionamiento.

Objetivos y Resultados

OBJETIVO GENERAL: desarrollar un sistema de procesamiento digital de imágenes en tiempo real, que permita la detección de malezas en barbecho agregando elementos de interés social y comunitario como puede ser el manejo eficiente de los recursos y el respeto a nuestra naturaleza en general.

OBJETIVOS ESPECÍFICOS: diseñar un sistema de procesamiento digital de imágenes que cumpla con los siguientes aspectos:

- Desarrollar un sistema que permita la detección de malezas a partir del procesamiento digital de imágenes.
- Utilizar la información brindada por el detector de malezas para la dosificación variable del herbicida.
- Comparar los resultados del prototipo con otros sistemas de detección.

Se desea que la propuesta de diseño de este sistema sea un primer paso para generar lineamientos de utilización de sistemas de procesamiento digital de imágenes en el ámbito agropecuario.

RESULTADOS: considerando que el proyecto se encuentra en su fase inicial, los resultados que se esperan, se pueden resumir como:

- ✓ Detección de las distintas especies de malezas que afectan el período previo a la siembra en el cultivo de soja (barbecho), caracterización de las mismas de acuerdo a sus porcentajes de incidencia en la región.

- ✓ Detección del método de visión artificial (estrategia de procesamiento) que más se adapte a las especies locales.
- ✓ Desarrollo del sistema de control de pulverización de herbicida sobre las malezas en barbecho.
- ✓ Diseñar un software amigable con el usuario.
- ✓ Planificar clases prácticas en asignaturas y cursos relacionados con la tecnología utilizada en la agricultura de precisión y la automatización en general, a fin de complementar la enseñanza en las carreras de grado de Ingeniería Electrónica, Mecánica, e Informática.
- ✓ Realizar transferencia de tecnología a empresas del sector, desde los mismos grupos de investigación en distintas regiones (centro a través de la UTN San Francisco y sur del país a través de la UTN Villa María).

Formación de Recursos Humanos

El director del proyecto, Mg. Ing. Costamagna Marcelo realizó la Maestría en Ingeniería en Calidad, en la Facultad Regional Villa María. En la actualidad es profesor Titular concursado en dicha Facultad, docente de Ingeniería Mecánica II, Metrología e Ingeniería en Calidad y Proyecto Final, integrante del Grupo de Estudios en Calidad en Mecatrónica, posee publicaciones en congresos relacionadas con el área Calidad y sistemas Mecánicos.

El Codirector del proyecto, Mg. Ing. Peretti Gastón Carlos realizó la Maestría en Ciencias de la Ingeniería, mención Telecomunicaciones en la Universidad Nacional de Córdoba. En la actualidad es docente de la carrera de Ing. Electrónica de la Universidad Tecnológica Nacional Facultad Regional San Francisco, en la cátedra de Dispositivos Electrónicos y posee publicaciones en congresos y libros referidas a las líneas de investigación sistemas de control y comunicaciones. El Ing. Felissia Sergio, se encuentra actualmente cursando la Maestría en Sistemas de Control aplicado en la UTN Facultad Regional Paraná y es docente de la misma universidad en las cátedras de de Informática I, Sistemas de Control Aplicado y Control de Procesos. El grupo de investigación también está conformado por alumnos avanzados de

la carrera Ingeniería Electrónica de la Facultad Regional San Francisco y de la Facultad Regional Villa María.

El impacto esperado del proyecto se basa en la fundación de un grupo de trabajo interdisciplinario orientado al control y los sistemas de comunicaciones con un objetivo único y general que es el de generar conciencia en el manejo eficiente de los recursos naturales en el ambiente agropecuario. Respecto al potencial humano que conforma el grupo de trabajo, está formado por docentes y alumnos de la Universidad Tecnológica Nacional de Villa María y Universidad Tecnológica Nacional Facultad Regional San Francisco (Pcia. de Córdoba).

Referencias

- [1] Vrindts, E. and J. De Baerdemaeker, "Feasibility of weed detection with optical reflection measurements", U.K. : British Crop Protection Council, 1996.
- [2] Moltoni A. F. and Moltoni L. A., "Análisis económico de la implementación de un sistema de aplicación selectiva de herbicidas", Buenos Aires: Ediciones INTA, 2005.
- [3] Shropshire, G. J., K. Von Bargen, and D. A. Mortensen, "Optical reflectance sensor for detecting plants", Bellinham, 1990.
- [4] Biller R. H., "Reduced input of herbicides by use of optoelectronic sensors", 1998.
- [5] Lee W. S. and D. C. Slaughter, "Robotic weed control system for tomatoes". Precision Agriculture, Vol. (1):95-113,1999.
- [6] Pérez A. J., López F., Benlloch J. V. and S. Christensen, "Colour and shape analysis techniques for weed detection in cereal fields. Computers and Electronics in Agriculture", 2000.
- [7] Meyer G. E., Mehta T., Kocher M. F., Mortensen D. A. and A. Samal, "Textural imaging and discriminant analysis for distinguishing weeds for spot spraying Transactions of the American Society of Agricultural Engineers", 1998.
- [8] Burks T. F., Shearer S. A., Gates R. S. and K. D. Donohue, " Backpropagation neural network design and evaluation for classifying weed species using color image texture Transactions of the American Society of Agricultural Engineers", 2000.
- [9] Tian L., "Development of a sensor-based precision herbicide application system". Computers and Electronics in Agriculture, 2002.

Implementaciones de GNSS RTK en Sistemas Embebidos Autónomos

José Hipólito Moyano Karina Cenci Jorge R. Ardenghi
 Laboratorio de Investigación en Sistemas Distribuidos
 Laboratorio de Investigación en Ingeniería de Software y Sistemas de Información
 Departamento de Ciencias e Ingeniería de la Computación
 Universidad Nacional del Sur
 Bahía Blanca - Buenos Aires - Argentina
 e-mail: {jose.moyano, kmc, jra} @cs.uns.edu.ar

RESUMEN

Esta línea de investigación tiene como objetivo general analizar y comparar diferentes tecnologías existentes en el área de *Global Navigation Satellite Systems* (GNSS), buscando destacar las ventajas y desventajas ofrecidas. Como objetivo particular, se buscará realizar el análisis y el desarrollo de GNSS de alta precisión, con dispositivos electrónicos de bajo costo, valiéndose de la biblioteca *Real Time Kinematic Library* (RTKLIB), la cual implementa algoritmos de *Real Time Kinematic* (RTK) para cualquier dispositivo embebido.

Palabras clave: RTK, high-precisión, GPS, GNSS, Sistemas Embebidos.

CONTEXTO

Esta línea de investigación y desarrollo se lleva adelante en el ámbito del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur (UNS). En particular, como parte de las tareas que se realizan en el Laboratorio de Investigación en Sistemas Distribuidos (LISIDI) y en Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información (LISSI). El proyecto se financia parcialmente con fondos asignados por la UNS a proyectos de investigación.

INTRODUCCIÓN

En función de cumplir el objetivo para el que fueron diseñados, los sistemas embebidos integran cuatro tipos de recursos: actuadores, sensores, comunicación e interfaces. Para muchas de las aplicaciones de estos sistemas, principalmente aquellas que implican

desplazamiento de un nodo o agente, es necesario incorporar sensores que permitan conocer la localización en el espacio. Los sensores predilectos para georreferenciar, es decir, posicionar al sistema en un mapa terrestre, son los que miden una o varias de las constelaciones de GNSS.

La medición GNSS logra precisiones del orden de metros. Se puede tomar, por ejemplo, el GNSS desplegado por USA, *Navigation System with Timing and Ranging, Global Positioning System* (NAVSTAR-GPS) popularmente conocido como GPS. Este sistema, en condiciones óptimas, logra una precisión que oscila los 7 metros [USGov08]. En situaciones promedio y peor caso, se considera entre 12 y 30 metros [USGov08].

Esta precisión resulta insuficiente para la mayoría de las aplicaciones, por lo que generalmente se utilizan sensores adicionales y estrategias algorítmicas para asistir y mejorar la georeferencia del sistema, como la fusión de datos de mapas, las técnicas de *dead reckoning* mediante sensores y procesamiento de imágenes [Mou13, GBR16].

Existen distintas técnicas para mejorar la precisión del sistema de posicionamiento GNSS, por ejemplo, *Real Time Kinematic* (RTK), *Precise Point Positioning* (PPP) o *Dinamic Global Positioning System* (DGPS). La técnica de RTK distingue dos elementos en el sistema de posicionamiento. Una estación base (*base*) que, en función de su posición conocida y la onda portadora de la señal del satélite, genera y transmite correcciones en tiempo real en los modelos de estimación, permitiendo a un nodo móvil (*rover*) corregir su propia estimación de georreferenciamiento a partir de la información recibida. La técnica de RTK mejora a los sistemas de GNSS hasta

lograr precisiones del orden de centímetros. Estos sistemas se denominan GNSS de alta precisión.

A medida que el *rover* se aleja de la *base*, la información de corrección pierde relevancia, y el error en el posicionamiento aumenta. En países desarrollados, se cuenta con programas e inversiones lo suficientemente grandes, como para instalar estaciones base que emiten información de corrección a quien la requiera, construyendo redes de estaciones base RTK. Lo que permite la red de *bases* es ampliar el territorio de alcance del servicio de RTK, dando continuidad al posicionamiento de alta precisión en territorios amplios. A medida que los *rovers* se desplazan en el terreno, cambian la fuente de la información de corrección a la *base* más cercana. Estas redes son conocidas como *Continuos Operating Reference Stations* (CORS) y son piezas fundamentales en el análisis topográfico y mapeo de extensiones nacionales. Las CORS permiten los análisis del ecosistema, proyectar la evolución de la expansión poblacional y la reacción ante emergencias y catástrofes [GSum15].

Los circuitos integrados (IC) GNSS comerciales que implementan el ajuste de su posición a partir de información de corrección tienen un costo levemente superior al de un IC de GNSS convencional. Sin embargo, las estaciones base que se ofrecen en el mercado tienen un costo prohibitivo (mayores a los 15.000 USD).

LÍNEAS DE DESARROLLO E INVESTIGACIÓN

La línea de investigación tiene como objetivo general el análisis y la comparación de las distintas tecnologías de GNSS, centrado en GNSS de alta precisión y en la técnica de RTK, implementada sobre sistemas embebidos.

Se buscará analizar la factibilidad y exactitud de un sistema de GNSS de alta precisión, implementado sobre dispositivos embebidos comerciales de bajo costo, basados en RTKLIB.

Como parte de la factibilidad, será necesario un análisis de las necesidades de

procesamiento, y de la capacidad ofrecida por un sistema embebido promedio, y como estas se ajustan a las necesidades de RTK.

Un objetivo particular será diseñar y desarrollar sistemas RTK donde sus nodos tengan la propiedad de convertirse de *rovers* a *bases*, y de *bases* a *rovers*. Con esta función, en lugar de contar con una estación base que nuclea toda la información de corrección en una zona, es posible construir una red dinámica de estaciones *base* y *rover*. Un nodo que comienza operando como *rover*, puede posicionarse en un punto alejado de la base, detenerse, y utilizando la información de corrección, determinar su posición con precisión. Una vez detenido, puede convertirse en una nueva *base*, emitiendo información de corrección en una nueva zona, ampliando el rango de cobertura de RTK.

RESULTADOS ESPERADOS

En un principio, se construirá un ambiente de desarrollo para las pruebas de experimentación. Estará conformado por un sistema embebido basado en Linux, sobre una placa de desarrollo que sea soportada por *Yocto Project* (una plataforma que permite construir sistemas basados en Linux a medida de un dispositivo embebido). A este sistema operativo se le dará soporte RTKLIB, compilando las bibliotecas con el *toolchain* correspondiente a la plataforma. La placa de desarrollo tendrá una conexión con un IC GNSS, que cuente con entradas para información de corrección RTK. Preferentemente se seleccionará un IC que soporte más de una constelación GNSS, para ampliar el rango de posibilidades de prueba.

Los experimentos se llevarán a cabo con implementaciones de software cruzadas, que se puedan ejecutar en el ambiente de desarrollo y recopilar datos, para ser contrastados con información topográfica conocida, tanto de posición geográfica como de altitud.

El ambiente de desarrollo con soporte RTKLIB, será el prototipo de un nodo *base* de un sistema RTK.

Además, se trabajará sobre la posibilidad de que cada nodo tenga la función dual y

dinámica de convertirse en *rover*. Con esta capacidad, sería factible crear redes dinámicas, que puedan expandirse o contraerse al cambiar la localización de las *bases*, para ajustarse a las necesidades del terreno y la aplicación.

Por la naturaleza de la implementación (costo, consumo, capacidad de cálculo), la precisión de estos sistemas puede no llegar a ser la misma que la de un sistema comercial costoso, pero sí muy superior a un sistema sin corrección, llegando a una precisión menor a los diez centímetros, útil para la mayoría de las aplicaciones [TY09, MOA15, WBM13].

Sin embargo, es importante considerar que el error de un sistema GNSS de alta precisión aumenta, al aumentar la distancia entre el base. Dado el bajo costo de las estaciones base propuestas, se espera poder plantear la creación de una red de estaciones base, conectadas entre sí en forma de *mesh*, permitiendo a cualquier *rover* conectarse siempre a la estación base más cercana, evitando una degradación significativa. O detenerse, y sumarse a la red de bases, para proveer servicio a otros *rovers*.

Una red de estaciones base de bajo costo combina los beneficios de alta precisión, con amplitud de alcance. Por otro lado, la dualidad de funcionamiento (*base/rover*) permite proyectar redes dinámicas de información de corrección, que puedan cambiar su forma para ocupar distintas áreas de terreno.

La solución obtenida permitiría a países en vías de desarrollo, contar con una solución de GNSS de alta precisión a bajo costo, que aún con precisiones no tan buenas como los sistemas comerciales existentes, permitiría resolver la mayoría de las necesidades de georreferenciamiento de precisión.

FORMACIÓN DE RECURSOS HUMANOS

En relación con la formación de los recursos humanos, los resultados de la investigación serán utilizados como parte del desarrollo de una tesis de Magister en Ciencias de la Computación. Además, esta línea de investigación permitirá la dirección de tesis y

trabajos finales de alumnos de grado. Asimismo, se podrá generar *courseware* para materias optativas a dictarse en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

BIBLIOGRAFÍA

[USGov08] GPS Standard Positioning System (SPS) Performance Standard, September 2008, US Government, 4th Ed.

[TY09] Takasu, T., & Yasuda, A. (2009, November). Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. In international symposium on GPS/GNSS (pp. 4-6). Jeju, Korea: International Convention Centre.

[MOA15] Matias, B., Oliveira, H., Almeida, J., Dias, A., Ferreira, H., Martins, A., & Silva, E. (2015, May). High-accuracy low-cost RTK-GPS for an unmanned surface vehicle. In OCEANS 2015-Genova (pp. 1-4). IEEE.

[WBM13] Wiśniewski, B., Bruniecki, K., & Moszyński, M. (2013). Evaluation of RTKLIB's Positioning Accuracy Using low-cost GNSS Receiver and ASG-EUPOS. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 7(1), 79-85.

[GSAge15] GSA, European GNSS Agency. GNSS Market Report. Issue 4. Luxembourg: Publications Office of the European Union, March 2015.

[KH05] Kaplan, E., & Hegarty, C. (2005). *Understanding GPS: principles and applications*. Artech house.

[SB11] Stempfhuber, W., & Buchholz, M. (2011, January). A precise, low-cost RTK GNSS system for UAV applications. In *Conference on Unmanned Aerial Vehicle in Geomatics, Zürich* (pp. 289-293).

[PHH14] Pesyna Jr, K. M., Heath Jr, R. W., & Humphreys, T. E. (2014). Centimeter positioning with a smartphone-quality GNSS

antenna. Proceedings of the ION GNSS, Tampa, FL.

[And12] Andrei, C. O. (2012, June). Cost-effective precise positioning using carrier phase navigation-grade receiver. In 2012 International Conference on Localization and GNSS (pp. 1-6). IEEE.

[GSum15] Report from the 2015 Geospatial Summit on Improving National Spatial Reference System.

[TYN14] Ryosuke Takai and Liangliang Yang and Noboru Noguchi. Development of a crawler-type robot tractor using RTK-GPS and IMU Engineering in Agriculture, Environment and Food, vol 7, no. 4, pp 143-147, 2014.

[TMFC15] Deodato Tapete, Stefano Morelli, Riccardo Fanti and Nicola Casagli. Localising deformation along the elevation of linear structures: An experiment with spaceborne InSAR and RTK GPS on the Roman Aqueducts in Rome, Italy. Applied Geography, vol 58, pp 65-83, 2015

[Hon12] Hongtao Xu. Application of GPS-RTK Technology in the Land Change Survey. Procedia Engineering In 2012 International Workshop on Information and Electronics Engineering.

[PP17] Jacek Paziewski and Pawel Wielgosz. Investigation of some selected strategies for multi-GNSS instantaneous RTK positioning. Advances in Space Research 2017.

[Ox17] Oxley, Alan. Uncertainties in GPS Positioning – Academic Press, pp 19-38 (2017)

[OFHS06] Manabu OMAE, Takehiko FUJIOKA, Naohisa HASHIMOTO and Hiroshi SHIMIZU. The application of RTK-GPS and Steer-By-Wire technology to the automatic driving of vehicles and an evaluation of driver behavior. 2006. IATSS Research.

[WKMBL07] René Warnant, Ivan Kutiev, Pencho Marinov, Michael Bavier and Sandrine Lejeune (2007). Ionospheric and geomagnetic conditions during periods of degraded GPS position accuracy: 2. RTK events during disturbed and quiet geomagnetic conditions. Advances in Space Research.

[Mou13] Hossein Mousazadeh. A technical review on navigation systems of agricultural autonomous off-road vehicles. 2013 Journal of Terramechanics.

[GBR16] Dominique Gruyer, Rachid Belaroussi and Marc Revilloud (2016). Accurate lateral positioning from map data and road marking detection. Expert Systems with Applications.

[MN13] Mustafa Berber and Niyazi Arslan (2013). Network RTK: A case study in Florida. Measurement vol. 46.

Modelos y Métodos Computacionales en Ingeniería (a17)

Javier Giacomantone¹, Oscar Bria¹, Luciano Lorenti¹, Armando De Giusti^{1,2}

¹Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – UNLP

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

La Plata, Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

{jog, onb, llorenti, degiusti}@lidi.info.unlp.edu.ar

Resumen

Este trabajo describe una línea de investigación y desarrollo (I/D) y los resultados esperados de la misma. El objetivo principal es analizar, generar y evaluar modelos matemáticos y métodos computacionales asociados en ingeniería. Se estudian problemas que por su nivel de complejidad requieren soluciones específicas. Los modelos abordados son dependientes del tipo de sistema estudiado, del fenómeno analizado y del área particular de ingeniería que originó el requerimiento. Determinar el tipo de sistema, el método para evaluar su rendimiento y las soluciones numéricas óptimas o sub-óptimas forma parte de los objetivos en esta línea de I/D.

Palabras Clave: métodos computacionales, análisis de sistemas, cálculo numérico, reconocimiento estadístico de patrones, análisis de imágenes, aprendizaje estadístico.

Contexto

Esta línea de I/D forma parte del proyecto 11/F018: “Arquitecturas multiprocesador en HPC: Software de base, Métricas y Aplicaciones” acreditado por el Ministerio de Educación.

1. Introducción

Modelar un problema en ingeniería requiere estudiar los detalles del proceso, del fenómeno, o del sistema que se pretende analizar o diseñar. Es necesaria la construcción de modelos matemáticos, métodos computacionales y técnicas numéricas que ofrezcan soluciones viables [1][2]. Las magnitudes que intervienen, escalares o vectoriales, eventualmente, cambian en el tiempo o espacio. Los sistemas estudiados pueden ser lineales o no lineales requiriendo modelos sofisticados [3][4]. Determinar si una solución es viable con fundamento científico, es una tarea compleja y dependiente del problema particular analizado [5][6]. Este proyecto tiene como primer objetivo analizar y proponer modelos computacionales, métodos y soluciones particulares derivadas de los mismos. Alcanzar los objetivos anteriores requiere estudiar los fundamentos que subyacen a cada modelo, evitando soluciones, que por su nivel de encapsulamiento, limiten una verdadera comprensión y abordaje científico de las mismas [7]. El tipo de problema de interés en esta línea de I/D requiere la integración de soluciones de tres áreas, ciencias de la computación, matemáticas aplicadas y un área de ingeniería o ciencia básica en particular. La evaluación de rendimiento es un aspecto fundamental para poder validar las soluciones propuestas o los modelos analizados [8]. Por lo tanto, otro aspecto fundamental de esta línea de I/D es el estudio de las métricas y paradigmas de desempeño en sistemas específicos.

En la sección 2 se presenta un breve resumen de los temas de I/D específicos en el período actual. La sección 3 enumera resultados obtenidos y esperados. Finalmente, la sección 4 resume los objetivos con respecto a la formación de recursos humanos en el contexto de esta línea de I/D.

2. Líneas de Investigación

2.1 Reconocimiento de patrones

El trabajo actual lo podemos clasificar en tres tópicos principales bien diferenciados correspondientes a sub-disciplinas dentro de reconocimiento estadístico de patrones [9][10][11]. La primera sub-disciplina es clasificación supervisada donde el énfasis de nuestro trabajo se centra en el estudio de métodos de clasificación basados en núcleos dispersos, en particular máquinas de soporte vectorial [12][13].

La segunda es clasificación no supervisada donde la principal línea de trabajo son las técnicas de agrupamiento. Actualmente con énfasis en detección de valores atípicos y métodos basados en teoría espectral de grafos [14][15][16]. El tercer tópico de fundamental importancia es el de reducción de dimensión en particular selección de características [17].

2.2 Análisis de Imágenes

Se estudian métodos de segmentación estáticos y modelos deformables. En particular en imágenes 2 ½ D obtenidas por cámaras de tiempo de vuelo, resonancia magnética funcional y otras modalidades con estructuras de datos similares [18][19]. El objetivo principal es mejorar la calidad de los descriptores obtenidos en función de su impacto en el sistema de clasificación [20][21][22]. En el caso de modelos deformables es posible estudiar el comportamiento temporal y medir magnitudes indirectamente. Se analizan métodos de generación de características a partir de señales en general y de imágenes digitales en particu-

lar de rango e intensidad. Se abordan sistemas de análisis de pseudo-imágenes a partir de campos vectoriales, series temporales, una modalidad particular de imagen o fusión de modalidades [23][24][25].

2.3 Desempeño de Sistemas de Posicionamiento, Navegación y Localización.

En los sistemas de posicionamiento, de navegación y de localización [26][27], el concepto de desempeño excede al habitual que está limitado a la calidad nominal de la estimación de ubicación y eventualmente a la confiabilidad [28][29]. En estos sistemas deben considerarse además los parámetros de integridad y continuidad que le garantizan al usuario que la información proporcionada por el sistema es correcta para que una operación crítica pueda realizarse en forma segura [30][31].

La integridad y la continuidad dependen en gran medida de la aplicación y del entorno específico y su aseguramiento afecta a otro parámetro de desempeño del sistema que es la disponibilidad [32].

Esta línea de trabajo se avoca al estudio de problemas puntuales de desempeño en los sistemas mencionados, utilizando criterios y métodos diversos de modelado, procesamiento y análisis [33][34][35][36].

3. Resultados y Objetivos

3.1 Resultados publicados recientemente

- Se estudiaron y propusieron métodos para detección en series temporales de fMRI [37][38].
- Se desarrollaron métodos de segmentación de imágenes de rango y supresión del plano de fondo [39][40][41].
- Se analizaron y propusieron alternativas para el agrupamiento de objetos en video [42].

3.2 Objetivos generales

- Desarrollar modelos y optimizar algoritmos particulares de clasificación supervisada y no supervisada.
- Evaluar métodos de análisis de desempeño y su aplicación sobre los clasificadores y conjuntos de datos particulares.
- Evaluar la monitorización de la integridad de los sistemas de ayuda a la navegación aérea basados en sistemas GNSS (Global Navigation Satellite Systems).
- Estudiar métodos de selección y extracción de características.
- Investigar modelos y métodos en procesamiento y análisis de imágenes. Kernels y algoritmos de optimización.

4. Formación de Recursos Humanos

En esta línea de I/D hay dos investigadores realizando su doctorado en temas relacionados. Se dictan asignaturas optativas en Ingeniería en Computación y cursos de postgrado que tienen por objetivo formar alumnos en temas específicos y fundamentos. Los alumnos tienen la posibilidad de realizar tesinas en esta línea. Además se promueve la interacción con otros grupos y líneas de I/D resultando en un mecanismo de permanente consulta y transferencia.

BIBLIOGRAFIA

1. Zalizniak V. Essentials of Scientific Computing – Numerical Methods for Science and Engineering. Woodhead Publishing, 2008.
2. Juergen G. Coupled Systems: Theory, Models, and Applications in Engineering. CRC, 2014.
3. Torokhti A., Howlett P. Computational Methods for Modelling of Nonlinear Systems. Elsevier, 2007.
4. Canuto C., et al. Spectral Methods. Evolution of Complex Geometries and Applications to Fluids Dynamics Scientific Computation. Springer, 2007.
5. Artlich S., et al. Scientific Computing in Chemical Engineering Springer, 1996.
6. Ciurpina G. Scientific Computing in Electrical Engineering Springer, 2007.
7. Gustafsson B. Fundamentals of Scientific Computing Springer, 2011.
8. Aslak T., et al. Elements of Scientific Computing. Springer, 2010.
9. Batagelj V, Bock H, Ferligoj A. Data Science and Classification. Springer, 2006.
10. Fukunaga K. Introduction to Statistical Pattern Recognition. Second Edition. Academic Press, 1990.
11. Devijer P, Kittler, J. Pattern Recognition: theory and applications. Springer, 1986.
12. Cortes C, Vapnik V, Support vector networks. Machine Learning v.20, pp.273- 297, 1995.
13. Vapnik, V. The Nature of Statistical Learning Theory. N. Y. Springer, 1995.
14. Von Luxburg U. A Tutorial on Spectral Clustering. Statistics and Computing, 17(4), 2007.
15. Shi J., Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 888-905, 2000.
16. Yang, P.; Huang, B. An Outlier Detection Algorithm Based on Spectral Clustering. Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol.1, pp. 507-510, 2008.
17. Aytug H. Feature selection for support vector machines using Generalized Benders Decomposition. European Journal of Operational Research, 244(1), 210-218, 2015.
18. Kim H.Y., Giacomantone J. O., Cho, Z. H. Robust Anisotropic Diffusion to Produce Enhanced Statistical Parametric Map, Computer Vision and Image Understanding, v.99, pp.435-452, 2005.
19. Chaovalitwongse W., Pardalos P. On the Time Series Support Vector Machine using Dynamic Time Warping Kernel for Brain Activity Classification, Cybernetics and Systems Analysis v.44 pp.125-138, 2008.
20. Anke Meyer-Base. Pattern Recognition for Medical Imaging. Academic Press, 2004.

21. Han Y., Feng X., Baciú G. Variational and PCA based natural image segmentation. *Pattern Recognition* 46, pp. 1971-1984, 2013.
22. Li S., Fevens L., Krzyzak A., Li S. Automatic Clinical Image Segmentation Using Pathological Modelling, PCA and SVM, *MLDN, LNAI 3587* pp.314-324, 2005.
23. Boekhorst R., Abnizova I., Wernich L. Discrimination of regulatory DNA by SVM on the basis of over and under-represented motifs, *ESANN* pp. 481-486, 2008.
24. Rüping S., SVM kernels for time series analysis, *G1-Workshop-Woche Lernen-Lehren-Wissen-Adaptivitet*, pp.43-50, 2001.
25. Yang K., Shahabi C. A pca-based kernel for kernel pca on multivariate time series, *IEEE Intern. Conf. on Data Mining*, 2005.
26. Partap Misra, Per Enge. *Global Positioning System: Signals, Measurements and Performance*, Ganga-Jamuna Press, 2010.
27. Hakan Koyuncu, Shuang Hua Yang. A Survey of Indoor Positioning and Object Locating Systems Indoor Positioning System, *International Journal of Computer Science*, 2010.
28. Petevelo Mark. Quantifying the performance of Navigation Systems and Standars for assisted-GNSS, *Inside GNSS*, 2008.
29. Morurikis A., Roumeliotis S. Performance Analysis of Multirobot Cooperative Localization, *IEEE*, 2005.
30. Murphy T., et. al., Fault Modeling for GBAS Airworthiness Assessments, *Navigation*, 2012.
31. Cosmen-Schortmann J., Azaola-Sáenz, Martínez-Olagüe M. A., Toledo-López M., Integrity in Urban and Road Environments and its use in Liability Critical Applications, *IEEE*, 2008.
32. Shuo-Ju Yeh1, Shau-Shiun, GBAS airport availability simulation tool, *GPS Solutions*, 2015.
33. Pengfei Duan, Maarten Uijy De Haa. Flight Test Results of a Measurement-Based ADS-B System for Separation Assurance, *Navigation*, 2013
34. Sam Pullen, Todd Walter, Per Enge. SBAS and GBAS Integrity for Non-Aviation Users: Moving Away from "Specific Risk," *International Technical Meeting of The Institute of Navigation*, 2011.
35. Zhipeng Wang, Christophe Macabiau, Jun Zhang, Anne-Christine Escher. Prediction and analysis of GBAS integrity monitoring availability at LinZhi airport, *Springer*, 2012.
36. Rinnan A., Gundersen N., Sigmond M., Nilsen J. K. *Operational GNSS Integrity*, *Dynamic Positioning Conference*, 2011.
37. Giacomantone J., Tarutina T. Diffuse Outlier Detection Technique for Functional Magnetic Resonance Imaging. *Computer Science and Technology Series. XVI Argentine Congress of Computer Science Selected Papers*. pp. 255-265, 2011.
38. Giacomantone J., De Giusti A. Detección de áreas de interés bajo la hipótesis de relación espacial de voxels activados en fMRI. *XX Congreso Argentino de Ciencias de la Computación*. San Justo. Argentina, 2014.
39. Lorenti L., Giacomantone J. Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo. *XI Workshop Computación Gráfica, Imágenes y Visualización*. pp. 430-439. Mar del Plata, Argentina, 2013.
40. Lorenti L., Giacomantone J. Time of flight image segmentation through co-regularized spectral clustering. *Computer Science & Technology Series. XX Argentine Congress of Computer Science. Selected papers*. La Plata, Editorial de la Universidad Nacional de La Plata, 2015.
41. Giacomantone J., et al. Supresión del plano de fondo en imágenes de tiempo de vuelo. *VII Workshop Procesamiento de Señales y Sistemas de Tiempo Real*, 2016.
42. Lorenti L., Giacomantone J., De Giusti A. Agrupamiento de trayectorias vía clustering espectral incremental. *XXII Congreso Argentino de Ciencias de la Computación*, pp. 222-231, 2016.

Monitor de Profundidad Anestésica

Coulombie, Diego; Orthusteguy, Fernando; Reyes, Agustín Ignacio; Ortalda, Federico; Nieva, Andres Edgardo; Fernández Nicolás Gonzalo; Rodríguez, Alejandro Edgar; Ricardo Ariel Quintana / Departamento de Ingeniería e Investigaciones Tecnológicas / Universidad Nacional de La Matanza

RESUMEN

El estado anestésico de un paciente no se puede evaluar directamente. Tradicionalmente se evalúan variables hemodinámicas. Estos parámetros reflejan de manera no objetiva, indirecta y tardía el estado del sistema nervioso central (SNC), pero su indicación numérica resulta útil para la toma de decisiones en el ámbito quirúrgico. Por otro lado la incidencia del fármaco sobre el SNC puede verse en un electroencefalograma, pero la técnica y dinámica de este estudio no se corresponde con la dinámica necesaria para el monitoreo de la profundidad de anestesia.

El objetivo principal es obtener un método de monitoreo de profundidad anestésica, definir sus características técnicas y tecnológicas, establecer una valoración de los riesgos asociados, determinar el grado de factibilidad para lograr su fabricación en nuestro país cumpliendo con el marco regulatorio actual, y de tener éxito en todas las etapas desarrollar un prototipo con miras a su transferencia a la aplicación industrial.

Los resultados alcanzados durante el primer año del proyecto son avances parciales o totales sobre las tareas que corresponden a la definición del funcionamiento esperado, el funcionamiento esencial y el planteo del análisis de riesgo, la técnica de aplicación en el quirófano, las necesidades ergonómicas del paciente y del profesional usuario.

Palabras clave:

Despertar intraoperatorio, neuromonitoreo, tecnología médica, sustitución de importaciones.

CONTEXTO

Existe un particular interés del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad nacional de La Matanza para llevar adelante proyectos de Ingeniería

Biomédica, a modo de basamento para establecer una futura carrera de grado esta área temática.

El proyecto está financiado en 2 tramos, uno con fondos propios de la universidad mediante el proyecto PROINCE C-196 y con recursos aportados por la Agencia nacional de Promoción Científica y Tecnológica mediante el proyecto PICTO 0086.

1. INTRODUCCIÓN

La misión del anestesiólogo en el quirófano es conseguir que el paciente no resulte afectado por la agresión que implica la cirugía empleando fármacos que permiten alcanzar un cierto nivel de protección. La capacidad de ajustar la dosis de agentes hipnóticos a las necesidades de cada paciente en cada momento de la intervención, mejora los resultados globales de las intervenciones quirúrgicas. La posibilidad de que el paciente esté consciente durante el desarrollo de la intervención quirúrgica, es lo que se denomina despertar intraoperatorio (DIO), o en inglés awareness. La incidencia de DIO en la población quirúrgica mundial está dentro del 0.13 %. Si bien es un número bajo, la cantidad de intervenciones que se realizan a diario representan un gran número de afectados. Las graves consecuencias sobre quienes lo padecen pueden requerir asistencia psicológica y tratamiento continuado durante años. Considerando esto, es comprensible que los sistemas de salud busquen disminuir la incidencia de DIO. Una solución es monitorizar el estado de profundidad anestésica durante la cirugía. Varias investigaciones demostraron que la incidencia de DIO disminuía en un 80% en el grupo monitorizado [1].

También En la Unidad de Cuidados Intensivos (UCI) se usa monitoreo para el control de la hipnosis y sedación, y ha

demostrado que es muy frecuente la sedación excesiva en los pacientes. Se estudió su utilidad como predictor de recuperación de la consciencia en pacientes con patología cerebral grave que persisten en estado de coma tras la retirada de la sedación. Se estableció una relación entre factores de profundidad de anestesia y la supervivencia de los pacientes tras la intervención quirúrgica. El tiempo transcurrido con valores por debajo de cierto nivel mínimo se correlaciona directamente con un aumento de la mortalidad transcurrido un año desde la intervención [2]. El monitoreo reduce un 20% el consumo de hipnóticos, lo que representa en menores costos directos quirúrgicos. Luego de la cirugía los pacientes se despiertan un 40% más rápido y llegan mejor al área de despertar desde donde pueden ser dados de alta con mayor precocidad, reduciendo costos de internación postoperatoria [3].

Aún con tantas ventajas para los pacientes y para el sistema de salud, el monitoreo de profundidad anestésica no es de uso masivo en los quirófanos argentinos. Los factores para analizar esta carencia son el costo de los insumos importados, la falta de una validación clínica que considere a las asociaciones de anestesiología locales y la imposibilidad de los importadores locales de garantizar la disponibilidad de insumos, repuestos y servicio técnico [4]. No existe un dispositivo de monitoreo de profundidad de anestesia de desarrollo y fabricación nacional. Generar localmente este tipo de equipamiento redundaría en un aporte de la soberanía tecnológica, disminuyendo en más de un 60% los costos de importación y fomentando la industria nacional de alta tecnología médica, plausible de ser exportada.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación y desarrollo competen a la adquisición y procesamiento de bioseñales neurológicas, la investigación sobre técnicas de monitoreo y los conceptos

dirigidos a la industrialización de este tipo de tecnologías.

3. RESULTADOS OBTENIDOS/ESPERADOS

Por tratarse de un proyecto de investigación que incluye el desarrollo de un eventual producto tecnológico y que en su justificación se menciona la sustitución de importaciones como un hecho relevante, el proyecto tuvo énfasis en la producción de material dedicado a la transferencia tecnológica. Como resultados parciales se han generado documentos necesarios para las etapas de industrialización del dispositivo a transferir. Al tratarse de un producto médico con un mercado regulado por el ministerio de salud, se deben cumplir con ciertas pautas normalizadas o reglamentos técnicos para la homologación del mismo. Estas pautas están fundamentalmente destinadas a probar la seguridad y eficacia del producto mediante el aporte de registros, análisis y ensayos que nacen desde el inicio mismo del diseño. El desarrollo de esta investigación se basa en el diseño de un dispositivo que de resultar exitoso podrá ser transferido a la industria de tecnología médica, por ese motivo se debe aportar documentación de diseño que sustente al informe técnico y que siga los lineamientos de la Disposición ANMAT 2318/02 (TO2004). La documentación generada hasta el momento en cada una de las tareas cumplidas corresponde a los contenidos requeridos por el informe técnico: definiciones del Funcionamiento Esperado, el Funcionamiento Esencial, las Condiciones de Uso y Ámbito y la Gestión de riesgos normalizada según ISO 14971:2012.

Los resultados esperados para el resto de la investigación se relacionan con: describir y definir las necesidades y condicionamientos tecnológicos y productivos especiales, definir las características y lineamientos del desarrollo del sistema de adquisición y procesamiento (diseño físico), plantear, definir y documentar las técnicas de procesamiento de señal y algoritmos (diseño lógico), diseñar y generar el prototipo de

investigación, verificar y validar el prototipo de investigación generado con señales sintéticas y de prueba para situaciones similares a las reales (prueba in Vitro). Finalmente establecer una comparativa con los registros de referencia y obtener conclusiones sobre su desempeño.

4. FORMACION DE RECURSOS HUMANOS

El proyecto es iniciático para varios integrantes del equipo y será parte de la formación complementaria al acervo de los profesionales y alumnos de grado que lo integran. La composición del equipo de trabajo es multidisciplinaria y busca la vinculación estrecha entre los departamentos de ingeniería (informática, electrónica) y salud de la Universidad Nacional de La Matanza. Aunque no hay hasta el momento una estrategia definida para incentivar la línea de investigación con este tipo de formación, es una aspiración futura que del proyecto surja alguna tesis.

5. BIBLIOGRAFÍA

- [1] J. Bruhn et al; "Depth of anaesthesia monitoring: what's available, what's validated and what's next?" *British Journal of Anaesthesia*,(2006), 97(1):85-94
- [2] PS. Myles, K. Leslie, J. McNeil, A. Forbes, MT. Chan. "Bispectral index monitoring to prevent awareness during anaesthesia: the B-Aware randomised controlled trial". *Lancet* (2004); 363: 1757-63.
- [3] A. Vakkuri, "Eeg monitoring in anaesthesia", *Euroanesthesia 2006 Madrid, Spain* (2006)
- [4] J. Cebeiro, M.J. Urcola, D. Craiem, "Estimación de la profundidad anestésica basada en índices espectrales", *SABI 2009 Rosario, Sociedad Argentina de Bioingeniería* (2009)

Procesamiento de Señales Vibro-Acústicas. Análisis de Casos de Estudio, Modelación, Prototipado y Experimentación

Guillermo Friedrich, Pablo Giron, Guillermo Reggiani, Adrián Azurro, Ricardo Coppo, Martín Sequeira, Patricia Baldini, Héctor Bambill, Ezequiel Cerda, Lorena Cofre y Gabriela Velasquez

Grupo de I+D en Sistemas y Tecnologías de la Información y las Comunicaciones (Grupo SiTIC)

Grupo de Análisis de Sistemas Mecánicos (Grupo GASM)

Universidad Tecnológica Nacional – Facultad Regional Bahía Blanca –

Departamento de Ingeniería Electrónica

11 de Abril 461 (8000) Bahía Blanca

{gfried, pgiron, ghreggiani, azzurro, rcoppo, martins, pnbaldi, hbambill}@frbb.utn.edu.ar, cerdaezequiel17,cofrelorena}@gmail.com, gabrielavelasquez2112@hotmail.com

RESUMEN

Los niveles excesivos de ruido y vibraciones presentan una gran variedad de efectos nocivos para el hombre y su entorno. Podría decirse, en forma general, que hay confort auditivo hasta los 65 dBA; niveles superiores presentan molestia e incluso, en función de las dosis de tiempo, pueden causar daño auditivo. El Control Activo de Ruido y/o Vibraciones (CAR/V) es la transformación favorable del campo vibro-acústico empleando medios electrónicos, sensores y actuadores. El trabajo a desarrollar incluye una profundización en el conocimiento del estado del arte y las tecnologías vinculadas al tema, mediante el relevamiento y estudio de bibliografía y publicaciones. En base al relevamiento preliminar efectuado en empresas del polo petroquímico de Bahía Blanca se han detectado algunos casos de estudio para los que sería de interés poder desarrollar soluciones innovadoras a fin de mitigar el ruido y sus consecuencias. Una parte del trabajo a realizar comprende tareas de análisis y simulación basadas en el uso de herramientas de software. Complementando al trabajo de análisis y simulación, se pretende también avanzar en el desarrollo de prototipos experimentales, que involucren hardware y su software asociado, con la finalidad de ensayar y evaluar no sólo la parte algorítmica sino también la eficacia de distintos dispositivos sensores y actuadores.

Palabras Clave: control activo de ruido; procesamiento digital de señales; tiempo real

CONTEXTO

La línea de investigación presentada se encuentra en ejecución en el marco del proyecto titulado: “*Procesamiento de Señales Vibro-Acústicas. Análisis de Casos de Estudio, Modelación, Prototipado y Experimentación*”. El mismo es financiado por la Universidad Tecnológica Nacional, está incorporado al Programa de Incentivos y su ámbito de realización es el Grupo SiTIC (Grupo de I+D en Sistemas y Tecnologías de la Información y las Comunicaciones) y GASM (Grupo de Análisis de Sistemas Mecánicos), de la UTN - Facultad Regional Bahía Blanca.

1. INTRODUCCION

Los niveles excesivos de ruido y vibraciones presentan una gran variedad de efectos nocivos para el hombre y su entorno. Podría decirse, en forma general, que hay confort auditivo hasta los 65 dBA; niveles superiores presentan molestia e incluso, en función de las dosis de tiempo, pueden causar daño auditivo.

Otro aspecto a destacar es la interferencia en la comunicación, lo que lleva a las personas a elevar la voz forzando sus cuerdas vocales [1]. Por otro lado, existe un gran número de efectos no auditivos del ruido, entre los cuales se hallan la hipertensión arterial pasajera, las cefaleas, el nerviosismo y estrés, la reducción del rendimiento físico y la pérdida de concentración y atención.

En particular, en determinados sectores de ciertas plantas industriales es frecuente encontrar niveles muy elevados de ruido, que obliga a reforzar las protecciones auditivas y limitar los tiempos de exposición, a fin de evitar la aparición de daños auditivos en los operarios. Para mitigar estos problemas se recurre a técnicas de control, que comprenden "el conjunto de medidas técnicas o estratégicas para corregir una situación en la cual el ruido sea o pueda ser un problema" [2].

La definición anterior puede generalizarse para incluir a las vibraciones y señales acústicas en general. Dentro del espectro de medidas técnicas de control se encuentran el control pasivo y el control activo. El control pasivo está compuesto por todas aquellas acciones que se orientan a la disminución de la contaminación vibro-acústica sin realizar modificaciones esenciales en las fuentes que la producen (barreras acústicas, encapsulamientos con materiales aislantes y absorbentes, silenciadores, montajes antivibratorios, etc.).

El Control Activo de Ruido y/o Vibraciones (CAR/V), a diferencia, es la transformación favorable del campo vibro-acústico empleando medios electrónicos, sensores y actuadores. El CAR/V surge como una técnica complementaria al control pasivo en el rango de las frecuencias bajas. Consiste en la generación de una interferencia destructiva entre el campo primario y otro secundario, en contrafase, generado electrónicamente.

Haciendo mención al caso del ruido, se cancela el ruido primario indeseado según el principio de superposición, mediante un ruido secundario o antiruido, que posee igual amplitud y fase opuesta al primario. Al sumarse ambas señales se obtiene la cancelación deseada. La idea del Control Activo de Ruido como interferencia destructiva, fue planteada por Lueg [3] en 1933, aunque su aplicación se postergó varios años debido a las limitaciones de las tecnologías existentes entonces. En el rango de frecuencias bajas es prácticamente inviable el uso de técnicas pasivas, lo que hace atractivo el estudio y experimentación de técnicas activas. Por otra parte, no existen antecedentes a nivel local y regional con respecto a la aplicación de este tipo de técnicas.

Teniendo en cuenta lo anterior, el principal objetivo del presente proyecto es estudiar la aplicabilidad de técnicas de control activo para solucionar problemas de ruido en plantas industriales. En tal sentido se propone trabajar a partir de casos de estudio que sean de interés en el ámbito local y/o regional, con la finalidad

última de obtener resultados de aplicación práctica. El trabajo a desarrollar incluye una profundización en el conocimiento del estado del arte y las tecnologías vinculadas al tema, mediante el relevamiento y estudio de bibliografía y publicaciones. En base al relevamiento preliminar efectuado en empresas del polo petroquímico de Bahía Blanca se han detectado algunos casos de estudio para los que sería de interés poder desarrollar soluciones innovadoras a fin de mitigar el ruido y sus consecuencias.

Una parte del trabajo a realizar comprende tareas de análisis y simulación basadas en el uso de herramientas de software. El grupo de trabajo dispone actualmente de distintos paquetes comerciales como Matlab, Soundplan y Labview. Sin embargo, se efectuará un relevamiento y ensayo de otras herramientas, a fin de evaluar su conveniencia de utilización. Las herramientas mencionadas serán utilizadas para evaluar y comparar distintas técnicas y algoritmos asociados al control activo de ruido. Eventualmente podrían desarrollarse herramientas de software adhoc, en función de necesidades particulares que surjan durante la ejecución del proyecto.

Complementando al trabajo de análisis y simulación, se pretende también avanzar en el desarrollo de prototipos experimentales, que involucren hardware y su software asociado, con la finalidad de ensayar y evaluar no sólo la parte algorítmica sino también la eficacia de distintos dispositivos sensores y actuadores.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Los integrantes del grupo de trabajo presentan diferentes trayectorias y experiencias, que contribuyen a los objetivos del Proyecto desde distintas vertientes en las cuales se puede mencionar:

Se trabajó sobre aspectos relacionados con el diagnóstico y control de la contaminación sonora en ambientes urbanos e industriales.

Se contribuyó a la caracterización objetiva del ruido en la ciudad de Bahía Blanca para que, de esta manera, se generen conocimientos relacionados aplicables a otras ciudades argentinas de características similares.

Se aplicaron metodologías de medición de ruido urbano e industrial coherentes con las normas de medición de fuentes fijas y móviles, y con el objetivo de ser aplicados para determinar el grado de contaminación sonora de un ambiente urbano o industrial y también para el desarrollo de modelos predictivos.

En este sentido, se generaron una serie de herramientas computacionales para la evaluación del impacto ambiental de ruido a fin de utilizarse como ayuda a procesos de planificación urbana.

De esta manera, se generaron algunos criterios preliminares para mitigar la contaminación sonora en dicha ciudad.

En la línea de ruido industrial se trabajó en el desarrollo de modelos teóricos y computacionales orientado a los problemas acústicos en recintos industriales [4], [5]. [6], [7]. [8].

Se ha trabajado en la temática de análisis de señales y control [9], [10]. [11], [12]. [13].

Se ha trabajado en la temática de redes, especialmente inalámbricas, con orientación a su aplicación en entornos con restricciones temporales. También se cuenta con experiencia en el diseño y desarrollo de hardware y software de sistemas embebidos, incluyendo aplicaciones al procesamiento digital de señales [14], [15]. [16], [17]. [18].

3. RESULTADOS ESPERADOS

Los objetivos del proyecto son:

Objetivos generales:

- Estudiar la aplicabilidad de técnicas de control activo para solucionar problemas de ruido en plantas industriales.

- Consolidar un grupo interdisciplinario de docentes-investigadores.

Objetivos específicos:

- Releva y caracterizar las fuentes de ruido y su entorno para los casos seleccionados.

- Evaluar distintas técnicas de control activo de ruido, con énfasis en su aplicación en ámbitos industriales.

- Desarrollar prototipos basados en técnicas de control activo de ruido para la realización de ensayos experimentales en laboratorio y en campo.

- Comparar eficacia entre soluciones pasivas y activas para cada caso seleccionado.

- Transferir conocimientos y resultados al medio socio-productivo. Se pretende obtener resultados que sean aplicables a problemas existentes en el ámbito industrial (compresores, soplantes, conductos, chimeneas, salas y cabinas de operadores, etc.) sobre fuentes sonoras para las que los métodos pasivos de control por sí solos son ineficaces y necesitan el complemento de técnicas activas.

En base al relevamiento preliminar efectuado en empresas del polo petroquímico de Bahía Blanca se han detectado algunos casos de

estudio para los que sería de interés poder desarrollar soluciones innovadoras para mitigar el ruido y sus consecuencias.

4. FORMACIÓN DE RECURSOS HUMANOS

En el presente proyecto se cuenta con la participación de dos becarios graduados y un becario alumno. También participó durante dos meses un becario alumno extranjero y se tiene prevista la participación de otro.

Otro impacto previsto en la formación de Recursos Humanos, está relacionada con la transferencia que se pueda realizar a los alumnos de las cátedras a cargo de los integrantes del proyecto. En este sentido, es de destacar que el enfoque interdisciplinario y la orientación a problemas del medio socio-productivo que se pretende para el Proyecto, representan un valor agregado a transferir a las aulas.

También es de esperar que otro resultado del proyecto sea la concreción de un curso / seminario de posgrado, que pueda ser ofrecido en las carreras de Especialización y Maestría en Ingeniería Ambiental, que se dicta en esta Facultad Regional. Finalmente, se espera transferir algunos resultados al medio socio-productivo, mediante cursos / seminarios de capacitación y/o divulgación.

5. BIBLIOGRAFIA

- [1] Miyara F., 2004. Acústica y Sistemas de Sonido. UNR Editora, Universidad Nacional de Rosario. Rosario, Argentina.
- [2] Miyara F., 1999. Control de Ruido. UNR Editora, Universidad Nacional de Rosario. Rosario, Argentina.
- [3] Lueg, P., 1936. Process of silencing sound oscillations, U.S Patent 2,043,416, June 9, 1936.
- [4] Estudio de Caracterización Sonora Objetiva de la Zona Industrial de Ingeniero White con particular énfasis en el Barrio 26 de Septiembre. Realizado para el Comité Técnico Ejecutivo de la Municipalidad de Bahía Blanca durante los meses de diciembre de 2003 a julio de 2004. Coautores: Dr. Víctor Cortínez, Ing. Adrián Azzurro, Ing. Mariano Tonini, Sr. Martín Sequeira, etc.
- [5] Convenio de colaboración conjunta entre el Centro de Investigación de Mecánica Teórica y Aplicada (CIMTA) de la UTN-FRBB y el Comité Técnico Ejecutivo de la Municipalidad de Bahía Blanca, sobre el

- monitoreo y control de la contaminación acústica en Bahía Blanca e Ingeniero White. Desde 2009 hasta 2011.
- [6] Proyecto Vale-Rio Colorado, mediciones de ruido y vibraciones en la planta a instalarse en la localidad de Ing. White, perteneciente a la empresa Vale. Desde agosto de 2011 hasta marzo de 2013.
- [8] Estudio de impacto acústico ambiental en los alrededores de la planta de Profertil (Bahía Blanca), generada por la ampliación de su capacidad productiva. Estudio de Impacto Ambiental realizado conjuntamente con la empresa Bahitek S.R.L. Mayo de 2013.
- [9] Introducción Temprana de Conceptos de Control Robusto: Experiencia Práctica y CAD". P. Baldini, G. Calandrini, P. Doñate y H. Bambill. Revista Iberoamericana de Tecnologías del Aprendizaje. IEEE-RITA (en prensa)
- [10] Medición de Vibraciones Mecánicas Mediante el uso de Acelerómetros MEMS. Damián Banfi, Guillermo Friedrich, Miguel Angel Banchieri, Patricia Baldini. VI Congreso de Microelectrónica Aplicada (uEA). La Matanza. Mayo 2015.
- [11] Experiencias de Laboratorio de Bajo Costo para el Aprendizaje de Sistemas de Control Embebido en Tiempo Real. Patricia Baldini, Guillermo Calandrini, Néstor Campos y Matías Frusto. XX Congreso Argentina de Ciencias de la Computación. XII WTIAE. 20 al 24 de octubre de 2014. La Matanza. Bs As. (ISBN 978-987-3806-05-6).
- [12] Controlador PID Diseñado bajo Control Robusto QFT implementado sobre Cortex M-4, Néstor Campos, Matías Frusto y Patricia Baldini. V Congreso de Microelectrónica Aplicada, Córdoba, Mayo de 2014 (ISBN 978-987-34680-5-2).
- [13] Metodologías para Loopshaping Automático en QFT, Pedro Doñate, Patricia Baldini, Guillermo Calandrini y Héctor Bambill. XVI Workshop de Investigadores en Ciencias de la Computación (WICC), Ushuaia, Mayo de 2014. (ISBN: 978-950-34-1084-4)
- [14] Guillermo Friedrich, Guillermo H. Reggiani, Lorena Cofré, Gabriela Velásquez y Walter Moreno, "Determinación de la zona óptima de trabajo de una red ZigBee", en VI Congreso de Microelectrónica Aplicada (uEA 2015), organizado por la Red Universitaria de Ingeniería Electrónica, y realizado en la Universidad Nacional de La Matanza, Argentina, del 27 al 29 de mayo de 2015.
- [15] Néstor Daniel Campos, Matías Guillermo Álvarez, Matías Frusto, Guillermo Friedrich y Adrián Laiuppa, "Experiencias sobre sistemas de control utilizando plataforma didáctica", en VI Congreso de Microelectrónica Aplicada (uEA 2015), organizado por la Red Universitaria de Ingeniería Electrónica, y realizado en la Universidad Nacional de La Matanza, Argentina, del 27 al 29 de mayo de 2015.
- [16] Damián Banfi, Guillermo Friedrich, Miguel Angel Banchieri y Patricia Baldini, "Medición de vibraciones mecánicas mediante el uso de acelerómetros MEMS", en VI Congreso de Microelectrónica Aplicada (uEA 2015), organizado por la Red Universitaria de Ingeniería Electrónica, y realizado en la Universidad Nacional de La Matanza, Argentina, del 27 al 29 de mayo de 2015.
- [17] Guillermo Friedrich, Guillermo Reggiani y Sergio Pellegrino, "Análisis para la selección de parámetros adecuados para una red inalámbrica en tiempo real", en XX Congreso Argentino de Ciencias de la Computación (CACIC 2014), IX Workshop de Arquitecturas, Redes y Sistemas Operativos (WARSO) 2014. Universidad Nacional de La Matanza, La Matanza, Argentina, 20 al 24 de octubre de 2014.
- [18] Guillermo Friedrich, Guillermo Reggiani, Ricardo Cayssials, Sergio Pellegrino, Gabriela Velásquez, Lorena Cofré y Walter Moreno, "Hacia una propuesta para la selección de parámetros adecuados para una red inalámbrica con restricciones temporales", en 43° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO), 15th Argentine Symposium on Technology (AST 2014). Universidad de Palermo, Buenos Aires, Argentina, del 1 al 5 de septiembre de 2014.

Reconocimiento de Patrones y Modelado en Señales de Electrocardiograma: Detección Temprana de Isquemia e Infarto de Miocardio

Sergio Liberzuc^{1,2}, María Lorena Bergamini¹, Anderson Rincón², Pedro D. Arini²

¹CAETI – Facultad de Tecnología Informática. Universidad Abierta Interamericana

Av. Montes de Oca 745. Ciudad de Buenos Aires

²IAM – Instituto Argentino de Matemática, Alberto P. Calderón, CONICET

Saavedra 15 3° piso. Ciudad de Buenos Aires

sliberczuk@uai.edu.ar - Maria.Bergamini@uai.edu.ar - ander.irs@gmail.com - pedro.arini@conicet.gov.ar

Resumen. El procesamiento de señales biomédicas tiene una importancia relevante en el diagnóstico temprano y prevención de enfermedades. El electrocardiograma es un estudio no invasivo, de bajo costo, que brinda información valiosa sobre la actividad eléctrica cardíaca. El análisis de esta señal estudia patrones que se asocian con condiciones anormales de funcionamiento. A partir de un modelo dinámico de la señal de ECG, nos proponemos diseñar modificaciones que contemplen la heterogeneidad de la despolarización y la repolarización ventricular latido a latido. Algoritmos de procesamiento de ECG con un enfoque Bayesiano, serán diseñados con el objetivo de sintonizar los parámetros del modelo que permitan la síntesis de señales de ECG registrables durante procesos de isquemia e infarto. Los índices y métricas utilizados en señales reales permitirán validar el modelo modificado.

Palabras clave: ECG, Filtrado Bayesiano, Isquemia, Infarto, Filtro de Kalman, Filtrado Particular

Contexto

La línea de investigación aquí descrita surge de la colaboración entre el Centro de Altos Estudios en Tecnología Informática (CAETI), de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana (UAI), conjuntamente con el Instituto Argentino de Matemática Alberto P. Calderón, (IAM) dependiente de CONICET y el Instituto de Ingeniería Biomédica (IIBM) de la Universidad de Buenos Aires.

El CAETI concentra proyectos de investigación básica y aplicada, que pretenden contribuir a la generación de herramientas informáticas y tecnológicas para dar solución a problemáticas sociales y humanas. Uno de los objetivos del centro es promover la inserción de la Tecnología Informática en asuntos multidisciplinarios de impacto social.

1. Introducción

Las enfermedades cardiovasculares (ECV) son una de las principales causas de muerte en el mundo. Según datos de la organización mundial de la salud (OMS), en el 2014 tan sólo en Argentina el 35% del total de las muertes registradas se debió a este tipo de enfermedades

[1]. De esta manera se entiende que la detección temprana y el control de las ECV tienen un fuerte impacto en la calidad de vida de miles de personas a nivel local y mundial.

La investigación biomédica se erige sobre dos pilares: la recolección de una enorme masa de datos clínicos, y el análisis de esos datos para determinar patrones que describan cuadros tanto fisiológicos como patológicos. Se deben desarrollar algoritmos específicos para llevar a cabo el análisis y el procesamiento de los datos y poder de ese modo obtener información tanto útil como transferible. Dicha información permitirá luego la semi-automatización de diagnósticos tempranos y más precisos, apoyados en dispositivos específicos diseñados para tal fin.

Señal de Electrocardiograma (ECG)

La señal de electrocardiograma registra la actividad eléctrica del corazón, mostrando la variación temporal del potencial eléctrico que se registra a partir de electrodos dispuestos convenientemente en la superficie del tórax. La morfología de este registro y su interpretación a partir de la detección de sus ondas características (los llamados puntos fiduciales que comprenden los principios, picos y finales de las ondas P, complejo QRS y onda T) así como diversos cálculos que surgen de la detección de dichos puntos característicos (segmento ST, intervalos QT, PR y otros) los cuales permiten el diagnóstico de distintas patologías: arritmias cardíacas, cardiopatías isquémicas, infarto y/o alteraciones de la conducción aurículo-ventricular [2]. Dichas patologías pueden ser detectadas con antelación mediante un adecuado análisis de la señal de ECG [2]. Por lo expuesto anteriormente se justifica que este tipo de análisis no invasivo y de bajo costo siga siendo una herramienta fundamental para la evaluación cardiovascular de los pacientes que llegan por demanda espontánea a las salas de emergencia de cualquier centro de salud. Cabe destacar que el ECG de superficie, además de los beneficios

mencionados, es la única herramienta No Invasiva de diagnóstico clínico capaz de monitorear la actividad eléctrica cardíaca. Otro tipo de estudios, como pueden ser Ecocardiograma o Angiografía, reflejan otro tipo de información más bien de tipo mecánica pero no eléctrica como es el caso del ECG.

En resumen, el estudio del ECG, sus amplitudes de onda, sus intervalos temporales y patrones ha sido objeto de intensas investigaciones, ya que proporciona información sustancial de la funcionalidad del corazón. La extracción de señales cardíacas de alta resolución de un electrocardiograma ruidoso sigue siendo un problema de interés para la comunidad de ingeniería biomédica. A pesar de la rica literatura en este campo, todavía hay muchas aplicaciones clínicas que carecen de herramientas confiables de procesamiento para extraer las componentes ricas en información del ECG, que se encuentran inmersas en potencias de ruido considerable por lo que poseen baja relación señal ruido (SNR), pero ofrecen mucha información relevante para el diagnóstico de diversas patologías cardíacas.

Herramientas de procesamiento

Para señales estacionarias, el filtro de Wiener es la técnica de filtrado lineal en el sentido del mínimo error cuadrático medio, aplicado de modo causal en el dominio temporal o de modo no causal en el dominio frecuencial. Sin embargo, no se espera que el filtrado de Wiener dé buenos resultados para un ECG ruidoso, debido a la naturaleza no estacionaria que posee la señal cardíaca. En estos casos, el filtro de ECG se basa esencialmente en el contenido de frecuencia y hasta cierto punto, en la localización de los picos en el tiempo.

Las técnicas estadísticas como el Análisis de Componentes Principales (PCA), Análisis de componentes independientes (ICA) y las redes neuronales (NNs) también se han utilizado para construir un modelo estadístico de la señal y del ruido, lo que permite la eliminación del ruido

en banda descartando las componentes correspondientes exclusivamente a dicho ruido. Aunque estos son esquemas poderosos de filtrado, el modelo utilizado es bastante arbitrario y pueden ser extremadamente sensibles a pequeños cambios tanto en la señal como en el ruido.

En algunos trabajos recientes se ha propuesto un paradigma de Filtrado Bayesiano para la compresión de la señal de ECG y el filtrado del ruido en la misma [8,9]. Este marco también puede ser utilizado eficazmente para la segmentación del latido de ECG y la extracción de sus puntos fiduciales. Pero de nuevo, se requiere un modelo adecuado de la dinámica eléctrica cardíaca para el funcionamiento de estas técnicas avanzadas de procesamiento.

Como puede observarse el área de ingeniería ha desarrollado y adaptado un gran número de métodos de procesamiento y análisis de señales de ECG, logrando avanzar considerablemente en la detección y diagnóstico temprano de las ECV. Sin embargo, la falta de una adecuada comprensión de la dinámica eléctrica cardíaca hace difícil optimizar las técnicas computacionales empleadas, obteniendo como consecuencia un aumento en el número de falsos positivos y por lo tanto una baja especificidad al momento de emitir un diagnóstico asistido por computadora. Una solución a este tipo de problema ha sido la creación de modelos matemáticos computacionales de la actividad eléctrica cardíaca, que permitan recrear diversas y complejas situaciones electrofisiológicas. De esta manera, se pueden sintetizar conjuntos de señales de ECG para entrenamiento y validación de los sistemas de diagnóstico asistido.

McSharry et al. [6], han desarrollado un modelo para simular con gran precisión diversos trazados de ECG correspondientes a condiciones fisiológicas y algunas condiciones patologías cardíacas. El modelo se basa en suponer que la señal consiste en una suma de funciones gaussianas con centro en los puntos fiduciales. Cada gaussiana está caracterizada

por su ubicación, su amplitud y su ancho. Así, el modelo queda dependiendo de 15 parámetros morfológicos. En su trabajo, McSharry reporta valores de los parámetros para un ECG fisiológico.

Clifford y et al. [11] proponen un método para hallar los parámetros que mejor reproducen el latido de un ECG real dado, logrando así la compresión (con pérdida) del mismo. El ajuste de los parámetros lo llevan a cabo aplicando optimización no lineal (gradiente descendente) para minimizar la distancia entre los datos y el modelo simulado.

El modelo entonces permite la representación 15-dimensional de cualquier ECG, fisiológico o patológico, con lo que luego puede utilizarse no solo en esquemas de filtrado que requieran un modelo sino en aplicaciones de compresión, clustering y/o clasificación de patrones en señales de ECG en el mencionado espacio multidimensional [11].

Entre los eventos cardiovasculares recreados por este modelo la literatura muestra estudios de fibrilación auricular, taquicardia sinusal, bradicardia sinusal y taquicardia ventricular [7]. Sin embargo, una revisión preliminar del estado del arte muestra que el modelo no ha sido utilizado para simular cambios en la heterogeneidad de la despolarización ventricular (DV) ni dispersión en la repolarización ventricular (RV) latido a latido, ligados a procesos isquémicos conducentes a infarto del miocardio.

2. Líneas de Investigación y Desarrollo

En trabajos anteriores se ha mostrado que, luego del infarto de miocardio, existe un vínculo entre el remodelado eléctrico (proceso de curación del tejido infartado) y remodelado inverso (cicatrizado del tejido cardíaco) con las variaciones morfológicas de la repolarización ventricular en el ECG de superficie [3]. Surge de este modo la hipótesis de que la varianza espectral medida recientemente en [3] ha sido

modulada durante el remodelado directo e inverso acorde a las modificaciones que sufrieron los potenciales de acción (PA) en la región infartada del miocardio.

En nuestro trabajo se esperan simular conjuntos de señales de ECG correspondientes a cambios específicos en los PA para las etapas post-infarto; reproduciendo así la dinámica electrofisiológica y consiguiendo la visualización de los cambios en el índice de varianza espectral latido a latido.

El desarrollo de esta investigación permitirá sintonizar y validar un modelo computacional del comportamiento electrofisiológico del corazón basado en el modelo de McSharry et al. [6], modificado por nuestro grupo tanto en sus ecuaciones diferenciales acopladas (ecuaciones de síntesis) como en la incorporación de modificaciones en los algoritmos que consideren y permitan recrear situaciones asociadas a cambios en la heterogeneidad de la DV y de la RV del miocardio durante procesos isquémicos y de infarto.

El modelo también será utilizado para validar otras investigaciones previas y en curso realizadas por nuestro grupo de investigación como por ejemplo la reproducción de las medidas de la varianza espectral durante la isquemia aguda [4]. Esto permitirá simular la dinámica de la heterogeneidad latido a latido durante la oclusión de alguna de las tres arterias coronarias principales [5].

Se pretende, establecer un lazo de realimentación entre los índices electrocardiográficos de riesgo medidos en las bases de datos reales y la generación del conjunto de señales de ECG a partir del modelo modificado. Basado en la hipótesis de que las señales generadas tendrán alto grado de similitud con señales reales de isquemia e infarto y con cualquiera de las señales de ECG utilizadas para el entrenamiento y validación del sistema.

Las técnicas de procesamiento que se utilizarán en esta línea se encuentran dentro del paradigma de filtrado bayesiano. Estos métodos de procesamiento como el filtrado de Kalman o

el filtrado particular, tienen alto rendimiento en señales con baja SNR, pero requieren un modelo adecuado de la dinámica del estado [12].

Las señales serán tomadas de bases de datos validadas internacionalmente como la Base de datos de ECG Physikalisch-Technische Bundesanstalt (PTB) disponible en Physio-Bank [10] que aporta gran cantidad de pacientes sanos, con isquemia e infarto. La misma contiene registros de 52 sujetos sanos y 148 pacientes con infarto de miocardio. Los ECG están digitalizados a 1Khz, con 16 bits de resolución sobre un rango de 16.384mV. Cada registro incluye las 12 derivaciones simultáneas y las derivaciones ortogonales de Franz. Se dispone de la historia clínica de los pacientes.

El adecuado procesamiento de las bases de datos en conjunto con los modelos computacionales utilizados, servirán como sistema de entrenamiento de algoritmos para filtrado, detección y clasificación de isquemias e infartos de miocardio en señales de ECG, aumentando la especificidad y la sensibilidad de los procedimientos utilizados hasta el momento para tratar este tipo de ECV.

3. Resultados esperados

Esta investigación espera lograr los siguientes resultados:

- Modelar el efecto de modulación que producen diferentes estadios del infarto (remodelado eléctrico y remodelado inverso) sobre la heterogeneidad tanto de la despolarización como de la repolarización ventricular.
- Comparar los índices medidos en bases de datos de ECGs reales con ECGs sintetizados a partir de los modelos propuestos. Validar y mejorar la estimación de dichos índices.
- Ajustar los modelos computacionales para hacer seguimiento de señales de ECG y poder detectar con mayor sensibilidad y especificidad la dinámica eléctrica cardíaca de la isquemia y el infarto de miocardio entre otras posibles patologías.

4. Formación de Recursos Humanos

El equipo de investigación llevará adelante las ideas aquí expuestas está conformado por los autores de este trabajo.

El Bioing. Sergio Liberczuk se encuentra desarrollando su tesis de Doctorado, bajo la dirección del Doctor Silvano Zanutto (Director del Instituto de Ingeniería Biomédica de la UBA) y la codirección del Dr. Pedro Arini en la Facultad de Ingeniería de la UBA. Por su parte el Mg. Anderson Rincón, está trabajando en la misma línea de investigación realizando su Doctorado en Ingeniería bajo la dirección del Dr. Pedro Arini en la Facultad de Ingeniería de la UBA.

Referencias

- [1] O. Mundial de la Salud, “Global status report of noncommunicable diseases 2014”. World Health Organization, Tech. Rep.
- [2] R. Martis, U. Acharya, and H. Adeli, (2014) Current methods in electrocardiogram characterization, *Comput. Biol. Med.*, 48 (1), pp. 133–149.
- [3] Arini, P.D., Valverde, E.R. (2016) Beat-to-beat electrocardiographic analysis of ventricular repolarization variability in patients after myocardial infarction. *Journal of Electrocardiology*, 49, pp. 206-213.
- [4] E. Valverde, G. Bertrán, , and P. Arini. (2013) Beat to beat ventricular repolarization variability evaluated during acute myocardial ischemia. *Biomed Signal Process Control*, 8 (6), pp. 869-875.
- [5] P. Arini, F. Baglivo, J. Martinez, and P. Laguna. (2014) Evaluation of ventricular repolarization dispersion during acute myocardial ischemia: Spatial and temporal ECG indices. *Med Biol Eng Comput*, 52 (4), pp. 375-391.
- [6] P. McSharry, G. Clifford, L. Tarassenko, and L. Smith. (2003) A dynamical model for generating synthetic electrocardiogram signals. *IEEE Trans. Biomed. Eng.*, 50 (3), pp. 289–294.
- [7] J. Lian, G. Clifford, and D. Muessig. (2007) Open source model for generating rr intervals in atrial fibrillation and beyond. *Biomed Eng Online*, 6 (9). doi:10.1186/1475-925X-6-9.
- [8] O. Sayadi, M. Shamsollahi, and G. Clifford. (2010) Robust detection of premature ventricular contractions using a wave-based Bayesian framework. *IEEE Trans. Biomed. Eng.*, 57 (2), pp. 353–362.
- [9] R. Sameni, M. Shamsollahi, C. Jutten, and G. Clifford (2007) A nonlinear Bayesian filtering framework for ECG denoising. *IEEE Trans. Biomed. Eng.*, 54 (12), pp. 2172–2185.
- [10] <https://www.physionet.org/physiobank/database>
- [11] Clifford, G. D., Shoeb, A., McSharry, P. E., & Janz, B. A. (2005) Model-based filtering, compression and classification of the ECG. *Int. J. Bioelectromag*, 7 (1), pp. 158-161.
- [12] Infante, S., Sánchez, L., & Cedeño, F. (2014) Nonlinear filters to reconstruct electrocardiogram signals. *Revista de Matemática Teoría y Aplicaciones*, 21 (2), pp. 199-226.

Red Inalámbrica de Sensores Aplicada a la Predicción del Comportamiento de Incendios Forestales

Rodrigo Atilio Elgueta¹, Miguel Méndez-Garabetti^{1,2}

¹Universidad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería
rodrigo.elgueta@um.edu.ar, miguel.mendez@um.edu.ar

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

RESUMEN

Las Redes Inalámbricas de Sensores poseen diversos campos de acción, uno de ellos está íntimamente vinculado con el seguimiento y prevención de catástrofes naturales. Los incendios forestales, como parte de éstas, generan grandes pérdidas y daños alrededor del mundo. En este contexto, el presente trabajo tiene como objetivo desarrollar y planificar el despliegue de una Red Inalámbrica de Sensores para la cuantificación de ciertas variables ambientales, permitiendo detectar la ocurrencia y propagación de incendios forestales. Se espera que las mismas puedan ser incorporadas como entradas en un sistema de predicción del comportamiento de incendios forestales, el cual, estará compuesto por el método de reducción de incertidumbre denominado ESS-IM (Sistema Estadístico Evolutivo con Modelo de Islas). Esta propuesta pretende servir como plataforma para el seguimiento de incendios y nutrir de información confiable a modelos predictivos de comportamiento del fuego, con el objeto de resultar una herramienta útil para minimizar los daños causados por este tipo de fenómenos.

Palabras clave: *Red Inalámbrica de Sensores, Teledetección de Incendios*

Forestales, WSN. Predicción de Incendios Forestales

CONTEXTO

El presente trabajo de I+D se desarrolla como proyecto de tesis de posgrado de la Maestría en Teleinformática, Dirección de Posgrado, perteneciente a la Facultad de Ingeniería de la Universidad de Mendoza, (Ciudad, Mendoza). El presente proyecto será presentado como propuesta de tesis.

1. INTRODUCCIÓN

Las causas que dan lugar a los incendios forestales pueden ser tanto de tipo estructural como inmediatas, dentro de estas últimas, las mismas pueden clasificarse en aquellas que derivan de comportamiento antrópico o de agentes naturales [1], siendo la primera, la que agrupa la mayor probabilidad de ocurrencias. Los incendios forestales transitan tres etapas: fase de iniciación, fase de propagación y etapa de extinción, finalizando por causas naturales (lluvia o falta de vegetación) o por acción humana (labores de extinción) [2][3]. Durante muchos años se ha trabajado en el desarrollo de diferentes herramientas para la prevención, detección y predicción de incendios forestales. Cada una de éstas

corresponde a distintas fases del proceso de lucha contra incendios [4][5][6][7].

La prevención tiene como objetivo lograr que los incendios no alcancen a materializarse; la detección se concentra en determinar la ubicación de los focos de incendio antes de que éstos adquieran una magnitud tal que lleve a perder el control sobre ellos; y por último la predicción, que tiene dos acepciones: predicción de incendios forestales y predicción del comportamiento de incendios forestales. La primera de ellas intenta predecir la ocurrencia de incendios antes de que estos sucedan, mientras que la segunda, busca determinar el posible comportamiento de un incendio forestal una vez que éste ya se ha iniciado, permitiendo tomar decisiones acertadas en el plan de acción a tomar [8].

Uno de los retos más importantes que enfrenta un modelo de predicción de incendios forestales es el de disminuir los efectos de la incertidumbre i.e., intentar conocer con el mayor grado de precisión posible los valores de los parámetros de entrada que determinan el comportamiento del modelo.

Las implementaciones computacionales de dichos modelos suelen realizarse con simuladores de comportamiento de incendios. Éstos suelen requerir el uso de grandes capacidades de cálculo, debido que generalmente se realiza un gran número de simulaciones por cada incendio en consideración. Debido al tamaño del conjunto de datos y la complejidad de las operaciones que deben efectuarse sobre los mismos, puede requerirse la utilización de sistemas de alto rendimiento (HPC, High performance Computing) para resolver el problema en el menor tiempo posible [9].

Las redes inalámbricas de sensores, además de utilizarse como sistema de detección de incendios en tiempo real, pueden ser un

complemento importante para reducir significativamente el tiempo de procesamiento necesario de los sistemas de predicción actuales. A fin de reducir la brecha de incertidumbre del modelo, podría alimentar al simulador con información real de un determinado parámetro, el cual, anteriormente era desconocido y calculado en base a estimaciones indirectas.

Las Redes Inalámbricas de Sensores (WSN, Wireless Sensor Networks) también se encuadran dentro de la llamada “Inteligencia Ambiental” y se encuentra ligada a los últimos avances en computación ubicua y los nuevos conceptos de interacción inteligente entre usuario y máquina. Desde el punto de vista práctico, consiste en la creación de una serie de objetos de uso cotidiano con cualidades interactivas suaves y no invasivas [10], siendo su objetivo fundamental, dotar a objetos de capacidades de adquisición de información, procesamiento y comunicación para ofrecer nuevos servicios a los usuarios.

Por lo tanto, las WSN se basan en dispositivos de bajo coste y consumo llamados nodos (motes), capaces de obtener información de su entorno, procesarla localmente, y comunicarla a través de enlaces inalámbricos hasta un nodo central de coordinación.

Los nodos actúan como elementos de la infraestructura de comunicaciones al reenviar los mensajes transmitidos por nodos más lejanos hacia al centro de coordinación.

La red de sensores inalámbricos está formada por numerosos dispositivos distribuidos espacialmente, que utilizan sensores para controlar diversas condiciones en distintos puntos, como temperatura, humedad y presión entre otras. Los dispositivos son unidades autónomas que constan de un microcontrolador, una fuente de energía, un radio-transceptor y un elemento sensor [11].

Las capacidades de autodiagnóstico, auto-configuración, auto-organización, auto-restauración y reparación, son propiedades que se han desarrollado para este tipo de redes para solventar problemas que no eran posibles con otras tecnologías.

Se caracterizan por ser redes desatendidas, con alta probabilidad de fallo en cada componente, habitualmente construidas ad-hoc para resolver un problema muy concreto, es decir, para ejecutar una única aplicación.

Existen diversas áreas de aplicación para las WSN, entre las que se destacan: monitoreo de energía, condiciones estructurales, transporte, monitoreo industrial y monitoreo ambiental [12]. Enmarcado en esta última, el presente trabajo desarrolla un proyecto de utilización de una WSN con componentes de bajo costo para la detección de incendios forestales y su conexión con modelos de predicción.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Definición de esquemas y despliegue:

El presente trabajo, considera un diseño con patrones en disposición de cuadrícula regulares de igual distribución con consumo energético [13]. Se ha establecido de esta manera en función de que además de facilitar la comunicación, proporcionaría la localización de los elementos, realizando una cuadrícula en el terreno evitando el uso de sensores GPS por nodo [14].

Definición de equipamiento a utilizar:

Debido a la existencia de variables mínimas que deben ser monitoreadas, el trabajo de tesis, cuenta con sistema de adquisición de datos mediante la utilización de sensores específicos. Un ejemplo de ello, es el sensor DHT11, utilizado en el anteproyecto de tesis

para obtener valores en tiempo real de temperatura y humedad de un lugar específico. Los sensores estarán conectados a un nodo que tendrá interfaces de Entrada/Salida para sensores, memoria y un procesador. Si bien existe una diversidad de componentes que poseen soluciones integradas entre este punto y el que viene a continuación, como el chip esp8266 o bien el mrf24j40ma que poseen placas de red incorporadas, se optó para este trabajo utilizar como mote placas “arduino nano”.

Protocolos y Primitivas de Conexión

Este punto se está estudiando cuidadosamente, ya que está en proceso de decisión basado en pruebas de laboratorio y costes, la utilización de diferentes tecnologías, entre ellas las más importantes son: IEEE 802.15.4, ZigBee eIEEE 802.11 b/g/n. Si bien ZigBee [15] [16] [17] está basada en IEEE

802.15.4 y por lo tanto posee un bajo consumo, acepta topologías tipo mesh y es de fácil integración, aún es una tecnología de alto costo para aplicaciones como la del presente trabajo, por lo que se está investigando alternativas para la utilización de componentes con el estándar IEEE 802.11 b/g/n de bajo consumo.

A pesar de la disposición de cuadrícula vista anteriormente, la topología lógica a utilizar será jerárquica ya que se designarían a algunos nodos como “cluster-head” con mayor responsabilidad para controlar a otros nodos. Se espera que esta característica, posea ventajas para la detección de incendios debido a que, ante cambios de condiciones, existiría un control que reaccionaría ante amenazas de fuego a fin de cambiar parámetros de: energía, ancho de banda, tiempos de recolección de datos, etc.

Por ello, se utilizará el modelo de ruteo de protocolos basados en clúster [16]. El despliegue del mismo en el escenario que se estudia, permitiría el balance de consumo energético y el envío de mensajes críticos lo más rápidamente posible mediante la utilización de protocolos de comunicación de entornos consistentes intra-clusters e inter-clusters.

Integración con método de predicción:

Se realizará una evaluación detallada por cada uno de los parámetros de entrada que el método ESS-IM utiliza para realizar los cálculos predictivos. Dicho análisis se realizará con el objetivo de garantizar el mayor grado de compatibilidad entre las variables medidas y la información que el método necesita para el cálculo de propagación del frente de fuego a fin de encontrar las que son de utilidad en las tareas de predicción.

3. RESULTADOS ESPERADOS

Actualmente, el trabajo en desarrollo, parte desde el enfoque de integración de todos los dispositivos de red, control, energía, etc., que formarán un nodo, para la implementación de una WSN para recolectar datos de los sensores y diseñada para poder ser utilizada como entrada del método ESS-IM.

Como objetivo de dicha WSN, se busca: poseer eficiencia energética ya que en grandes extensiones con un gran número de sensores puede ser costoso e incluso imposible su reemplazo, por consiguiente, también se busca que sea auto regenerativa para que continúe funcionando ante el fallo de algún nodo. También se necesita que la misma posea la capacidad de adaptarse a ambientes hostiles y que a su vez permita optimizar los

recursos y la vida útil de los sensores mediante la definición de una arquitectura y protocolo de comunicación que lo permitan.

De esta forma, la incorporación de WSN para la lucha contra el fuego mediante la captura de parámetros medidos en tiempo real, podría detectar la ocurrencia de incendios forestales, colaborar en su prevención y alimentar modelos de predicción de su comportamiento en función de las variables de entrada del método ESS-IM.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de posgrado por parte del Ing. Rodrigo Atilio Elgueta, quien es estudiante de la Maestría en Teleinformática de la Universidad de Mendoza.

5. BIBLIOGRAFÍA

1. Comisión Nacional Forestal (2010). Guía práctica para comunicadores - Incendios Forestales. San Juan de Ocatán. Guadalajara. México
2. Pausas, J. G. 2012. Incendios forestales. Ed Catarata-CSIC
3. Molina, DM; Grillo-Delgado, F; Garcia-Marco, D, (2006). Uso del fuego prescrito para la creación de rodales cortafuegos: estudio del caso "Las Mesas de Ana López", Vega de San Mateo, Gran Canaria, España. InvestAgrar: SistRecurFor (2006) 15(3), 271-276, Madrid
4. P.N. (2005) Omni, ForestFires: A Reference Handbook. ContemporaryWorldIssues.
5. E.A. Johnson, K. Miyanishi (2001), Forest Fires: Behavior and Ecological Effects. Academic Press.
6. J.D. Lowe (2000). Wildland Firefighting Practices. Delmar Thomson Learning.

7. Frausto, Juan Manuel y Landa, Rossana (2005) Incendios forestales. Definiendo el problema, ecología y manejo, participación social, fortalecimiento de capacidades. Educación y divulgación. Mundi Prensa.
8. Méndez-Garabetti Miguel, Tardivo María Laura, Bianchini Germán, y Caymes-Scutari Paola (2014). Predicción del Comportamiento de Incendios Forestales mediante un Método de Reducción de Incertidumbre basado en HPC y Evolución Diferencial. XVI WICC 2014.
9. Méndez-Garabetti Miguel, Bianchini Germán, Caymes-Scutari Paola, y Tardivo María Laura (2016). Increase in the quality of the prediction of a computational wildfire behavior method through improvement of the internal metaheuristic. *Fire Safety Journal*, 82 (2016), 49-62.
10. Waqas Ali, Abdullah, y Ishfaq-ur-rashid. (2016). A Survey on WSN-based Forest Fire Detection Techniques. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. 1(1), 60-65.
11. Fernandez Barcell Manuel (2015). *Wireless Sensor Network*. Recuperado de: <http://www.mfbarcell.es/conferencias/wsn.pdf>
12. Ni.com. (2017). *Innovations - National Instruments*. [online] Recuperado de: <http://www.ni.com/wsn/applications/esa>
13. Yunus Emre Aslan, Ibrahim Korpeoglu, y Özgür Ulusoy (2012). A framework for use of wireless sensor networks in forestfire detection and monitoring. *Computers, Environment and Urban Systems*, 36 (2012) 614–625.
14. Serna M Ángeles, Bermúdez Aurelio, y Casado Rafael (septiembre de 2012). Modelado de incendios forestales con WSNs mediante múltiples envoltentes. En Manzoni (moderador). *Jornadas de la Sociedad de Arquitectura y Tecnología de Computadores (SARTECO)*. Sesión 3C de las Jornadas SarteCoElx, llevada a cabo en Elche, Valencia, España.
15. Solobera Javier. (2010). *Libelium: Detecting Forest Fires using Wireless Sensor Networks*. Libelium. Zaragoza, España.
16. Garbarino Jimena (2011). *Tesis: Protocolos para redes inalámbricas de sensores*. Facultad de Ingeniería. Universidad de Buenos Aires.
17. Erazo Jennifer. Hervas Carlos (2014). *Sistema de detección de incendios forestales mediante redes sensoriales inalámbricas (Zigbee)*. MASKANA, I+D+ingeniería 2014 Ecuador.

Redes de Sensores Inteligentes para Monitoreo de Datos Remotos

Valdez J.C., Pandolfi D., Villagra A.
 Laboratorio de Tecnologías Emergentes (LabTEM)
 Instituto de Tecnología Aplicada (Caleta Olivia)
 Universidad Nacional de la Patagonia Austral
 {jcvaldez, dpandolfi, avillagra}@uaco.unpa.edu.ar

Resumen

Una red de sensores inalámbricos (en inglés *Wireless Sensor Network, WSN*) es un conjunto de nodos sensores y actuadores desplegados sobre la región de actuación. Los sistemas SCADA (Supervisión, Control y Adquisición de Datos) se aplican a la administración de la recolección de información de los procesos de controles industriales, tales como las redes eléctricas, redes de distribución de agua, transporte, etc.). Los sistemas Multi-Agente (en inglés *Multi-Agent System, MAS*) aparecieron como un nuevo enfoque arquitectónico en aplicaciones de ingeniería complejas y altamente dinámicas. Además, los MAS pueden incluir agentes reactivos como Arduinos y agentes inteligentes tales como Redes Neuronales Artificiales (en inglés *Neural Network artificial, NNa*), o Lógica Difusa (en inglés *Fuzzy Logic, FL*).

En este trabajo, se propone construir un sistema tipo SCADA flexible e interoperable, basado en la integración de protocolos de arquitecturas MAS. El sistema propuesto será implementado para administrar algunos de los siguientes posibles escenarios: monitoreo de una red de transporte público de pasajeros o tránsito de una ciudad, monitoreo de una red de distribución de agua potable, y monitoreo de

una red eléctrica en edificios públicos.

Palabras clave: Red de sensores inalámbricos, agentes inteligentes, sistemas multiagentes, Arduino.

Contexto

La línea de investigación presentada en este documento se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM) en el marco del Programa de Investigación en Ciencia y Tecnología de la Universidad Nacional de la Patagonia Austral, Unidad Académica Caleta Olivia. En el ámbito de una convocatoria de proyectos de I&D UNPA, proyecto denominado: “Desarrollo de una planta piloto de desalación de agua de mar”.

Introducción

Los sensores son dispositivos económicos, de bajo consumo de energía con recursos limitados. Una *WSN* se compone de gran número de nodos de sensores con capacidad limitada de computación, almacenamiento y comunicación. Los ambientes, donde se despliegan los nodos sensores, pueden ser aplicados en ambientes controlados (tales

como el hogar, oficina, almacén, bosque, etc.) o no controlados (tales como áreas hostiles o de desastre, regiones tóxicas, etc.). Las *WSN*, representan una de las áreas de investigación más promisorias debido al completísimo campo de aplicación y el desarrollado de nuevas tecnologías de controladores electrónicos o electromecánicos, y los avances en la tecnología de comunicación inalámbrica [7]. El diseño de una *WSN* debe hacer frente a la ocurrencia de defectos de alguna o múltiples componentes en la red y por lo tanto es necesario la aplicación de técnicas robustas que puedan tolerar los fallos, sin resentir el buen funcionamiento de la red. Los problemas y las necesidades de mecanismos de mejora de la fiabilidad dependen de los recursos y aplicaciones disponibles para el que se aplica la *WSN* [20].

Los rápidos avances en las tecnologías electrónicas han dado lugar al desarrollo de una variedad de nuevas y baratas placas electrónicas con capacidades de detección, monitoreo y control que ofrecen oportunidades para su implementación en múltiples áreas de aplicación. La placa Arduino se desarrolló originalmente en el año 2005 en el *Interaction Design Institute*, de Ivrea Italia, como una plataforma de hardware de código abierto. Estudios sobre Arduino, muestran su utilidad en las mediciones automatizadas, y ofrecen una guía para otros investigadores en el desarrollo de sensores de bajo costo y sistemas de monitoreo y automatización en diversos campos de aplicación [17].

RFID (Identificación por Radio Frecuencia) describe un sistema de identificación remota [5]. RFID se basa en almacenar y recuperar de forma remota información o datos y se compone de una etiqueta, un lector y de una base de datos back-end [19]. RFID no es una tecnología nueva y ha pasado por muchas décadas de uso en áreas como bibliotecas,

seguridad, salud, deporte, granjas de animales, aplicaciones militares y otras áreas. Las industrias utilizan tecnologías RFID para diversas aplicaciones tales como el control personal / control de acceso de vehículos, seguridad de la tiendas, seguimiento de los equipos, equipaje, establecimientos de comida rápida, logística, etc. [2].

La comunicación entre los sensores y el sistema de control, la *WSN* y RFID son componentes importantes de la computación ubicua. Ambas, son tecnologías inalámbricas que tienen gran variedad de aplicaciones y proporcionan potencialidades futuras ilimitadas. La integración de las tecnologías RFID con un sistema de control u otros nodos, permite que la información pueda ser recogida fácilmente a partir de múltiples etiquetas RFID [10].

Según [1], un sistema SCADA es responsable de recopilar información y datos en tiempo real de la variedad de procesos o ambiente y proporcionar estos datos a los operadores situados en cualquier lugar y en cualquier momento. Desafortunadamente, muchos de los sistemas SCADA convencionales no son capaces de proporcionar la gestión de la información y enfoques inteligentes de alto nivel. Los MAS proponen soluciones a problemas distribuidos en los dominios computacionales dinámicos y abiertos. Según Luck, [15] los sistemas informáticos, especialmente los relacionados con las aplicaciones industriales modernas, tales como los sistemas SCADA, se están volviendo cada vez más interconectados y más difícil en mantener. Los enfoques basados en agentes (MAS) parecen ser una solución prometedora y el rápido desarrollo del campo de los sistemas basados en agentes ofrece un nuevo paradigma interesante para el desarrollo de programas sofisticados en entornos dinámicos y

abiertos.

Debido a los grandes avances realizados en las tecnologías RFID y *WSN*, se ha estudiado en varios tipos de aplicaciones que involucran vehículos, tales como el pago automático de peaje, gestión de flotas, seguridad de la navegación y sistemas de transporte inteligentes [18]. Además, para mejorar la eficiencia y la seguridad en los sistemas de transporte, se está utilizando tecnología RFID en diferentes tipos de aplicaciones tales como estacionamiento inteligente, los billetes electrónicos, el cobro de peajes, seguimiento de la ruta virtual y control de semáforos digital [6] y [8]. La tecnología RFID se puede utilizar en la identificación automática de vehículos para obtener información en tiempo real de las rutas, y simplemente consiste en colocar etiquetas RFID pasivas en varios lugares de la ruta elegida, y lectores RFID en el autobús [21]. Araar y Khali [4], propusieron la incorporación de las tecnologías RFID y *WSN* con el fin de identificar y rastrear vehículos sustituyendo sistemas basados en GPS (Sistema de posicionamiento global).

El uso de *WSNs* para la auditoría y la gestión del consumo de energía en un edificio es un área de investigación emergente [11]. En [16] se presenta un sistema de control y gestión para la conservación de la energía en grandes edificios públicos diseñado para conectar el sistema de medición de distribución de energía, una red de sensores inalámbricas, y la automatización de edificios. Corucci et al. [9], proponen GreenBuilding, un sistema basado en sensores para la administración de energía automatizada de aparatos eléctricos en un edificio capaz de proporcionar un importante ahorro energético mediante el uso de estrategias de conservación de la energía apropiados adaptados a aparatos específicos.

Las *WSNs* han emergido como una

tecnología eficaz para el control de la infraestructura crítica, tales como tuberías de agua, petróleo y gas. En [3] se muestra un diseño escalable y la simulación de un sistema de monitoreo de fugas de tuberías de agua utilizando la Identificación por Radio Frecuencia (RFID) y tecnología *WSN*. Huang et. al [13] proponen un enfoque viable y eficaz en el uso de una *WSN* para monitorear la calidad del agua de un lago y además se propone un motor de razonamiento basado en reglas que se utiliza para llevar a cabo un sistema de soporte de decisiones a través de técnicas de razonamiento y sensibilidad al contexto. En [14] se propone un sistema de monitoreo ambiental del agua basado en una red de sensores inalámbricos. Se compone de tres partes: nodos de control de datos, estaciones base de datos y un centro de monitorización remota.

Resultados obtenidos/esperados

El sistema de administración de redes de sensores inteligentes para procesamiento de datos remotos se propone con las siguientes ventajas:

- a) simplicidad de operación;
- b) flexible y ser capaz de adaptarse a su ambiente cambios dinámicos;
- c) interoperable;
- d) de bajo costo y fuente abierta y
- e) tolerantes a fallas.

Su aplicación a diversos escenarios remotos permitirá validar las características funcionales propuestas. Los escenarios propuestos para redes remotas de monitoreo se ajustarán a: redes de transporte o tráfico, redes de agua, y redes de energías en

edificios públicos.

Formación de recursos humanos

Un integrante del proyecto está desarrollando su tesis de Maestría orientada a esta línea de investigación.

Referencias

- [1] H. Abbas, S. Shaheen, M. Amin; “*Simple, Flexible, and Interoperable SCADA System Based on Agent Technology*”; Journal of Intelligent Control and Automation, 2015, 6, 184-199. Published Online August 2015 in SciRes.
- [2] K. Ahsan, H. Shah and P. Kingston; “*RFID Applications: An Introductory and Exploratory Study*”; IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 1, No. 3, January 2010 ISSN (Online): 1694-0784 ISSN (Print): 1694-0814.
- [3] A. Almazayad, Y. Seddiq, A. Alotaibi, A. Al-Nasheri, M. BenSaleh, A. Obeid, S. Qasim; “*Proposed Scalable Design and Simulation of Wireless Sensor Network-Based Long-Distance Water Pipeline Leakage Monitoring System*”; Journal of Sensors 2014, 14(2), 3557-3577.
- [4] A. Araar, A. and Khali, H.; “*Investigating RFID Tags Fusion in a Clustered WSN Applied to Vehicle Identification and Tracking*”; (2012) International Journal of Computing and Information Technology, 1, 17-24.
- [5] J. Bohn, “*Prototypical implementation of location-aware services based on a middleware architecture for super-distributed RFID tag infrastructures*”, Pers Ubiquit computing, (2008) Journal 12:155-166.
- [6] J. Caffery, and L. Gordon: “*Subscriber Location in CDMA Cellular Network*”; (2010) IEEE Transactions on Vehicular Technology, 47, 406-416.
- [7] E.H. Callaway, “*Wireless Sensor Networks: Architectures and Protocols*”; (2004) CRC Press, Boca Raton, 1-40 .
- [8] P. O. Charlebois, “*Radio Frequency Indentification (RFID): Principles and Applications. Electromagnetic Fields and Waves*”; (2004) McGill University, Montreal.
- [9] F. Corucci, G. Anastasi and F. Marcelloni, “*A WSN-based testbed for energy efficiency in buildings, Computers and Communications*”. (ISCC), 2011 IEEE Symposium on, Kerkyra, 2011, pp.990-993. doi: 10.1109/ISCC.2011.5983971.
- [10] C. Englund, and H. Wallin, “*RFID in Wireless Sensor Network,*” Technical Report, Department of Signals and Systems, Chalmers University of Technology, Sweden, April 2004.
- [11] A. Guerrieri, A. Ruzzelli, G. Fortino, and G. O’Hare; “*A WSN-based building management framework to support energy-saving applications in buildings*”. (2011); Advancements in Distributed Computing and Internet Technologies: Trends and Issues, Al-Sakib Khan Pathan, Mukaddim Pathan, Hae Young Lee, eds, 161-174.
- [12] A. 12 and T. Koiv, “*Energy Consumption Monitoring Analysis for Residential, Educational and Public Buildings*”; (2012) Smart Grid and Renewable Energy, Vol. 3 No. 3, 2012, pp. 231-238.

- [13] X. Huang, J. Yi , S. Chen and X. Zhu; “*A Wireless Sensor Network-Based Approach with Decision Support for Monitoring Lake Water Quality*”; (2015) Journal of Sensors, ISSN 1424- 8220 15(11), 29273-29296.
- [14] P. Jiang, H. Xia, Z. He and Z. Wang; “*Design of a Water Environment Monitoring System Based on Wireless Sensor Networks*”; (2009); Journal of Sensors, 9(8), 6411-6434.
- [15] M. Luck, “*From Definition to Deployment: What Next for Agent-Based Systems?*”; The Knowledge Engineering Review, (1999) 14, 119-124.
- [16] X. Ma, R. Cui, Y. Sun, C. Peng and Z. Wu, “*Supervisory and Energy Management System of large public buildings,*” Mechatronics and Automation (ICMA), 2010 International Conference on, Xi'an, 2010, pp. 928-933. doi: 10.1109/ICMA.2010.558996.
- [17] E. Melgar y C. Diez, *Arduino and Kinect Projects: Design, Build, Blow Their Minds*; Apress Berkely, CA, USA ©2012 ISBN:1430241675 9781430241676.
- [18] S. P. Narayan, M. H. Assaf, S. K. Prasad, “*Wireless Sensor Enabled Public Transportation System*”; International Journal Communications, Network and System Sciences, 2015, 8, 187-196 Published Online May 2015 in SciRes.
- [19] J. Schwierer, G. Vossen, “*A Design and Development Methodology for Mobile RFID Applications based on the ID-Services Middleware Architecture*”, IEEE Computer Society, (2009), Tenth International Conference on Mobile Data Management: Systems, Service and Middleware.
- [20] L. Venkatesan, S. Shanmugavel, C. Subramaniam; “*A Survey on Modeling and Enhancing Reliability of Wireless Sensor Network,* Journal of Wireless Sensor Network; (2013), 5, 41-51 <http://dx.doi.org/10.4236/wsn.2013.53006>.
- [21] Y. Zhang, “*RFID-Based Tracking in Supporting Real-Time Urban Traffic Information*”. (2009), IEEE 5th International Joint Conference on INC, IMS and IDC, Seoul, 25-27 August 2009, 657-65

Sistema de Control Electrónico de Estabilidad para Estudiar Datos de Interés Estratégico sobre Cultivos

Peretti Gastón¹, Gallina Sergio², Felissia Francisco¹, Bernardi Emanuel¹, Pipino Hugo¹, Depetris Leonardo¹, Cervetto Mayco¹, Depetris Lorenzo¹

- 1) Departamento de Electrónica de la Facultad Regional San Francisco de la Universidad Tecnológica Nacional
- 2) Departamento de Electrónica Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca

Av. De la Universidad 50- San Francisco (Pcia. de Córdoba) – CP (2400) Tel.:
03564-421147 / e-mails: gastonperetti@gmail.com

Resumen

En este proyecto se propone desarrollar un sistema de control electrónico de estabilidad factible de ser montado sobre un aeromodelo, que nos permita obtener el enfoque y la resolución necesaria para capturar imágenes de muy alta resolución y luego mediante las mismas, extraer datos de interés estratégico del cultivo como tipos de cultivos, rendimientos, estimación de pérdidas, cantidad de superficies inundadas, necesidades hídricas (cantidad de agua que necesitan), etc. El sistema será montado sobre dos tipos de aeromodelos (cuadricóptero y un planeador). El motivo de montar el sistema sobre distintos aeromodelos, es analizar y evaluar el comportamiento del sistema de estabilización electrónico a desarrollar, y por otra parte evaluar la eficiencia de ambos tipos de aeromodelos en función de la variable que se desee analizar.

Palabras clave: Control de estabilidad - Agricultura de precisión – Estimación de datos de cultivos.

Contexto

La investigación está inserta dentro de la línea de control y procesamiento de imágenes. El proyecto de

investigación se lleva a cabo en el ámbito del Departamento de Electrónica de la Facultad Regional San Francisco y el Departamento de Electrónica de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca.

La Institución que acredita el proyecto de Investigación y desarrollo (PID) es Rectorado de la Universidad Tecnológica Nacional como proyecto de investigación y desarrollo mediante la financiación del mismo con fecha de aprobación en el mes de Abril del 2016.

Introducción

Un sistema con sensores o videocámaras aerotransportados se presenta como una poderosa herramienta para el estudio de recursos naturales de nuestro país, y para la medición de datos estratégicos de cultivos. Los sistemas de adquisición de imágenes representan una gran posibilidad para cubrir las necesidades de numerosas líneas de investigación en las ciencias del suelo y del medioambiente.

Las cámaras o diversos tipos de sensores que se instalan en aeromodelos u otros medios de elevación, aportan una visión diferente a la de los sensores terrestres, es decir tienen la ventaja de obtener una lectura mucho más cercana a la realidad. Los mayormente utilizados son los sensores de radiación [1], sensores multi-espectrales [2], [3] [4], sensores electromagnéticos [5], entre muchos otros [6], todos

ellos ayudan al entendimiento de fenómenos que se manifiestan sobre la superficie de la tierra, e inclusive sirven para conocer lo que sucede por debajo de esta. Las ciencias que se benefician de la utilización de los diversos sensores son la: geología, agronomía, ecología, hidrología, biología, micro-meteorología, sociología, ciencias ambientales, entre muchas otras.

Esta información generalmente sirve como soporte a mediciones efectuadas desde otros puntos de vista como son aquellas en que se utilizan sensores terrestres o satelitales y mediante las mismas se busca diagnosticar la salud de la cobertura terrestre, o determinar cualquier factor que influya en ella de alguna manera.

Los productores agropecuarios se benefician con la información que brindan este tipo de sistemas permitiendo realizar un diagnóstico general de toda la superficie que tienen cultivada [6], en busca de evidencias de alguna posible deficiencia hídrica o nutricional, o cuantificar el daño en la superficie por algún fenómeno o plaga, o simplemente para conocer lo saludable que se encuentran sus cultivos. Para el caso particular de imágenes multi-espectrales, de amplia cobertura como son las imágenes satelitales poseen limitaciones en la resolución espacial, la resolución espectral, el período de revisita o frecuencia temporal de pasaje del satélite y el costo de la información en términos generales. Es por este motivo que cada día son mayormente utilizados los sistemas que poseen aeromodelos ya sea tripulados o no tripulados (UAV) para captar imágenes y estimar mediante las mismas datos de cultivos [7]. Se debe tener en cuenta que un sistema de este tipo podría ser utilizado prácticamente en el momento que sea requerido.

La utilización de sensores aerotransportados tienen amplio espectro de aplicaciones dentro de múltiples líneas de investigación actuales. Por ejemplo la adquisición de datos de flujo de gases con efecto invernadero es una temática activa en aquellas naciones que firmaron el protocolo de Kyoto. Por ende todas las investigaciones que relacionen estas temáticas tendrán gran utilidad y aplicación futura.

Durante el año 2015, en la Facultad Regional San Francisco, se ha ensamblado un pequeño planeador con capacidad para montar una cámara go-pro sobre su fuselaje. También se ha probado con éxito el transmisor de 2,4 Ghz para el control del vuelo, como

así también el sistema de transmisión de video en tiempo real. Si bien existen hoy en día desarrollos sobre aeromodelos no tripulados, con sistemas automáticos de vuelo, no existen desarrollos que combinen un sistema de control de estabilidad de tres ejes electrónico que permita montar una cámara go-pro de alta definición, sobre un planeador y cuadricóptero. Se prevee analizar la eficiencia del control electrónico de estabilidad a desarrollar sobre ambos tipos de aeromodelos.

Por otra parte el control de estabilidad de la cámara go-pro, le permitiría al sistema obtener fotografías y videos de mejor calidad, pudiéndose de esta manera realizar un análisis más minucioso de los datos extraídos y estimar datos de interés como rendimientos, pérdidas o cantidad de superficies inundadas con mucha mayor precisión a la hora de aplicar procesamiento digital de imágenes sobre las mismas. También se prevee en el desarrollo del sistema de estabilidad electrónico de tres ejes controlar la posición de la cámara y el momento del disparo en tiempo real, de modo que el sistema sea novedoso, innovador y de gran utilidad para la aplicación concreta en que se desea aplicar.

Otro aporte de investigación es la comparación entre las distintas imágenes extraídas y el análisis de los datos relevados mediante procesamiento de imágenes, que permitirían en principio comparar las virtudes y defectos de ambos aeromodelos (planeador y cuadricóptero) en función de la variable estratégica del cultivo que se desea estimar.

La principal contribución al avance científico es la estimación de datos de interés estratégico sobre los cultivos, los cuales permitiría en principio la caracterización de los rendimientos en función del tipo de cultivo, la zona y variables locales como temperaturas, humedad, cantidad de lluvias, etc. Por otra parte también permite al productor agropecuario una estimación rápida de ganancias y pérdidas después de eventos no esperados como inundaciones por ejemplo, y evaluar en forma rápida y concreta posibles soluciones para tomar decisiones en cuanto a la logística de la cosecha.

En el ámbito tecnológico, se presenta la posibilidad de implementar distintos desarrollos de hardware y software. Es de esperar que la solución que se pretende desarrollar sea referente en cuanto a la implementación de un sistema de corrección entiempos real, de bajo peso y de gran estabilidad para poder ayudar a la

obtención de imágenes y videos de mayor estabilidad y calidad. Por otra parte un sistema de este tipo abre un abanico de posibilidades concretas de desarrollo futuro en cuanto a la utilización por ejemplo de cámaras espectrales para estimar distintos índices de vegetación como el índice normalizado diferencial de vegetación (NVDI), que está relacionado con la actividad fotosintética de la planta y la estructura foliar, permitiendo determinar la vigorosidad de la planta.

En cuanto a transferencia al medio, el sistema pretende tener aplicación inmediata para que sea utilizado en un principio, para el productor para la toma de decisiones concretas, sobre todo en momentos en que el acceso a la zona a estudiar se hace inviable por inundaciones o caminos anegados por ejemplo.

Línea de Investigación y Desarrollo

Nuestro proyecto consta de los siguientes ejes:

1. Determinación de las necesidades de hardware y software que se requieren para el desarrollo del sistema de estabilidad en general.
2. Estudio de los dispositivos que pueden ser integrados en el sistema, en base a la disponibilidad y oferta en Argentina.
3. Definición de la estrategia de procesamiento digital de imágenes a utilizarse en función del objetivo planteado (detección de porcentaje de superficies inundadas, estimación de pérdidas, rendimientos, etc).
4. Desarrollo del software de control y hardware para el sistema de control de estabilidad de la cámara.

Objetivos y Resultados

OBJETIVO GENERAL: desarrollar un sistema de control electrónico de estabilidad de tres ejes con el fin de mantener estable una cámara de alta definición go-pro, que permita contrastar las virtudes y defectos de ambos aeromodelos en función de la variable estratégica del cultivo que se desea estudiar utilizando las imágenes extraídas en el relevamiento.

OBJETIVOS ESPECÍFICOS:

- Evaluar la eficiencia del control electrónico de estabilidad de tres ejes para cámara de alta definición go-pro en función de las imágenes y videos obtenidos en ambos aeromodelos.
- Estimar mediante procesamiento digital de imágenes tipo de cultivo, rendimientos, cantidad de superficie inundada, etc.
- Procesar la información estimada con la información real del sitio elegido a modo de calibrar el sistema estimar mediante procesamiento digital de imágenes tipo de cultivo, rendimientos, cantidad de superficie inundada, etc.
- Contrastar las virtudes y defectos de ambos aeromodelos en función de la variable que se desea estimar estimar mediante procesamiento digital de imágenes tipo de cultivo, rendimientos, cantidad de superficie inundada, etc.
- Aportar los resultados de la investigación a la región de modo de concientizar la utilización de aeromodelos no tripulados para la toma de decisiones en el sector agropecuario.
- Sembrar las bases necesarias para el desarrollo de un aeromodelo no tripulado en la Facultad Regional San Francisco, a modo de profundizar cuestiones académicas, profesionales y de extensión en el ámbito agropecuario y aportar los resultados de la investigación a la región de modo de concientizar la utilización de aeromodelos no tripulados para la toma de decisiones en el sector agropecuario.

RESULTADOS: considerando que el proyecto se encuentra en su fase inicial, los resultados que se esperan, se pueden resumir como:

- ✓ Estabilización de la cámara go-pro con un sistema de estabilización electrónico frente a posibles movimientos de los aeromodelos y vibraciones.
- ✓ Estimación de los distintos datos estratégicos de los cultivos tales como tipos de cultivos, rendimientos, estimación de pérdidas, cantidad de superficies inundadas, necesidades hídricas (cantidad de agua que necesitan), etc.

- ✓ Planificar clases prácticas en asignaturas y cursos relacionados con la tecnología utilizada en la agricultura de precisión y la automatización en general, a fin de complementar la enseñanza en las carreras de grado de Ingeniería Electrónica, Mecánica, e Informática.
- ✓ Realizar transferencia de tecnología a empresas del sector agropecuario.

Formación de Recursos Humanos

El director del proyecto, Mg. Esp. Ing. Peretti Gastón Carlos realizó la Maestría en Ciencias de la Ingeniería, mención Telecomunicaciones en la Universidad Nacional de Córdoba. En la actualidad es docente de la carrera de Ing. Electrónica de la Universidad Tecnológica Nacional Facultad Regional San Francisco, en la cátedra de Dispositivos Electrónicos y posee publicaciones en congresos y libros referidas a las líneas de investigación sistemas de control y comunicaciones. El Codirector del proyecto, Ing. Sergio H. Gallina ha cursado los módulos correspondientes a la Maestría en Ingeniería de software y a la especialidad en Gestión Estratégica de los Servicios de Telecomunicaciones, actualmente trabaja en la tesis de la especialización manteniendo una estrecha relación con otras instituciones tales como el Dpto. Electrónica de la Facultad de Ciencias Exactas de la Universidad Nacional de Córdoba y el Dpto. Electrónica de la UTN Facultad Regional San Francisco. El Ing. Felissia Sergio, se encuentra actualmente cursando la Maestría en Sistemas de Control aplicado en la UTN Facultad Regional Paraná y es docente de la misma universidad en las cátedras de de Informática I, Sistemas de Control Aplicado y Control de Procesos. El grupo de investigación también está conformado por alumnos avanzados de la carrera Ingeniería Electrónica de la Facultad Regional San Francisco y la Universidad Nacional de Catamarca.

El impacto esperado del proyecto se basa en la fundación de un grupo de trabajo interdisciplinario orientado al control y los sistemas de comunicaciones con un objetivo único y general que es el de generar conciencia en el manejo eficiente de los recursos naturales en el ambiente agropecuario.

Referencias

- [1] Berni A., Zarco P., Suarez P., Fereres E., "Thermal and Narrow-band Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle", IEEE Transactions on Geoscience and Remote Sensing, 2008.
- [2] Suarez, L., Zarco P.J., Gonzalez V., Berni J.A., Fereres E. Septiembre, "Detección de stress hídrico y calidad de frutos en cultivos mediante el índice PRI a través de imágenes de alta resolución espacial", Calatayud, Salomón Montesinos Aranda, pp. 29-32, 2009.
- [3] Melchiori A., "Sensores aerotransportados para el estudio de los recursos naturales. Instituto de clima y agua". CNIA-INTA, 2009.
- [4] Berni A., Zarco P., Sepulcro G, Federes E, Villalobos F. Septiembre, "Estimación de conductancia estomática y detección de estrés hídrico en vegetación mediante imágenes térmicas de alta resolución espacial obtenidas con un vehículo aéreo no tripulado (UAV)". Calatayud, Salomón Montesinos Aranda, pp. 25-28, 2009.
- [5] Schellberg J., Hill M., Rothmund M., Braun M., "Precision agriculture in grassland: Applications, perspectives and constraints". European journal of Agronomy, vol 80, pp. 59-71, 2008.
- [6] Jensen T., Apan A., Young F., Zeller L., "Detecting the attributes of a wheat crop of awheat crop using digital imagery acquired from a low- altitude platform", Computers and Electronics in Agriculture, vol 59, pp 66-77, 2007.
- [7] Kymberly Tuck., "Tilt sensing using linear accelerometers", Acelerometer Systems and Applications Systems. Rev. 6, 2007.

Sistema de Monitoreo Continuo de Niveles de Densidad de Potencia Electromagnética Presentes en el Medio Ambiente

Juan Cruz Guidi, Lorenzo De Pasquale, Miguel Angel Banchieri, Guillermo Reggiani, Sergio Pellegrino y Mauricio Jose Mancini
 Departamento Electrónica, Facultad Regional Bahía Blanca,
 Universidad Tecnológica Nacional
 11 de Abril 461, (B8000LMI) Bahía Blanca
pasquale@frbb.utn.edu.ar, mbanch@frbb.utn.edu.ar, ghreggiani@frbb.utn.edu.ar,
spellegrino@frbb.utn.edu.ar, guidiutn@hotmail.com.ar,
mancini_mauricio@hotmail.com

RESUMEN

En últimos años el crecimiento de los sistemas inalámbricos a puesto de manifiesto la necesidad de desarrollar herramientas tecnológicas que permitan definir políticas de planificación y control que protejan a la población de la energía electromagnética presente en el ambiente. En el presente proyecto se pretende desarrollar un sistema de monitoreo de la densidad de potencia irradiada por las diversas fuentes emisoras, empleando para su estudio a la Descomposición Wavelet Packet (WPD) basada en la Transformada Wavelet Discreta (DWT) para segmentar y detectar las anomalías. Su procesado posterior se hará mediante periodograma, Transformada Discreta de Fourier (DFT) empleando el algoritmo para realizar la Transformada Rápida de Fourier (FFT) o la Transformada Z Chirp (CZT). La primera etapa de este proyecto será depurar el desarrollo del conjunto SDR- Computadora (el cual fue publicado en el catálogo como Producto Innovador con el N° 18471 de la Undécima Edición del Concurso Nacional de Innovaciones INNOVAR 2015) y el sistema de análisis y comparación con valores normalizados. La segunda etapa sería depurar y robustecer el análisis de la señal adquirida y clasificada por la WPD mediante periodograma, FFT o CZT. La tercer etapa sería la construcción de un prototipo para el montaje en campo de forma de poder realizar las pruebas correspondientes.

Palabras Claves: WPD, DWT, inalámbrico, analisis.

CONTEXTO

La línea de investigación presentada se encuentra en ejecución en el marco del proyecto titulado: “Sistema de monitoreo continuo de niveles de densidad de potencia electromagnética presentes en el medio ambiente”. El mismo es financiado por la Universidad Tecnológica Nacional, y su ámbito de realización es el Grupo SiTIC (Grupo de I+D en Sistemas y Tecnologías de la Información y las Comunicaciones) de la UTN - Facultad Regional Bahía Blanca.

1. INTRODUCCIÓN

Debido a la elevada y creciente proliferación de sistemas y equipos de comunicaciones inalámbricos vista en los últimos años, y la incertidumbre que existe en torno a ellos con respecto a potenciales efectos nocivos sobre el medio ambiente y, específicamente sobre la salud humana, resulta necesario desarrollar herramientas tecnológicas que posibiliten definir políticas de planificación y control que protejan a la población de la energía electromagnética presente en el ambiente. En el universo de equipos que emiten energía electromagnética, merece especial atención los sistemas de comunicación de radio frecuencia y en especial los inalámbricos de telefonía móvil. Para conocer el nivel de energía electromagnética o específicamente la densidad de potencia que irradian las Estaciones Emisoras, y determinar si cumplen con la normativa vigente, se efectúan mediciones en

banda angosta o en banda ancha. Las mediciones directas en campo representan costos muy elevados por el equipamiento a utilizar, insumen mucho tiempo cuando se trata cubrir grandes espacios, y deben ser realizadas por personal altamente capacitado. Una forma de tener información de manera eficaz es monitorear en forma continua independizándose del operador. En el presente proyecto se pretende desarrollar un sistema de monitoreo de la densidad de potencia irradiada por las diversas fuentes emisoras, en la ciudad de Bahía Blanca, empleando para su estudio a la Descomposición Wavelet Packet (WPD) basada en la Transformada Wavelet Discreta (DWT) para segmentar y detectar las anomalías. Su procesamiento posterior se hará mediante periodograma, Transformada Discreta de Fourier (DFT) empleando el algoritmo para realizar la Transformada Rápida de Fourier (FFT) o la Transformada Z Chirp (CZT) Para poder determinar la densidad de potencia se utilizará un sistema de medición compuesto por un módulo receptor que cubra el ancho de banda de los canales de radiofrecuencia de bajada de la estación base, que para este PID será una Radio Definida por Software (SDR) unida a placa de microcomputadora cuyo hardware es libre, de bajo consumo y desarrollada para emplear software de código abierto. Las capturas realizadas por el conjunto descripto será enviada vía Internet a un centro de monitoreo en el cual se realizaría el análisis off-line de la emisiones detectadas, fuera de rango, mediante la WPD para determinar tiempo, frecuencia y niveles de energía en dicho ancho de banda. La WPD resulta ser una aproximación más flexible para determinar con mayor exactitud el tiempo y el rango de frecuencia por medio de la variación del tamaño de las ventanas y los umbrales de detección. Las mediciones fuera de rango, en lo que respecta a densidad de potencia, se analizarán mediante periodograma, FFT o CZT para determinar fehacientemente la frecuencia y valor de la densidad de potencia de la anomalía. La primera etapa de este proyecto será depurar el desarrollo del conjunto SDR-Computadora (el cual fue publicado en el catálogo como Producto Innovador con el N° 18471 de la

Undécima Edición del Concurso Nacional de Innovaciones INNOVAR 2015) y el sistema de análisis y comparación con valores normalizados. La segunda etapa sería depurar y robustecer el análisis de la señal adquirida y clasificada por la WPD mediante periodograma, FFT o CZT. La tercer etapa sería la construcción de un prototipo para el montaje en campo de forma de poder realizar las pruebas correspondientes.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los posibles efectos nocivos sobre la salud humana que pueden generar las estaciones de radio frecuencia y en particular las de comunicaciones inalámbricas hacen necesario desarrollar diversas herramientas tecnológicas para poder definir políticas de planificación y control que protejan a la población. Los organismos vivos se hallan sometidos diariamente en los ambientes urbanos a radiaciones electromagnéticas causadas por las emisiones de las antenas de los sistemas de comunicaciones. Los niveles de exposición resultan variables dependiendo de la distancia a las fuentes, la presencia de estructuras pasivas que intensifican o reducen las amplitudes de las ondas, la posición relativa a la orientación de la antena, entre otros factores.

Las radiaciones de microondas pulsadas de baja intensidad pueden interferir los campos electromagnéticos intrínsecos de las estructuras biológicas de los seres vivos y producir respuestas anómalas en “ventanas” de frecuencia específicas, que resultan fuertemente dependientes de las intensidades y los tiempos de exposición.

Numerosas investigaciones han reportado efectos nocivos a nivel molecular, celular de procesos inmunes y sobre los sistemas nervioso, endocrino, reproductivo.

Dado que las mediciones directas en campo representan costos muy elevados, además de insumir mucho tiempo cuando se trata de abarcar grandes espacios, la alternativa inmediata resulta el monitoreo continuo de las

emisiones para su análisis y comparación con los valores normalizados.

El procesado con la Transformada Discreta Wavelet se ha detectado en las publicaciones [1], [2], [3] y [4], [5], [6], como aplicaciones de radio frecuencia. Así esta aplicación será utilizada como una posibilidad de proponer una determinación en el tiempo y niveles de densidad de potencia electromagnética.

Al momento en el país para medir la densidad de potencia se utilizan instrumentos que deben ser operados por personal calificado. Los datos obtenidos son los del instante de la medición y no se tienen los registros en todo tiempo, los cuales permitirían realizar estudios estadísticos para determinar las posibles fluctuaciones de los niveles de densidad de potencia de acuerdo a la franja horaria, época del año, condiciones ambientales y otras.

Con la implementación de la estación de monitoreo se podrá contar con una colección de datos en forma continua sobre el comportamiento de las emisoras seleccionadas. Esto permitirá evaluar los resultados obtenidos, para posteriormente conocer los desvíos y calcular la incertidumbre.

De los estudios realizados por este grupo de trabajo [7], [8], [9], [10] se determinó que el empleo de la Descomposición Wavelet Packet (WPD) basada en la Transformada Discreta Wavelet (DWT) es una buena forma de detectar el rango de frecuencia donde se producen densidades de potencia fuera del umbral preestablecido.

Las consideraciones anteriores nos permiten concluir que el estudio de las emisiones electromagnéticas con la WPD y el posterior análisis de la segmentación obtenida con Transformadas Rápida de Fourier (FFT) o Z Chirp, resulta ser un método innovador que reduciría los costos de medición en los distintos ámbitos que se requiera obtener información acerca de los niveles de densidad de potencia electromagnética, así también como lograr un sistema de medición económica y con menor personal necesario para su desarrollo.

3. RESULTADOS OBTENIDOS

Dentro de la temática y objetivos de este proyecto, podemos mencionar que hemos desarrollado una experiencia desde hace unos años plasmada a través de diversos trabajos publicados entre los cuales podemos mencionar:

“Determinación de los cambios de potencia irradiada en señales moduladas en frecuencia mediante la transformada discreta wavelet”, Macchi, Marcos, De Pasquale, Lorenzo, Banchieri, Miguel Angel. uEA 2013, IV Congreso de MicroElectrónica Aplicada, 25, 26 y 27 de Septiembre de 2013, Universidad Tecnológica Nacional, Facultad Regional Bahía Blanca, Argentina.

“La Transformada Wavelet Packet en las Técnicas de Sensado Espectral”, Macchi Konrad, J. M., De Pasquale, L., Banchieri, M. A., V Congreso de Microelectrónica Aplicada 2014, 14 al 16 de mayo de 2014, Instituto Universitario Aeronáutico, Córdoba. ISBN 978-987-34680-5-2

“Wavelet Hardware Processing Unit for Transient Signal Detection”, Juan Marcos Macchi Konrad, Lorenzo De Pasquale, Miguel Banchieri, Guillermo Reggiani, Ricardo Cayssials and Edgardo Ferro Proceedings of the IX IEEE Southern Programmable Logic Conference, pp.:3-8, November 5-7, 2014, Buenos Aires, Argentina, ISBN - 978-1-4799-6848-0.

“Sistema de Medición de Potencia de RF Empleando un Software Defined Radio (SDR,)” Damián Banfi, Juan Marcos Macchi Konrad, Lorenzo De Pasquale, Miguel Angel Banchieri, Matías Alvarez, VI Congreso de Microelectrónica Aplicada 2015, 27 al 29 de mayo de 2015, Universidad Nacional de La Matanza. ISBN 978-987-3806-24-7.

“Determinación de Transitorios en Sistemas Físicos usando la Transformada Discreta Wavelet”, Juan Marcos Macchi Konrad, Lorenzo De Pasquale, Miguel Angel Banchieri, Leandro Nereo Ortiz, WICC 2014, XVI Workshop de Investigadores en Ciencias de la Computación, 7 y 8 de Mayo 2014, Universidad Nacional de Tierra del Fuego, Instituto de Desarrollo Económico e Innovación, Ushuaia, Argentina. RedUNCI, ISBN: 978-950-34-1084-4.

“Aplicaciones de la Transformada Wavelet a Sistemas de Comunicaciones”, Juan Marcos Macchi Konrad, Lorenzo De Pasquale, Miguel Angel Banchieri, VI Congreso de Microelectrónica Aplicada 2015, 27 al 29 de mayo de 2015, Universidad Nacional de la Matanza. ISBN 978-987-3806-24-7

Los objetivos del proyecto son:

Objetivo general:

Desarrollar un sistema de monitoreo para determinar el valor de la densidad de potencia entregada por las estaciones emisoras de radio frecuencia y comunicaciones inalámbricas.

Objetivos específicos:

- Segmentar las densidades de potencia, en los distintos rangos de frecuencia, de las señales emitidas por las estaciones emisoras de radio frecuencia y comunicaciones inalámbricas por medio de la Descomposición Wavelet Packet (WPD)

- Analizar las señales segmentadas que excedan los umbrales de densidad de potencia prefijados en la WPD por medio de las Transformadas Discreta de Fourier a través de un algoritmo FFT o Z Chirp (CZT)

- Determinar desvíos e incertidumbres con respecto los métodos de medición empleados comúnmente.

- Desarrollar un posible diagrama de un sistema de monitoreo integrado para la ciudad de Bahía Blanca.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro del marco del proyecto y de lo que se pretende llevar a cabo, se incorporan alumnos para que se inicien en tareas de investigación. Esto permitirá los alumnos con un adecuado perfil académico puedan en el futuro seguir adelante con un posgrado. Como recursos disponibles dentro de la UTN se cuenta con la posibilidad de Becas de Iniciación a la Investigación (BINID) tanto para alumnos como para graduados y becas de investigación de la Secretaría de Asuntos Estudiantiles (Becas SAE) para alumnos.

5. BIBLIOGRAFIA

- [1] Said E. El-Khamy, Mohamed S. El-Mahallawy, El-Nasser S. Youssef, "Improved Wideband Spectrum Sensing Techniques Using Wavelet-Based Edge Detection for Cognitive Radio", 2013 International Conference on Computing, Networking and Communications, Wireless Communications Symposium, pp. 418-423.
- [2] D.D.Ariananda, M.K.Lakshmanan and H.Nikookar, "A Wavelet Packet Transceiver for Spectral Analysis and Dynamic Spectrum Access ", IEEE, 2011.
- [3] D.D.Ariananda, M.K.Lakshmanan and H.Nikookar, "A Study on the Application of Wavelet Packet Transforms to Cognitive Radio Spectrum Estimation", Proceedings of the 4th International Conference on CROWNCOM 2009.
- [4] Tevfik Yücek, Hüseyin Arslan, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications", IEEE Communications Surveys & Tutorials, VOL. 11, NO. 1, FIRST QUARTER 2009,
- [5] K.C. Ho, H. Liu, L. Hong, "On improving the accuracy of a wavelet based identifier to classify CDMA signal and GSM signal", ISCAS '99. Proceedings of the 1999 IEEE International Symposium on Circuit and Systems, Vol. 4, pp.564-567.
- [6] H. Liu, K.C. Ho, "Identification of CDMA Signal and GSM signal using the wavelet transform", 42nd Midwest Symposium on Circuit and Systems, 1999, vol. 2, pp. 678-681.
- [7] Catálogo como Producto Innovador con el N° 18471 de la Undécima Edición del Concurso Nacional de Innovaciones INNOVAR 2015
- [8] Macchi Konrad, J. M., De Pasquale, L., Banchieri, M. A. “Análisis Basado en la Transformada Wavelet para Determinar Duración de Transitorios en Señales”, IV Congreso de Microelectrónica Aplicada 2013, 25 al 27 de septiembre de 2013, Facultad Regional Bahía Blanca de la UTN.
- [9] Macchi Konrad, J. M., De Pasquale, L., Banchieri, M. A., “La Transformada Wavelet Packet en las Técnicas de Sensado Espectral”, V Congreso de Microelectrónica Aplicada

2014, 14 al 16 de mayo de 2014, Instituto Universitario Aeronáutico, Córdoba.

[10] J. M. Macchi Konrad, L. De Pasquale, M. A. Banchieri, “Análisis Basado en la Transformada Wavelet Packet para Determinar Bandas de Frecuencias Vacantes”, V Congreso de Microelectrónica Aplicada, 2014.

Procesamiento Distribuido y Paralelo

Algoritmos, Estrategias y Análisis de Arquitecturas Orientados al Manejo de Datos Masivos

Rubén Apolloni, Mercedes Barrionuevo, Mariela Lopresti, Natalia Miranda, Cristian Perez-Monte, Fabiana Piccoli, Marcela Printista, Cristian Tissera

LIDIC- Univ. Nacional de San Luís

San Luís, Argentina

{rubenga, mdbarrio, omlopres, ncmiran, mpiccoli, mprinti, [ptissera](mailto:ptissera@unsl.edu.ar)}@unsl.edu.ar

Resumen

En la vida cotidiana, existen problemas cuya solución requiere trabajar con gran cantidad de datos. Algunos de estos problemas incluyen la detección de anomalías en el tráfico en redes, el desarrollo de algoritmos eficientes en la toma de decisiones usando modelos de simulación y el desarrollo de infraestructura para mejorar aspectos de consumo y generación de calor.

En este trabajo se expone distintas líneas de trabajo a seguir teniendo como objetivo desarrollar técnicas de Computación de Alto Desempeño para resolver este tipo de problemas.

Palabras clave: Computación de Alto Desempeño, Datos masivos, Arquitecturas Multicore y Manycore.

Contexto

Esta propuesta de trabajo se lleva a cabo dentro del proyecto de investigación “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” y del proyecto binacional CAPG-BA 66/13 entre la

Universidad Nacional de San Luis y la Universidad de Pernambuco, Recife, Brasil.

El proyecto de investigación se desarrolla en el marco del Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC), de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis y el Centro de Informática de la UFPE.

Introducción

El uso masivo de Internet, el surgimiento de nuevas tecnologías y el crecimiento en la velocidad de transmisión de datos originó nuevos conceptos tal como Big Data. Éste [MCJ13] es un conjunto de grandes volúmenes, diversos o no estructurados, complejos, longitudinales o distribuidos de datos, generados desde transacciones en Internet, sensores, instrumentos, vídeos, mails, redes sociales y una variedad de fuentes digitales disponible en la actualidad y también futuras. El conjunto de datos es tan grande y complejo que los medios tradicionales de procesamiento son ineficaces. Por lo cual es un desafío analizar, capturar, recolectar, buscar, compartir, almacenar, transferir, visualizar, etc., cantidades masivas

de información, obtener conocimiento y realizar toda su gestión en un tiempo razonable [N13].

Todo lo expuesto anteriormente, nos lleva a tener la necesidad de utilizar nuevas técnicas y arquitecturas para contribuir a mejorar el procesamiento y los tiempos de respuesta. Las técnicas de computación de altas prestaciones (HPC) permitirán resolver con eficiencia cada uno de los objetivos a plantear.

Uno de los campos de aplicación de Big Data es la detección de anomalías en redes de datos, la cual consiste en la identificación de patrones que se desvían del comportamiento normal del tráfico en una red [BLMP16]. Detectar posibles ataques en la red requiere contar con tecnologías para su clasificación, asociando flujos de datos con las aplicaciones que los generan. Uno de los desafíos actuales es trabajar con un conjunto de datos, los cuales crecen a mayor velocidad que su capacidad de procesamiento. Por ejemplo, utilizar y procesar imágenes para representar el tráfico de red a fin de detectar tráfico anómalo, tiene como ventajas no sólo contar con una herramienta de visualización de tráfico, sino también con las propiedades de las imágenes y su procesamiento: técnicas bien conocidas y naturaleza paralela de las computaciones.

Otro campo de aplicación está relacionado con la recuperación y análisis de grandes volúmenes de datos para la toma de decisiones basados en técnicas de simulación. Este tipo de sistemas normalmente utilizan datos generados en tiempo real provenientes de distintas fuentes, los cuales son usados para desarrollar simulaciones orientadas a reducir la incertidumbre en los escenarios abordados

De acuerdo a todo lo expuesto, el procesamiento de grandes volúmenes de datos nos introduce en una nueva era de la computación, debido a que genera mayores demandas del procesador, de la memoria en todos los niveles (tanto a memoria principal y memoria cache) [HP08], de los dispositivos de almacenamiento, y también requiere nuevas soluciones de software, ejemplo de ellos son MapReduce[DG04], Hive[CWR12] e Impala[R13], los cuales permiten procesar terabytes de información sin necesidad de cambiar las estructuras de datos subyacentes.

Entre los requerimientos de hardware, se encuentra la necesidad de mayor cantidad de almacenamiento para los datos, introduciendo nuevos desafíos tanto en las investigaciones como en los desarrollos. Además, estas aplicaciones requieren mayor capacidad de memoria, esperándose un incremento en la demanda. Otro aspecto a considerar es el consumo de energía, criterio muy importante a tener en cuenta en el diseño y desarrollo de sistemas de computadoras ya que está directamente relacionado con el consumo de energía total de la infraestructura computacional.

Con el continuo crecimiento de la Ley de Moore [G65], se observa una constante reducción del tamaño de los transistores, lo que permite diseñar procesadores más potentes, con mayor cantidad de núcleos capaces de empaquetar más datos dentro de una pastilla.

Con los actuales sistemas de computación, el paralelismo se hace omnipresente a todos los niveles. A nivel micro, el paralelismo es explotado desde los circuitos, el paralelismo a nivel de pipeline e instrucciones sobre procesadores multicore. A nivel macro, se promueve el paralelismo desde múltiples máquinas en un rack a muchos rack en un centro de datos, hasta llegar a infraestructuras

globales basadas en Internet [RR11].

La presente propuesta tiene como objetivo aplicar técnicas HPC en las etapas del proceso de obtención de información a partir de datos masivos considerando arquitecturas multi y manycore como arquitecturas subyacentes, así como la búsqueda de soluciones a los diferentes problemas que se plantean en la siguiente sección.

Líneas de Investigación, Desarrollo e Innovación

Mejorar el trabajo con Big Data implica considerar diferentes áreas, estas constituyen sendas líneas de investigación. Para lograrlo nos planteamos las siguientes:

- Detectar anomalías en redes, consiste en la identificación de patrones que se desvían del comportamiento normal de tráfico. Con el fin de descubrir comportamientos anormales, se deben utilizar modelos de tráfico precisos y estables para describir un comportamiento de tráfico libre de anomalías. Este es un paso crítico en su detección, ya que un modelo de tráfico incorrecto o inestable causaría un alto número de falsas alarmas.

Modelar el tráfico de red es realizado mediante imágenes, debido a que facilita la comprensión de las características tanto a gran como a pequeña escala de los datos, permitiendo revelar propiedades no sólo relacionada a los datos en sí, sino a la forma en la cual fueron recolectados.

En esta línea, el objetivo es detectar posibles anomalías en el tráfico de una red haciendo uso de una combinación de técnicas de análisis de tráfico de red, procesamiento de imágenes y HPC.

- Algoritmos y Estrategias de recuperación de datos para soporte en la toma de decisiones: Actualmente existe una

tendencia a desarrollar sistemas de soporte a la toma de decisiones basados en técnicas de simulación. Estos sistemas generalmente usan datos generados en tiempo real por diferentes tipos de sensores o dispositivos móviles para realizar la simulación del sistema a modelar e incluso estos datos son utilizados para corregir posibles desviaciones en la ejecución en curso. Ejemplos de estos sistemas son, la monitorización de individuos para el desarrollo de estrategias de evacuación, mitigar el impacto de enfermedades infecciosas [CT13, FC16] o el estudio de cuencas de ríos [AG16] con la finalidad de lanzar alertas tempranas ante inundaciones. En estos casos, además de contar con un modelo de simulación, es importante desarrollar algoritmos y estrategias de alto desempeño que nos permitan trabajar con grandes volúmenes de datos heterogéneos, con el objetivo de que puedan ser procesados por los sistemas de simulación para la toma de decisiones.

- El análisis de las arquitecturas de procesadores y de las jerarquías de memoria es importante para determinar el desempeño de un sistema HPC, y aún más dado que los volúmenes de datos actuales requieren mayor capacidad de memoria. Una manera de abordar los problemas de densidad, consumo, desempeño y escalabilidad de las tecnologías de memorias y almacenamientos tradicionales, es empleando las Memorias No Volátiles (NVM). A pesar de que se avizoran nuevas tecnologías NVM, también se introducen nuevos desafíos a ser abordados, tales como limitada durabilidad y alta latencia de las escrituras. Otro de los aspectos es considerar el manejo de la cache de último nivel (LCC), la cual es compartida por todos los núcleos del procesador y presenta inconvenientes cuando el número de núcleos aumenta: la

contención producida por las aplicaciones que la comparten se incrementa, el rendimiento de estos sistemas estará influenciado por la eficiencia del manejo de esta cache.

Todas las líneas de investigación mencionadas tienen en cuenta la portabilidad de los desarrollos a pesar de las características propias de cada uno de los datos no estructurados.

Resultados y Objetivos

Como objetivos de las líneas de investigación nos planteamos facilitar el desarrollo de soluciones paralelas portables, de costo predecible y bajo consumo, capaces de explotar las ventajas de modernos ambientes de HPC a través de herramientas y “frameworks de computación” de alto nivel. Para ello será necesario proponer nuevas metodologías a ser aplicadas en cada una de las fases del tratamiento de datos masivos.

Formación de Recursos Humanos

Los resultados esperados respecto a la formación de recursos humanos son hasta el momento el desarrollo de 6 tesis doctorales y 4 tesis de maestría. Además se están ejecutando varias tesinas de grado.

Referencias

[AG16] A. Gaudiani, E. Luque, P. Garcia, M. Naiouf, A. De Giusti. “*Optimización y computación paralela aplicadas a mejorar la predicción de un simulador de cauce de ríos*”, XXII Congreso Argentino de Ciencias de la Computación. CACIC 2016. Pp. 179-188. Octubre 2016, San Luis, Argentina

[BLMP16] Mercedes Barrionuevo, Mariela Lopresti, Natalia Miranda, Fabiana Piccoli. “*Un enfoque para la detección de anomalías en el tráfico de red usando imágenes y técnicas de Computación de Alto Desempeño*”. XXII Congreso Argentino De Ciencias de la Computación. CACIC 2016. Pp. 1166-1175. Octubre 2016, San Luis, Argentina

[CT13] P.C. Tissera, M. Printista, E. Luque. “*Simulating behaviors to face up an emergency evacuation*”, International Journal of Soft Computing and Software Engineering- JSCSE. Volumen 3. Pp. 857-863. 2013

[CWR12] E. Capriolo , D. Wampler , J. Rutherglen. “*Programming Hive: Data Warehouse and Query Language for Hadoop*”. O'Reilly Media. 2012.

[DG04] J. Dean and S. Ghemawat: “*MapReduce: Simplified Data Processing on Large Clusters*”. Proc. Sixth Symposium on Operating System Design and Implementation, 2004.

[FC16] F. Casares, P.C.Tissera, F. Piccoli. “*A parallel proposal for SEIR model using Cellular Automata*”. XXII Congreso Argentino de Ciencias de la Computación. CACIC 2016. Pp. 208-219. Octubre 2016, San Luis, Argentina.

[G65] G. E. Moore. “*Cramming More Components onto Integrated Circuits. Electronics*”. Proceedings of the IEEE, Volume 38, No. 8, pp 114-117. April 19, 1965.

[HP08] J. L. Hennessy and D. A. Patterson. “*Computer Organization & Design - The Hardware/Software Interface*”. Morgan Kaufmann, 4th edition, 2008.

[MCJ13] V. Mayer-Schönberger, K. Cukier. A.I. Jurado. “*Big data: La revolución de los datos masivos*”. Turner. 2013.

[N13] J. Needham. “*Disruptive Possibilities: How Big Data Changes Everything*”. Kindle Edition. O'Reilly Media Inc. 2013.

[R13] J.Rusell. “*Cloudera Impala*”. O'Reilly Media, Inc.2013.

[RR11] T. Rauber, G. Runger. “*Parallel Programming for multicore and Cluster Systems*”. Springer. 2011.

Arquitecturas Multiprocesador en Computación de Alto Desempeño: Software, Métricas, Modelos y Aplicaciones

De Giusti Armando ^{1,2}, Tinetti Fernando ^{1,3}, Naiouf Marcelo¹, Chichizola Franco¹, De Giusti Laura^{1,3}, Villagarcía Horacio^{1,3}, Montezanti Diego¹, Encinas Diego¹, Pousa Adrián¹, Rodríguez Ismael¹, Rodríguez Eguren Sebastián¹, Iglesias Luciano¹, Paniego Juan Manuel¹, Pi Puig Martín¹, Dell'Oso Matías¹, Mendez Mariano^{1,4}

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires

Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

² CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CIC - Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴ Facultad de Ingeniería - Universidad de Buenos Aires

{degiusti, fernando, mnaiouf, francoch, ldgiusti, hvw, dmontezanti, dencinas, apousa, ismael, seguren, li, mmendez, jmpaniego, mpipuig, mdelloso}@lidi.info.unlp.edu.ar, marianomendez@gmail.com

Con la colaboración en la dirección de Tesis de Posgrado de la Universidad Autónoma de Barcelona (España) y la Universidad Complutense de Madrid (España).

RESUMEN

Caracterizar las arquitecturas multiprocesador distribuidas enfocadas especialmente a cluster y cloud computing, con énfasis en las que utilizan procesadores de múltiples núcleos (multicores, GPUs y Xeon Phi), con el objetivo de modelizarlas, estudiar su escalabilidad, analizar y predecir performance de aplicaciones paralelas, estudiar el consumo energético y su impacto en la performance así como desarrollar esquemas para detección y tolerancia a fallas en las mismas.

Profundizar el estudio de arquitecturas basadas en GPUs y su comparación con clusters de multicores, así como el empleo combinado de GPUs y multicores en computadoras de alta performance. Iniciar investigación experimental con arquitecturas paralelas basadas en FPGAs. En particular estudiar performance en Clusters “híbridos”. Analizar y desarrollar software de base para clusters, tratando de optimizar el rendimiento.

Investigar arquitecturas multicores asimétricas, desarrollar algoritmos de planificación en el software de sistema operativo para permitir la optimización del

rendimiento y consumo energético en aplicaciones de propósito general.

Estudiar clases de aplicaciones inteligentes en tiempo real, en particular el trabajo colaborativo de robots conectados a un cloud y procesamiento de Big Data.

Es de hacer notar que este proyecto se coordina con otros proyectos en curso en el III-LIDI, relacionados con Computación de Alto Desempeño, Algoritmos Paralelos, Sistemas Distribuidos y Sistemas de Tiempo Real.

Palabras claves: *Sistemas Paralelos – Multicore – GPU – FPGAs - Cluster y Cloud Computing - Cluster híbridos - Performance y eficiencia energética - Tolerancia a fallas – Planificación – Scheduling - Cloud Robotics - Big Data - Simulación de Modelos de Hardware.*

CONTEXTO

Esta línea de Investigación está dentro del proyecto 11/F018: “Arquitecturas multiprocesador en HPC: Software de base, Métricas y Aplicaciones” acreditado por el Ministerio de Educación y de proyectos específicos apoyados por organismos nacionales e internacionales. También del proyecto “Computación de Alto Desempeño,

Minería de Datos y Aplicaciones de interés social en la Provincia de Buenos Aires” financiado por la CIC PBA dentro de la convocatoria a Proyectos de Innovación y Transferencia en Areas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA).

El III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MINCYT y en esta línea de I/D hay cooperación con varias Universidades de Argentina, de América Latina y Europa en proyectos con financiación nacional e internacional.

En la Facultad de Informática de la UNLP (a partir del equipo del proyecto) se han incorporado asignaturas optativas en la currícula de grado de las carreras de Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación relacionadas con Cloud Computing, Programación sobre GPGPUs y procesamiento de Big Data. Además, la Facultad aprobó y financia el proyecto “Transformación de Algoritmos para Nuevas Arquitecturas Multiprocesador”.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática, así como el desarrollo de la Maestría y Especialización en Computación de Altas Prestaciones, acreditadas por CONEAU.

Por último, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, INTEL, AMAZON AWS) en las temáticas de Cloud Computing y Big Data.

1. INTRODUCCIÓN

La investigación en Paralelismo (a partir de arquitecturas multiprocesador distribuidas o concentradas en supercomputadoras) es una de las líneas de mayor desarrollo en la Ciencia Informática actual [GRA03]. La utilización de clusters, multiclusters, grids y clouds, soportadas por redes de diferentes características y topologías se ha generalizado, tanto para el desarrollo de algoritmos paralelos orientados a HPC como

para el manejo de aplicaciones distribuidas y/o servicios WEB concurrentes [GRA03][MCC12].

El cambio tecnológico, fundamentalmente a partir de los procesadores multicore, ha impuesto la necesidad de investigar en paradigmas "híbridos", en los cuales coexisten esquemas de memoria compartida con mensajes [LEI12]. Asimismo la utilización de aceleradores (GPU, FPGA, Xeon Phi) presenta una alternativa para alcanzar un alto speedup en determinadas clases de aplicaciones [KIN09][SIN12]. Debe notarse que el modelo de programación orientado a estas arquitecturas cambia sensiblemente y la optimización de código paralelo requiere nuevos recursos.

Este cambio también se refleja en la aparición de los procesadores multicore asimétricos (AMPs) que integran en un mismo chip diversos tipos de cores con distintas características (frecuencia, microarquitectura o consumo), pero con el mismo repertorio de instrucciones. Los cuales ofrecen un mayor rendimiento por watt y unidad de área que los multicores simétricos [SAE10][ANN12].

Por su lado, las arquitecturas tipo "Cloud" se presentan como una evolución natural del concepto de *Clusters* y *Grids*, integrando grandes conjuntos de recursos virtuales (hardware, plataformas de desarrollo y/o servicios), fácilmente accesibles y utilizables por usuarios distribuidos, vía WEB [VAQ09]. Estos recursos pueden ser dinámicamente reconfigurados para adaptarse a una carga variable, permitiendo optimizar su uso. Al enfocarnos en Cloud Computing aparecen problemas clásicos de la Ciencia Informática, extendidos para este nuevo modelo de arquitectura: planificación, virtualización, asignación dinámica de recursos, migración de datos y procesos [ARD09][VAZ09]. En el proyecto se ha abierto una línea específica dedicada a Cloud Computing y su uso para aplicaciones de *Big Data* o en aquellas que requieren centralizar el accionar de "robots" distribuidos en tiempo real (Cloud Robotics) [MAY13].

Los avances tecnológicos y la búsqueda continua de mayor eficiencia hacen necesario investigar diferentes componentes de las arquitecturas. En particular, las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con herramientas que permitan predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores. Para esto se puede utilizar el modelado basado en agentes y simulación (Agent-Based Modeling and Simulation, ABMS) [MAC06].

La importancia que ha adquirido el uso del paralelismo para aplicaciones científicas de gran duración, requiere la necesidad de estudiar los problemas de detección y tolerancia a fallos en arquitecturas paralelas, debido al alto costo de relanzar la ejecución desde el comienzo en caso de resultados incorrectos. Esto se debe lograr tratando de minimizar el overhead temporal y de aprovechar la redundancia de recursos de hardware que caracteriza a estas arquitecturas [GOL09][FIA11]. El manejo de fallos es una preocupación creciente en HPC, se esperan crecimientos en las tasas de errores, mayores latencias de detección y elevadas cantidades de fallos silenciosos con capacidad de corromper los resultados de las aplicaciones. Por ello, se han desarrollado estrategias de detección y recuperación de fallos transitorios basadas en replicación de software, detectando divergencias en las comunicaciones entre réplicas para evitar que la corrupción se propague a otros procesos, restringiendo así la latencia de detección [MON14][MON15]. De esta forma se permiten obtener ejecuciones fiables con resultados correctos o conducir al sistema a una parada segura. La recuperación puede lograrse mediante múltiples checkpoints de nivel de sistema o de un único checkpoint de capa de aplicación.

Desde otro punto de vista, interesan también los problemas que significan integración de redes de sensores con modelos del mundo real (por ej. modelos meteorológicos, hídricos o de terreno) para prevención de emergencias [SIN06][GAU16]. En esta línea,

el eje del proyecto sigue estando en la problemática del paralelismo combinado con sistemas de tiempo real, pudiendo contribuir a proyectos multidisciplinarios, en particular por temas de emergencias hídricas, exploración de recursos naturales y temas de atención sanitaria y evacuación de edificios en situaciones de emergencia.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio y caracterización de arquitecturas paralelas: clusters, grids, clouds, aceleradores (GPU, FPGA, Xeon Phi) e híbridos.
- Desarrollo de algoritmos de planificación de procesos orientado a procesadores asimétricos para optimizar el rendimiento general. Análisis en los diferentes niveles: sistema operativo, compiladores, técnicas de programación.
- Desarrollo de aplicaciones concretas (numéricas y no numéricas) sobre diferentes máquinas paralelas utilizando técnicas de optimización adecuadas a cada arquitectura.
- Desarrollo de técnicas de tolerancia a fallas en sistemas paralelos y distribuidos, lo cual supone una mejora en el aprovechamiento de la redundancia de recursos que no resultan eficientemente utilizadas en dichas arquitecturas.
- Desarrollo de herramientas para la transformación de código heredado, buscando su optimización sobre arquitecturas paralelas.
- Integración de métricas de rendimiento computacional y energético. Predicción de performance de aplicaciones paralelas.
- Cloud Computing. Software de base. Desarrollo de aplicaciones de HPC (principalmente de big data).
- Sistemas inteligentes distribuidos de tiempo real aprovechando la potencia de cómputo del Cloud (Cloud Robotics).
- Utilización de ABMS para desarrollar un modelo de Entrada/Salida en HPC que permita predecir cómo cambios realizados en los diferentes componentes del mismo

afectan a la funcionalidad y el rendimiento del sistema.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Estudiar modelos complejos, que integren redes de sensores en tiempo real y cómputo paralelo. Estrategias de predicción de catástrofes (inundaciones, incendios por ejemplo) se basan en estos modelos con alta capacidad de procesamiento y monitoreo de señales en tiempo real [GAU16].
- Se han desarrollado diferentes aplicaciones adaptadas para diferentes arquitecturas "híbridas" (que combinan clusters, multicores y aceleradores), y analizado/comparado el rendimiento obtenido [RUC16][POU15][MON16].
- Se está trabajando en técnicas de recuperación a partir de múltiples checkpoints de nivel de sistema, que sirvan para garantizar la correcta finalización de aplicaciones científicas sobre sistemas de HPC, que resultan afectadas por la ocurrencia de fallas transitorias externas y aleatorias, integrando esta solución con las herramientas de detección desarrolladas previamente [MON14][MON15].
- Se han desarrollado nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE15].
- Desarrollo de un modelo de la Entrada/Salida en HPC por medio de ABMS (Agent-Based Modeling and Simulation) que permita predecir cómo cambios realizados en los diferentes componentes del modelo afectan a la funcionalidad y el rendimiento del sistema [ENC15].
- Desarrollo de aplicaciones vinculadas con "Big Data", especialmente para resolver en Cloud Computing [BAS15]. (en relación con los otros proyectos del III- LIDI).
- Optimización de algoritmos paralelos para controlar el comportamiento de múltiples robots que trabajan colaborativamente, considerando la distribución de su capacidad de procesamiento "local" y la coordinación con la potencia de cómputo y capacidad de almacenamiento (datos y conocimiento) de un Cloud.
- Actualización y modernización de código fuente de Sistemas Heredados (Legacy Systems) de Cómputo Científico a través de la aplicación de un proceso de desarrollo iterativo e incremental dirigido por transformaciones de código fuente, apoyado fuertemente en las herramientas de desarrollo. Dichas transformaciones se implementan para ser aplicadas automáticamente en un entorno integrado de desarrollo [MEN14][TIN15].
- Trabajar en la implementación de transformaciones que ayuden a la paralelización del código fuente, así como también en herramientas de análisis estático de Código fuente [TIN13][TIN15].
- Adaptar las técnicas de scheduling y mapeo de procesos a procesadores de acuerdo a los objetivos actuales (en particular los relacionados con el consumo), considerando la migración dinámica de datos y procesos en función de rendimiento y consumo [GRA03][DEG10]. Se debe incluir la utilización de los registros de hardware de los procesadores para la toma de diferentes decisiones en tiempo de ejecución.

4. FORMACIÓN DE RECURSOS HUMANOS

En cooperación con Universidades iberoamericanas se ha implementado la Maestría en Cómputo de Altas Prestaciones y se continúa dictando la Especialización en Cómputo de altas Prestaciones y Tecnología GRID. Asimismo se tiene un importante número de doctorandos (del país y del exterior) realizando el Doctorado en Ciencias Informáticas de la UNLP.

Se han organizado las IV Jornadas de Cloud Computing & Big Data (JCC&BD), y se realizarán en junio de 2017 las V JCC&BD, integrando una Escuela con cursos de Posgrado relacionados con la temática de las líneas de investigación presentadas.

Existe cooperación a nivel nacional e internacional y dentro de la temática del proyecto se espera alcanzar 5 Tesis de Doctorado y 6 Tesis de Maestría en los próximos 3 años, en el país. Al menos tener 3 Doctorandos en el exterior/mixtos en el mismo período.

En 2016 se aprobaron 2 Tesis Doctorales [RUC16][MEN16] y otra está entregada para su evaluación. También se aprobaron 3 trabajos de Especialista, 1 Tesis de Magister y 2 Tesinas de grado.

5. BIBLIOGRAFÍA

- [ANN12] Annamalai A., Rodrigues R., Koren I., Kundu S., "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt," 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.
- [ARD09] Ardisson L., Goy A., Petrone G., Segnan M. "From Service Clouds to User-centric Personal Clouds". 2009 IEEE Second International Conference on Cloud Computing.
- [BAS15] Basgall M. J., Aquino G., Lanzarini L., Naiouf M. "Un Enfoque Dinámico para la Detección de Relaciones entre Tópicos en Textos provenientes de Redes Sociales". III Jornadas de Cloud Computing y Big Data. Facultad de Informática. UNLP. 2015.
- [DEG10] De Giusti L., Chichizola F., Naiouf M., De Giusti A.E., Luque E. "Automatic Mapping Tasks to Cores - Evaluating AMTHA Algorithm in Multicore Architectures". IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010. ISSN (Online): 1694-0784. ISSN (Print): 1694-0814. Págs. 1-6.
- [ENC15] D. Encinas et al., "Modeling I/O System in HPC: An ABMS Approach". The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015
- [FIA11] Fialho L. "Fault Tolerance configuration for uncoordinated checkpoints". Ph.D. Thesis, Universidad Autónoma de Barcelona, Barcelona, España, Julio 2011.
- [GAU16] Adriana Gaudiani, Emilio Luque, Pablo García, Mariano Re, Marcelo Naiouf, Armando De Giusti. "How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation". Advances and New Trends in Environmental and Energy Informatics (part V). ISBN 978-3-319-23455-7. Pp337-351. 2016.
- [GOL09] Golander A., Weiss S., Ronen R. "Synchronizing Redundant Cores in a Dynamic DMR Multicore Architecture". IEEE Transactions on Circuits and Systems II: Express Briefs Volume 56, Issue 6, 474-478. 2009.
- [GRA03] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. Pearson Addison Wesley, 2003.
- [KIN09] Kindratenko, V.V.; Enos, J.J.; Guochun Shi; Showerman, M.T.; Arnold, G.W.; Stone, J.E.; Phillips, J.C.; Wen-Mei Hwu, "GPU clusters for high-performance computing," Cluster Computing and Workshops, 2009. CLUSTER '09. IEEE International Conference on, vol., no., pp.1,8, Aug. 31 2009-Sept. 4 2009
- [LEI12] Leibovich F., Chichizola F., De Giusti L., Naiouf M., Tirado Fernández F., De Giusti A. "Programación híbrida en clusters de multicore. Análisis del impacto de la jerarquía de memoria". XII Workshop de Procesamiento Distribuido y Paralelo. CACIC2012. ISBN: 978987-1648-34-4. Pág. 306-315. 2012.
- [MAC06] C. Macal, M. North, "Tutorial on agent-based modeling and simulation part 2: how to model with agents", in: Proceedings of the Winter Simulation Conference, 2006.
- [MAY13] Mayer-Schönberger V., Cukier K., "Big Data: A revolution that will transform

- how we live, work and think” Houghton Mifflin Harcourt 2013.
- [MCC12] McCool M., Robison A., Reinders J. “Structured Parallel Programming” Elsevier-Morgan Kaufmann, 2012.
- [MEN14] Méndez M., Tinetti F. “Integrating Software Metrics for Fortran Legacy into an IDE”. XI Workshop Ingeniería de Software – CACIC 2014. ISBN 978-987-3806-05-6. Pág. 771-780. San Justo, Buenos Aires, Argentina. Octubre 2014.
- [MEN16] Méndez M. “Aplicaciones de Cómputo Científico: Mantenimiento del Software Heredado”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016
- [MON14] Diego Montezanti, Enzo Rucci, Dolores Rexachs, Emilio Luque, Marcelo Naiouf, Armando De Giusti. “A tool for detecting transient faults in execution of parallel scientific applications on multicore clusters”. *Journal of Computer Science & Technology (JCS&T)*, Vol. 14, Nro. 1 (Abril 2014). Páginas: 32 a 38. ISSN: 1666-6038.
- [MON15] Diego Montezanti, Dolores Rexachs, Enzo Rucci, Emilio Luque, Marcelo Naiouf, Armando De Giusti. “Caracterización de una estrategia de detección de fallos transitorios en HPC”. XXI CACIC 2015 (Congreso de Ciencias de la Computación) - XV Workshop de Procesamiento Paralelo y Distribuido (WPDP). ISBN: 978-987-3724-37-4.
- [MON16] E. Montes de Oca, L. C. De Giusti, F. Chichizola, A. E. De Giusti, M. Naiouf. “Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo”. *Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – Workshop de Procesamiento Distribuido y Paralelo*. Octubre 2016. Pp 159-168.
- [POU15] Adrian Pousa, Victoria Sanz, Armando De Giusti. Comparación de rendimiento de algoritmos de cómputo intensivo y de acceso intensivo a memoria sobre arquitecturas multicore. Aplicación al algoritmo de criptografía AES. CACIC 2015.
- [RUC16] Rucci, Enzo. “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
- [SAE10] Sáenz J. C. "Planificación de Procesos en Sistemas Multicore Asimétricos". Ph.D. Thesis, Universidad Complutense de Madrid, España, 2010.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias. ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems. ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SIN06] Singh B. D. "Influence of Digital Elevation Models Derived from Remote Sensing on SpatioTemporal Modelling of Hydrologic and Erosion Processes". Ph.D. Thesis. pp 370. Cranfield University Siloe England. 2006.
- [SIN12] Sinha, R.; Prakash, A.; Patel, H.D., "Parallel simulation of mixed-abstraction SystemC models on GPUs and multicore CPUs," *Design Automation Conference (ASP-DAC)*, 2012 17th Asia and South Pacific, pp.455,460. 2012.
- [TIN13] F. G. Tinetti, M. Méndez, A. De Giusti. “Restructuring Fortran legacy applications for parallel computing in multiprocessors”. *The Journal of Supercomputing*, May 2013, Volume 64, Issue 2, ISSN 0920-8542, DOI 10.1007/s11227-012-0863-x, pp 638-659.
- [TIN15] M. Mendez, F. G. Tinetti. “Integrating Software Metrics for Fortran Legacy into an IDE” *Computer Science & Technology Series – XX Argentine Congress of Computer Science, Selected Papers.*, ISBN: 978-987-1985-71-5, Red UNCI, págs. 126-134, 2015.
- [VAQ09] Vaquero L.M. et al. "A Break in the Clouds: Towards a Cloud Definition". *ACM SIGCOMM Computer Communication Review*, vol. 39, num. 1, páginas 50-55, ISSN 0146-4833. Enero 2009.
- [VAZ09] Vázquez Blanco C., Huedo E., Montero R. S., Llorente I. M. "Elastic Management of Cluster-based Services in the Cloud". *Proceedings* pp. 19-24, *ACM Digital Library* 2009. ISBN 978-1-60558-564-2.

Cómputo de Altas Prestaciones. Fundamentos de Cómputo Paralelo y Distribuido. Construcción y Evaluación de Aplicaciones

Marcelo Naiouf⁽¹⁾, Armando De Giusti⁽¹⁾⁽²⁾, Laura De Giusti⁽¹⁾⁽³⁾, Franco Chichizola⁽¹⁾, Victoria Sanz⁽¹⁾⁽²⁾⁽³⁾,
Adrián Pousa⁽¹⁾, Enzo Rucci⁽¹⁾⁽²⁾, Silvana Gallo⁽¹⁾⁽²⁾, Erica Montes de Oca⁽¹⁾, Emmanuel Frati⁽¹⁾,
Mariano Sánchez⁽¹⁾, María José Basgall⁽¹⁾⁽²⁾, Adriana Gaudiani⁽⁴⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires
Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

² CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
³CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires

⁴Universidad Nacional de General Sarmiento
{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, sgallo, emontesdeoca, fefrati,
msanchez, mjbasgall}@lidi.info.unlp.edu.ar, agaudi@ungs.edu.ar

RESUMEN

El eje central de la línea presentada son los temas de procesamiento paralelo y distribuido para HPC (fundamentos y aplicaciones). Interesa la construcción, evaluación y optimización de soluciones con algoritmos concurrentes, paralelos y distribuidos sobre diferentes plataformas de software y arquitecturas con múltiples procesadores (multicore, clusters de multicore, cloud y aceleradores como GPU, FPGA y Xeon Phi), los lenguajes y paradigmas de programación paralela (puros e híbridos), los modelos de representación de aplicaciones paralelas, los algoritmos de mapping y scheduling, el balance de carga, las métricas de evaluación de complejidad y rendimiento (speedup, eficiencia, escalabilidad, consumo energético), y la construcción de ambientes para la enseñanza de la programación concurrente y paralela.

Se propone aplicar los conceptos en problemas numéricos y no numéricos de cómputo intensivo y/o sobre grandes volúmenes de datos (búsquedas, simulaciones, n-body, imágenes, big data, reconocimiento de patrones, bioinformática, etc), con el fin de obtener soluciones de alto rendimiento.

En la dirección de tesis de postgrado existe colaboración con el grupo HPC4EAS (High Performance Computing for Efficient Applications and Simulation) del Dpto. de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona, y con el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid, entre otros.

Palabras clave: Cómputo paralelo y distribuido de altas prestaciones. Algoritmos paralelos y distribuidos. Clusters. Multicore. GPU. Consumo energético. Balance de carga. Aplicaciones. Evaluación de performance.

CONTEXTO

La línea de I/D que se presenta en este trabajo es parte del Proyecto 11/F017 “Cómputo Paralelo de Altas Prestaciones. Fundamentos y Evaluación de rendimiento

en HPC. Aplicaciones a Sistemas Inteligentes, Simulación y Tratamiento de Imágenes” del III-LIDI acreditado por el Ministerio de Educación, y de proyectos acreditados y subsidiados por la Facultad de Informática de la UNLP. Además, existe cooperación con Universidades de Argentina, Latinoamérica y Europa a través de proyectos acreditados por AECID, CyTeD, OEI y CIC y becas de Telefónica de Argentina. Asimismo, el Instituto forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD).

1. INTRODUCCIÓN

El área de procesamiento paralelo se ha convertido en clave dentro de las Ciencias de la Computación, debido al creciente interés por el desarrollo de soluciones a problemas con muy alta demanda computacional y de almacenamiento, produciendo transformaciones profundas en las líneas de I/D [RAU10][KIR12].

El desafío se centra en cómo aprovechar las prestaciones obtenidas a partir de la evolución de las arquitecturas físicas. En esta línea de I/D la mayor importancia está en los algoritmos paralelos y en los métodos utilizados para su construcción y análisis a fin de optimizarlos.

Uno de los cambios de mayor impacto ha sido el uso de manera masiva de procesadores con más de un núcleo (*multicore*), produciendo plataformas distribuidas híbridas (memoria compartida y distribuida) y generando la necesidad de desarrollar sistemas operativos, lenguajes y algoritmos que las usen adecuadamente. También creció la incorporación de placas aceleradoras a los sistemas multicore constituyendo plataformas paralelas de memoria compartida con paradigma de programación propio asociado. Asimismo, los entornos de computación cloud introducen un nuevo foco desde el punto de vista del HPC, brindando un soporte “a medida” para la ejecución de aplicaciones sin la necesidad de adquirir el hardware.

La creación de algoritmos paralelos en arquitecturas multiprocesador no es un proceso directo [MCC12]. El costo puede ser alto en términos del esfuerzo de programación y el manejo de la concurrencia adquiere un

rol central en el desarrollo. Los pasos básicos para diseñar aplicaciones paralelas incluyen particionamiento, comunicación, aglomeración y mapeo de procesos a procesadores, y si bien en las primeras etapas el diseñador puede abstraerse de la máquina sobre la que ejecutará el algoritmo, para obtener buen rendimiento debe tenerse en cuenta la plataforma de destino. En las máquinas multiprocesador, se deben identificar las capacidades de procesamiento, interconexión, sincronización y escalabilidad [PAR09]. La caracterización y estudio de rendimiento del sistema de comunicaciones es de interés para la predicción y optimización de performance, así como la homogeneidad o heterogeneidad de los procesadores.

Muchos problemas algorítmicos se vieron impactados por las máquinas multicore y el uso de clusters de multicore. A partir de incorporar varios chips multicore dentro de un nodo y conectar múltiples nodos vía red, se puede crear una arquitectura NUMA, de modo que los cores en un chip compartan memoria principal, y puedan acceder remotamente a la memoria dedicada de otro chip, aunque ese acceso sea más costoso, surgiendo así varios niveles de comunicación. Esto impacta sobre el desarrollo de algoritmos que aprovechen adecuadamente las arquitecturas, y motiva el estudio de performance en sistemas híbridos. Además, es necesario estudiar la utilización de diferentes lenguajes ya que aún no se cuenta con un standard, aunque puede mencionarse el uso de MPI, OpenMP y Pthreads [MUR11].

Para algunos problemas ha crecido la utilización de placas aceleradoras, como pueden ser las unidades de procesamiento gráfico (GPU, graphic processing unit) de NVIDIA y AMD o los coprocesadores Xeon Phi de Intel [KIR12][JEF13]. Esto se debe a la capacidad que tienen estos dispositivos de alcanzar picos de rendimiento y cocientes de eficiencia energética superiores a los de la CPU a un menor costo. Por otro lado, el uso de FPGAs (Field Programmable Gate Array) se ha vuelto atractivo para HPC debido a la evolución en su capacidad de cómputo, su bajo consumo energético y al desarrollo de nuevas herramientas de programación más familiares para el área [SET13].

La combinación de arquitecturas con múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características [RUC16b]. Más allá del tipo de acelerador utilizado, la programación de esta clase de plataformas representa un verdadero desafío. Para lograr aplicaciones de alto rendimiento, los programadores deben enfrentar diversas dificultades como pueden ser: estudiar características específicas de cada arquitectura y aplicar técnicas de programación y optimización particulares para cada una de ellas, lograr un balance de carga adecuado entre los diferentes dispositivos de procesamiento y afrontar la ausencia de estándares y de herramientas avanzadas para este tipo de sistemas.

Por otra parte, los avances en las tecnologías de virtualización han dado origen al paradigma de Cloud Computing, que se presenta como una alternativa a los

tradicionales sistemas de cluster [EC213]. El uso de cloud para HPC presenta desafíos atractivos, brindando un entorno reconfigurable dinámicamente sin la necesidad de adquirir hardware, y es una excelente plataforma para testear escalabilidad de algoritmos aunque queda mucho por hacer en cuanto al diseño, lenguajes y programación

Métricas de evaluación del rendimiento y balance de carga

La diversidad de opciones vuelve complejo el análisis de performance de los Sistemas Paralelos, ya que los ejes sobre los cuales pueden compararse dos sistemas son varios. Existe un gran número de métricas para evaluar el rendimiento, siendo las tradicionales: tiempo de ejecución, speedup, eficiencia.

Por su parte, la *escalabilidad* permite capturar características de un algoritmo paralelo y la arquitectura en que se lo implementa. Posibilita testear la performance de un programa sobre pocos procesadores y predecirla en un número mayor, así como caracterizar la cantidad de paralelismo inherente en un algoritmo.

El objetivo principal del cómputo paralelo es reducir el tiempo de ejecución haciendo uso eficiente de los recursos.

El *balance de carga* es un aspecto central y consiste en, dado un conjunto de tareas que comprenden un algoritmo y un conjunto de procesadores, encontrar el mapeo (asignación) de tareas a procesadores tal que cada una tenga una cantidad de trabajo que demande aproximadamente el mismo tiempo, y esto es más complejo si hay heterogeneidad. Dado que el problema general de mapping es *NP*-completo, pueden usarse enfoques que dan soluciones subóptimas aceptables. Las técnicas de planificación a nivel micro (dentro de cada procesador) y macro (en un cluster) deben ser capaces de obtener buen balance de carga. Existen técnicas estáticas y dinámicas cuyo uso depende del conocimiento que se tenga sobre las tareas que componen la aplicación [DUM08].

Un aspecto de interés que se ha sumado como métrica es el del consumo energético requerido [BAL13]. Entre las aplicaciones de interés se encuentran las numéricas y no numéricas con alta demanda de cómputo.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Paralelización de algoritmos secuenciales. Diseño y optimización de algoritmos.
- Comparación de lenguajes y bibliotecas para procesamiento paralelo y distribuido.
 - Estudio de complejidad de algoritmos paralelos, considerando multicore y heterogeneidad.
- Modelos y paradigmas de computación paralela. Modelo Map-reduce.
 - Arquitecturas multicore y many-core. Arquitecturas FPGA.
- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores y GPUs) y arquitecturas heterogéneas.

- Programación sobre modelos híbridos: pasaje de mensajes y memoria compartida en cluster de multicore, clusters de GPU, clusters multicore-GPU.
- Técnicas de programación sobre arquitecturas many-core (GPU y Xeon Phi) y FPGA.
- Técnicas para soluciones de HPC en cloud.
- Lenguajes y Estructuras de Datos para nuevas arquitecturas de cómputo paralelo.
- Mapping y scheduling de aplicaciones paralelas sobre distintas arquitecturas multiprocesador. Balance de carga estático y dinámico. Técnicas.
- Desarrollo de soluciones paralelas a problemas de cómputo intensivo y/o con grandes volúmenes de datos (búsquedas, simulaciones, n-body, aplicaciones científicas, big data, bioinformática) sobre diferentes modelos de arquitectura homogéneas y heterogéneas (multicore, clusters, clusters de multicore, GPU, Xeon Phi, FPGA y cloud).
- Evaluación de rendimiento, eficiencia energética y costo de programación de las diferentes soluciones implementadas teniendo en cuenta las arquitecturas y las herramientas de programación utilizadas.
- Ambientes para la enseñanza de programación concurrente

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollar y optimizar algoritmos paralelos sobre diferentes modelos de arquitectura. En particular, en aplicaciones numéricas y no numéricas de cómputo intensivo y tratamiento de grandes volúmenes de datos (big data).
- Utilizar arquitecturas híbridas que combinan memoria compartida y pasaje de mensajes, evaluando performance para distintos modelos de comunicación.
- Estudiar y comparar los lenguajes sobre las plataformas multiprocesador para diferentes modelos de interacción entre procesos.
- Investigar la paralelización en plataformas que combinan clusters, multicore y aceleradores. Comparar estrategias de distribución de trabajo teniendo en cuenta las diferencias en potencias de cómputo y comunicación, dependencia de datos y memoria requerida.
- Evaluar la performance (speedup, eficiencia, escalabilidad, consumo energético) de las soluciones propuestas. Analizar el rendimiento de soluciones paralelas a problemas con diferentes características (dependencia de datos, relación cómputo / comunicación, memoria requerida).
- Mejorar y adecuar las técnicas disponibles para el balance de carga (estático y dinámico) entre procesos a las arquitecturas consideradas.

En este marco, pueden mencionarse los siguientes resultados:

- Para la experimentación se han utilizado y analizado diferentes arquitecturas homogéneas o heterogéneas, incluyendo multicore, cluster de multicore (con 128 núcleos), GPU y cluster de GPU, Xeon Phi y FPGA.

- Se experimentó la paralelización en arquitecturas híbridas, con el objetivo de estudiar el impacto del mapeo de datos y procesos, así como de los lenguajes y librerías utilizadas.

- Respecto de las aplicaciones estudiadas y algoritmos implementados, se trabajó fundamentalmente con los siguientes problemas:

➤ **Best-first search (BFS) paralelo sobre multicore y cluster de multicore.**

El algoritmo de búsqueda BFS es utilizado para resolver problemas combinatorios, para los cuales se requiere encontrar una secuencia de acciones que permita transformar una configuración inicial (problema) en una configuración final (solución). En especial, A* es una variante de BFS la cual permite encontrar soluciones de costo óptimo. Estos algoritmos requieren una alta capacidad de cómputo y gran cantidad de memoria, por esto su paralelización es imprescindible. En los últimos años se ha re-impulsado el desarrollo de algoritmos paralelos BFS con el objetivo de aprovechar: (a) la potencia de cómputo de los procesadores *multicore* (b) la gran cantidad de RAM y potencia de cómputo de los *clusters de multicore*. En este sentido, HDA* [KIS13] El algoritmo de búsqueda BFS es utilizado para resolver problemas combinatorios, para los cuales se requiere encontrar una secuencia de acciones que permita transformar una configuración inicial (problema) en una configuración final (solución). En especial, A* es una variante de BFS la cual permite encontrar soluciones de costo óptimo. Estos algoritmos requieren una alta capacidad de cómputo y gran cantidad de memoria, por esto su paralelización es imprescindible. En los últimos años se ha re-impulsado el desarrollo de algoritmos paralelos BFS con el objetivo de aprovechar: (a) la potencia de cómputo de los procesadores *multicore* (b) la gran cantidad de RAM y potencia de cómputo de los *clusters de multicore*. En este sentido, HDA* [KIS13] paraleliza A* sobre clusters utilizando MPI: cada procesador realiza una búsqueda cuasi-independiente utilizando un esquema de distribución de nodos basado en una función hash estándar. Otros autores [BUR10] adaptaron HDA* para máquinas multicore utilizando Pthreads: esta versión elimina el overhead extra del paso de mensajes entre procesadores (threads) en una arquitectura de memoria compartida y utiliza menor cantidad de memoria, ya que los threads comparten estructuras de datos comunes. Los mismos autores notaron que utilizar una función hash estándar para asignar nodos a procesadores genera alto overhead de comunicación entre threads. Por ende, propusieron AHDA*, una versión de HDA* para máquinas multicore que utiliza una función especial (función de abstracción) para asignar bloques de nodos a procesadores, en vez de nodos individuales. En una primera etapa de desarrollo, implementamos versiones propias de HDA* (HDA* MPI) [SAN16a] y HDA* para máquinas multicore (HDA* Pthreads) [SAN14]

[SAN15]. Ambas versiones incluyen técnicas para mejorar el rendimiento, respecto a las versiones originales. Luego presentamos Optimized AHDA* [SAN16b], una versión propia de AHDA* que incorpora una técnica para optimizar las transferencias de nodos entre threads, la cual permitió reducir la comunicación y contención. Con el objetivo de explotar eficientemente los recursos de un *cluster de multicore*, desarrollamos Hybrid HDA* (HHDA*), una versión híbrida de HDA* programada con MPI+Pthreads. El trabajo experimental demostró que HHDA* alcanza un rendimiento superior y consume menor cantidad de memoria, comparado con HDA* MPI. Estas mejoras permitieron a HHDA* resolver instancias más complejas del caso de estudio. Como línea de trabajo futuro se plantea adaptar los algoritmos antes mencionados para encontrar soluciones en menor tiempo comprometiendo la calidad de la solución.

➤ **Criptografía de grandes volúmenes de datos.** El encriptado y desencriptado de datos requiere un tiempo de cómputo adicional, que dependiendo de su tamaño puede ser muy alto. AES (Advanced Encryption Standard), es un algoritmo de cifrado simétrico por bloques que se ha convertido en estándar en 2002, y actualmente es el más ampliamente usado para codificar información. Este algoritmo se caracteriza por ser simple, rápido y por consumir pocos recursos. Sin embargo el tiempo de cifrar y descifrar grandes cantidades de datos es importante por lo que es oportuno aprovechar las posibilidades que brindan las arquitecturas multicore para reducir este tiempo. Para este propósito, las arquitecturas multicore actuales, como clusters de multicore y GPUs, proporcionan una forma de acelerar el cómputo de encriptar y desencriptar información logrando un excelente rendimiento [POU15].

➤ **Alineamiento de secuencias biológicas.** Esta operación consiste en comparar dos o más secuencias biológicas, como pueden ser las de ADN o las de proteínas, y resulta fundamental en investigaciones de la bioinformática y la biología molecular. El algoritmo de Smith-Waterman es considerado el método de alineamiento más preciso. Desafortunadamente, este algoritmo resulta costoso debido a su complejidad computacional cuadrática mientras que la situación se agrava aún más a causa del crecimiento exponencial de datos biológicos en los últimos años [RUC16a]. El reciente surgimiento de aceleradores en HPC (GPU, Xeon Phi, FPGA, entre otros) da la oportunidad de acelerar los alineamientos sobre hardware comúnmente disponible a un costo accesible. En primer lugar, se continuó el estudio preliminar del empleo del modelo de programación OpenCL sobre FPGAs para acelerar el alineamiento de secuencias de proteínas [RUC15]. Se desarrollaron diferentes soluciones paralelas para arquitecturas heterogéneas basadas en FPGA de forma

de encontrar la configuración más beneficiosa. Se analizó el rendimiento, la eficiencia energética y el costo de programación de cada solución y se los comparó con el de otras implementaciones basadas en multicores, Xeon Phi y GPUs [RUC16b]. Aprovechando el conocimiento adquirido, se adaptó el código desarrollado para computar alineamientos de secuencias largas de ADN [RUC17]. A futuro, interesa explorar el uso de nuevas generaciones de procesadores y aceleradores para esta aplicación, como pueden ser los procesadores Xeon Phi Knights Landing de Intel.

➤ **Simulación distribuida de modelos orientados al individuo.** El modelado orientado al individuo resulta de gran interés ya que permite analizar y extraer conclusiones acerca de un sistema a través de la simulación de la interacción de sus individuos. No obstante, a medida que los modelos incorporan más características del sistema se vuelven más complejos y en consecuencia se necesita mayor cantidad de cómputo y comunicación para lograr resultados significativos. La aparición de arquitecturas distribuidas y procesadores con varios núcleos ha favorecido el desarrollo de dichos modelos a gran escala utilizando técnicas de simulación distribuida. Lograr mejoras en los tiempos de ejecución y en la eficiencia de dichas aplicaciones es esencial. Con el fin de mejorar la distribución de la carga entre los procesos lógicos de la simulación se pretende analizar el desempeño del simulador en cuanto al balance de cómputo, ya que al trabajar con grandes cantidades de individuos, y debido a la cantidad de cómputo asociada a la selección de vecinos para realizar los desplazamientos, se producen desbalances en cuanto a la cantidad de trabajo de cada uno de los diferentes procesos lógicos. Dichos desbalances obstaculizan el aprovechamiento de la arquitectura disponible, por lo que se deben estudiar diferentes técnicas de balance de cómputo y desarrollar las soluciones que aporten dichas mejoras.

➤ **Problemas de tipo N-body.** Se utilizaron las plataformas de memoria compartida GPU y cluster de multicore para la resolución de problemas con alta demanda computacional del tipo N-body. Se emplearon diferentes modelos de comunicación: memoria compartida (Pthreads en CPU y CUDA en GPU), pasaje de mensajes (MPI) y soluciones híbridas (MPI-Pthreads). Se han mostrado los beneficios del uso de la GPU en problemas con características similares al caso planteado. El trabajo experimental ha dado como resultado una buena aceleración obtenida utilizando cluster de GPU. Además, se observó claramente que el uso del cluster de GPU logró una aceleración proporcional al speedup conseguido con el cluster de CPU pero con tiempos de ejecución significativamente menores [MON14]. También se han desarrollado diferentes alternativas de distribución de trabajado usando un Cluster de GPUs heterogéneas [MON16].

Actualmente, se ha puesto énfasis en el análisis del consumo energético de esta experimentación.

➤ **Problemas de simulación relacionados con fenómenos naturales (inundaciones).** Análisis de diferentes soluciones para la paralelización de este tipo de aplicaciones que son intensivas en cómputo; y el tiempo de ejecución y la performance alcanzable son críticas dado que los resultados que se esperan determinarán alertas y toma de decisiones. La utilización de escenarios de simulación en entornos donde interesa estudiar el comportamiento en situaciones de desastres producidos por fenómenos naturales como las inundaciones. En este ámbito se avanza en dos temas: (1) La implementación de un método de sintonización de un simulador de inundaciones en ríos de llanura, mediante la técnica de simulación paramétrica. El proceso requiere lanzar miles de escenarios de simulación hasta encontrar un conjunto ajustado de parámetros de entrada del simulador. La experimentación se lleva a cabo con un modelo master-worker sobre un cluster [GAU15]. (2) En colaboración con el Laboratorio de Hidrología de la UNLP se comenzó con la paralelización de la simulación de inundaciones producidas por lluvias (en particular en el ámbito de la ciudad de La Plata, donde una corrida “standard” es del orden de las 8 hs), a fin de reducir el tiempo de ejecución a pocos minutos y permitir establecer un sistema de alertas [GAU16].

➤ **Ambientes para la enseñanza de concurrencia.** Se desarrolló el entorno R-INFO para la enseñanza de programación concurrente y paralela a partir de cursos iniciales en carreras de Informática. Incluye un entorno visual que representa una ciudad en la que pueden definirse varios robots que interactúan. Combina aspectos de memoria compartida y distribuida mediante instrucciones para bloquear y liberar esquinas de la ciudad y el concepto de pasaje de mensajes a través de primitivas de envío y recepción. Además, se incluyen los conceptos de heterogeneidad (diferentes velocidades de los robots) y consumo energético [DEG16a]. Se ha integrado con el uso de robots físicos (Lego Mindstorm 3.0) que ejecutan en tiempo real las mismas instrucciones que los robots virtuales y se comunican con el entorno mediante bluetooth [DEG14]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [DEG16b].

➤ **Aplicaciones en Big Data.** En esta línea se está trabajando en la aplicación de técnicas de Machine Learning sobre grandes volúmenes de datos utilizando el framework para el procesamiento paralelo y distribuido de grandes cantidades de datos, llamado MapReduce. Se está llevando a cabo la evaluación de rendimiento de los diferentes algoritmos implementados para las soluciones de procesamiento en batch y en streaming aplicados a problemas de Big Data

[BAS16a][BAS16b]. Se trata de una línea incipiente en el grupo.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D se concluyó 1 tesis doctorales, 3 Trabajos Finales de Especialización y 2 Tesinas de Grado de Licenciatura. Se encuentran en curso en el marco del proyecto 7 tesis doctorales, 2 de maestría, 3 trabajos de Especialización y 2 Tesinas.

Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática UNLP, por lo que potencialmente pueden generarse más Tesis y Trabajos Finales.

Hay cooperación con grupos de otras Universidades del país y del exterior, y tesistas de diferentes Universidades realizan su trabajo con el equipo del proyecto.

5. BIBLIOGRAFÍA

[BAL13] Ballardini J., Rucci E., De Giusti A., Naiouf M., Suppi R., Rexachs D., Luque E. “Power Characterisation of Shared-Memory HPC Systems”. Computer Science & Technology Series – XVIII Argentine Congress of Computer Science Selected Papers. ISBN 978-987-1985-20-3. Pp. 53-65. EDULP, La Plata (Argentina), 2013

[BAS16a] M. J. Basgall, W. Hasperué, M. Naiouf. “Data stream treatment using sliding windows with MapReduce”. Journal of Computer Science & Technology; vol. 16, no. 2. Noviembre 2016. Pp. 76-83. ISSN: 1666-6038.

[BAS16b] M. J. Basgall, W. Hasperué, C. A. Estrebou, M. Naiouf. “Clustering de un flujo de datos usando MapReduce”. Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – XIII Workshop Bases de datos y Minería de Datos. Octubre 2016. Pp 682-691.

[BUR10] Burns E, Lemons S, Ruml W, Zhou R. “Best First Heuristic Search for Multicore Machines”. Journal of Artificial Intelligence Research, Vol.39, No.1, pp. 689-743, 2010.

[DEG14] De Giusti L., Leibovich F., Sanchez M., Chichizola F., Naiouf M., De Giusti A. "Herramienta interactiva para la enseñanza temprana de Concurrencia y Paralelismo: un caso de estudio", Procs XX Congreso Argentino de Ciencias de la Computación – Workshop de Innovación en Educación. Octubre 2014. Pp 133-140. ISBN: 978-987-3806-05-6

[DEG16a] L. C. De Giusti, F. Leibovich, F. Chichizola, M. Naiouf. “Teaching Concurrency and Parallelism Concepts with CMRE”. Journal of Computer Science & Technology; vol. 16, no. 2. Noviembre 2016. Pp 95-100. ISSN: 1666-6038.

[DEG16b] L. C. De Giusti, F. Chichizola, S. Rodriguez Eguren, M. Sanchez, J. M. Paniego, A. E. De Giusti. “Introduciendo conceptos de Cloud Computing utilizando el entorno CMRE”. Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)

- Workshop de Innovación en Educación en Informática. Octubre 2016. Pp 1357-1365.
- [DUM08] Dummler J., Ruaber T., Runger G., Mapping Algorithms for Multiprocessor Tasks on Multi-Core Clusters, Proc. 2008 International Conference on Parallel Processing IEEE CS 2008.
- [EC213] Amazon Elastic Compute Cloud (Amazon EC2). <http://aws.amazon.com/es/ec2/>. Febrero 2013.
- [GAU15] Adriana Gaudiani. “Simulación y optimización como metodología para mejorar la calidad de la predicción en un entorno de simulación hidrológica”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2015.
- [GAU16] Adriana Gaudiani, Emilio Luque, Pablo García, Mariano Re, Marcelo Naiouf, Armando De Giusti. “How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation”. *Advances and New Trends in Environmental and Energy Informatics (part V)*. ISBN 978-3-319-23455-7. Pp337- 351. 2016.
- [JEF13] Jeffers, James; Reinders, James. “Intel Xeon Phi Coprocessor High Performance Programming”, Morgan Kaufmann, 2013.
- [KIR12] Kirk D., Hwu W. “Programming Massively Parallel Processors, second edition: A Hands-on Approach. Morgan-Kaufmann. 2012.
- [KIS13] A. Kishimoto, A. Fukunaga, and A. Botea, "Evaluation of a simple, scalable, parallel best-first search strategy," *Artificial Intelligence*, vol. 195, pp. 222–248, 2013.
- [MCC12] McCool, Michael. “Structured Parallel Programming: Patterns for Efficient Computation”, Morgan Kaufmann, 2012
- [MON14] E. Montes de Oca, L. De Giusti, F. Chichizola, A. De Giusti, M. Naiouf. "Utilización de Cluster de GPU en HPC. Un caso de estudio". Proceedings del XX Congreso Argentino de Ciencias de la Computación (CACIC 2014), pp. 1220-1227, 2014.
- [MON16] E. Montes de Oca, L. C. De Giusti, F. Chichizola, A. E. De Giusti, M. Naiouf. “Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo”. Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), pp. 159-168, 2016.
- [MUR11] Muresano Cáceres R. “Metodología para la aplicación eficiente de aplicaciones SPMD en clústers con procesadores multicore” Ph.D. Thesis, UAB, Barcelona, España, Julio 2011.
- [PAR09] Parashar M., Li Xiaolin, Chandra Sumir, “Advanced Computational Infrastructures for Parallel and Distributed Applications”, Wiley-Interscience, 2009 ISBN-10: 0470072946.
- [POU15] Adrian Pousa, Victoria Sanz, Armando De Giusti. “Comparación de rendimiento de algoritmos de cómputo intensivo y de acceso intensivo a memoria sobre arquitecturas multicore. Aplicación al algoritmo de criptografía AES”. Actas del XXI Congreso Argentino de Ciencias de la Computación, pp. 163-172, 2015.
- [RAU10] Rauber T., Rüniger G. “Parallel programming for multicore and cluster systems”. Springer. 2010.
- [RUC15] “Smith-Waterman Protein Search with OpenCL on FPGA”. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. Proceedings of 2014 IEEE Symposium on Parallel and Distributed Processing with Applications (ISPA). 20 al 22 de Agosto de 2015. Helsinki, Finlandia. ISBN: 978-1-4673-7952-6. Pp. 208-213. DOI: 10.1109/Trustcom.2015.634.
- [RUC16a] “State-of-the-art in Smith-Waterman Protein Database Search”. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. *Big Data Analytics in Genomics*. Springer, pp. 197-223, 2016.
- [RUC16b] “OSWALD: OpenCL Smith-Waterman on Altera’s FPGA for Large Protein Databases”. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. *International Journal of High Performance Computing Applications*, Online first, 2016.
- [RUC16c] Rucci, Enzo. “Evaluación de rendimiento y eficiencia energética de sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas, Facultad de Informática, UNLP. 2016.
- [RUC17] “Accelerating Smith-Waterman Alignment of Long DNA Sequences with OpenCL on FPGA”. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. *Lecture Notes in Computer Science*, En prensa, 2017.
- [SAN14] V. Sanz, A. De Giusti, and M. Naiouf, "On the Optimization of HDA* for Multicore Machines. Performance Analysis," in PDPTA'14 (The 2014 International Conference on Parallel and Distributed Processing Techniques and Applications), Las Vegas, 2014, pp. 625-631.
- [SAN15] V. Sanz, A. De Giusti, and M. Naiouf, "Performance tuning of the HDA* algorithm for multicore machines." in *Computer Science & Technology Series - XX Argentine Congress of Computer Science - Selected Papers*. Argentina: EDULP, 2015, pp. 51-62. [SAN16b] V. Sanz, A. De Giusti, and M. Naiouf, "Improving Hash Distributed A* for Shared Memory Architectures Using Abstraction," in ICA3PP-2016 - The 16th International Conference on Algorithms and Architectures for Parallel Processing, Granada, 2016, pp. 431 - 439.
- [SAN16a] V. Sanz, A. De Giusti, and M. Naiouf, "Scalability analysis of Hash Distributed A* on commodity cluster: results on the 15-puzzle problem.," in PDPTA'16 - The 22nd International Conference on Parallel and Distributed Processing Techniques and Applications, Las Vegas, 2016, pp. 221-227.
- [SET13] High-performance Dynamic Programming on FPGAs with OpenCL. Sean Settle. 2013 IEEE High Performance Extreme Computing Conf (HPEC '13), 2013.

Estudio de las Mejoras de Aplicar Fog Computing en la Distribución de Servicios en Cloud Computing

Nelson Rodríguez¹, María Murazzo², Susana Chávez³, Adriana Martín⁴, Diego Medel⁵, Jorge Mercado⁶, Matías Montiveros⁷

^{1-5,7}Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

⁶Departamento de Matemática - Facultad de Ingeniería - UNSJ

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

¹nelson@iinfo.unsj.edu.ar ²marite@unsj-cuim.edu.ar ³schavez@iinfo.unsj.edu.ar

⁴arianamartinsj@gmail.com ⁵vdiego.unsj@hotmail.com ⁶jorgenmp@gmail.com,

⁷nicolasmontivero.nw273@gmail.com

Resumen

La amplia adopción de Internet de las cosas (IoT) por parte de los usuarios y la industria, sumada a las estimaciones de crecimiento por parte de las consultoras, anticipa la generación de cantidades excesivas de datos hacia el Cloud. Sin bien el modelo tradicional de Cloud Computing ha dado respuesta a IT y ha tenido en crecimiento notable en los últimos años, está mostrando dificultades con estas nuevas tendencias. En consecuencia aparecieron nuevos requerimientos de IoT que Cloud no puede cumplir como son: restricciones de ancho de banda, alta latencia y baja capacidad de los dispositivos, interrupciones momentáneas de Internet, que junto a nuevos desafíos en seguridad, obligan a redefinir el modelo de Cloud centralizado. Mover parte importante del cómputo y otras funcionalidades al borde de la red, permitirá optimizar los recursos. A consecuencia de esto, surge Fog computing, que es un nuevo modelo que concentra gran parte del almacenamiento, procesamiento y otras tareas en el borde o sea fuera del Cloud. Pero debido a ser un nuevo paradigma, que incluye demasiados componentes y redefine el Cloud tradicional, presenta muchos desafíos y problemas sin resolver. Esta línea de investigación

plantea un estudio profundo de las mejoras que introduce de Fog Computing, comparado con el modelo de Cloud Computing tradicional centralizado.

Palabras clave: *Distributed Computing, Cloud Computing, Fog Computing, Edge Computing, IoT*

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y se enmarca dentro del proyecto de investigación: Evaluación de arquitecturas distribuidas de bajo costo basadas en software libre, que fue aprobado satisfactoriamente para el período 2016-2017. Asimismo el grupo de investigación viene trabajando en proyectos afines relacionados con la computación distribuida y de alta performance desde hace más de 17 años. Las unidades ejecutoras para dichos proyectos son el Departamento e Instituto de Informática de la FCEFYN de la UNSJ, y como resultados de los mismos, se produjeron numerosas publicaciones, transferencias y formación de recursos humanos en el área.

Introducción

Cloud Computing es un modelo tecnológico que permite el acceso ubicuo, adaptado y bajo demanda en red a un conjunto compartido de recursos de computación configurables (por ejemplo: redes, servidores, almacenamiento, aplicaciones y servicios), que pueden ser rápidamente aprovisionados y liberados con un esfuerzo de gestión reducido o interacción mínima con el proveedor del servicio [1].

Cloud ha dominado el discurso de la Informática en la década pasada, debido a que tiene una doble propuesta de valor. En primer lugar la centralización explota las economías de escala para reducir costos de administración y operación de los sistemas. En segundo lugar, las organizaciones pueden evitar el gasto de capital de crear un centro de datos consumiendo recursos informáticos a través de Internet de un gran proveedor de servicios.

Estas consideraciones han llevado a la consolidación de la capacidad de cómputo en múltiples y grandes centros de datos distribuidos por todo el mundo.

Por otro lado, Internet de las cosas (IoT) es una tendencia reciente de la computación distribuida que integra aspectos de la vida real a escalas masivas. A pesar de comunicarse con redes inalámbricas, presentan características diferentes de las redes de sensores en general: no existen estándares, la variedad de tecnologías es importante, existen varias arquitecturas de software, las velocidades de transferencia en general son bajas y las unidades de transferencia son menores que las redes IP en general. En 1999, Kevin Ashton del Instituto Tecnológico de Massachusetts (MIT) acuñó el término Internet de las Cosas. En ese momento, las tecnologías de automatización comenzaban a pasar de la fábrica a nuevos entornos como hospitales, bancos y oficinas. A medida

que las primeras implementaciones de M2M (machine-to-machine) se hicieron cada vez más sofisticadas, las máquinas comenzaron a ser conectadas a otros tipos de dispositivos como servidores, y estos servidores se trasladaron a centro de datos y al Cloud. Hoy en día, Internet de las Cosas puede incluir productos industriales y comerciales, productos cotidianos como lavavajillas y termostatos, y redes locales de sensores para vigilar granjas y ciudades [2].

Esta soluciones que ofrece IoT promueve la incorporación de los dispositivos a la red y se pronostica que entre 20 y 50 millones de los mismos se añadirán a Internet para 2020, creando una economía de más de 3 billones de dólares [3]; En consecuencia, 43 billones de gigabytes de datos serán generados y necesitarán ser procesados en los centros de datos de Cloud. Las aplicaciones que generan datos en dispositivos de usuario, como teléfonos inteligentes, tablets y dispositivos portátiles, usan actualmente Cloud como un servidor centralizado, pero pronto se convertirá en un modelo informático insostenible [4].

La importancia de IoT y las capacidades de Cloud, imponen que esta asociación sea por demás necesaria.

Sin embargo, en la actualidad Cloud Computing está encontrando serias dificultades para satisfacer los requerimientos de Internet de las cosas (IoT). En consecuencia, surgió, como propuesta de solución llevar el almacenamiento, las funciones de red y gran parte del procesamiento hacia el borde de la red, lo que resultó en un nuevo modelo llamado fog computing.

Fog Computing

Fog es una arquitectura emergente para computar almacenamiento, control y servicio de red que distribuye estos

servicios a usuarios finales entre Cloud y las cosas [5].

A veces el término fog es usado intercambiándose con edge, aunque fog es más amplio que la típica noción de edge. Este último, se conoce como mist, que consiste en que las computaciones se realizan en el borde mismo: el sensor y los controladores del actuador. Extender la computación hasta el borde puede tener mucho sentido dependiendo de la topología de red que se necesite.

Fog computing es una arquitectura horizontal a nivel de sistema que distribuye recursos y servicios de computación, almacenamiento, control y red en cualquier lugar a desde Cloud a las cosas.

Soporta múltiples dominios de aplicaciones, ofreciendo inteligencia y servicios a usuarios y empresas

Habilita servicios y aplicaciones que se distribuyen más cerca de las cosas, y en cualquier lugar a lo largo del continuo entre Cloud las cosas [6].

Agrega una jerarquía de elementos entre Cloud y los dispositivos de borde, y entre dispositivos y gateway, para satisfacer estos desafíos de alta performance, abierto e interoperable.

A su vez, han surgido tecnologías que ayudan a que este modelo sea factible. Estas tecnologías y la jerarquía se muestra en la figura 1 [7].

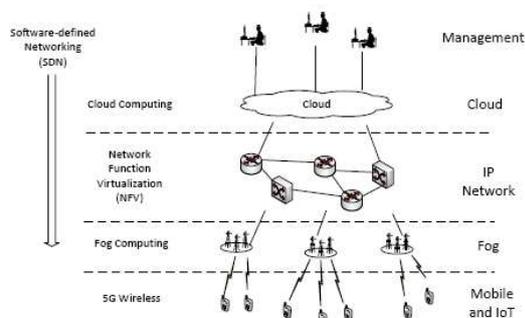


Figura 1. Niveles de jerarquía de Fog

Observando la figura se puede apreciar que en Cloud y fog sucede lo siguiente:

Nodos fog:

Reciben datos de dispositivos IoT usando cualquier protocolo, en tiempo real, Ejecuta aplicaciones habilitadas para IoT para control y análisis en tiempo real, con tiempo de respuesta de milisegundos, proporciona almacenamiento transitorio, a menudo cada 1 o 2 horas y envía resúmenes de datos periódicos al Cloud.

Por su parte la plataforma Cloud: recibe y agrega resúmenes de datos desde los nodos fog, realiza el análisis sobre los datos de IoT y datos de otras fuentes para obtener una visión del negocio y puede enviar nuevas reglas de aplicación a los nodos fog, basadas en los nuevos conocimientos adquiridos.

Otra definición de Fog afirma: que: es un escenario donde un gran número de dispositivos heterogéneos, ubicuos y descentralizados (inalámbricos y posiblemente autónomos) se comunican y potencialmente cooperan entre sí y con la red para realizar tareas de almacenamiento y procesamiento sin la intervención de terceros. Estas tareas pueden ser para el soporte de funciones de red básica o para nuevos servicios y aplicaciones que corren en un ambiente protegido. Los usuarios que ofrecen parte de sus dispositivos para hospedar estos servicios reciben incentivos para hacerlo [8].

Los sectores de gobierno y de la industria que se encuentran beneficiados con estas tecnologías van desde industrial, salud, control vehicular, gobiernos y negocios entre otras. Las aplicaciones de automatización de edificios, casa inteligente, provisión de energía inteligente, cadenas de proveedores inteligentes, autos conectados, realidad aumentada, gaming, cuidados de la salud, sistemas de conocimiento de contexto, el uso de dispositivos especiales como

google glass o relojes, dispositivos embebidos en general, sensores de redes de área de cuerpo y realidad virtual, se asocian en una gran red que solicitan servicios al cloud o suben datos y otros recursos.

Desafío de Fog Computing

Si bien, Fog computing presenta muchas ventajas, existen muchos problemas a solucionar para que este tipo de sistema distribuido ofrezca lo que promete: procesar eficientemente los recursos, lograr un adecuado balanceo de carga para la distribución entre el borde y el cloud, las API's y la administración y compartición de servicios y comunicaciones en redes definidas por el software y virtualizadas [9].

Para la realización del análisis en tiempo real se deben administrar eficientemente los recursos para determinar cual tarea de análisis se debe realizar en los recursos de fog o ser enviado al cloud. Se debe mejorar los modelos de programación y arquitecturas que favorezcan la escalabilidad y flexibilidad dado que las arquitecturas están basadas en configuraciones estáticas, la seguridad debido a que cuenta con múltiples servicios, usuarios, proveedores y recursos y simples soluciones de autenticación no resultan suficientes, la confiabilidad y tolerancia a fallos y el consumo de energía. Todos estos desafíos llevan a minimizar la latencia y maximizar el throughput.

Otro de los problemas a resolver es determinar cuándo y cómo distribuir la computación, la comunicación, el control y el almacenamiento a lo largo de la continuidad de cloud a las cosas. Por ejemplo como descomponer y recomponer las tareas de computación sobre un conjunto heterogéneo de nodos fog con la capacidad suficiente y

disponibilidad con conexiones inalámbricas inestables y limitaciones de ancho de banda y energía.

Fog no va a sustituir a Cloud, ambos presentan sus ventajas. Cloud ofrece el almacenamiento masivo, la computación "pesada", la coordinación global, la conectividad de área extensa y además puede escalar a los requerimientos de IoT, mientras que fog será utilizada para procesamiento en tiempo real, innovación rápida, servicios centrados en el usuario y agrupación de recursos del borde.

Fog y Cloud se complementan uno con otro de forma de continuar el servicio entre el Cloud y los puntos finales proveyendo beneficios mutuos y servicios independientes para hacer posible la computación, el almacenamiento, el control y la comunicación a lo largo de la continuidad.

Líneas de Investigación, Desarrollo e Innovación

Las tareas de investigación se realizarán sobre determinados escenarios, dado que la variedad de protocolos, dispositivos, arquitecturas de los mismos y redes de comunicación, dificultan la obtención de resultados totalmente genéricos. Se analizando la distribución de servicios y otros recursos y se compararán con el Cloud tradicional, en entornos simulados.

Resultados y Objetivos

Resultados Obtenidos

Si bien el grupo de investigación trabaja desde hace 17 años, los tareas de investigación relacionadas a esta línea son más recientes. Durante los últimos seis años se trabajó en proyectos sobre Cloud Computing y en particular durante los últimos tres años sobre Cloud híbridos y Cloud privados, Cluster de commodity y

arquitecturas distribuidas de bajo costo. Dicha experiencia motivó esta línea de investigación. El grupo ha realizado publicaciones en el área durante el último año: nueve trabajos de investigación en diferentes Congresos y Jornadas y cuatro publicaciones en revistas científicas y actividades de divulgación. Se han aprobado cinco tesinas de grado y un trabajo de especialización.

Objetivos

El objetivo del grupo de investigación es realizar el estudio de las mejoras que ofrece un ecosistema de fog computing para la distribución de servicios en la continuidad a Cloud, comparándolo con el modelo de Cloud centralizado tradicional.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los seis docentes-investigadores que figuran en este trabajo y cinco alumnos. Se están realizando tres tesinas de licenciatura una sobre evaluación de algoritmos de algebra lineal sobre arquitecturas diversas, otra sobre Cloud Computing Privado, otra sobre arquitecturas de seguridad sobre Cloud Privado. Se espera realizar también una tesis de maestría sobre Metodologías de desarrollo aplicadas a SaaS y otra sobre control topológico para reducción de interferencia en redes IoT. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otros organismos del estado.

Referencias

- [1] Mell, P., Grance, T.. "The NIST definition of cloud computing". NIST Special Publication 800 – 145 (2011).
- [2] Joe Biron and Jonathan Follett, Foundational Elements of an IoT Solution The Edge, The Cloud, and Application Development. O'Reilly Media, Inc. (2016).
- [3] Garnet: Gartner Says 6.4 Billion Connected "Things" Will Be in Use in 2016, Up 30 Percent From 2015. <http://www.gartner.com/newsroom/id/3165317>.
- [4] Blesson Varghese, Nan Wang, Dimitrios S. Nikolopoulos, Rajkumar Buyya - Feasibility of Fog Computing - arXiv:1701.05451v1cs.DC
- [5] Mung Chiang - Fog and IoT: An Overview of Research Opportunities (Invited Paper)- IEEE Internet of Things Journal, Vol. 3, No. 6, December 2016.
- [6] <https://www.openfogconsortium.org/>
- [7] Tom H. Luan_, Longxiang Gao_, Zhi Liz, Yang Xiang_, Guiyi Wey, and Limin Sunz - Fog Computing: Focusing on Mobile Users at the Edge - arXiv:1502.01815v3 [cs.NI] 30 Mar 2016.
- [8] Luis M. Vaquero, Luis Roderó-Merino - Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing - ACM SIGCOMM Computer Communication Review- Volume 44, Number 5, October 2014 (2014).
- [9] Amir Vahid Dastjerdi and Rajkumar Buyya - Fog Computing: Helping the Internet of Things Realize Its Potential – COMPUTER, IEEE COMPUTER SOCIETY - 0018-91 62 / 16 /2016.

Evolución y Tendencias en Sistemas Paralelos y Distribuidos

Marcelo Naiouf⁽¹⁾, Franco Chichizola⁽¹⁾, Laura De Giusti⁽¹⁾⁽³⁾, Enzo Rucci⁽¹⁾⁽²⁾, Adrián Pousa⁽¹⁾, Ismael Rodríguez⁽¹⁾, Sebastián Rodríguez Eguren⁽¹⁾, Erica Montes de Oca⁽¹⁾, Juan Manuel Paniego⁽¹⁾, Martín Pi Puig⁽¹⁾, César Estrebow⁽¹⁾, Leandro Libutti⁽¹⁾, Javier Balladini⁽⁴⁾, Armando De Giusti⁽¹⁾⁽²⁾

¹Instituto de Investigación en Informática LIDI (III-LIDI),
Facultad de Informática, Universidad Nacional de La Plata – Comisión de Investigaciones Científicas de la
Provincia de Buenos Aires

²CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴Universidad Nacional del Comahue

{mnaiouf,francoch,ldgiusti,erucci,apousa,ismael,seguren,emontesdeoca,jmpaniego,mpipuig,cesarest,llibutti,
degiusti}@lidi.info.unlp.edu.ar; javier.balladini@gmail.com

Resumen

El eje de esta línea de I/D lo constituye el estudio de tendencias actuales en las áreas de arquitecturas y algoritmos paralelos. Incluye como temas centrales:

- Arquitecturas many-core (GPU, procesadores MIC), FPGAs, híbridas (diferentes combinaciones de multicores y aceleradores), y asimétricas.
- Cloud Computing para HPC (especialmente para aplicaciones de Big Data) y sistemas distribuidos de tiempo real (Cloud Robotics).
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento energético y computacional.

Palabras clave: *Sistemas Paralelos. Clusters. Arquitecturas asimétricas. GPU, MIC, FPGA. Cloud Computing. Cloud robotics. Performance y eficiencia energética.*

Contexto

Se presenta una línea de Investigación que es parte de los Proyectos 11/F018 “Arquitecturas multiprocesador en HPC: Software de Base, Métricas y Aplicaciones” y 11/F017 “Cómputo Paralelo de Altas Prestaciones. Fundamentos y Evaluación de Rendimiento en HPC. Aplicaciones a Sistemas Inteligentes, Simulación y Tratamiento de Imágenes” del III-LIDI acreditados por el Ministerio de Educación y de proyectos

específicos apoyados por organismos nacionales e internacionales. También del proyecto “Transformación de algoritmos para nuevas arquitecturas multiprocesador” financiado por la Facultad de Informática de la UNLP.

En el tema hay cooperación con varias Universidades de Argentina y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por CyTED, AECID y la OEI (Organización de Estados Iberoamericanos).

Por otra parte, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, Intel) en la temática de Cloud Computing.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Asimismo el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MinCyT.

Introducción

Una de las áreas de creciente interés lo constituye el cómputo de altas prestaciones, en el cual el rendimiento está relacionado con dos aspectos: por un lado las arquitecturas de soporte y por el otro los algoritmos que hacen uso de las mismas.

A la aparición de arquitecturas *many-core* (como las GPU o los procesadores MIC), se ha sumado el uso de FPGAs debido a su

potencia de cómputo y rendimiento energético. Su combinación en sistemas HPC da lugar a plataformas híbridas con diferentes características [RUC16].

Lógicamente, esto trae aparejado una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los lenguajes mismos de programación y el software de base), así como la evaluación de las soluciones que éstos implementan. También resulta necesario investigar las estrategias de distribución de datos y procesos a fin de optimizar la performance.

Además de las evaluaciones clásicas de rendimiento prestacional como el speedup y la eficiencia, otros aspectos comienzan a ser de interés, tales como el estudio del consumo y la eficiencia energética de tales sistemas paralelos [BAL13].

Por otra parte, los avances en las tecnologías de virtualización y cómputo distribuido han dado origen al paradigma de Cloud Computing, que se presenta como una alternativa a los tradicionales sistemas de Clusters y Multicluster para ambientes de HPC [ROD07][BER08]. A su vez, este concepto se puede ampliar a sistemas distribuidos de tiempo real, en particular sistemas inteligentes como son los robots que pueden trabajar en paralelo utilizando su propia capacidad de procesamiento y al mismo tiempo conectándose con la potencia de un servidor en la nube (*Cloud Robotics*) [LOR13][GUO12][KEH15].

En esta línea de I/D se trabaja sobre aspectos que marcan tendencias en el área.

GPUs y Cluster de GPUs

Las GPUs son el acelerador dominante en la comunidad de HPC al día de hoy por su alto rendimiento y bajo costo de adquisición. En la actualidad, tanto NVIDIA como AMD trabajan especialmente en mejorar la eficiencia energética de sus placas y disminuir el alto costo de programación.

La combinación de GPUs con otras plataformas paralelas como clusters y multicores, brindan un vasto conjunto de posibilidades de investigación en arquitecturas

híbridas, a partir de diferentes combinaciones a saber:

- Máquinas multicore con más de una GPU, que combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.
- Cluster de máquinas multicore cada una con una o más placas de GPU, lo que permite combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas a fin de optimizar el rendimiento de las soluciones.

MIC

En forma reciente Intel brinda una alternativa a partir de la arquitectura MIC (*Many Integrated Core Architecture*). Esta arquitectura permite utilizar métodos y herramientas estándar de programación con altas prestaciones (lo que los distingue especialmente de las GPUs). De esta forma, se remueven barreras de entrenamiento y se permite focalizar en el problema más que en la ingeniería del software. Xeon Phi es el nombre elegido por Intel para su serie de procesadores many-core. Recientemente, Intel ha lanzado Knights Landing (KNL), la segunda generación de Xeon Phi. A diferencia de sus predecesores que operaban como co-procesador a través del puerto PCI, los procesadores KNL pueden operar en forma autónoma. Además, integran las nuevas extensiones vectoriales AVX-512 y tecnología de memoria 3D, entre otras características avanzadas [REI16].

FPGAs

Una FPGA (*Field Programmable Gate Array*) es una clase de acelerador basado en circuitos integrados reconfigurables. La capacidad de adaptar sus instrucciones de acuerdo a la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones. Tradicionalmente fueron utilizadas para el procesamiento digital de señales. Sin embargo, en los últimos años,

existen dos tendencias claras para extender su uso a otros dominios. En primer lugar, el establecimiento de alianzas estratégicas entre fabricantes de procesadores y de FPGAs para integrar estos dispositivos en arquitecturas híbridas (Intel con Altera; IBM con Xilinx) [IBM15][INT16]. En segundo lugar, el desarrollo de nuevas herramientas de programación para FPGAs empleando estándares familiares para HPC, con las cuales se espera reducir los tradicionales tiempos y costos de programación [SEA13][XIL15].

Consumo energético

Un aspecto de interés creciente en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores, es el del consumo energético que los mismos producen.

Muchos esfuerzos están orientados a tratar el consumo como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirlo.

Entre los puntos de interés pueden mencionarse:

- Análisis de metodologías y herramientas para medir consumo energético.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo a las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas.
- Optimización de la eficiencia energética. A partir de los valores de energía que brindan los contadores hardware es posible definir estrategias de programación que lleven a reducir el consumo, manteniendo a su vez el rendimiento en valores aceptables [SAE17].

Cloud Computing

Cloud Computing, proporciona grandes conjuntos de recursos físicos y lógicos (como pueden ser infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), fácilmente accesibles y utilizables por medio de una interfaz de administración web, con un modelo de arquitectura “virtualizada” [SHA10][XIN12]. Estos recursos son proporcionados como servicios (“as a service”) y pueden ser dinámicamente

reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), logrando una mejor utilización y evitando el sobre o sub dimensionamiento (elasticidad) [VEL09][VAQ09].

Más allá de las potenciales características y beneficios que brinda un Cloud, de por sí atractivas, es de gran interés estudiar el despliegue de entornos de ejecución para cómputo paralelo y distribuido (Clusters Virtuales), como así también realizar I/D en la portabilidad de las aplicaciones de HPC en el Cloud [DOE11][ROD11].

Por otro lado, Cloud Robotics es una de las áreas más prometedoras de la investigación informática actual en la cual se cuenta con “robots” dotados de diferentes sensores y capacidades, conectados a un Cloud vía Internet. Los temas de investigación derivados son múltiples: sensores, redes de sensores e inteligencia distribuida; robótica y sistemas colaborativos de tiempo real basados en robots; aplicaciones críticas (por ej. en ciudades inteligentes o en el ámbito industrial).

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [RAS17] u Odroid [ODR17] que poseen múltiples núcleos simples. Asimismo, existen diversos dispositivos móviles con capacidades similares. Es de interés estudiar como explotar el paralelismo en estos dispositivos para mejorar el rendimiento y/o consumo energético de las aplicaciones.

Líneas de Investigación, Desarrollo e Innovación

- Arquitecturas many-core (procesadores MIC y GPU) y FPGA. Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado.
- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs). Diseño de algoritmos paralelos sobre las mismas.

- Exploración de nuevos lenguajes y modelos de programación para HPC.
- Cloud Computing para realizar HPC. Evaluación de performance en este tipo de arquitectura. Análisis del overhead por el software de administración del Cloud.
- Sistemas inteligentes de tiempo real distribuidos (Cloud Robotics).
- Interconexión de Brokers de mensajes QTT sobre Cloud Públicos y Cloud Privados.
- Consumo energético en las diferentes arquitecturas paralelas, en particular en relación a los algoritmos paralelos y la configuración de la arquitectura. Análisis de metodologías y herramientas de medición.

Resultados y Objetivos

Investigación experimental a realizar

- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas. Análisis de rendimiento, eficiencia energética y costo de programación.
- Análisis del overhead introducido por el sistema gestor del Cloud en un entorno de HPC para aplicaciones científicas de Big Data. Comparar el rendimiento entre Cloud y Cluster Computing.
- Analizar y comparar las diferentes estrategias para interconectar brokers de mensajes QTT tanto sobre cloud públicos como privados.
- Realizar el desarrollo de nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE15][SAE17].
- Optimización de algoritmos paralelos para controlar el comportamiento de múltiples robots que trabajan colaborativamente, considerando la distribución de su capacidad de procesamiento “local” y la coordinación con la potencia de cómputo y capacidad de almacenamiento (datos y conocimiento) de un Cloud.

Resultados obtenidos

- En la línea de planificación de procesos sobre multicores asimétricos se evaluaron nuevos algoritmos de planificación sobre los

kernel Solaris y Linux, se analizaron las mejoras en rendimiento y consumo energético para un conjunto diverso de cargas de trabajo [SAE15].

- Se ha trabajado en el desarrollo de una herramienta de monitorización a nivel de sistema operativo que reporte información valiosa a la hora de desarrollar algoritmos de planificación a través del uso de contadores de hardware [SAE17].

- Se ha finalizado una tesis de doctorado focalizada en evaluación de rendimiento y eficiencia energética de diferentes plataformas heterogéneas (integradas por GPUs, FPGAs y Xeon Phi), en la cual se desarrollaron diferentes soluciones paralelas para un problema del área bioinformática con alta demanda computacional [RUC16]. Se analizaron sus rendimientos y el consumo energético y costo de programación asociados.

- Evaluación de estrategias de distribución de carga en clusters heterogéneos de CPUs-GPUs [MON16][POU16].

- Análisis del despliegue de un sistema de múltiples robot interconectados por Wi-Fi a través de un servicio de cloud público para un problema de riego artificial [ROD16].

Organización de Eventos

En el año 2016 se han organizado las IV Jornadas de Cloud Computing y Big Data (JCC&BD 2016) en Argentina, con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing. En junio de 2017 se organizarán las V Jornadas de Cloud Computing y Big Data (JCC&BD 2017).

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyeron: 2 tesis doctorales, 1 trabajos de Especialización y 1 tesina de grado. Al mismo tiempo se encuentran en curso 3 tesis de Doctorado en Ciencias Informáticas y 2 tesis de Maestría.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad

de Informática de la UNLP (acreditadas por la CONEAU con categoría A, B y A, respectivamente), por lo que potencialmente pueden generarse nuevas Tesis de Doctorado y Maestría y Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesis de diferentes Universidades realizando su Tesis con el equipo del proyecto.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación tres materias directamente relacionadas con los temas de la misma: “Taller de Programación sobre GPUs”, “Cloud Computing y Cloud Robotics” y “Conceptos y Aplicaciones en Big Data”.

Referencias

- [BAL13] Balladini, J., Rucci, E., De Giusti, A., Naiouf, M., Suppi, R., Rexachs, D., Luque, E. Computer Science & Technology Series – XVIII Argentine Congress of Computer Science Selected Papers. ISBN 978-987-1985-20-3. Págs. 53-65. 2013.
- [BER08] Bertogna, M., Grosclaude, E., Naiouf, M., De Giusti, A., Luque, E.: “Dynamic on Demand Virtual Clusters in Grids”. In: 3rd Workshop on Virtualization in High-Performance Cluster and Grid Computing (VHPC 08). España. (2008).
- [BOR05] S. Y. Borkar, P. Dubey, K. C. Kahn, D. J. Kuck, H. Mulder, S. S. Pawlowski, y J. R. Rattner: “Platform 2015: Intel® Processor and Platform Evolution for the Next Decade”. In: Intel Corporation, White Paper, 2005.
- [DOE11] Doelitzcher, F., Held, M., Sulistio, A., Reich, C. ViteraaS: “Virtual Cluster as a Service”. In: 3rd IEEE International Conference on Cloud Computing Technology and Science. Grecia (2011).
- [FED09] Alexandra Fedorova, Juan Carlos Saez, Daniel Shelepov, Manuel Prieto: “Maximizing Power Efficiency with Asymmetric Multicore Systems”. In: Communications of the ACM, Vol. 52 (12), pp 48-57. 2009.
- [GUO12] Guoqiang Hu, Wee Peng Tay, Yonggang Wen: “Cloud robotics: architecture, challenges and applications”. In: Network, IEEE, vol.26, no.3, pp.21-28. 2012.
- [IBM15] IBM. “IBM and Xilinx Announce Strategic Collaboration to Accelerate Data Center Applications”. Disponible en <https://www-03.ibm.com/press/us/en/pressrelease/48074.wss>
- [INT16] Intel. “Intel Acquisition of Altera”. Disponible en <http://intelacquiresaltera.transactionannouncement.com>
- [KEH15] Kehoe B., Patil S., Abbeel P., Goldberg K.: “A Survey of Research on Cloud Robotics and Automation”. In: IEEE Transactions on Automation Science and Engineering (T-ASE): Special Issue on Cloud Robotics and Automation. Vol. 12, no. 2. 2015.
- [LOR13] Lorencik D., Sincak P.: “Cloud robotics: Current trends and possible use as a service”. In: Applied Machine Intelligence and Informatics (SAMII), 2013 IEEE 11th International Symposium on , vol., no., pp.85-88. 2013.
- [MON16] Montes de Oca, E., De Giusti, L., Chichizola, F., De Giusti, A., Naiouf, M.: “Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo”. Actas del XXII Congreso Argentino de Ciencias de la Computación. Argentina (CACIC 2016). Págs. 169-168.
- [ODR16] Odroid <http://www.hardkernel.com> Accedido 21 de Marzo de 2016.
- [POU16] Pousa, A., Sanz, V., De Giusti, A.: “Estructurando código paralelos en clusters heterogéneos de CPUs/GPUs”. Actas del XXII Congreso Argentino de Ciencias de la Computación. Argentina (CACIC 2016). Págs. 139-148.
- [RAS16] Raspberry PI. <https://www.raspberrypi.org/> Accedido 21 de Marzo de 2016.
- [REI16] Reinders, J., Jeffers, J., Sodani, A. “Intel Xeon Phi Processor High Performance Programming Knights Landing Edition”. Morgan Kaufmann Publishers Inc., Boston, MA, USA, 2016
- [ROD07] Rodriguez, I. P., Pousa, A., Pettoruti, J.E., Chichizola, F., Naiouf, M., De Giusti, L., De Giusti, A.: “Estudio del overhead en la migración de algoritmos paralelos de cluster y multicluster a GRID”. In: Proceedings del XIII Congreso Argentino de Ciencias de la Computación. Argentina (2007).
- [ROD11] Rodriguez, I., Pettoruti, J.E., Chichizola, F., De Giusti, A.: “Despliegue de un Cloud Privado para entornos de cómputo científico”. In: Proceedings del XVII Congreso Argentino de Ciencias de la Computación. Argentina (2011).
- [ROD16] Rodriguez, I., Paniego, J. M., Rodriguez Eguren, S., Estrebou, C., De Giusti, A. “Cloud Robotics: Sistema Multi-Robot conectado al Cloud público AWS”. Actas del XXII Congreso Argentino de Ciencias de la Computación. Argentina (CACIC 2016). Págs. 189-198.
- [RUC16] Rucci, Enzo: “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias: “ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems”. In: ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SAE17] Saez, J.C., Pousa, A., Rodríguez-Rodríguez, R., Castro, F., Prieto-Matias, M. “PMCTrack: Delivering performance monitoring counter support to the OS scheduler”. The computer journal Volume 60, Issue 1 January 2017.

- [SEA13] Sean Settle: “High-performance Dynamic Programming on FPGAs with OpenCL”. In: IEEE High Performance Extreme Computing Conference. 2013.
- [SHA10] Shafer, J.: “I/O virtualization bottlenecks in cloud computing today”. In: Proceedings of the 2nd conference on I/O virtualization (VIOV10). USA (2010).
- [VAQ09] Vaquero, L. M., Rodero-Merino, L., Caceres, J., Lindner, M.: “A Break in the Clouds: Towards a Cloud Definition”. In: ACM SIGCOMM Computer Communication Review, Volume 39, Issue 1, pp. 50--55. USA. (2009).
- [VEL09] Velte, A.T., Velte, T.J., Elsenpeter, R.: “Cloud Computing: A Practical Approach”. McGraw Hill Professional. 2009.
- [XIL15] Xilinx Inc. “SDAccel Development Environment”. [Online]. Disponible en <http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html>
- [XIN12] Xing, Y., Zhan, Y.: “Virtualization and Cloud Computing”. In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.

Hive: Framework de Sincronización de Objetos en la Nube para Sistemas Distribuidos Multiplataforma

Matías Teragni, Gonzalo Zabala, Ricardo Morán, Sebastián Blanco.
 Centro de Altos Estudios en Tecnología Informática
 Facultad de Tecnología Informática
 Universidad Abierta Interamericana
 Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina
 (+54 11) 4301-5323; 4301-5240; 4301-5248
 { Matias.Teragni, Gonzalo.Zabala, Ricardo.Moran, Sebastian.Blanco}@uai.edu.ar

Resumen

El propósito de este proyecto es diseñar e implementar un conjunto de librerías que permitan compartir su estado de ejecución a distintos programas que se encuentran corriendo en distintas máquinas. Esta funcionalidad será extendida a diversos lenguajes ampliamente utilizados, permitiendo la interoperabilidad entre distintos tipos de aplicaciones (web, desktop, mobile) y diversas plataformas.

Palabras clave: Programación

Distribuida, Internet of things, Cloud robotics, multiplataforma

Contexto

El presente proyecto será radicado en el Centro de Altos Estudios en Tecnología Informática, dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación “Nuevas Tecnologías para Internet”. El financiamiento está compartido entre el CONICET y la misma Universidad por partes iguales.

Introducción

Se entiende a un sistema distribuido como un conjunto de computadoras

independientes, cada una con su propia memoria y capacidad de procesamiento, que se encuentran interconectadas por medio de una red, y sobre las cuales opera un conjunto de software que colabora para lograr un comportamiento complejo como un todo.

Los sistemas distribuidos pueden proveer, dada su naturaleza, un conjunto de beneficios, desde resistencia a fallos mediante replicación, hasta la capacidad de ejecutar procesos sumamente complejos en computadoras económicas mediante la partición del problema en partes más simples.

Debido a estos beneficios, y que algunos sistemas requieren por su dominio del problema estar distribuidos, se volvieron populares un conjunto de tecnologías con el objetivo de facilitar la construcción de este tipo de sistemas interconectados, admitiendo abiertamente la llamada a procedimientos remotos (corba, web services, http, etc). A los sistemas que admiten esta diferencia se los considera fuertemente acoplados. Esto implica que a la hora de crear, o programar estos sistemas, no solo se requiere trabajo adicional, sino que también aumenta la complejidad del sistema, dificultando los procesos de testing, debugging, y mantenimiento del software.

En paralelo a estos desarrollos se idearon formas de interconectar sistemas de forma transparente, es decir, que comparten variables, memoria, o estado, sin la necesidad de que el programador cambie como se utilizan las mismas en su ámbito particular. Estos métodos de distribución se conocen como bajamente acoplados, y presentan como principal beneficio que no genera trabajo adicional para los desarrolladores.

Estos últimos no fueron exitosos por dos grandes motivos. En principio la velocidad de las redes de telecomunicación era insuficiente para reflejar la velocidad y extensión de los cambios en la memoria de un programa. Y para empeorar el problema, la mayoría de las implementaciones se basan en la definición de lenguajes de programación propietarios, que definen reglas de interacción arbitrarias. Es decir, se requiere desarrollar nuevas versiones del sistema, contratando a programadores expertos en estos lenguajes, para lograr que un software ya desarrollado se pueda utilizar en un ambiente distribuido.

El presente trabajo tiene como objetivo desarrollar un conjunto de librerías que permitan mantener las ventajas de las arquitecturas distribuidas sin tener que utilizar semánticas especiales y complejas para su especificación.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto en particular tiene dos fuentes de complejidad. En primer lugar el objetivo de proveer la capacidad de distribuir un ambiente de objetos, y en segundo el que esto se pueda realizar en múltiples plataformas, lenguajes y de forma transparente para el programador.

Dado que las posibles ramificaciones de esta tecnología son múltiples, plantea un proceso iterativo e incremental. Estando cada iteración compuesta por las siguientes actividades:

- Identificación de problemas no resueltos: Se definirán un conjunto de escenarios o funcionalidades que se desee explorar, o errores encontrados en iteraciones anteriores, generando un listado que indicará los objetivos que se desean cumplir con la iteración
- Planteo teórico sobre una solución: Se evaluará para el listado de problemas no resueltos cuáles se abordarán en esta iteración, y se diagramará una solución teórica para estas dificultades.
- Definición de casos de prueba: Se definirán formas de probar las funcionalidades a ser abordadas durante esta iteración.
- Implementación de un prototipo incremental: Se implementará el código que lleve a cabo las soluciones teóricas planteadas, realizando pruebas unitarias sobre el mismo y de haberse realizado iteraciones previas se incorporará a aquello ya desarrollado.
- Evaluación de avances alcanzados: Se tomarán los datos recaudados de las pruebas, para concluir hasta qué punto se lograron resolver los objetivos planteados para esta iteración. Esto permitirá definir cuál es el estado actual de la solución, y planteará la base para la próxima iteración.

Los hitos a cumplir en pos de realizar dicha solución son:

- Crear un sistema que permita compartir estado entre aplicaciones en un solo lenguaje de programación

- Extender la funcionalidad incorporando otros lenguajes y plataformas de uso masivo
- Demostrar el uso y las ventajas del proyecto mediante la implementación de casos de ejemplo basados en sistemas reales
- Incorporar la sincronización de comportamiento a lo ya existente.
- Demostrar las ventajas a la hora de construir sistemas complejos mediante la colaboración provista por esta herramienta mediante el uso de ejemplos basados en sistemas reales

Resultados y objetivos

Diseñar e implementar un esquema de persistencia cloud que permita mantener el estado de una aplicación orientada a objetos

Construir las librerías y APIs que permitan sincronizar transparentemente el estado en aplicaciones en javascript

Extender la funcionalidad a lenguajes ampliamente utilizados (C#, Java), permitiendo la interoperabilidad entre distintos tipos de aplicaciones (web, desktop) y diferentes plataformas.

Formación de Recursos Humanos

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, dos doctorandos, y un ayudante alumno de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

Referencias

- Coulouris, George; Jean Dollimore; Tim Kindberg; Gordon Blair (2011). Distributed Systems: Concepts and Design (5th Edition). Boston: Addison-Wesley.
- LeLann, G. (1977). "Distributed systems - toward a formal approach,". Information Processing. 77: 155·160. – via Elsevier.
- Andrews, Gregory R. (2000), Foundations of Multithreaded, Parallel, and Distributed Programming, Addison-Wesley, ISBN 0-201-35752-6.
- Christian Cachin; Rachid Guerraoui; Luís Rodrigues (2011), Introduction to Reliable and Secure Distributed Programming (2. ed.), Springer, ISBN 978-3-642-15259-7

Implementación de Plataforma HPC Dinámica para la Resolución de Problemas de Alta Complejidad Computacional

Miguel Méndez-Garabetti^{1,2}, Javier Rosenstein¹, María A. Murazzo³, Ailin Carribero¹, Pedro Orellana¹, Nelson R. Rodríguez³, Miguel Guevara³ y Pablo Gómez³

¹Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat. Belgrano 721, 5501 Godoy Cruz, Mendoza, Mendoza, Argentina. {mendez-garabettimiguel}, {rosensteinjavier}, {orellanapedro}@uch.edu.ar, ailin@carribero.com.ar

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

³Departamento de Informática – FCEFy N, UNSJ - CUIM (Complejo Islas Malvinas) Rivadavia, San Juan, Argentina marite@unsj-cuim.edu.ar, nelson@iinfo.unsj.edu.ar, migueljoseguevaratencio@gmail.com, pablo.gomez.allende@gmail.com

RESUMEN

En la actualidad existe gran cantidad de problemas que son resueltos de forma aceptable mediante la utilización de algoritmos secuenciales. Sin embargo, a medida que la complejidad del problema incrementa, las implementaciones secuenciales suelen volverse ineficientes hasta incluso obsoletas. Siendo necesario recurrir a la utilización de herramientas que permitan ofrecer resultados más eficientes. En este tipo de escenarios, la computación paralela se ha convertido en la forma tradicional de resolver gran variedad de problemas de alta complejidad computacional. Si bien existen diferentes niveles de paralelismo, la utilización de esquemas basados en múltiples computadoras es la más adecuada para tratar problemas de este tipo. Sin embargo, el acceso a una infraestructura de alto rendimiento no siempre suele estar al alcance de todo tipo de instituciones. Debido a esto, en este trabajo se diseña una implementación de reutilización de equipamiento informático existente, para conformar un clúster HPC. La implementación de esta plataforma será la herramienta fundamental para poder trabajar de forma activa

en el área de HPC, posibilitando al desarrollo y evaluación de nuevas técnicas de optimización.

Palabras clave: computo paralelo, HPC, cluster, optimización, metaheurísticas

CONTEXTO

El presente proyecto se desarrolla en el Instituto de Investigaciones de la Facultad de Informática y Diseño de la Universidad Champagnat (Godoy Cruz, Mendoza), en el marco de la Licenciatura en Sistemas de Información; en cooperación con el Departamento de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan (San Juan). Este trabajo es parte del proyecto de investigación que dio inicio en agosto de 2016 denominado “Diseño e implementación de una plataforma de computación paralela, dinámica y heterogénea, para el desarrollo de estrategias avanzadas de optimización aplicadas a la resolución de problemas de alta complejidad (FID-002/16)”.

1. INTRODUCCIÓN

El objetivo principal de la presente línea de investigación consiste en el desarrollo y mejora de técnicas de optimización para ser aplicadas en la resolución de problemas de alta complejidad computacional, tales como análisis, procesamiento y visualización de grandes volúmenes de datos. Trabajar con datos masivos, implica un gran desafío debido a la necesidad de explorar un universo de nuevas tecnologías, las cuales no sólo hacen posible la obtención y procesamiento de los datos sino también realizan su gestión en un tiempo razonable [1]. El tratamiento de datos a gran escala posee varias etapas, las cuales deben ser resueltas de manera eficiente y eficaz a fin de obtener información útil que de solución a los problemas. Cada una de estas etapas constituyen en si problemas computacional- mente costosos, para los cuales considerar el uso de nuevas técnicas y arquitecturas contribuirá a mejorar su rendimiento. Es por ello que la búsqueda y selección de técnicas de computación de altas prestaciones (HPC, [2]) en cada una de las etapas o procesos involucrados, permitirá resolver con eficiencia cada uno de sus objetivos.

HPC es la evolución de los sistemas de cómputo convencional, los cuales permiten realizar operaciones de cómputo intensivo y mejorar la velocidad de procesamiento. Esto involucra diferentes tecnologías, tal como los sistemas distribuidos y los sistemas paralelos; incluyendo clúster de computadoras, cloud computing, tarjetas gráficas y computadoras masivamente paralelas. Todos estos entornos son ideales para resolver aplicaciones científicas, computacionalmente costosas con manejo de grandes cantidades de datos, a fin de lograr resultados en menor tiempo.

Según [3], un clúster HPC es una colección de estaciones de trabajo autónomas o PCs, interconectadas entre sí por una red de alta velocidad que trabajan conjuntamente como

un solo recurso informático integrado. Partiendo de esta definición podemos entender que un clúster posee tanto, componentes hardware: nodos o PCs y elementos de comunicación; y componentes software: como los sistemas operativos y las aplicaciones que conforman el entorno de programación [4],[5],[6]. Existen diferentes tipos de implementaciones HPC, aunque dentro de las soluciones que priman el aspecto económico, aquella que ofrece el menor costo es sin duda la que consiste en “reutilizar” el equipamiento existente. Es importante remarcar que con reutilización nos referimos a dar una segunda utilidad o función a una computadora que ya tiene un rol asignado. Por ejemplo, un conjunto de computadoras de un laboratorio de una universidad, el cual es utilizado para realizar operaciones de cálculo científico en la franja horaria que el mismo se encuentra ocioso. Si bien este tipo de soluciones no es el escenario ideal y habitual, presente en los grandes centros de HPC, suele ser una solución muy utilizada por instituciones que no pueden justificar el costo de adquirir un clúster HPC dedicado, ya sea por falta de presupuesto o por un bajo volumen de trabajo (e.g. generalmente esto ocurre en aquellos grupos de investigación que se están iniciando en el uso o aplicación de HPC). Si bien la reutilización de equipamiento informático es la solución más viable económicamente, trae consigo diversos retos y complicaciones, tanto desde el punto de vista logístico, administrativo, político, como así también técnico.

Construir una arquitectura de bajo costo y reutilizable, que sea eficiente en tiempo y uso de recursos, no es tarea sencilla. La complejidad de una arquitectura distribuida está dada debido a la diversidad de componentes que la misma incluye como son: servidores, almacenamiento, arquitecturas de software en capas, variedad de middlewares y redes, que deben trabajar conjuntamente de la forma más óptima

posible y que además deben ser configurados adecuadamente para ser utilizados bajo demanda.

Una vez que la plataforma de cálculo se encuentre implementada, será utilizada para diseñar y evaluar el rendimiento de nuevas técnicas de optimización, las cuales operen exclusivamente en ambientes paralelos. Como técnicas de optimización nos referimos al uso de metaheurísticas paralelas y sus diferentes esquemas de colaboración.

Según [7], las metaheurísticas son estrategias inteligentes, de propósito general, que tienen como objetivo diseñar y mejorar procedimientos heurísticos para resolver problemas de alta complejidad. Si bien las metaheurísticas implementadas de manera tradicional pueden resolver de forma eficiente gran cantidad de problemas, en ciertos casos es necesario recurrir a estrategias avanzadas. Como el caso de la hibridación de metaheurísticas utilizando esquemas paralelos, permitiendo de esta forma sumar las bondades y minimizar sus deficiencias, en pos de obtener resultados de mayor calidad.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto está compuesto por dos etapas o fases de trabajo. Una de ellas corresponde al diseño e implementación técnica de la plataforma HPC, la cual será la base necesaria para poder realizar las actividades propias de la segunda etapa. Esta última corresponde a la línea de investigación (i.e., línea I+D) propiamente dicha. Una breve descripción de cada una de las etapas se describe a continuación:

1. La primer parte consiste en la implementación de un clúster que permita a cada equipamiento de uso administrativo y/o educativo de la institución, formar parte del conjunto de nodos de cálculo del

clúster. Para ello se deberá determinar una adecuada configuración que permita operar con equipamiento completamente heterogéneo y de forma dinámica, sin afectar el fin principal que cada terminal tiene definido.

2. La segunda etapa consiste en el desarrollo y mejora de nuevas técnicas de optimización existentes. Los desarrollos consistirán en el estudio de diferentes metaheurísticas y sus diferentes combinaciones, con el objetivo de incrementar su robustez y capacidad de rendimiento. Dichas técnicas deberán ser evaluadas mediante su aplicación en diferentes problemas de alto costo computacional, que permitan evidenciar las capacidades de dichas técnicas.

3. RESULTADOS ESPERADOS

Como resultados esperados, los mismos pueden dividirse en dos grandes grupos:

1. *Plataforma HPC*, se espera obtener como producto final un esquema de configuración de equipos de red que permita utilizar cualquier equipamiento informático de la Universidad Champagant para conformar uno o más clusters HPC utilizando principalmente: MPI [8], OpenMP [9], C y C++ [10]. Inicialmente se trabajará solamente con el equipamiento informático de dos laboratorios de informática, los cuales poseen las siguientes características:
 - a. **Laboratorio 1:** 7 PCs con procesadores AMD Athlon 64x2 4600 con 2Gb RAM.
 - b. **Laboratorio 2:** 18 PCs con procesadores i7 4770 con 8GB de memoria RAM y ATI Radeon HD6450 1Gb.

También está previsto trabajar con ecosistemas distribuidos sobre cluster, tales como Hadoop [11] y Spark [12], los cuales usan el paradigma MapReduce [13] para el procesamiento de datos masivos. Además, se prevé el estudio de bases de datos avanzadas (NoSQL [14] y NewSQL [15]) para el almacenamiento de los grandes volúmenes de datos.

2. *Técnicas de optimización*, en líneas generales se espera identificar las estrategias de colaboración y las distintas hibridaciones que ofrecen más eficiencia tras ser aplicados a problemas de alto costo computacional [16].

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesina de grado, por parte del estudiante de la Licenciatura en Sistemas de Información de la Universidad Champagnat, Pedro Orellana. Dicha tesina se centra en el análisis e implementación de alternativas para la plataforma de HPC. Por parte de la misma institución, también se cuenta con la reciente participación de la estudiante de segundo año de la carrera: Ailin Carribero, quien pertenece al programa de Iniciación la Investigación Científica de la UCH. Dicho programa invita a participar en tareas de investigación a estudiantes desde los primeros años de la carrera.

Por parte del Departamento de Informática de la FCEFyN de la UNSJ se cuenta con la participación de dos estudiantes de grado, en instancia de tesis, ellos son: Miguel Guevara y Pablo Gomez, ambos de la carrera Licenciatura en Sistemas de Información.

Finalmente, una vez que la plataforma se encuentre implementada, la misma será utilizada como recurso para el dictado de talleres de computación paralela tanto para

estudiantes de la universidad, como así también para aquellos interesados externos a la UCH.

5. BIBLIOGRAFÍA

1. Marz, N., & Warren, J. (2015). *Big Data: Principles and best practices of scalable realtime data systems*. Manning Publications Co..
2. Hager, G., & Wellein, G. (2010). *Introduction to high performance computing for scientists and engineers*. CRC Press.
3. Jin, H., Buyya, R., & Baker, M. (n.d.). *Cluster Computing Tools, Applications, and Australian Initiatives for Low Cost Supercomputing*.
4. Gropp, Lusk and Sterling, (2003). "Beowulf Cluster Computing with Linux". The MIT Press, second edition,
5. Brown. "What's a Beowulf?". *Engineering a Beowulf-style Compute Cluster*. Physics Department. Duke University. http://www.phy.duke.edu/~rgb/brahma//beowulf_book/node9.html, (2003). Accedido en marzo de 2017.
6. Ferreira, Kettmann, Thomasch, Silcocks, Chen, Daunois, Ihamo, Harada, Hill, Bernocchi and Ford. (2001). "Linux HPC Cluster Installation". IBM Redbooks, first edition (ISBN: 9780738422787).
7. Glover F. (1986). Future paths for integer programming and artificial intelligence. *Computers & Operations Research*, 13(5):533–549.
8. Message P Forum. (1994). *Mpi: a Message-Passing Interface Standard*. Technical Report. University of Tennessee, Knoxville, TN, USA.
9. Dagum L. and Menon R. (1998). OpenMP: An Industry-Standard API for Shared-Memory Programming. *IEEE Comput. Sci. Eng.* 5, 46-55.

DOI=<http://dx.doi.org/10.1109/99.660313>

10. Schildt H. (2000). *C/C++ Programmer's Reference* (2nd ed.). McGraw-Hill, Inc., New York, NY, USA.
11. White, T. (2012). *Hadoop: The definitive guide*. O'Reilly Media, Inc.
12. Spark, A. (2016). *Apache spark: A fast and general engine for large-scale data processing*. <http://spark.apache.org>
13. Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.
14. Hecht, R., & Jablonski, S. (2011). NoSQL evaluation: A use case oriented survey. In *Cloud and Service Computing (CSC), 2011 International Conference on* (pp. 336-341). IEEE.
15. Stonebraker, M. (2012). Newsql: An alternative to nosql and old sql for new oltp apps. *Communications of the ACM*. Retrieved, 07-06.
16. Engineering and Physical Sciences Research Council, "International Review of Research Using HPC in the UK", Engineering and Physical Sciences Research Council, (ISBN 1-904425-54-2), (2005).

Método de Reducción de Incertidumbre Basado en HPC y Metaheurísticas Híbridas Aplicado a la Predicción de Incendios Forestales

Méndez-Garabetti Miguel^{1,2}, Bianchini Germán¹, Caymes-Scutari Paola^{1,2}
y Tardivo María Laura^{1,2,3}

¹Laboratorio de Investigación en Cómputo Paralelo/Distribuido
Departamento de Ingeniería en Sistemas de Información
Facultad Regional Mendoza, Universidad Tecnológica Nacional
Mendoza, Argentina

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Departamento de Computación, Universidad Nacional de Río Cuarto
Río Cuarto, Córdoba, Argentina
mmendez@mendoza-conicet.gob.ar, gbianchini@frm.utn.edu.ar,
pcaymesscutari@frm.utn.edu.ar, lauratardivo@dc.exa.unrc.edu.ar

RESUMEN

La predicción del comportamiento de incendios forestales no es una tarea sencilla ya que dicho proceso se ve afectado por la falta de precisión o incertidumbre en los parámetros de entrada. En base a esto, resulta importante desarrollar métodos que permitan tratar la incertidumbre posibilitando la obtención de predicciones más precisas y confiables. En el presente trabajo se presenta un proyecto de desarrollo de un método de reducción de incertidumbre denominado Sistema Estadístico Evolutivo Híbrido con Modelo de Islas (HESS-IM). HESS-IM es un método que combina las bondades de tres metaheurísticas poblacionales evolutivas: Algoritmos Evolutivos, Evolución Diferencial y Optimización por Cúmulo de Partículas bajo un esquema de combinación colaborativa basado en migración mediante modelo de islas y HPC. Es importante mencionar que si bien el método se encuentra aún en desarrollo, existe una versión preliminar la cual ha sido implementada con Algoritmos Evolutivos y Evolución Diferencial. Ésta,

ha obtenido resultados alentadores, ya que ha superado resultados obtenidos por metodologías previamente desarrolladas, sin considerar que aún resta realizar estudios más exhaustivos; los cuales podrían permitir obtener resultados de mayor calidad.

Palabras clave: predicción, incendios forestales, reducción de incertidumbre, metaheurísticas evolutivas poblacionales, HPC, metaheurísticas híbridas.

CONTEXTO

El presente proyecto de I+D se desarrolla en el Laboratorio de Investigación en Cómputo Paralelo/Distribuido (LICPaD) de la UTN-FRM, puntualmente pertenece a la línea de investigación denominada: “Desarrollo de Aplicaciones Paralelo/Distribuidas orientadas a la Predicción de Incendios Forestales”.

Para el desarrollo del mismo se cuenta con el financiamiento de la UTN bajo el proyecto código EIUTIME0003939TC. Ade-

más, se cuenta con el apoyo del CONICET a través de dos becas internas doctorales.

1. INTRODUCCIÓN

Los incendios forestales afectan anualmente poco más de 350 millones de hectáreas, causando todo tipo de pérdidas y daños: flora, fauna, propiedad, ecosistemas e incluso vidas humanas [1]. Como hecho reciente podemos citar al gran incendio forestal ocurrido en Chile, donde cerca de 140 focos afectaron 380.000 hectáreas en seis regiones de dicho país, causando 11 víctimas fatales y destruyendo más de 1000 viviendas [2].

Debido al gran potencial destructivo que poseen los incendios, desde años se ha visto necesario trabajar en el desarrollo de diferentes herramientas que permitan reducir sus efectos negativos, tal es el caso de los sistemas de cómputo que permiten predecir su comportamiento.

La predicción del comportamiento de los incendios forestales es considerada una herramienta muy útil ya que permite determinar las zonas que serán afectadas por el fuego, permitiendo tomar medidas de evacuación y de gestión de recursos de forma eficiente. La predicción de cualquier fenómeno natural es una tarea compleja, tanto por el diseño de los modelos involucrados como así también por la presencia de incertidumbre en la información con la que dicho modelo es alimentado. En base a esto, la predicción de un incendio forestal no escapa a esta situación, ya que su comportamiento está determinado por una serie de parámetros que usualmente no pueden ser medidos en tiempo real. Esta falta de precisión en los parámetros de entrada impacta directamente en la salida del modelo, impidiendo a las técnicas de predicción clásicas [3][4][5] obtener predicciones aceptables. En este contexto, se debe recurrir a métodos de reducción de incertidumbre que permitan generar predicciones aceptables independientemente de la imprecisión en los parámetros de entrada.

El grupo de investigación LICPaD viene trabajando en esta problemática desde hace años; producto de este trabajo se han desarrollado diferentes métodos de reducción de incertidumbre [6][7][8][9].

El presente proyecto parte de dichos trabajos previos, donde se ha evidenciado que la utilización de metaheurísticas evolutivas paralelas [10] junto con análisis estadístico [11] y HPC [12] pueden formar métodos robustos mejorando considerablemente la calidad de predicción.

El método propuesto utiliza como técnica de optimización, una metaheurística híbrida basada en Algoritmos Evolutivos (EA, Evolutionary Algorithms) [13], Evolución Diferencial (DE, Differential Evolution) [14] y Optimización por Cúmulo de Partículas (PSO, Particle Swarm Optimization) [15], bajo un esquema de integración colaborativa. El nuevo método se ha denominado Sistema Estadístico Evolutivo Híbrido con Modelo de Islas (HESS-IM, Hybrid-Evolutionary Statistical System with Island Model).

Hasta el momento, la primera aproximación de HESS-IM ha sido implementada haciendo uso sólo de EA+DE, si bien los resultados superan a metodologías previas, en ciertos pasos de predicción la diferencia no es muy significativa, (una de las hipótesis es que al sumar las bondades de PSO la calidad de predicción del método será incrementada).

Es importante mencionar que HESS-IM opera con un esquema de paralelización basado en islas con doble jerarquía master-worker, lo que permite implementar las diferentes metaheurísticas con múltiples poblaciones y migración.

En la actualidad se está realizando la implementación de HESS-IM con EA+DE+PSO, lo que implica someter al método a un estudio de rendimiento y comparación en función de las diferentes configuraciones de parámetros que cada metaheurística tiene involucrados.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto centra su investigación en el análisis, diseño y desarrollo de métodos de reducción de incertidumbre para ser aplicados principalmente a modelos ambientales, como es el caso de los incendios forestales. Respecto a la investigación actual, por un lado, gran parte de los esfuerzos se encuentran concentrados en el análisis de las diferentes estrategias de colaboración entre metaheurísticas poblacionales evolutivas. Estas estrategias son las que permiten incrementar la robustez de búsqueda del método, impactando directamente en la calidad de predicción del método en cuestión. Por el otro, también se está trabajando en la optimización en términos de rendimiento del método, para ello ya se ha implementado ciertas funciones del método haciendo uso de GPUs, lo que podrá permitir obtener ahorros considerables de tiempo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Tal como se ha mencionado anteriormente, HESS-IM ha sido implementado en una primera instancia con una metaheurística híbrida basada en EA+DE. Dicha implementación ha permitido incrementar el nivel de calidad de predicción en contraste con ESS-IM [8][9], una metodología previamente desarrollada. Los resultados obtenidos hasta la fecha han sido publicados en [16][17].

Como resultados a futuro, esperamos que una vez implementado de manera completa HESS-IM (i.e., con las tres metaheurísticas: EA, DE y PSO), y tras realizar las experimentaciones correspondientes, podamos evaluar el grado de mejora que esta combinación ofrece en relación a las implementaciones de menor grado de complejidad. También esperamos trabajar para optimizar el método en términos de utilización de recursos.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto cuenta con la dirección del Dr. Germán Bianchini (UTN-FRM) y la codirección de la Dra. Paola Caymes-Scutari (UTN-FRM, CONICET). En lo que hace a estudiantes de doctorado, esta línea de investigación cuenta con dos tesis de posgrado en curso, una perteneciente al Ing. Miguel Méndez Garabetti, cuyo plan de tesis doctoral versa específicamente dentro de este proyecto y de esta línea de investigación, y de la Lic. María Laura Tardivo, cuya temática se vincula estrechamente, dado que se especializa en distintos métodos y posibilidades referentes a DE. Ambos cursan el doctorado en Ciencias de la Computación de la Universidad Nacional de San Luis, y son becarios internos doctorales de CONICET.

5. BIBLIOGRAFÍA

- [1] FAO - Noticias: El problema de los incendios forestales. (n.d.). Recuperado en Marzo 19 de 2017, desde <http://www.fao.org/news/story/es/item/29145/icode/>
- [2] Chile: Pérdidas forestales por incendios ascenderían a US\$ 350 millones | Imagen | Multimedia | Gestion.pe. (n.d.). Recuperado en Marzo 19 de 2017, desde <http://gestion.pe/multimedia/imagen/2181045/200422>
- [3] Wallace, G. (1993). A Numerical Fire Simulation-Model. *International Journal of Wildland Fire*, 3(2), 111. <https://doi.org/10.1071/WF9930111>
- [4] Finney, M. A., & Finney, M. A. (n.d.). FARSITE: Fire Area Simulator-model development and evaluation.
- [5] Lopes, A. M. G., Cruz, M. G., Viegas, D. X., & Lopes, A. M. G. (n.d.). FireStation -An integrated software system for the numerical simulation of fire spread on complex topography. Retrieved from <http://adai.dem.uc.pt>
- [6] Bianchini, G., Denham, M., Cortés, A., Margalef, T., & Luque, E. (2010). Wildland fire growth prediction meth-

- od based on Multiple Overlapping Solution. *Journal of Computational Science*, 1(4), 229–237. <https://doi.org/10.1016/j.jocs.2010.07.005>
- [7] Bianchini, G., Caymes-Scutari, P., & Méndez-Garabetti, M. (2015). Evolutionary-Statistical System: A parallel method for improving forest fire spread prediction. *Journal of Computational Science*, 6(1), 58–66. <https://doi.org/10.1016/j.jocs.2014.12.001>
- [8] Méndez-Garabetti, M., Bianchini, G., Tardivo, M. L., & Caymes-Scutari, P. (2015). Comparative analysis of performance and quality of prediction between ESS and ESS-IM. *Electronic Notes in Theoretical Computer Science*, 314, 45–60. <https://doi.org/10.1016/j.entcs.2015.05.004>
- [9] Méndez-Garabetti, M., Bianchini, G., Caymes-Scutari, P., & Tardivo, M. L. (2016). Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic. *Fire Safety Journal*, 82, 49–62. <https://doi.org/10.1016/j.firesaf.2016.03.002>
- [10] Talbi E., *Metaheuristics: From Design to Implementation*, University of Lille – CNRS
- [11] Montgomery D.C, Runger G.C.: *Applied Statistics and Probability for Engineers*. 6th Edition. Limusa Wiley & Sons, New Jersey (2014)
- [12] Buyya, R. (1999). *High Performance Cluster Computing: Architectures and Systems*. Prentice Hall, PTR, NJ, USA.
- [13] Alba E., Tomassini M.: Parallelism and evolutionary algorithms, in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 5, 443–462 (2002)
- [14] Storn, R., & Price, K. (1995). Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces. Berkeley, CA.
- [15] Kennedy J., Eberhart R. *Particle Swarm Optimization Proceedings of IEEE International Conference on Neural Networks (1995)*, pp. 1942–1948
- [16] Méndez-Garabetti, M., Bianchini, G., Caymes-Scutari, P., & Tardivo, M. L. (2016). Método híbrido de reducción de incertidumbre aplicado a la predicción del comportamiento de incendios forestales. *Libro de Actas XXII CA-CIC 2016, XXII Congreso Argentino de Ciencias de La Computación*, 1, 159–168.
- [17] Méndez-Garabetti, M. A., Bianchini, G., Caymes-Scutari, P. y Tardivo, M. L. (2016). Método Híbrido Paralelo de Reducción de Incertidumbre Aplicado a la Predicción del Comportamiento de Incendios Forestales. *XXII Congreso Sobre Métodos Numéricos Y Sus Aplicaciones*, XXXIV, 2857–2869.

Paralelismo Híbrido Aplicado a Soluciones de Problemas de Datos Masivos

Maria A. Murazzo*, Maria Fabiana Piccoli[#], Nelson R. Rodriguez*, Diego Medel*, Jorge N. Mercado[&], Federico Sanchez**, Ana Laura Molina***, Martin Tello**

*Departamento de Informática – FCEFY N, UNSJ.

[#]Departamento de Informática – FCFMy N, UNSL.

[&]Departamento de Matemática – FI, UNSJ.

**Alumno Avanzado de la Carrera Licenciatura en Ciencias de la Computación.

***Alumno Avanzado de la Carrera Licenciatura en Sistemas de Información.

marite@unsj-cuim.edu.ar, mpiccoli@unsl.edu.ar, nelson@iinfo.unsj.edu.ar, mdiego88@gmail.com,
jorgenmp@gmail.com, fedegsancheza@gmail.com, lauramolina@outlook.com,
martinl.tello@gmail.com

Resumen

Con el uso masivo de Internet, se está en presencia de un fenómeno donde la aceleración tanto del crecimiento del volumen de datos capturados y almacenados, como la creciente variación en los tipos de datos, hace que las técnicas tradicionales para el procesamiento, análisis y obtención de información útil deban ser redefinidas para formular nuevas metodologías.

Este trabajo aborda las líneas de investigación relacionadas con el objetivo de definir técnicas o metodologías computacionales para mejorar tanto los tiempos de respuesta y la escalabilidad de los sistemas desarrollados, como así también solucionar los inconvenientes presentes en las soluciones existentes. Esto posibilita la transferencia de los logros y experiencias adquiridos, permitiendo, además, verificar la aplicabilidad de los métodos o técnicas desarrollados en problemas reales con uso de datos masivos.

Palabras clave: Datos Masivos, Computación de Alto Desempeño, Arquitecturas Multiprocesadores, Sistemas Distribuidas.

Contexto

El presente trabajo se encuadra dentro del área de I/D “Procesamiento Distribuido y Paralelo” y en particular dentro del proyecto de investigación "Evaluación de arquitecturas distribuidas de commodity basadas en software libre", código E1038, el cual ha sido aprobado en la última convocatoria de CICITCA (duración dos años y unidad ejecutora al Departamento de Informática de la FCEFYN de la UNSJ).

Introducción

Trabajar *con grandes volúmenes de datos*, implica un gran desafío debido a la necesidad de explorar un universo de nuevas tecnologías, las cuales no sólo hacen posible la obtención y procesamiento de los datos sino también realizan su gestión en un tiempo razonable, lo que permite contar con una fuente inagotable de problemas para aplicar de técnicas computacionales de alto desempeño.

El crecimiento de la cantidad de datos es algo cotidiano y obedece a la proliferación de diferentes fuentes de generación de información, como son la web, aplicaciones de imagen y vídeo, redes sociales, dispositivos móviles, sensores, Internet de las cosas, etc., todas ellas capaces de generar según IBM [1], más de 2.5 quintillones de bytes diarios. Este

aumento en la cantidad de datos demanda nuevas estrategias para su almacenamiento, procesamiento y análisis, conllevando a un cambio de paradigma en las arquitecturas de cómputo, los algoritmos y los mecanismos de procesamiento.

Ejemplos cotidianos de datos masivos son el número de imágenes subidas diariamente a las redes sociales (300 millones en Facebook, 45 millones en Instagram), los videos vistos por día en YouTube (2 billones), la cantidad de mensajes de texto enviados, la cantidad mensual de búsquedas en Twitter, el tráfico mundial en Internet, entre otros [2]. Esto no sólo es aplicable a las actividades desarrolladas diariamente en Internet, sino también en aquellas relacionadas a fenómenos naturales como el clima o datos sismográficos, entornos referidos a la salud, bioinformática, seguridad o al ámbito empresarial.

Además de la gestión del volumen de datos, gran parte de la información requerida para la toma de decisiones y la resolución de problemas de índole general proviene de información no estructurada, almacenada o accedida no necesariamente en estructuras clásicas de almacenamiento como matrices, registros de bases de datos, etc.

Restringirse al uso de información estructurada lleva, muchas veces, a representar una visión parcial del problema y dejar fuera de consideración información de gran importancia para la resolución efectiva del mismo.

Frente a esta problemática se ha popularizado el término Big Data [3], el cual es usado para describir grandes conjuntos de datos, que exhiben las propiedades de variedad, volumen, velocidad, variabilidad, valor y complejidad. Hablar de Big Data es hacer referencia a datos multidimensionales, estructurados o no estructurados.

Big data es un área de investigación focalizada en recolectar, examinar y procesar grandes conjuntos de datos con el objeto de descubrir patrones, correlaciones y extraer información de ellos [4]. Por lo general, esta tarea se implementa mediante el uso de

diferentes técnicas, entre ellas se encuentra machine learning supervisado y no supervisado[5], técnicas computacionalmente muy costosas, ya sea en la fase de aprendizaje o en la de predicción llegando a ser intratables de manera secuencial cuando involucra grandes volúmenes de datos [6].

Estos aspectos hacen que los sistemas de cómputo convencionales sean muchas veces inapropiados para lograr un procesamiento adecuado, por lo que una alternativa llega a considerar técnicas de computación de alta prestaciones (HPC) con el fin de aumentar la velocidad de procesamiento [7,8].

Generalmente, HPC es referenciado como una evolución de los sistemas de cómputo convencionales, los cuales permiten realizar operaciones de cómputo intensivo y mejorar la velocidad de procesamiento. HPC involucra diferentes tecnologías tal como los sistemas distribuidos y los sistemas paralelos; incluyendo a los cluster de computadoras, cloud computing, tarjetas gráficas y computadoras masivamente paralelas. Todos estos entornos son ideales para resolver aplicaciones científicas, computacionalmente costosas con manejo de grandes cantidades de datos, a fin de lograr resultados en menor tiempo.

La constante demanda de mayores prestaciones hizo que la industria de los procesadores se encontrara en una situación límite respecto al cumplimiento de la ya conocida Ley de Moore sobre rendimiento del hardware [9]. La evolución de los sistemas de computación con multiprocesadores ha seguido dos líneas de desarrollo: las arquitecturas multi-core (multi-núcleos) y las arquitecturas many-cores (muchos-núcleos o muchos-cores). En el primer caso, los avances se centraron en el desarrollo de mejoras con el objetivo de acelerar las aplicaciones, generalmente secuenciales, con por ejemplo la incorporación de varios núcleos de procesamiento.

La industria tomó la idea de las supercomputadoras existentes e incorporó más procesadores a sus desarrollos, surgiendo así computadoras con 2, 3, 4, 8 o más

procesadores por unidad central (multi-core). En el caso de las arquitecturas many-cores, los desarrollos se centraron en optimizar el desempeño de aplicaciones paralelas. Dentro de este tipo de arquitectura se encuentran las tarjetas gráficas o GPUs [10]. La característica más relevante de este tipo de arquitectura es la capacidad de ofrecer cores simples y rápidos con acceso a una jerarquía de memoria compartida.

Sin embargo, los sistemas de memoria compartida tienen como inconveniente principal la cantidad de memoria disponible, la cual es limitada. Una alternativa a estas arquitecturas son las arquitecturas con memoria distribuida, las cuales permiten incrementar el espacio de almacenamiento (principal y secundario) aunque deben pagar el precio de la latencia de la red para llevar a cabo las comunicaciones [11]. Ejemplos de sistemas con memoria distribuida son los cluster y cloud computing [12]. Estos sistemas permiten conectar un gran número de máquinas (nodos) y utilizar la memoria perteneciente a diferentes procesadores, logrando eliminar el problema de la limitación de memoria RAM.

Una tercera alternativa son los sistemas híbrido, los cuales permiten combinar las características de ambos sistemas: Memoria compartida y Memoria Distribuida, e incrementar la capacidad y poder de cómputo de los sistemas computacionales. Esto posibilita la ejecución en paralelo de múltiples procesos y threads con distintas administraciones de memoria.

La presente propuesta tiene como objetivo desarrollar y aplicar técnicas computacionales híbridas de alto desempeño para la adquisición, tratamiento y análisis de datos masivos en ambientes mixtos de computación.

Líneas de Investigación, Desarrollo e Innovación

En función de lo explicado anteriormente, la línea de investigación pretende el desarrollo de herramientas de software adecuadas para

resolver problemas de datos masivos en ambiente paralelos híbridos. Dichas herramientas tendrán como objetivo mejorar el desempeño de cada una de las etapas involucradas en la solución de este tipo de problemas: adquisición, análisis y visualización de los resultados.

El desempeño de cada una de las técnicas y/o herramientas propuestas será evaluado tanto en las soluciones computacionales a los problemas específicos planteados como así también en problemas de datos masivos reales. Para ello se pretende realizar un análisis respecto a:

1. *Rendimiento*: En este caso no solo se consideraran las métricas de rendimiento estándares como aceleración, eficiencia y costo, sino también plantear otras, las cuales estarán relacionadas a las características de los ambientes de computación híbridos, por ejemplo referidas a la distribución y asignación de trabajo, y a los factores limitantes como el desempeño de la red subyacente, las comunicaciones, sincronizaciones y gránulo de la computación.
2. *Calidad*: En este caso, la evaluación se centrará en las propiedades de los resultados obtenidos, es su Relevancia y Precisión.

Además se analizarán otros factores como la portabilidad (tanto de ambiente de computación como de problema), escalabilidad y robustez de cada una de las técnicas desarrolladas.

Resultados y Objetivos

Como objetivo de la investigación se ha planteado el análisis, diseño e implementación de soluciones computacionales eficientes a problemas de datos masivos mediante la aplicación de modelos de programación y técnicas de Computación de Alto Desempeño en ambientes híbridos.

Los ambientes de computación híbridos están formados por arquitecturas multiprocesador (multi-core y many-core) y

arquitecturas distribuidas (clusters y cloud) como así también distintos modelos de memoria. Trabajar en ambientes híbridos, permite aplicar una estrategia de paralelización más efectiva mediante múltiples niveles de paralelismo y reducción del overhead de comunicación. Esto es importante cuando se trabaja con grandes volúmenes de datos debido a la necesidad de contar con una arquitectura escalable.

Por ello, es necesario investigar sobre:

- Arquitecturas Híbridas: Características básicas, adaptabilidad a problemas de diferente naturaleza.
- Modelos de programación estándares e híbridos existentes.
- Problemas de Datos Masivos: Características, etapas involucradas, aspectos paralelos de cada una.
- Herramientas existentes para resolver problemas de datos masivos: análisis de sus características, ventajas y desventajas, factibilidad de aplicación en ambientes computacionales híbridos.
- Problemas reales con uso de datos masivos.
- Análisis de desempeño en ambientes homogéneos y heterogéneos, parámetros de evaluación.

En particular con respecto a los datos masivos es necesario:

- Analizar las soluciones existentes para cada una de las etapas involucradas en problemas de datos masivos, evaluando sus limitaciones e inconvenientes de uso en ambientes híbridos.
- Analizar las características asociadas a resolver problemas con información estructurada y no estructurada.
- Diseñar e implementar nuevas técnicas para las etapas de adquisición, tratamiento y análisis de datos masivos a fin de proveer de una/s herramienta/s para resolver problemas reales aplicando técnicas de computación de

alto desempeño en ambientes computacionales híbridos.

- Analizar y elaborar métricas de rendimiento para evaluar el desempeño de los desarrollos en ambientes de computación híbridos.

Para realizar las investigaciones de esta línea, el equipamiento necesario estará en diferentes ubicaciones, en la UNSJ se cuenta con un cluster multi-core y en la UNSL se cuenta con una cluster de características híbridas, un cluster con varios multi-cores y many-cores. Además se cuenta con acceso a equipamiento perteneciente a diferentes unidades académicas y a centros de computación de alto desempeño como aquellos nucleados en el Sistema Nacional de Computación de Alto Desempeño dependiente del Ministerio de Ciencia, Tecnología e Innovación de la Nación.

Formación de Recursos Humanos

Los resultados esperados respecto a la formación de recursos humanos son hasta el momento el desarrollo de 1 tesis doctoral y 2 tesis de maestría. Respecto a las carreras de grado, se están ejecutando 4 tesinas de grado.

Además se prevé la divulgación de los temas investigados, tanto a través del dictado de cursos de postgrado/actualización, como de publicaciones en diferentes congresos y revistas del ámbito nacional e internacional.

Referencias

- [1] “¿Qué es Big Data?” [Online]. Available: <https://www.ibm.com/developerworks/ssa/loc-al/im/que-es-big-data/>.
- [2] “Blog Cisco Cansac » Tráfico de datos móviles crecerá casi 10 veces en los próximos cinco años, predice estudio Cisco Visual Networking Index (VNI).” [Online]. Available: <http://gblogs.cisco.com/cansac/trafico-de-datos-moviles-crecera-casi-10-veces-en-los-proximos-cinco-anos-predice-estudio-cisco-visual-networking-index-vni/>?

- [doing_wp_cron=1473174616.1210250854492187500000](http://dx.doi.org/10.1109/2187500000).
- [3] A. McAfee, E. Brynjolfsson, and H. Dreyfus, "Big Data: The Management Revolution SPOTLIGHT ON BIG DATA," 2012.
- [4] Y. Zhai, Y.-S. Ong, and I. W. Tsang, "The Emerging Dimensionality of Big Data," *IEEE Comput. Intell. Mag.*, vol. 9, no. 3, pp. 14–26, Aug. 2014.
- [5] Xindong Wu, Xingquan Zhu, Gong-Qing Wu, and Wei Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 1, pp. 97–107, Jan. 2014.
- [6] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann. 2016.
- [7] Y. You, S. L. Song, H. Fu, A. Marquez, M. M. Dehnavi, K. Barker, K. W. Cameron, A. P. Randles, and G. Yang, "MIC-SVM: Designing a Highly Efficient Support Vector Machine for Advanced Modern Multi-core and Many-Core Architectures," in *2014 IEEE 28th International Parallel and Distributed Processing Symposium*, 2014, pp. 809–818.
- [8] M. Alexander, W. Gardner, B. Wilkinson, M. J. Sottile, T. G. Mattson, C. E. Rasmussen, Y. Robert, and F. Vivien, "Introduction to High Performance Computing for Scientists and Engineers Chapman & Hall/CRC Computational Science Series."
- [9] C. A. Mack, "Fifty Years of Moore's Law," *IEEE Trans. Semicond. Manuf.*, vol. 24, no. 2, pp. 202–207, May 2011.
- [10] W. Hwu, K. Keutzer, and T. G. Mattson, "The Concurrency Challenge," *IEEE Des. Test Comput.*, vol. 25, no. 4, pp. 312–320, Jul. 2008.
- [11] N. Miranda, "Cálculo en tiempo real de identificadores robustos para objetos multimedia mediante una arquitectura paralela CPU-GPU," 2016.
- [12] K. Kaur and A. K. Rai, "A Comparative Analysis: Grid, Cluster and Cloud Computing," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 3, no. 3, pp. 2278–1021, 2014.

Performance de Arquitecturas Multiprocesador: Técnicas de Modelado y Simulación en HPC y Cloud Computing

Diego Encinas^{1,2}, Jimena Jara¹, David Rosatto¹, Román Bond¹, Andrea Bermudez¹, Martín Morales^{1,3}

¹Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

²Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática
- UNLP

³Laboratorio de Ingeniería en Sistemas de Información - FRLP - UTN

dencinas@unaj.edu.ar, elchejime@gmail.com, danielrosatto@gmail.com,
roman.alejandro.b@gmail.com, abermudez@unaj.edu.ar, martin.morales@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de la performance de las arquitecturas multiprocesador y Cloud Computing a través de modelos de simulación. Enfocando a la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios y reconfigurar el sistema físico. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o reconfigurarse.

Palabras clave: *Arquitecturas Multiprocesador. Simulación. Sistema E/S paralela. Modelado y Simulación basado en agentes (Agent-Based Modeling and Simulation, ABMS). Cloud Computing. CloudSim.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación “Modelado y Simulación en Cómputo de Altas Prestaciones (HPC). Aplicaciones en arquitecturas multiprocesador, sistemas paralelos y redes de datos” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 186/15.

En el tema existe un convenio de colaboración en actividades de Investigación y Postgrado con el Instituto de Investigación en Informática – LIDI de la Universidad Nacional de La Plata.

Por otra parte, se tiene financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de varios proyectos aprobados en la UNAJ.

Introducción

El crecimiento sostenido en la demanda del poder de cómputo remarca la necesidad de sistemas con enfoques de

paralelización masiva y cómputo de alta performance (HPC, High Performance Computing) [1]. Los clusters se han convertido en uno de los enfoques principales para lograr paralelismo a bajo costo. Una noción extendida lo constituye la utilización de gridcomputing y más recientemente cloudcomputing. Independientemente de la solución, estos sistemas constan de un gran número de componentes incluyendo nodos de procesamiento, bancos de memoria, discos, entre otros.

En cuanto a las herramientas de simulación, CloudSim es un framework desarrollado en Java que provee las APIs necesarias para que el usuario genere una simulación de un Data Center funcionando como servidor de nube capaz de simular la ejecución de CloudLets.

Sistemas de E/S Paralela

Las exigencias en los sistemas de E/S paralelos se han incrementado debido al aumento en número, velocidad y potencia de las unidades de procesamiento en los clusters. También las aplicaciones científicas que utilizan cómputo de altas prestaciones acrecientan estos requerimientos.

En muchos casos, el cuello de botella de los sistemas paralelos es la E/S de estos sistemas dada las exigencias que debe afrontar [2]. La E/S Paralela es esencial para emparejar el avance de las arquitecturas de los procesadores y el rápido crecimiento de la capacidad computacional. Aunque la arquitectura jerárquica de memoria multinivel puede evitar grandes pérdidas de prestaciones debido a los retardos de acceso a disco, la capacidad de memoria es limitada. Además, como la capacidad computacional aumentará, la disponibilidad de memoria por core decrecerá, especialmente si la escala de

los sistemas de HPC se proyecta a millones de cores o más. Varias simulaciones científicas y de ingeniería de áreas críticas de investigación, tales como la nanotecnología, astrofísica, clima y energía física están convirtiéndose en aplicaciones intensivas de datos. Para poder disminuir la brecha entre CPUs-E/S se deben identificar los factores que influyen en las prestaciones y proponer nuevas soluciones [3] [4].

En el área de tolerancia a fallas en sistemas de cómputo de alta prestaciones se puede notar la importancia de la unidad de E/S en las arquitecturas paralelas como un punto a mejorar para lograr cubrir las exigencias de las aplicaciones que utilizan HPC. Una manera de llevar a cabo este trabajo es utilizar técnicas de simulación para evaluar el efecto de los cambios de los factores con mayores influencias en las prestaciones del sistema de E/S paralelo.

Se puede disminuir la complejidad y la probabilidad de errores en la generación de sistemas híbridos desarrollando una simulación específica de éstos utilizando diferentes frameworks [5] [6] [7].

Las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con una herramienta que permita predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores de aplicaciones paralelas. Por otro lado, ABMS ha sido utilizado para modelar problemas y sistemas complejos en diversas áreas de la ciencia.

Evaluar las prestaciones del subsistema de E/S con diferentes configuraciones y la misma aplicación permite adaptar la configuración de E/S teniendo en cuenta el patrón de acceso de la aplicación Pero también puede ser una gran ventaja analizar las necesidades de las aplicaciones antes de configurar el

sistema físico. Una manera de predecir el comportamiento de las aplicaciones en el sistema de cómputo ante distintas configuraciones, es utilizando técnicas de modelado y simulación.

Se propone modelar y simular la arquitectura de E/S paralela, por medio de técnicas de simulación basadas en agentes o Sistemas Multi-Agente, (MAS-MultiAgentSystems), para evaluar el efecto de dimensionar el sistema de E/S o cambiar componentes como la red de almacenamiento, dispositivos de E/S, entre otros [8].

Simulación de arquitecturas de Cloud Computing

CloudSim [9] es un Framework de simulación generalizado y extensible que permite el modelado, la simulación y la experimentación de diferentes infraestructuras y servicios de aplicaciones de Cloud Computing. Un ejemplo de utilización es la simulación de muchos centros de datos.

Su arquitectura consiste en entidades específicas que se representan como clases Java que pueden ser heredadas o variadas para simular experimentos. Estas clases representan centros de datos, hosts físicos, máquinas virtuales, servicios a ejecutar en los centros de datos, servicios en la nube de usuarios, redes internas centro de datos y consumo de energía de los hosts físicos y elementos de los centros de datos. Además, CloudSim soporta la inserción dinámica de los elementos de simulación y proporciona aplicaciones de paso de mensajes y la topología de la red del centro de datos.

Una definición importante de CloudSim es la de entidad. Una entidad es una instancia de un componente, que es una clase o un conjunto de clases que representan un modelo CloudSim (datcenter, host). El motor de simulación

es capaz de simular el tiempo de ejecución de las apps ingresadas como Cloudlets con información básica [10].

La versatilidad de CloudSim es la principal ventaja del sistema. La integración de nuevos parámetros y conceptos de la simulación es implementada desde abstracciones preestablecidas convenientemente por los autores. Las abstracciones principales son SimEvent [11], SimEntity [12], DataCenterCharacteristics y Vm.

En cada una de estas instancias es posible establecer el detalle del comportamiento y el estado de la simulación teniendo en cuenta el alcance del modelado que se planifica.

Se propone una simulación que obtenga estadísticas de entrada/salida SAAS, PAAS o IAAS desarrollada con objetos que heredan de SimEvent y de SimEntity pero con sus métodos y estados únicos en cada simulación. En éste desarrollo se tendrán más detalles a la integración de éstos elementos en las características del objeto Vm (Virtual Machine) que es el principal actor en los sistemas Cloud. Es decir que si una Vm deriva de una SimEntity debe contener otras entidades de software como el proceso init que está formado por el SimEvent “fork” entre otros capaces de llamar a la creación de nuevos procesos.

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore,

- Clusters, Clusters de multicore. Grid. Cloud.
- Arquitectura de E/S paralela considerando el software, hardware, comunicaciones entre módulos y dispositivos de almacenamiento.
- Modelado y simulación basada en agentes.
- Simulaciones con CloudSim para el análisis de la performance del sistema de archivos en arquitecturas de Cloud Computing.

Resultados y Objetivos

Investigación experimental

- Diseño y desarrollo de modelos mediante técnicas de Modelado y simulación basada en agentes (ABMS)
- Utilización de agentes para generar la funcionalidad de los elementos físicos (procesadores, memoria, buses, drivers, entre otros) como así también de las interfaces en las arquitecturas de E/S.
- Análisis y modelado de librerías de archivos para aplicaciones que utilizan computo de altas prestaciones-HPC.
- Desarrollo de pruebas de conceptos con el modelo inicial del sistema de archivos paralelos utilizando un entorno de programación MAS.
- Utilización de CloudSim como herramienta de modelado para la simulación de la arquitectura utilizada pudiendo verificar estadísticas temporales.

- Implementar un caso de estudio con CloudSim que ofrezca el comportamiento de diferentes sistemas de archivos.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería en Informática de la UNAJ. También aportan trabajos de alumnos de las materias Redes de Computadoras 1 y 2, Programación en Tiempo Real y Tráfico en Redes.

En 2016 se obtuvo una beca para alumnos (Estimulo a las Vocaciones Científicas del CIN). Además, se han realizado 2 publicaciones nacionales y 2 internacionales.

En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay 2 investigadores realizando su Doctorado y 3 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". SecondEdition. Pearson Addison Wesley, 2003.
2. H Hennessy, J. L., Patterson, and D. A., Computer Architecture, Fourth Edition: A Quantitative Approach. San Francisco, CA, USA: Morgan KaufmannPublishers Inc., 2006.
3. J. M. May, Parallel I/O for high performance computing. San Francisco, CA, USA: Morgan KaufmannPublishers Inc., 2001.
4. V. Balaji, Earth system modelling – Volume 4. IO and Postprocessing. Springer, 2013.
5. D. Encinas, Utilización de un reloj global para el modelado de un ambiente simulado distribuido. XVIII Congreso Argentino de

- Ciencias de la Computación. 2012
6. D. Encinas, Simulación de una red CAN para dimensionar las comunicaciones de una IMU. VII Congreso Argentino de Tecnología Espacial. 2013.
 7. D. Black, SystemC: From the Ground Up. Second Edition, Springer, 2010.
 8. D. Encinas et al., Modeling I/O System in HPC: An ABMS Approach. The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015.
 9. R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose and R. Buyya “CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms” Published online 24 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.995.
 10. <http://www.cloudbus.org/cloudsim>. 2017
 11. <http://www.icsa.inf.ed.ac.uk/research/groups/hase/simjava>. 2017
 12. F. Howell, R Mc Nab. A discrete event simulation library for java. International Conference on Web-Based Modeling and Simulation. 1998.

Performance de Cloud Computing para HPC: Despliegue y Seguridad

Brian Galarza¹, Gonzalo Zaccardi¹, Maximiliano Belizán¹, David Duarte¹,
Martin Morales^{1,2}, Diego Encinas^{1,3}

¹Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

²Laboratorio de Ingeniería en Sistemas de Información - FRLP - UTN

³Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática
- UNLP

briangalarza@hotmail.com, gonzalozaccardi@gmail.com,
maximiliano.h.belizan@gmail.com, davidjulianduarte@gmail.com,
martin.morales@unaj.edu.ar, dencinas@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de la performance de las arquitecturas tipo cloud a través del despliegue de IaaS y utilización de IaaS públicos, en particular en el área de cómputo paralelo de altas prestaciones (HPC). Enfocando a la obtención de herramientas que permitan predecir la eficiencia del sistema ante posibles escenarios. Analizando los diferentes componentes del sistema que pueden influir en las prestaciones significativamente y pueden llegar a modelarse y/o configurarse.

Palabras clave: *Arquitecturas Multiprocesador. Cloud Computing. OpenStack. Seguridad en Cloud Computing.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación

“Modelado y Simulación en Cómputo de Altas Prestaciones (HPC). Aplicaciones en arquitecturas multiprocesador, sistemas paralelos y redes de datos” de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 186/15.

Por otra parte, se tiene financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través de varios proyectos aprobados en la UNAJ.

Introducción

Cloud Computing es un paradigma que está en constante crecimiento durante estos últimos años, cada vez más compañías y grupos de investigación trabajan en conjunto con el fin de explotar las oportunidades ofrecidas por el mismo [1]. Dicho paradigma ofrece muchas ventajas, tales como el bajo costo de implementación, ya que no se necesitan computadoras de última tecnología

debido a que éstas trabajan conjuntamente (Clustering) con la posibilidad de escalar horizontalmente de manera sencilla. Además, hay software Open Source disponible para los nodos en el clúster como las infraestructuras Eucalyptus, OpenNebula, CloudStack u OpenStack integradas con GNU/Linux y compatibles, por ejemplo, con Amazon WebServices.

Aunque Cloud Computing tiene muchos beneficios, el uso inadecuado puede conllevar a pérdidas de datos sensibles para el usuario y el proveedor.

Despliegue de IaaS

Las comunicaciones en Cloud Computing son una parte fundamental del paradigma que consisten en utilizar distintos nodos y lograr hacerlos funcionar conjunta y sincronizadamente.

Para lograr una comunicación entre estos nodos se propone utilizar OpenStack [2].

OpenStack es un software de código abierto que permite la implementación de, por ejemplo, una “Infraestructure as a Service” (IaaS) a través de múltiples servicios que, de manera coordinada, cumplen diferentes propósitos para lograr el correcto funcionamiento de dicha infraestructura. Algunos de los servicios ofrecidos por OpenStack son: hypervisor (Nova), autenticación (Keystone), Imágenes (Glance), Dashboard (Horizon) y block storage (Cinder). Según las necesidades se pueden requerir de ciertos servicios u otros. La Arquitectura básicamente consiste en dos tipos de nodos: “Compute Node” y “Controller Node”. Se llaman Compute Node a todos aquellos que se encargan del procesamiento de servicios específicos mientras que Controller Node es aquel que comunica a cada uno de los anteriores [3] [4] [5].

La implementación de estas infraestructuras ofrece ventajas en las cuales los clústers virtualizados trabajan en conjunto ofreciendo un buen rendimiento a bajos costos y con posibilidad de escalabilidad al poder agregar mayor cantidad de nodos para procesamiento de manera sencilla.

Como consecuencia de la dificultad de instalación y configuración de OpenStack, seguido de la constante actualización de versiones, se opta por utilizar Fuel. Ésta es una herramienta desarrollada por Mirantis en la cual se ejecuta un script que permite configurar, de manera más amigable los recursos que se desean otorgar a la infraestructura, desde la cantidad de nodos, los núcleos de procesador, la memoria RAM, entre otros [6].

Fuel trabaja con un nodo master el cual es el encargado de controlar a los nodos slaves que contendrán la infraestructura OpenStack. Es decir, desde el nodo Fuel Master se indican qué paquetes se van a instalar en cada nodo slave (Glance, Nova-Compute, Keystone, etc.) para luego en los slaves tener armados los nodos compute y controller, sin necesidad de realizar configuraciones manuales en cada uno de los mismos.

Seguridad en arquitecturas de Cloud Computing

Entre los principales riesgos y problemas de seguridad/privacidad se puede destacar los siguientes: Falta de control de datos, Ambigüedad de responsabilidad entre el usuario y el proveedor, autenticación y autorización, error de aislamiento, cumplimiento y riesgos legales, manejo de incidentes de seguridad, vulnerabilidad de la interfaz de administración, protección de aplicaciones y de datos, indisponibilidad del servicio, bloqueo del proveedor,

eliminación de los datos inseguros o incompletos, visibilidad y auditoría y seguridad [7] con la virtualización (ataques entre Máquinas Virtuales (MVs), MVs inactivas desactualizadas, falta de monitoreo y registro de MVs, Host con MVs seguras e inseguras y ataques DoS por uso no autorizado de recursos compartidos) [8].

Es importante crear un modelo de seguridad de datos, para prevenir los problemas antes mencionados, que proporcionen seguridad en los: datos para procesar (encriptación homomórfica, ejemplo: Unpadded RSA), datos de transmisión (certificados SSL, ejemplo: HTTPS), datos de almacenamiento (encriptación, ejemplo: AES), administración de claves (ejemplo: HSM), Autenticación y autorización (autenticación de dos factores, ejemplo: Token), protección de usuario (alentar el encriptado de datos sensibles antes de subir) y controles de uso (Firewalls) [9]. OpenStack cuenta con un servicio llamado Keystone que ayuda en la autenticación y autorización distribuida de clientes a través de la implementación de Tokens, esto con apoyo de la encriptación de la base de datos, comunicación sobre capas seguras y el uso de Firewall, como Iptables, puede ayudar a proveer un Cloud más confiable [10].

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- Arquitecturas multiprocesador para procesamiento paralelo: multiprocesador de memoria compartida, multiprocesador on-chip de memoria distribuida. Multicore, Clusters, Clusters de multicore. Grid. Cloud.
- Plataformas de software para implementar y administrar Clouds públicos, privados e híbridos.
- Seguridad y privacidad en Cloud Computing.

Resultados y Objetivos

Investigación experimental

- Implementación de un IaaS encargado de realizar operaciones en procesamiento paralelo aumentando la eficiencia y reduciendo los costes generados.
- Implementación de OpenStack Dashboard y de un sistema propio desarrollado para poder controlar/administrar de manera visual (web) y más básica cada uno de los servicios.
- Utilización de Fuel para administrar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y 1 controller.
- Implementación de un modelo de seguridad en un IaaS.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ. También aportan trabajos de alumnos de las materias Sistemas Operativos 1, Redes de Computadoras 2 y Programación en Tiempo Real.

En 2016 se obtuvieron dos becas para alumnos (Estimulo a las Vocaciones Científicas del CIN y de Entrenamiento de la CIC) y se aprobaron 2 Prácticas Profesionales Supervisadas. Además, se han realizado 2 publicaciones nacionales y 2 internacionales.

En esta línea de I/D existe cooperación a nivel nacional. Hay 2 investigadores realizando su Doctorado y 4 alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Kondo, D., Javadi, B., Malecot, P., Cappello, F., Anderson, D. P.: "Cost-benefit analysis of Cloud Computing versus desktop grids". In: IPDPS '09 Proceedings. IEEE International Symposium on Parallel and Distributed Processing. Washington, USA (2009).
2. OpenStack Cloud Software: Open source software for building private and public clouds. <http://www.openstack.org>. Febrero 2015.
3. Galarza, B.; Tuamá, C.; Zaccardi, G.; Encinas, D.; Morales, M. "Implementaciones de Cloud Computing y aplicaciones en el ámbito universitario". I Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2013). Ciudad de Córdoba, Argentina.
4. Zaccardi, G.; Galarza, B.; Encinas, D.; Morales, M. "Implementación de Cloud Computing utilizando OpenStack". II Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.
5. Galarza, B.; Zaccardi, G.; Encinas, D.; Morales, M. "Análisis de despliegue de una IaaS utilizando Openstack". XXI Congreso Argentino de Ciencias de la Computación (CACIC 2015). Ciudad de Junín, Argentina.
6. OpenStack Deployment Fuel. <https://www.mirantis.com/products/mirantis-openstack-software/openstack-deployment-fuel/>. Febrero 2016
7. Cloud Standards Consumer Council <http://www.cloud-council.org/deliverables/CSCC-Security-for-Cloud-Computing-10-Steps-to-Ensure-Success.pdf>. Marzo 2017
8. Mishra, A.; Mathur, R.; Jain, S.; Rathore, J. "Cloud Computing Security". International Journal on Recent and Innovation Trends in Computing and Communication ISSN 2321-8169 Volume: 1 Issue: 1 36 – 39. 2013
9. Mohamed, E.; Abdelkader, H.; El-Etriby, S. "Data Security Model for Cloud Computing". Journal of Communication and Compute. ISSN: 1047-1062. 2013
10. Balu, V.; Mary, L. "A Model of Security Architecture on Private Cloud Using OpenStack". International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 2 587-590. 2015

Sistemas de Procesamiento de Altas Prestaciones Reconfigurables

Eduardo Kunysz^{1,2}, Martín Morales^{1,2}, José Rapallini¹, Jorge Osio¹

¹Centro de Técnicas Analógico Digitales (CeTAD) – Universidad Nacional de La Plata (UNLP)

²Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

ekunysz@unaj.edu.ar, mmorales@unaj.edu.ar, jrapallini@gmail.com, josio@unaj.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio de nuevas técnicas de procesamiento basadas en el procesamiento paralelo para aplicaciones específicas. Dicho sistema utiliza unidades de lógica programable como núcleos de procesamiento, permitiendo implementar funciones de software paralelas directamente en hardware dedicado.

Palabras clave: *Arquitecturas Multiprocesador. Plataformas reconfigurables para procesamiento paralelo. Lógica programable.*

Contexto

Se presenta una línea de Investigación “Aplicaciones de sistemas reconfigurables y microtecnologías en plataformas para cálculo de alta performance”, que es parte del Proyecto de Investigación Científico-Tecnológico *Microtecnologías y Nuevos Recursos de Sistemas Digitales*, Proyectos acreditado por la UNLP y el Programa de Incentivos,

I 212 (01/01/2016 – 31/12/2019) desarrollado en la UIDET CeTAD perteneciente a la Facultad de Ingeniería de la Universidad Nacional de La Plata (UNLP).

En colaboración con actividades desarrolladas por docentes en la Universidad Nacional Arturo Jauretche UNAJ relacionados en el marco de convenio entre ambas instituciones.

Introducción

La tecnología de arreglo de compuertas o FPGA, ha venido siendo utilizada en aplicaciones de diversas áreas, ya sea en investigación como en la industria. Algunos ejemplos pueden ser sistemas de criptografía, exploración de arquitecturas de computadores, procesamiento multimedia, simulaciones financieras, físicas y emuladores.[2] [5]

Cuando se realizan desarrollos de altas prestaciones, como puede ser el diseño de una supercomputadora, hay que analizar las ventajas que podría tener la implementación de uno de estos sistemas con FPGAs. En los últimos años se ha venido intensificando la tendencia a la utilización de FPGAs en vez de procesadores de propósitos generales y ASICs, sobretodo en pequeñas y

medianas empresas que carecen de los recursos necesarios para realizar implementaciones en ASIC .

La mayor ventaja de plataformas reconfigurables, por sobre procesadores de propósitos generales y ASICs es el balance entre el diseño de circuitos electrónicos especializados, o dedicados y la flexibilidad de programación.

Estas ventajas las podemos resumir en los siguientes puntos:

1. El coste de realizar un diseño en una FPGA es mucho menor que en un ASIC.
2. Las herramientas de diseño para FPGA son más baratas que las correspondientes para ASIC.
3. El tiempo de desarrollo hasta la llegada al mercado de un diseño basado en FPGA es mucho menor que el equivalente basado en ASIC.
4. En una FPGA es posible corregir un error de hardware en el diseño, incluso cuando ya ha sido lanzado el producto final.

El crecimiento sostenido en la demanda del poder de cómputo remarca la necesidad de sistemas con enfoques de paralelización masiva y cómputo de alta performance (HPC, High Performance Computing) [1]. La paralelización a nivel de hardware reconfigurable, brinda una herramienta enorme a la hora de procesar grandes volúmenes de datos. [2]

Implementación

El grupo de investigación está explorando técnicas de procesamiento de HPRC con una placa que está compuesta por una FPGA (Xilinx Spartan 6), 3 memorias SRAM para maximizar el

throughput de acceso aleatorio y 2 memorias del tipo DDR2 para almacenamiento de datos. Esto, junto a las fuentes de alimentación e interfaces de comunicación están integrados en un circuito impreso de 12 capas y de un tamaño de 8x12cm.

Sobre la base de esta placa se está trabajando en el diseño de una versión prototipo para evolucionar hacia tecnologías superiores con varias FPGAs por placa.

Para el diseño del nuevo sistema HPRC a medida, en una primera etapa se realizó la planificación del sistema completo, seleccionando cuidadosamente cada uno de los elementos y definiendo las características de la interacción entre los mismos. Para el proyecto se eligió una topología con dos FPGAs (Arreglo de compuertas programables por campo), comunicadas entre sí por medio de una memoria compartida y puertos GTX (Gigabit transceiver (6 Gbps), formado por una línea de clock y una interfaz de datos paralela) y GPIO (puertos de entrada / salida genéricos). Cada FPGA se encuentra conexionada a su vez con diferentes periféricos, estableciendo una topología asimétrica debido a la limitación de recursos disponibles.

En la Fig 1, se observa la arquitectura propuesta para esta variante.

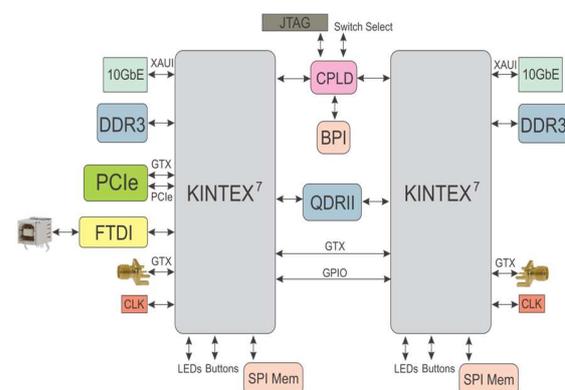


Fig. 1. Diseño de dos FPGAS

El objetivo final es el de obtener unidades de procesamiento compuestas por 4 FPGAs, dichas unidades de procesamiento interconectadas entre si forman un único equipo de procesos de alta performance.

Con el objeto de estudiar distintas topologías de interconexión que permiten flexibilidad de proyectos y aplicaciones específicas se están estudiando topologías como las que se muestran en las siguientes figuras.

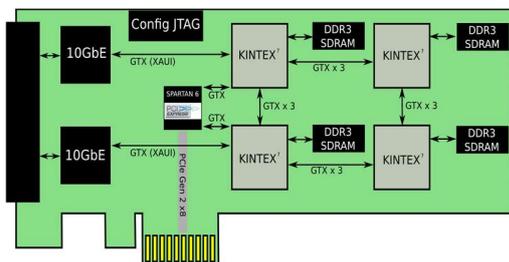


Fig. 2. Topología Backplane / Gigabit ethernet

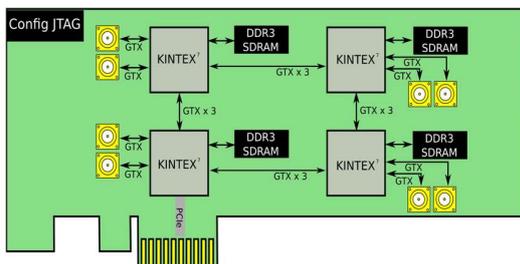


Fig. 3. Topología cubo utilizando buses de alta velocidad propios de las FPGAs de generación 7

Líneas de Investigación, Desarrollo e Innovación

Temas de Estudio e Investigación

- ♣ Arquitecturas multiprocesamiento, para procesamiento paralelo embebido en lógica programable.
- Arquitectura de E/S paralela considerando el software, hardware,

comunicaciones entre módulos y dispositivos de almacenamiento.

- ♣ Técnicas de comunicación entre dispositivos de lógica programable para optimizar el rendimiento y consumo de las HPRC (High Performance Reconfigurable Computing).

Resultados y Objetivos

Resultados y Discusión

Los principales resultados de la propuesta permiten afirmar que se ha diseñado una plataforma de bajo costo que permitirá realizar cómputo de altas prestaciones oem (a medida). Cada uno de los componentes y la topología de comunicación fueron cuidadosamente seleccionados y planificados para lograr una herramienta de cómputo potente y versátil.

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D se participa en el dictado de la carrera de Ingeniería Informática de la UNAJ e Ing. Electrónica y Ing. en Computación de la UNLP. Aportando trabajos de alumnos de las materias Proyectos Final Electrónica, Circuitos Digitales y microprocesadores y Codiseño Hardware Software

Hay 2 investigadores realizando su Maestría, dos investigadores de la UIDET CeTAD, y tres alumnos realizando el proyecto de graduación en la temática.

Referencias

- 1 "CUBE: a 512-FPGA Cluster", Oskar Mencer, Kuen Hung Tsoi, Stephen Cramer, Timothy Todman and Wayne Luk, Ming Yee Wong and Philip Heng Wai Leong , Dept. of Computing, Imperial College London , Dept. of Computer Science and Engineering The Chinese University of Hong Kong .
- 2 "The NAPA Adaptive Processing Architecture", Charlé R. Rupp, Ph.D., Mark Landguth, Tim Garverick, Edson Gomersall, Harry Holt , National Semiconductor Corporation .
- 3 "grape-4 : A Massively Parallel Special-Purpose Computer for Collisional N-Body Simulations", Junichiro Makino , Makoto Taiji, Toshikazu Ebisuzaki, and Daiichiro Sugimoto . IEEE
- 4 "Implementation trade-offs of Triple DES in the SRC-6e Reconfigurable Computing Environment " ,Osman Devrim Fidanci, Hatim Diab, Tarek El-Ghazawi, Kris Gaj and Nikitas Alexandridis
- 5 "FPGAs vs. CPUs: Trends in Peak Floating-Point Performance ", Keith Underwood , Sandia National Laboratories .
- 6 "High Performance Reconfigurable Computing for Science and Engineering Applications ", Peter Leonard McMahon
- 7 "Implementing Simulink Designs on SRC-6 System", David Meixner, Volodymyr Kindratenko, David Pointer, Innovative Systems Laboratory, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign (<http://www.srccomp.com/carte-programming-environment>)
- 8 "Using Hardware Libraries with Impulse C", Ralph Bodenner, Director of Product Development, Impulse Accelerated Technologies, Inc. Application Note (<http://www.impulseaccelerated.com/>)
- 9 "Low Power Hybrid Computing for Efficient Software Acceleration", Mitronics, White Paper. (<http://www.mitronics.com/>)
- 10 "DK4 - Handel-C Language Reference Manual", Celoxica, (<http://www.celoxica.com/>)
- 11 "CoreFire™ Design Suite", Annapolis Micro Systems, Inc., DataSheet (<http://www.annapmicro.com/corefire.html>)
- 12 "Model-Based Design with Simulink, HDL Coder, and Xilinx System Generator for DSP", Kiran Kintali and Yongfeng Gu, MathWorks, White Paper (<http://www.mathworks.com/fpga-design/simulink-with-xilinx-system-generator-for-dsp.html>)
- 13 Customizing Virtual Networks with Partial FPGA Reconfiguration, Dong Yin , Deepak Unnikrishnan, Yong Liao, Lixin Gao and Russell Tessier . Dept. of Electrical and Computer Engineering University of Massachusetts
- 14 "High Performance Biological Pairwise Sequence Alignment: FPGA versus GPU versus Cell BE versus GPP", Khaled Benkrid, Ali Akoglu, Cheng Ling, Yang Song, Ying Liu, Xiand Tian. International Journal of Reconfigurable Computing 2012.

Seguridad Informática

Algoritmo de Cifrado para Sistemas Móviles

Castro Lechtaler, Antonio^{1,2}; Cipriano, Marcelo¹; García, Edith¹,
Liporace, Julio¹; Maiorano, Ariel¹; Malvacio, Eduardo¹; Tapia, Néstor¹;

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad de Ingeniería. Instituto Universitario del Ejército.

² CISTIC/FCE - Universidad de Buenos Aires.

acastro@est.iue.edu.ar , marcelocipriano@est.iue.edu.ar,
{edithgarcia; jcliporace; maiorano; edumalvacio; tapianestor87}@gmail.com

RESUMEN.

Este proyecto de investigación persigue elaborar el diseño y desarrollo de un **Algoritmo de Cifrado para Sistemas Móviles** que por sus propiedades de velocidad, compactibilidad y robustez; pueda ser implementado en equipos de comunicaciones que funcionan sobre Sistemas Móviles.

Se esperan obtener resultados teóricos, prácticos y la realización de un desarrollo experimental.

Los modernos algoritmos criptológicos responden a principios y filosofías diferentes a las que se llevaban a cabo antaño. Uno de los principios de diseño es que el algoritmo debe demostrar su resistencia a los ataques conocidos. Para ello deben contemplarse instancias o funciones, desde la mismísima etapa de diseño que demuestren su capacidad de resistir tal o cual ataque. Así demostrar la robustez y resistencia del algoritmo frente a un conjunto conocido de ataques.

Por ello los diseñadores deben estar en conocimiento y mantenerse actualizados en cuanto a los avances que se efectúen en Criptoanálisis.

Palabras Clave:

Criptografía. Criptosistemas de Clave Privada, Stream Ciphers. Sistemas Móviles

CONTEXTO.

El *Grupo de Investigación en Criptología y Seguridad Informática (GICSI)* pertenece al *Laboratorio de Investigación en Técnicas*

Criptográficas y Seguridad Teleinformática (CriptoLab) pertenece a los *Laboratorios de*

Informática (InforLabs) de la *Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)*, dependiente de la *Facultad del Ejército, Universidad Nacional de la Defensa (UNDEF)*. El mismo se enmarca en el área de la carrera de grado de *Ingeniería en Informática* y del posgrado en *Criptografía y Seguridad Teleinformática* que se dictan en esta institución.

1. INTRODUCCIÓN.

Se entiende aquí por Sistemas Móviles (SM) a aquellos Sistemas de permiten la realización de comunicaciones en posiciones fijas, como también en movimiento: como equipos de tipo VHF¹ y telefonía móvil que requieren de enlaces confidenciales y deben recurrir a la criptografía para obtener tales servicios. Sin embargo no todo sistema de cifrado que ofrezca seguridad puede ser montado sobre tales plataformas.

Estos dispositivos no cuentan con los mismos recursos de hardware y software que otros que están fijos o no tiene sus limitaciones.

El uso eficiente de los recursos de los que el Sistema Móvil disponga será preponderante. Es por ello que si un criptosistema puede demandar una cantidad de recursos mayor a lo disponible (espacio de carga útil, tiempo de

¹ Very High Frequency: rango de frecuencias de 30 MHz a 300 MHz. Empleado por sistemas satelitales, televisión, radiodifusoras de FM, bandas aéreas y marítimas, entre otras.

ejecución o retardos en la implementación, energía consumida, memoria requerida, etc.) puede que atente contra el sistema que pretende proteger.

Tal es el caso de diversos dispositivos cuyas misiones dependen del uso eficiente de sus recursos. En particular los vehículos aéreos no tripulados del Proyecto LIPAM del Ejército Argentino, los cascos de RAIOM² proyecto que lleva adelante CITEDEF³ (en ambos proyectos se realizaron aportes desde el Cripto-Lab).⁴

También se pueden mencionar otros sistemas y vehículos, como el PANHARD francés que el Ejército y otras fuerzas poseen y que le fue encomendado a la EST para su modernización.

Es por ello que el diseño de un criptosistema compacto, veloz y austero en el consumo de los recursos se hace indispensable para dar respuesta a la seguridad de los canales de datos, comando y control o cualquier otro que se precise proteger y dotar de confidencialidad.

Otros temas a investigar son la existencia o no de claves débiles que generan ciclos cortos o debilidades en el cifrado. Y demostrar ser inmune a los ataques criptoanalíticos conocidos, como el Criptoanálisis Diferencial, Lineal, Algebraico, Cube Attack, entre otros[1-5].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN.

Hemos dividido el proyecto en 4 etapas de investigación y desarrollo:

- a) Estudio y análisis de algoritmos que satisfacen los requerimientos y condiciones de entorno del proyecto.
- b) Personalización, diseño y desarrollo del algoritmo:
 - Estudio de sus vulnerabilidades y ataques conocidos.

² RAIOM: Realidad Aumentada para la Identificación de Objetivos Militares.

³ CITEDEF: El Instituto de Investigaciones Científicas y Técnicas para la Defensa; ex Instituto de Investigaciones Científicas y Técnicas de las Fuerzas Armadas (CITEFA)

⁴ El Laboratorio de Criptografía y Seguridad Teleinformática realizó algunos aportes a ambos proyectos.

- Implementación y pruebas del algoritmo.
- c) Determinación de las propiedades criptológicas:
 - Estudio de las propiedades.
 - Experiencias de laboratorio.
 - d) Ejecución de los test y demás pruebas de robustez criptológica.
 - Diseño y programación de los test.
 - Diseño e implementación de los ataques.
 - Análisis de los resultados obtenidos.
 - Redacción del informe final
 - Puesta a punto del algoritmo a entregar

3. RESULTADOS Y OBJETIVOS.

Garantizar la seguridad de las comunicaciones mediante el diseño de un esquema de cifrado y descifrado bajo la modalidad Stream Cipher o Cifrado en Cadena, para que por medio de una Clave Privada pueda dotar de confidencialidad a uno o varios canales de comunicaciones de un Sistema Móvil.

El mismo deberá, ante todo, demostrar su robustez por medio de sus propiedades matemáticas pertinentes. A su vez, la Secuencia Cifrante (Key Bit Stream)[6] que de él se obtenga, deberá satisfacer todos los requisitos aceptados por la comunidad científica que deben tener las Secuencias Seudo-Aleatorias: Test de Golomb, de NIST, Die Hard y demás, estudio de la longitud de recursión, complejidad lineal y período.

Ahorrar recursos económicos al realizar un desarrollo propio y nacional, frente a los costos en equipos y algoritmos comprados en el exterior y en moneda extranjera.

El incremento del Know-How que tendrá el equipo a lo largo de la vida del proyecto será una económica Formación de Recursos Humanos en beneficio de los alumnos del equipo.

4. FORMACIÓN DE RECURSOS HUMANOS.

Los docentes investigadores de este proyecto se encuentran dictando las asignaturas *Matemática Discreta*, *Paradigmas de Programación I, II y Criptografía y Seguridad Telein-*

formática. Desde allí se invita a los alumnos a participar en los proyectos de investigación que se llevan adelante. Es por ello que los alumnos LEIRAS, F. MIGLIARDI A., MONTANARO, L. ROMERO, E. y UVIEDO, G. han demostrado su interés y se han sumado en calidad de colaboradores.

El Cap. Pérez, P. integra el equipo de investigación desde el año 2015 y se espera que este año realice su Proyecto Final de Carrera en un tema afín con este proyecto de investigación.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

- [1] Ding C.; *The differential cryptanalysis and design of natural stream ciphers*. In: Anderson R. (eds.) Fast Software Encryption. FSE 1993. Lecture Notes in Computer Science, vol. 809. Springer Berlin, Heidelberg.
- [2] Wu H., Preneel B. *Differential Cryptanalysis of the Stream Ciphers Py, Py6 and Pypy*. In: Naor M. (eds.) Advances in Cryptology. EUROCRYPT 2007. Lecture Notes in Computer Science, vol. 4515. Springer Berlin, Heidelberg. 2007.
- [3] Muller F., Peyrin T. *Linear Cryptanalysis of the TSC Family of Stream Ciphers*. In: Roy B. (eds.) Advances in Cryptology - ASIACRYPT 2007. Lecture Notes in Computer Science, vol. 3788. Springer, Berlin, Heidelberg. 2005.
- [4] Dinur I., Shamir A. *Cube Attacks on Tweakable Black Box Polynomials*. Advances in Cryptology - EUROCRYPT 2009. Lecture Notes in Computer Science, vol 5479. Springer, Berlin, Heidelberg. 2009
- [5] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; IEEE Transactions on Information Theory. Vol. 55 Ed.7º, 2009.
- [6] Biryukov A., Shamir A. (2000) Cryptanalytic Time/Memory/Data Tradeoffs for Stream Ciphers. In: Okamoto T. (eds) Advances in Cryptology — ASIACRYPT 2000. ASIACRYPT 2000. Lecture Notes in Computer Science, vol 1976. Springer, Berlin, Heidelberg.

Análisis Comparativo de Distintas Toolkits para el Reconocimiento Biométrico de Personas Mediante Voz

Silvia Ruiz, Ernesto Miranda, Mauro Herlein, Graciela Etchart, Carlos Alvez.

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

sruiz@fcad.uner.edu.ar, emiranda@fcad.uner.edu.ar, herlein.mauro@gmail.com, getchart@fcad.uner.edu.ar,
caralv@fcad.uner.edu.ar,

Resumen

El objetivo de este trabajo es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz. Hoy en día los sistemas de identificación de personas se han convertido en una necesidad para la sociedad. A medida que avanza la tecnología y la aplicación de la misma en entornos tanto de ocio como de seguridad, la evolución en desarrollo biométrico es muy grande. Los sistemas de identificación o verificación tradicionales (tarjetas o claves) se han ido sustituyendo por sistemas automáticos de reconocimiento biométrico. La biometría de voz es muy utilizada en aplicaciones o sistemas relacionados con la seguridad, ya que cada individuo tiene características físicas diferentes. En este sentido, el presente trabajo propone realizar un estudio comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz en base a criterios de evaluación previamente definidos que permitirán determinar la/s herramientas consideradas más eficaces.

Palabras clave: biometría, rasgos biométricos, voz, seguridad.

Contexto

Este trabajo surge en el marco del proyecto de investigación *PID 07/G044 "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales"*, que da continuidad al Proyecto *PID 07/G035 "Identificación de personas mediante Sistemas Biométricos. Estudio de factibilidad y su implementación en organismos estatales"* [1][2][3][4][5].

El mismo constituye un futuro trabajo de tesis de la Maestría en Sistemas de Información dictada en la Facultad de Ciencias de la Administración perteneciente a la Universidad Nacional de Entre Ríos. El objetivo general es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz, evaluando y seleccionando la/s herramienta/s consideradas más adecuadas para la implementación de un sistema de reconocimiento por voz.

Introducción

La importancia del uso de la tecnologías basadas en biometrías, se encuentra en su aplicación como mecanismo de control a través de una serie de medidas de características específicas que permite el reconocimiento de personas, para superar los problemas de vulnerabilidad e inconvenientes (gastos en su creación, control, administración, posibles extravíos, olvidos, etc.) del sistema de autenticación por contraseñas, tarjetas de acceso, entre otros.

Los sistemas de reconocimiento biométrico usan características fisiológicas o de comportamiento propias de cada individuo para identificarlo, es decir, se reconoce al usuario por lo que es en lugar de por lo que tiene o sabe [6]. Los rasgos fisiológicos presentan una reducida variabilidad a lo largo del tiempo, pese a que su adquisición es más invasiva y requiere de la cooperación de los sujetos. Por el contrario, los rasgos de comportamiento resultan menos invasivos aunque la exactitud de la identificación es menor debido a la variabilidad de los patrones de comportamiento.

De esta forma, el objetivo de todos ellos será obtener, a partir de la captura de un rasgo biométrico, una representación de cada individuo que resulte lo suficientemente discriminante respecto a las de los demás usuarios del sistema; de modo que, mediante técnicas de reconocimiento de patrones, el sistema sea capaz de determinar la identidad del usuario que está intentando acceder al sistema o simplemente si el usuario está registrado o no en él, dependiendo del tipo de aplicación para la que se utilice.

El reconocimiento de la voz es el proceso de reconocer automáticamente quién está hablando utilizando la información dada por las ondas de sonido emitidas [7][8]. En general, los sistemas de reconocimiento de voz tienen dos módulos principales: extracción de características y comparación de características. La extracción de características es el proceso por el cual extraemos una pequeña cantidad de datos de

la señal de voz que puede ser usada para representar a cada persona. La comparación de características involucra el proceso de identificar a la persona desconocida comparando las características extraídas de su voz, con las previamente obtenidas, que corresponden a las personas conocidas por sistema.

En biometrías de voz existen diferentes tipos:

- Dependiente de texto (modo limitado).
- Independiente del texto (modo ilimitado)

El principal tipo de aplicación de la interacción por voz son los conocidos como Sistemas de Diálogo Hablado (dependientes de texto). A la hora de implementarlos, principalmente se pueden optar por dos tipos de estrategias. La primera de ellas consiste en el uso de un lenguaje basado en comandos o palabras aisladas. Así, el usuario podría utilizar un comando como usar las palabras "Sí" o "No" en un diálogo dirigido, o utilizar como comandos una serie de números para seleccionar una opción en un sistema de asistencia telefónica en vez del sistema DTMF (Dual Tone Multiple Frequency) tradicional. La segunda opción es utilizar el lenguaje natural, o uno restringido a ciertas estructuras gramaticales.

Por otro lado, la Biometría independiente del texto resulta atractiva debido a que se puede aplicar a sistemas de seguridad, realmente hay ocasiones en las que se necesita mayor seguridad para ciertas actividades o simplemente es que se nos puede olvidar una clave o puede ser sustraída. Por eso se ha ido desarrollando proyectos biométricos, además de por la comodidad que le aporta al usuario, ya que no es necesario que recuerde una contraseña o que lleve una tarjeta o una identificación consigo.

La utilización de la voz como rasgo biométrico, además de cumplir propiedades como: universalidad, distintividad, estabilidad, evaluabilidad, no-intrusivo, etc.,

es un rasgo de fácil adquisición. Además, este tipo de biometría no necesita nada que la mayoría de los usuarios no tengan. Para poder utilizar este tipo de sistemas simplemente es necesario algo que permita realizar capturas de señales de voz, como un micrófono, ya sea en el móvil, en la tablet o en una computadora.

A través de los años se han desarrollado diferentes tipos de programas para el reconocimiento de voz que utilizan múltiples herramientas matemáticas como lo son: Dinamic Time Warping (DTW), modelos ocultos de Markov (HMM), redes neuronales, entre otros. Existen una serie de herramientas en la actualidad de bastante uso en el área de reconocimiento de voz tales como: HTK [9], CMU Sphinx [10], CSLU toolkit [11], entre otros.

En este sentido, en el presente trabajo se realizará un estudio comparativo de distintas toolkits para el reconocimiento de voz en base a los criterios de evaluación definidos, analizando sus fortalezas y debilidades.

Líneas de Investigación y Desarrollo

Una de las principales líneas de investigación de este trabajo son los sistemas de reconocimiento de personas mediante voz, utilizados ampliamente en la actualidad en aplicaciones o sistemas relacionados con la seguridad, ya que cada individuo tiene características físicas diferentes. No resulta sencillo realizar este tipo de aplicaciones ya que pueden existir inconvenientes tales como el ruido de fondo, enfermedades, edad, estados de ánimo en los cuales pueden cambiar la voz y también pueden existir diferencias entre sistemas de captación de la señal, ajenos al locutor.

El uso del procesamiento de la voz como un modo de interacción con los sistemas o de acceso a ellos, ofrece innumerables ventajas respecto a las interfaces clásicas de comunicación con los sistemas como por

ejemplo: bajo costo, facilidad de uso, aceptación por parte de los usuarios, captura y transmisión de manera simple, etc.

Resultados y Objetivos

El objetivo general de este trabajo es realizar un análisis comparativo de distintas toolkits para el reconocimiento biométrico de personas mediante voz. En principio, se realizará una investigación exploratoria de las distintas herramientas para el reconocimiento biométrico de personas mediante voz utilizadas en la actualidad. Para ello, se analizarán los sitios webs oficiales de las herramientas, manuales, foros, artículos, etc. que permitan conocer las principales características que presentan las distintas herramientas. Luego, se definirán los criterios de evaluación que permitirán realizar el análisis comparativo y finalmente se seleccionarán la/s herramienta/s consideradas más adecuadas para la implementación de un sistema de reconocimiento por voz.

Formación de Recursos Humanos

La estructura del equipo de investigación está formada por el Director, Co-Director y cuatro integrantes docentes. El Director del proyecto dirige la tesis de Maestría en Sistemas de Información (MSI) de la Facultad de Ciencias de la Administración de la UNER de tres integrantes y la Co-Directora, donde realizarán sus trabajos de investigación en el área del proyecto.

Además, el proyecto cuenta con un becario de Iniciación en la Investigación cuyas tareas están relacionadas con la captura, registración y almacenamiento de datos biométricos y un integrante interno alumno de la Licenciatura en Sistemas,

Referencias

1. Carlos E. Alvez, Marcelo G. Benedetto. "Los Sistemas Biométricos y su Factibilidad de Aplicación en los Organismos Estatales". XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010), El Calafate, Santa Cruz Argentina, 5 y 6 de Mayo de 2010, Universidad Nacional de la Patagonia San Juan Bosco. Páginas 247-251.
2. Graciela Etchart, Lucas Luna, Carlos Leal, Marcelo Benedetto, Carlos Alvez. Sistemas de reconocimiento biométricos, importancia del uso de estándares en entes estatales. CGIV - XIII Workshop de Investigadores en Ciencias de la Computación (WICC 2011), 5 y 6 de Mayo de 2011. Universidad Nacional de Rosario. Rosario – Argentina. Páginas 339-343.
3. Graciela Etchart, Lucas Luna, Rafael Leal, Marcelo Benedetto, Carlos Alvez. "Sistema adecuado a estándares de reconocimiento de personas mediante el iris". CGIV - XIV Workshop de Investigadores en Ciencias de la Computación (WICC 2012), 25 y 26 de Abril de 2012. Universidad Nacional de Misiones. Posadas – Argentina. Páginas 321-325.
4. Graciela Etchart, Carlos Alvez, Marcelo Benedetto "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales" XV Workshop de Investigadores en Ciencias de la Computación (WICC 2013). pp. 97-101. Paraná Entre Ríos.
5. Silvia Ruiz, Graciela Etchart, Carlos Alvez, Ernesto Miranda, Marcelo Benedetto, Juan José Aguirre " Representación e interoperabilidad de imágenes biométricas" XVII Workshop de Investigadores en Ciencias de la Computación (WICC 2015). Salta.
6. Anil K. Jain, Arun Ross, and Salil Prabhakar. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 2004.
7. Doddington, G. R. (1985). Speaker Recognition – Identifying people by their Voices. Proc. of the IEEE.
8. Furui, S. (1981). "Cepstral analysis technique for automatic speaker verification." IEEE Transactions on acoustic, speech and signal processing 29(2): 254- 277.
9. Cambridge University Engineering Department (CUED). Hidden Markov Model Toolkit (HTK), accedido 15 de marzo de 2017. <http://htk.eng.cam.ac.uk/>.
10. Carnegie Mellon University. CMUSphinx, accedido 15 de marzo de 2017. <http://cmusphinx.sourceforge.net/>.
11. Technology and Research Collaborations, Oregon Health & Science University CSLUtoolkit, accedido 15 de marzo de 2017. <http://www.cslu.ogi.edu/toolkit/>.

Análisis de Metodologías de Recolección de Datos Digitales

Mónica D. Tugnarelli (1), Mauro F. Fornaroli (1) , Sonia R. Santana (1), Eduardo Jacobo (1), Javier Díaz (2)

⁽¹⁾ Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos

⁽²⁾ Facultad de Informática – Universidad Nacional de La Plata
e-mail: montug, maufor [@fcad.uner.edu.ar]

Resumen

Una arquitectura de seguridad informática bien definida debe ofrecer un plan y un conjunto de políticas que describan tanto los servicios de seguridad ofrecidos a los usuarios como los componentes del sistema requeridos para implementar dichos servicios. Cuando se produce un incidente o amenaza de seguridad, en el cual un recurso del sistema queda comprometido o potencialmente expuesto a accesos no autorizados, esta arquitectura de seguridad se ve vulnerada.

Considerando la fragilidad y volatilidad de un evento digital, las técnicas y metodologías de forensia informática deben asegurar que se pueda determinar adecuadamente el *qué, quién, cuándo y cómo sucedió* el incidente de seguridad, así como también ocuparse del correcto aseguramiento y preservación de los datos recolectados

Los objetivos establecidos en este proyecto permitirán obtener información sobre la performance de dos metodologías de recolección de datos y analizar comparativamente sus prestaciones en base a determinados criterios y puntos de control establecidos sobre servidores web HTTP y HTTP/2.

Palabras clave: seguridad, incidente, forensia digital, evidencia digital, servidores web, HTTP.

Contexto

El presente PID 7052 se encuadra en una de las líneas de investigación establecidas como prioritarias para su fomento, de la carrera Licenciatura en Sistemas de la Facultad de Ciencias de la Administración correspondiente a la línea "Arquitectura, Sistemas Operativos y Redes". Se adecua además, a las prioridades de la UNER considerando que es un proyecto aplicado a la investigación sobre Tecnologías de la Información y la Comunicación. Asimismo, se continúan y profundizan líneas de trabajo planteadas en cursos de posgrado cursados sobre la temática y en la tesis doctoral presentada por el Dr. Darío Piccirilli [1]

Introducción

Si una arquitectura de seguridad informática está correctamente definida para un sistema, debe ofrecer un plan y un conjunto de políticas que describan tanto los servicios de seguridad ofrecidos a los usuarios como los componentes del sistema requeridos para implementar dichos

servicios. Cuando se produce un incidente o amenaza de seguridad, en el cual un recurso del sistema queda comprometido o potencialmente expuesto a accesos no autorizados, esta arquitectura de seguridad se ve vulnerada.

A modo general, como amenazas del entorno, deben considerarse aspectos que incluyan desde la seguridad administrativa, la seguridad de las comunicaciones, la seguridad informática, la seguridad ambiental hasta la seguridad física. La arquitectura de seguridad debe poder afrontar tanto amenazas intencionales como accidentales, como así también lograr una registración adecuada de los incidentes o eventos de seguridad que ocurran en el sistema.

Diariamente cientos de equipos se encuentran expuestos a potenciales incidentes, consideremos como ejemplo el avance de Internet de las Cosas (IoT) y sus características de trabajo para llegar a dimensionar el grado de posibilidad y el riesgo de ocurrencia de un incidente y su consecuente impacto [2],[3]

En este entorno tecnológico las técnicas y metodologías de forensia informática deben asegurar que se pueda determinar adecuadamente el *qué, quién, cuándo y cómo sucedió* en relación a ese incidente de seguridad, así como también ocuparse de la correcta preservación y trazabilidad de los datos recolectados.

La definición brindada por la primera *Digital Forensics Research Workshop (DFRWS)* celebrada en Nueva York en 2001, acuerda que el análisis forense digital o forensia informática es “*El uso de métodos científicamente probados y derivados hacia la preservación,*

recolección, validación, identificación, análisis, interpretación, documentación y presentación de evidencia digital derivada de fuentes digitales con el fin de facilitar o promover la reconstrucción de los hechos, que pueden constituirse en evidencia legal, o ayudando a anticipar acciones no autorizadas que han demostrado ser perjudiciales para operaciones planeadas.” [4]

Las fuentes digitales proveedoras de los datos a analizar son numerosas, abarcan desde computadoras, teléfonos celulares, tarjetas de cámaras digitales, chips embebidos hasta snapshots de memoria, es decir cualquier tipo de dispositivo que produzca datos digitales.

El análisis forense digital requiere aplicar métodos científicos, técnicas y herramientas para cumplimentar etapas relacionadas con la identificación, preservación y análisis de la evidencia digital, la cual llegado el caso puede ser considerada legalmente en un proceso judicial.

Un aspecto importante es la recolección de esta evidencia y la manera en que se asegura la calidad los datos recolectados.

Actualmente, las metodologías de recolección se concentran mayormente en dos enfoques:

1.- **Recolección de datos a priori de un evento de seguridad:** también conocido como *Forensic Readiness*. Este enfoque introduce el concepto de resguardar la posible evidencia antes de que ocurra el incidente para cubrir dos objetivos: maximizar la capacidad del entorno para reunir evidencia digital confiable y minimizar el costo forense durante la respuesta a un incidente. [5],[6], [7]

2.- Recolección de datos a posteriori de un evento de seguridad. Este enfoque recupera la evidencia luego de que se haya detectado el incidente de seguridad con el objetivo de realizar un análisis forense para determinar lo ocurrido.

En este trabajo se analizarán ambos enfoques metodológicos aplicados a servidores web, específicamente analizando información del protocolo HTTP en sus versiones 1.1 y 2. [8],[9],[10].

Las características funcionales del protocolo HTTP demandan que la recolección de evidencia sea realizada con herramientas o toolkits de análisis forense que proporcionen un entorno adecuado para asegurar la calidad de los datos, su trazabilidad y su eventual admisibilidad como prueba legal.

En este proyecto, para la ejecución de pruebas y adquisición de datos se usarán herramientas de forensia con licenciamiento libre [11],[12],[13] tales como CAINE [14], Xplico [15] y BACKTRACK [16] para las cuales se han configurado entornos de testing. Como guía general para las pruebas se considerará lo establecido en el OSSTMM (Open Source Security Testing Methodology Manual) [17].

Líneas de Investigación, Desarrollo e Innovación

Con este proyecto de investigación se espera conformar una base de conocimiento acerca de la forensia informática en relación a metodologías de recolección de datos digitales. Además del aseguramiento de la evidencia digital, un tema no menor, es

poder determinar la calidad de datos obtenidos, la trazabilidad de los mismos y el volumen de información que se recopila con ambas metodologías. Este volumen de datos está directamente relacionado con el análisis de los tiempos de respuesta a incidentes y la capacidad de acciones inmediatas en tal sentido. Se utilizarán herramientas open source de forensia para determinar las más adecuadas para cada metodología y una guía práctica de aplicación de las mismas.

Resultados y Objetivos

Las actividades propuestas en este proyecto se sustentan en la necesidad de arribar a conclusiones generales y comparativas acerca de dos enfoques de recolección de datos digitales. Puntualmente, el objetivo primario es analizar la performance de ambas metodologías en entornos de servidores web.

Como resultados principales se espera lograr una matriz comparativa que permita:

1. Identificar y describir puntos de control en protocolos HTTP y HTTP/2
2. Identificar puntos de comparación entre enfoques de recolección de evidencia digital
3. Definir y configurar entornos de testing.
4. Describir procedimientos para la recolección de pruebas.
5. Determinar la correcta aplicación de herramientas/toolkits de análisis forenses considerando el entorno y el enfoque de recolección.

Formación de Recursos Humanos

Este proyecto prevé la formación e iniciación en actividades de investigación de cuatro docentes de la carrera Licenciatura en Sistemas, también la incorporación de un becario estudiante, el desarrollo de, al menos, dos proyectos de Trabajo Final de la carrera Licenciatura en Sistemas y la realización de una tesis de maestría correspondiente a la Maestría en Redes de Datos de la Facultad de Informática de la UNLP.

Referencias

1. Piccirilli, Dario. (2016). *Protocolos a aplicar en la forensia informática en el marco de las nuevas tecnologías (pericia – forensia y cibercrimen)*. Tesis de doctorado. Facultad de Informática. Universidad Nacional de La Plata. <http://hdl.handle.net/10915/52212>
2. Internet Crime Complaint Center (IC3). *Annual Report 2015*. <http://www.ic3.gov/media/annualreports.aspx>
3. Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires. *CyberCrime Informe Final 2013 - Delitos Informáticos*. <http://delitosinformaticos.fiscalias.gob.ar/wp-content/uploads/2014/02/CyberCrime-Informe-Final-2013-flip.pdf>
4. Digital Forensic Research Workshop (DFRWS). <http://www.dfrws.org/>
5. TAN, John. (2001). *Forensic Readiness*. http://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf
6. Rowlingson, Robert. *A Ten Step for Forensic Readiness*. (2004) International Journal of Digital Evidence. Volume 2, Issue 3.
7. Poee, A. , Labuschagne, L. *A conceptual model for digital forensic readiness* (2012) <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6320452>
8. RFC 1945 Hypertext Transfer Protocol - HTTP/1.0 <http://tools.ietf.org/html/rfc1945>
9. RFC 2616 Hypertext Transfer Protocol - HTTP/1.1 <http://tools.ietf.org/html/rfc2616>
10. Draft Hypertext Transfer Protocol version 2.0 draft-ietf-httpbis-http2-04 <https://tools.ietf.org/html/draft-ietf-httpbis-http2-04>
11. *Digital Forensic with Open Tools*. (2011). DOI: 10.1016/B978-1-59749-586-8.00001-7. Elsevier.Inc
12. Díaz, Francisco Javier.Venosa, Paula.| Macía, Nicolás. Lanfranco, Einar Felipe. Sabolansky, Alejandro Javier. Rubio, Damián. *Análisis digital forense utilizando herramientas de software libre* . Atículo presentado en Workshop de Investigadores en Ciencias de la Computación (2016) <http://sedici.unlp.edu.ar/handle/10915/52766>
13. Tugnarelli, M.; Fornaroli, M.; Pacifico, C. *Análisis de prestaciones de herramientas de software libre para la recolección a priori de evidencia digital en servidores web*. Artículo presentado en Workshop de Investigadores en Ciencias de la Computación (WICC 2015). ISBN 978-987-633-134-0
14. Computer Aided Investigative Environment <http://www.caine-live.net/>
15. Open Source Network Forensic Analysis Tool (NFAT). <http://www.xplico.org/>
16. Penetration Testing and Security Auditing Linux Distribution. www.backtrack-linux.org/
17. Open Source Security Testing Methodology Manual (OSSTMM) <http://www.isecom.org/mirror/OSSTM.M.3.pdf>
18. . U.S. Department of Justice. *Electronic Crime Scene Investigation: A Guide for*

- First Responders, Second Edition.*
<https://www.ncjrs.gov/pdffiles1/nij/219941.pdf>
19. Forte, D. *Principles of digital evidence Collection* (2003)
<http://www.sciencedirect.com/science/article/pii/S1353485803000060>
 20. World Wide Web Consortium (W3C).
<http://www.w3.org/>
 21. Stallman, Richard: "*Free Software, Free Society: Selected Enssays*". (2002). GNU Press, Boston Massachusetts,
 22. RFC 3227 Guidelines for Evidence Collection and Archiving.
<https://www.ietf.org/rfc/rfc3227.txt>
 23. The *Open Web Application Security Project* (OWASP).
<https://www.owasp.org>
 24. Ley 26.388 "Ley de Delitos Informáticos".
<http://www.infoleg.gob.ar/infolegInternet/anexos/140000-144999/141790/norma.htm>
 25. Caracciolo Claudio, Rodriguez Marcelo, Sallis Ezequiel. (2010). *Ethical Hacking - un enfoque metodológico para profesionales*
 26. López Rivera, Rafael. (2012). *Peritaje Informático y Tecnológico.*
 27. Piattini, Mario. del Peso, Emilio. *Auditoria Informática*, 2da.edicion (2001) Editorial Alfaomega
 28. Computer Forensics. (2008) Volume 56, Number 1. U.S. Department of Justice
 29. FBI Cyber Crime.
<http://www.fbi.gov/about-us/investigate/cyber>
 30. Jarrett, Marshall. Bailie, Michael W. *Electronic Evidence in Criminal Investigations.* Computer Crime and Intellectual Property Section
 31. del Peso Navarro, Emilio y colaboradores. (2001) *Peritajes Informáticos.* Editorial Díaz de Santos
 32. Northcutt, Stephen. Novak, Judy. *Detección de Intrusos* 2da. Edición. (2001). Editorial Prentice Hall
 33. Código Procesal Penal de la Nación Argentina. <http://www.infojus.gov.ar>
 34. Código Procesal Civil de la Nación Argentina. <http://www.infojus.gov.ar>
 35. Ley 25236, Habeas Data. <http://www.infojus.gov.ar>
 36. Noblett, M., Pollitt, M., Presley, L. (2000). *Recovering and Examining Computer Forensic Evidence.* Forensic Science Communications. Volume 2, Number 4. U.S. Department of Justice. Federal Bureau of Investigation (FBI)
 37. *Digital Evidence and Computer Crime.* Forensic Science, Computers and Internet. Third Edition (2011). Eoghan Casey. Elsevier Inc.

Análisis de Seguridad en Redes Wireless Utilizando Dispositivos Móviles

Lic. Paula Venosa - Lic. Nicolás Macia - Lic. Einar Lanfranco - Lic. Alejandro Sabolansky
[pvenosa | nmacia | einar | asabolansky] at linti.unlp.edu.ar

LINTI (Laboratorio de Investigación en Nuevas Tecnologías Informáticas)
Facultad de Informática - UNLP
Calle 50 y 120 – 2do piso – La Plata, Buenos Aires, Argentina

1. Resumen

En la actualidad el mundo es digital y la mayoría de las organizaciones utilizan redes inalámbricas como parte de su infraestructura, extendiendo así sus posibilidades de conectividad, tanto para brindar servicios a sus usuarios como a terceros. Dada esta proliferación resulta necesario realizar una evaluación de las redes Wireless en el marco de una auditoría de la seguridad de una organización.

La implementación de este tipo de redes y su posterior interconexión con la infraestructura cableada de la organización en muchos casos es una tarea sencilla, incluso realizable por usuarios no administradores. Es por esto que al momento de auditar la seguridad de una red no sólo se deben considerar las condiciones en que se brindan los servicios sino también la existencia de redes Wireless “no oficiales o no declaradas” en el ámbito de la organización en cuestión. Entre los alcances esperados de esta línea de I/D/I se busca adquirir experiencia en lo relacionado al campo de investigación de redes Wireless, en particular lo referente a la seguridad de las mismas. Para ello, se pretende identificar y evaluar herramientas de software libre que den soporte a la auditoría de seguridad y que puedan ser utilizados desde dispositivos móviles como un smartphone o una tablet, facilitando la tarea del auditor dentro de las instalaciones de la organización.

Palabras clave: seguridad de la información, Wireless, Mobile, Smartphone, IoT

2. Contexto

En el Laboratorio de Investigación en Nuevas Tecnologías Informáticas (LINTI) [1] de la Facultad de

informática de la Universidad Nacional de La Plata [2], un grupo de docentes/investigadores se dedica a estudiar temas relacionados con la seguridad y privacidad de la información, aplicando los conocimientos en los distintos proyectos en los que participan.

En el marco del proyecto de incentivos “Internet del Futuro: Ciudades Digitales Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de Aprendizaje del Futuro”, este grupo investiga vulnerabilidades de seguridad actuales que afectan a sistemas, redes y servicios.

En líneas anteriores de investigación desarrolladas por el mismo grupo se han tratado las problemáticas relacionadas a dispositivos móviles, principalmente en torno al fenómeno conocido como BYOD [9].

Consideramos que esta línea de investigación resulta fundamental en el marco del proyecto ya que, a través de una conexión inalámbrica es, como en la mayoría de los casos, las cosas empiezan a conectarse a Internet. Se cree que a medida que haya más dispositivos conectados en nuestra red las vulnerabilidades seguirán apareciendo y lo harán en forma exponencial. Esto hará que la auditoría continua y la investigación asociada en busca de soluciones se vuelvan indispensables.

3. Introducción

En la actualidad el mundo es digital y la mayoría de las organizaciones utilizan redes inalámbricas como parte de su infraestructura, extendiendo así sus posibilidades de conectividad, tanto para brindar servicios a sus usuarios como a terceros.

Las redes inalámbricas son una extensión del

perímetro de la infraestructura de las organizaciones.

Dada esta proliferación se torna necesario realizar una evaluación de las mismas dentro del alcance de las auditorías de seguridad de una organización.

Esta evaluación es importante tanto por lo que los atacantes puedan obtener como valor de activo de información de la propia víctima como lo que estos puedan hacer utilizando los recursos de la organización para afectar a terceros.

Un caso recurrente es la utilización del enlace para realizar ataques de denegación de servicio, denominado habitualmente DoS [3]. Para ejemplificarlo basta simplemente pensar que el atacante logra acceder a la infraestructura de la Empresa X, a través de una vulnerabilidad de alguno de sus activos. Dado que X tiene un enlace con capacidad de 20Mbps, y se decide utilizar a la misma para atacar a la Empresa Y que tiene 5Mbps contratados, no hay que ser un experto para adivinar el resultado; es evidente que la Empresa Y se quedará sin posibilidades de utilizar su enlace, ya que estará saturado por el tráfico entrante.

La superficie a defender incrementa su tamaño, ya que potencialmente hay muchos más objetivos, pudiendo listar entre los factores responsables:

- La disponibilidad de conexión a Internet, sobre todo con la masificación del acceso a través de banda ancha.
- La aparición de dispositivos móviles, en particular los denominados smartphones o teléfonos inteligentes.
- La llegada de Internet de las Cosas (IoT), donde muchos de los dispositivos que se conectan carecen de la posibilidad de ser actualizados sumado al fenómeno denominado Plug It and Forget It [4].
- La necesidad de las personas, tanto técnicos como no técnicos, de contar con acceso a la red en cualquier lugar.
- La facilidad de conectar puntos de acceso de tipo Rogue, es decir no autorizados por la organización, que pueden ser instalados por los empleados sin respetar las normas de seguridad de la organización.

En los procesos de auditoría se suele incluir el análisis de las redes wireless disponibles, tarea que habitualmente se realizan desde un escritorio, una sala de racks o alguna oficina de las instalaciones que se facilita a los analistas para ubicar su computadora y hacer su trabajo.

En este nuevo enfoque y con las necesidades de evaluar todas las secciones de la organización objetivo, entendemos que el camino a seguir es un estudio más parecido al wardriving [5], donde se realiza un recorrido en busca de relevar las redes existentes y no sólo las conocidas por los administradores de la red bajo evaluación. Como objetivos adicionales consideramos minimizar los requerimientos de hardware y aplicar mayor grado de inteligencia al proceso de relevamiento.

Como experiencia previa, en el año 2009 investigadores de este equipo realizamos un wardriving por las calles de la ciudad de La Plata obteniendo como resultado un mapa de las redes inalámbricas disponibles halladas al recorrer las calles de la ciudad en un vehículo [6]. En esa oportunidad fue necesario conectar un analizador de redes a un navegador satelital (GPS) y todo ello a una notebook.

En particular, la línea que se presenta en este trabajo se enfoca en proponer mejoras en la metodología y en los procesos para medir los niveles de seguridad de las redes wireless a través del uso de dispositivos móviles.

De esta forma, en nuestro trabajo actual, los dispositivos móviles que hasta ahora han sido un campo de estudio en cuanto a las vulnerabilidades que presentan y las posibles soluciones para mitigarlos [7], [8], adquieren ahora un rol principal ya que en la presente línea aprovecharemos las ventajas de la movilidad que los mismos nos brindan para su utilización en las primeras etapas de la auditoría de redes Wireless.

4. Líneas de Investigación, Desarrollo e Innovación

Sobre los ejes de investigación, inicialmente planteados:

- Realizamos una recopilación de bibliografía para conocer el estado del arte actual.

- Se dirigió alumnos en el desarrollo de sus tesis relacionada con la temática.
- Asistimos a distintas charlas de seguridad en general; en particular en la conferencia Ekoparty se han realizado diversas presentaciones sobre esta temática.
- En el marco de la operatoria diaria del CERTUNLP, hemos realizado varios pentests en distintas organizaciones.

5. Resultados y Objetivos

5.1. Objetivo General

Se espera continuar adquiriendo experiencia en lo relativo al campo de investigación de la seguridad y auditoría de redes y servicios, aprovechando los beneficios de los dispositivos móviles, en este caso como herramienta para facilitar las tareas de testeo.

Se busca lograr tanto una mejora tanto en el resultado de los tests de penetración como llegar a facilitar la tarea de los analistas.

En esta línea es de investigación es que buscamos que eso mismo pueda realizarse utilizando un solo equipo no especial, es decir un smartphone de uso habitual del investigador agregando otras funciones como ser captura de handshake de autenticación, escaner de puertos o generador de informes.

5.2. Objetivos Específicos

- Abordar un análisis de herramientas preexistentes para auditar la seguridad de las redes wireless a fin de poder identificar y entender los problemas que poseen, poniendo especial énfasis en aquellas que sean de software libre.
- Desarrollar una herramienta que permita hacer uso de un smartphone para llevar a cabo la etapa de descubrimiento y relevamiento en el marco de un test de penetración a redes inalámbricas.
- Probar la herramienta en campo, utilizándola en pruebas de intrusión reales.
- Formar RRHH que retroalimenten al grupo de investigadores convirtiéndolo en un referente en el tema.

- Fomentar el uso y la mejora continua de la aplicación a partir de las observaciones que se puedan realizar en campo y ante la evolución propia de los sistemas operativos.

Algunas de las características no negociables en esta primer versión de la solución a implementar:

- El sistema operativo a utilizar para el desarrollo será Android
- El sistema operativo del equipo a utilizar no deben ser modificado ya que esto podría ocasionar un problema de seguridad en el software mismo. En otros términos, el equipo a utilizar no debe ser rooteado para poder utilizar la aplicación.
- Los resultados obtenidos a partir del uso de la aplicación deben poder utilizarse para complementar el resultado de la auditoria general.
- Se espera obtener más información que simplemente relevamientos de redes disponibles y sus configuraciones de seguridad.
- No es necesario que el dispositivo realice todo el procesamiento de la información, lo que recolecte puede ser trasladado a un recurso externo de mayor poder de computo, como un servidor. El postprocesamiento permitirá obtener datos adicionales a los procesados insitu.

6. Formación de Recursos Humanos

La línea de investigación seguridad en redes wireless está siendo abordada por los alumnos Juan Ignacio Bernal y Alejandro Zurita en el marco de la realización de su tesina de grado de la Licenciatura en Sistemas, en conjunto con los docentes Alejandro Sabolansky, Nicolás Macia, Einar Lanfranco y Paula Venosa quienes también forman parte del grupo de seguridad del LINTI de la Facultad de Informática de la UNLP, el CERTUNLP y las cátedras de grado y postgrado Seguridad y privacidad en redes.

En cuanto a las experiencias en gestión y despliegue de redes inalámbricas, la Facultad de Informática de la UNLP cuenta con dos soluciones de administración centralizada, una de las cuales es administrada en forma íntegra por el grupo de investigadores y la otra es de gestión compartida, ya que parte

del control se realiza vía una controladora alojada en el centro de cómputos de la UNLP.

Esta solución implementada dispone de múltiples SSID que entregan direcciones de red en distintas VLAN, contando actualmente con más de 20 dispositivos conectados. Cada una de estas redes posee diferentes mecanismos de autenticación, entre los cuales podemos mencionar WPA2 y WPA2 combinado con un portal captivo, similar al que se encuentra en redes públicas como las disponibles en aeropuertos o locales de comidas rápidas. El portal captivo fue adaptado para cumplir con la identidad institucional y fue complementado con un software que permite manipular los tokens de acceso en forma automática.

El grupo de seguridad del LINTI de la Facultad de Informática de la UNLP trabaja desde el año 2000 con distintas experiencias relacionadas con la Seguridad de la Información: auditorías de seguridad, implementación de infraestructuras de seguridad, consultoría, desarrollo e implementación de Software Libre y concientización.

Cabe destacar que este grupo de investigadores representa a la UNLP en el Centro de excelencia [10] en el tema “Ciberseguridad” de la UIT, desde el año 2015 [11].

Referencias

- [1] Laboratorio de Investigación de Nuevas Tecnologías Informáticas - LINTI. Facultad de Informática: <https://www.linti.unlp.edu.ar>
- [2] Facultad de Informática: <https://info.unlp.edu.ar>
- [3] <https://www.us-cert.gov/ncas/tips/ST04-015>
- [4] Plug IT and Forget IT <https://insights.ubuntu.com/2016/12/15/research-consumers-are-terrible-at-updating-their-connected-devices/>
- [5] Definición de Wardriving: <https://es.wikipedia.org/wiki/Wardriving>
- [6] Wardriving: an experience in the city of La Plata. Autores: Javier Díaz. Matías Robles. Nicolás Macia. Paula Venosa. Germán Vodopivec. CA- CIC 2008. ISBN: 978-987-24611-0-2
- [7] Uso de dispositivos móviles y BYOD: Su impacto en la seguridad. Autores: Nicolás Macia, Einar Lanfranco, Paula Venosa, Alejandro Sabo-lansky, Carlos Damián Piazza Orlando, Sebastian Exequiel Pacheco Veliz. WICC 2015. ISBN 978-987-633-134-0.
- [8] Dispositivos móviles y el fenómeno del BYOD . Su impacto en la seguridad de las organizaciones Autores: Venosa, Paula; Macia, Nicolás; Piazza Orlando, Carlos Damián; Pacheco Veliz, Sebastián. CACIC 2016.
- [9] BYOD: <https://www.sans.org/reading-room/whitepapers/leadership/managing-implementation-byod-policy-34217>
- [10] Centro de excelencia en Ciberseguridad. <http://www.itu.int/en/ITU-D/Regional-Presence/Americas/Documents/EVENTS/2015/0225-BR-COE/Agenda-EN.pdf>
- [11] ITU. <http://www.itu.int>

Análisis del Estado del Arte de Sistemas de Soporte para el Estudio de Vulnerabilidades en Sistemas Web

Juan C. Cuevas, Roberto M. Muñoz, M. Alejandra Di Gionantonio,
Iris N. Gastañaga, Fabián A. Gibellini, Germán Parisi, Diego Barrionuevo,
Milagros Zea Cárdenas.

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas / Universidad
Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja S/N, 5016

juancarloscue@hotmail.com, robertomunioz@gmail.com, ing.alejandravg@gmail.com,
irisg@ciec.com.ar, fgibellini@bbs.frc.utn.edu.ar, germanparisi@gmail.com,
santosdiegob@gmail.com, milyzc@gmail.com

Resumen

Las aplicaciones web contienen vulnerabilidades, las cuales pueden conducir a serias brechas de seguridad tal como el robo de información confidencial.

A través del presente se busca exponer el estado del arte actual referente a la seguridad de la información en sistemas web en producción (sistema objetivo), teniendo en cuenta que un profesional de seguridad busca detectar vulnerabilidades en el sistema objetivo, previamente acordado con la empresa propietaria del sistema. Se presentarán las metodologías existentes, técnicas y herramientas con las que se llevan a cabo diversas pruebas manuales, para luego buscar la repetición de dichas pruebas ante un mismo sistema objetivo generando su automatización.

Por último a partir de experiencias del equipo se nombran las necesidades identificadas durante la realización de pruebas de seguridad.

Palabras clave: Seguridad web;
Auditorías; Ethical Hacking;
Vulnerabilidad; Pentesting; PT, White hat
hacking.

Contexto

Este proyecto de investigación se lleva a cabo en el ámbito del Laboratorio de Sistemas (LabSis) de la Universidad Tecnológica Nacional – Facultad Regional Córdoba (UTN - FRC).

Se enmarca dentro de las líneas Seguridad Informática y forma parte del proyecto de investigación Sistema Integrado de Soporte para Análisis de Vulnerabilidades en Sistemas Web (S.I.S.A.V.S.W.). Código: EIUTNCO0004084. Acreditado por la Secretaría de Ciencia, Tecnología y Postgrado, y financiado por la UTN – FRC.

Introducción

La seguridad de la información describe actividades relativas a la protección de la información y los activos de la infraestructura de la información contra riesgos de pérdida, uso inadecuado, revelación o daño.

Los riesgos de estos activos pueden ser calculados mediante el análisis de las siguientes cuestiones:

- Amenazas a sus activos: Eventos no deseados que pueden causar pérdida,

daño o uso inadecuado de los activos en forma deliberada o accidental.

- Vulnerabilidades: Se refiere a cuán susceptibles son sus activos a ataques.
- Impacto: La magnitud de la pérdida potencial o la seriedad del evento.

Existen disponibles normas, modelos y estándares para asistir a las organizaciones en la implementación de programas y controles apropiados para mitigar estos riesgos, como por ejemplo, la norma ISO 27.001 [1] y los modelos ITIL [2], COBIT [3] y el NIST [4].

El NIST, además de definir una evaluación de la seguridad de la información, plantea tres métodos de evaluación que pueden ser usados para alcanzarla: las pruebas, el examen y las entrevistas [4].

El panorama de las amenazas para la seguridad de las aplicaciones cambia constantemente, los factores claves en esta evolución son los avances hechos por los atacantes, la liberación de nuevas tecnologías con nuevas debilidades, como así también en defensas más integradas y el despliegue de sistemas con complejidad que se incrementa. Los atacantes pueden usar potencialmente diferentes pasos a través de la aplicación web para dañar el negocio o la organización. Juntos estos factores determinan el riesgo total. [5]

En lo referido a las vulnerabilidades diferentes autores abordan la temática. Xie et al postulan que muchas de las vulnerabilidades de seguridad en las aplicaciones actuales son introducidas por los desarrolladores de software al escribir código inseguro los cuales se pueden deber a la falta de comprensión de programación segura [6].

Bates et al postulan que en los últimos años, los fabricantes de navegadores e investigadores han tratado de desarrollar filtros del lado del cliente para mitigar

estos ataques. Pero algunos de estos filtros podrían introducir vulnerabilidades en sitios que antes estaban libres de errores. [7]

Tripp et al postulan que los autores de publicaciones recientes sugieren que las aplicaciones web son altamente vulnerables a ataques de seguridad. En referencia a lo cual citan un reporte reciente de la WASC que proveen estadísticas de seguridad sobre 12186 sitios web en producción, listando un total de 97554 vulnerabilidades detectadas en estos sitios web. Más severamente aún es que cerca del 49% de los sitios analizados se encontró que contenían vulnerabilidades de alto riesgo [8].

Según Shahriar las aplicaciones web contienen vulnerabilidades, las cuales pueden conducir a serias brechas de seguridad tal como el robo de información confidencial [9].

Grossman en un reporte indica que las aplicaciones web de diferentes dominios de uso más frecuente (por ejemplo: banca, salud, TI, educación, redes sociales) son más propensas a ser vulneradas. [10]

Ha habido una gran puja por incluir la seguridad durante el ciclo de desarrollo de software y formalizar la especificación de requerimientos de una manera estándar. Además de un gran incremento en la cantidad de organizaciones dedicados a la seguridad de aplicaciones como la Open Web Application Security Project (OWASP) [11]. Sin embargo, todavía hay aplicaciones web descaradamente vulnerables, principalmente porque los programadores están más concentrados en la funcionalidad que en la seguridad [12].

En lo que se refiere a pruebas de penetración se encontraron, en este estudio, algunos postulados que se referencian en los siguientes párrafos.

Las pruebas de penetración son pruebas de seguridad en las cuales los evaluadores imitan ataques reales para

identificar los pasos realizados por el atacante a los fines de eludir la seguridad de una aplicación, sistema o red. La mayoría de las pruebas implican buscar combinaciones de vulnerabilidades en uno o más sistemas [4].

Una prueba de seguridad puede ser definida como un intento legal y autorizado para ubicar y explotar sistemas de información con el propósito de hacer un sistema más seguro. El proceso incluye pruebas de vulnerabilidad que deben terminar con recomendaciones específicas para arreglar los problemas descubiertos durante las mismas. Estas pruebas de seguridad son conocidas como: pentesting, PT, hacking, ethical hacking, white hat hacking. [13]

Las nociones definidas por las palabras penetración y vulnerabilidad son complementarias a cada una de la otra. Si una penetración o intrusión fue detectada esto sucedió debido a la vulnerabilidad (agujero) [14].

En referencia a los ataques, según el último top 10 de las OWASP realizado cada 3 años, esta organización lista los ataques más usados para explotar vulnerabilidades. Estos son: Inyección (SQL, OS y LDAP), pérdida de autenticación y gestión de sesiones, cross site scripting (XSS), referencia directa e insegura a objetos, configuración de seguridad Incorrecta, exposición de datos sensibles, inexistente control de acceso a funcionalidades, falsificación de peticiones en sitios cruzados (CSRF), uso de componentes con vulnerabilidades conocidas, reenvíos y redirecciones no válidas. [5]

El individuo que lleva a cabo estas pruebas es llamado tester de penetración o pentester. Las pruebas son ejecutadas usando tecnologías automatizadas y manuales para comprometer servidores, sistemas, aplicaciones web, redes inalámbricas, dispositivos de redes,

móviles y otros puntos de exposición. [15]

Las organizaciones, deseando garantizar la seguridad de sus sistemas, pueden mirar hacia la adopción de las medidas adecuadas para proteger contra posibles fallos de seguridad. Una de estas medidas es contratar los servicios de pruebas de penetración para encontrar vulnerabilidades presentes en la red de la organización, y proporcionar recomendaciones a los medios para atenuar esos riesgos. [16]

Por otro lado y en relación a los temas de seguridad de la información, vulnerabilidades y pruebas de penetración, en el presente estudio no se puede separar lo antes expuesto de la existencia de metodologías, técnicas y herramientas específicas para el abordaje de la temática antes expuesta.

Respecto a las metodologías que son pasibles de ser utilizadas, existen mejores prácticas sobre las cuales basar la realización de este tipo de evaluación, aunque en general, cada profesional puede incorporar sus variantes. Algunos ejemplos pueden ser: el documento del NIST [4], el documento del Open Source Security Testing Methodology Manual (OSSTMM) [17], el marco de trabajo denominado Information System Security Assessment Framework (ISAAF) [18] y el Open Web Application Security Project (OWASP) [11].

Además existen herramientas que facilitan a los pentester realizar ataques de Injection, particularmente, existen varias relacionadas al ataque SQL Injection como por ejemplo: SQLmap, Havij y V1p3R. Estas herramientas se caracterizan principalmente por contener vectores de ataques más usados permitiendo automatizar un ataque [18] [19].

Cross-Site Scripting (XSS) aborda otro tipo de ataque común en aplicaciones web

que consisten en aprovechar las características de los lenguajes que se ejecutan en el navegador, tales como el lenguaje JavaScript. Yusof y Pathan referencian tres tipos de ataques XSS: persistente, no persistente y basado en el DOM [20]. Hay varias maneras de prevenir estos ataques XSS: mediante el “sanitizado” de todas las entradas de información al sistema [20] o un método que haga inútil la cookie que pueda robar el atacante [21].

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se inscribe dentro de esta línea de investigación de seguridad de la información, más específicamente pruebas de penetración en sistemas web en producción, de forma que permitan identificar las vulnerabilidades existentes dentro de un ambiente controlado.

Resultados y Objetivos

Como resultado del estudio del estado del arte surgen algunas reflexiones primigenias: la cantidad de publicaciones relativas a pruebas de penetración son menores a las referidas al tema de vulnerabilidades, las cuales a su vez también son menores referidas al tema de seguridad de la información; de la documentación analizada se identificó que sólo una de ellas integra los temas de seguridad de la información, vulnerabilidad y pruebas de penetración, a saber el estándar del NIST; la mayoría de las publicaciones analizadas abordan la temática para aplicaciones web en producción con algunas excepciones orientadas al ciclo de vida de desarrollo de la aplicación; y por último podemos reflexionar que algunas publicaciones hacen referencia a que las vulnerabilidades son generadas por los

programadores o desarrolladores. Lo anteriormente publicado genera la necesidad de profundizar el conocimiento en el estado del arte en lo referente a seguridad de la información en general y de riesgos, vulnerabilidades y pruebas de penetración en particular.

En la práctica, a partir de experiencias de los integrantes del equipo se identificó la necesidad de gestionar la ejecución de múltiples pruebas de penetración en el contexto de la seguridad de la información de sistemas web en producción, basándose en metodologías abiertas, para identificar y analizar sus vulnerabilidades. Para lo que es necesario:

- Permitir a los pentesters identificar vulnerabilidades y automatizar el proceso de identificación.
- Lograr que el sistema emita un diagnóstico respecto a las vulnerabilidades del sistema web analizado.
- Crear una base de datos que facilite al pentester vincular metodologías, técnicas y herramientas para abordar la evaluación de vulnerabilidades de un sistema web determinado.
- Generar un sistema que contribuya y facilite el desarrollo de las actividades del pentester.
- Obtener un producto (sistema) que sea simple de utilizar por los profesionales de la seguridad.

Formación de Recursos Humanos

En el equipo trabajarán estudiantes avanzados de la carrera de Ingeniería en Sistemas de Información, los cuales actualmente se desempeñan en el Laboratorio de Sistemas (LabSis), con la finalidad de que inicien su formación en investigación científica y tecnológica profundizando sus conocimientos en temas significativos en la seguridad de la

información. Los estudiantes podrán realizar la Práctica Supervisada.

A los fines de facilitar el acceso a este producto a los docentes, estudiantes y graduados, el mismo se instalará en el Laboratorio de Sistemas (LabSis).

Referencias

[1] ISO/IEC 27001. "Tecnología de la información". Técnicas de la seguridad. Sistemas de gestión de la seguridad de la información (SGSI). Requisitos. ISO Ginebra, Suiza 2013.

[2] ITIL. Information Technology Infrastructure Library. Vs. 3 - 2001.

[3] COBIT Control Objectives for Information and Technology. Vs. 5 - 2012.

[4] Technical Guide to Information Security Testing and Assessment. SP 800-115. NIST National Institute of Standards and Technology. 2008.

[5] "OWASP top 10 2013 Project". Open Web Application Security Project. https://www.owasp.org/index.php/Top_10_2013-Top_10. 2013.

[6] XIE, J.; Chu B.; Lipfort, H. R.; Melton, J. T.: "ASIDE: IDE Support for Web Application Security". ACSAC '11, Orlando, Florida USA. Dec. 5-9/2011

[7] Bates,D; Barth, A.; Jackson, C.: "Regular Expressions Considered Harmful in Client-Side XSS Filters", Raleigh, NC, USA. April 26-30/2010.

[8] Tripp, O.; Weisman, O. y Guy, L.: "Finding Your Way in the Testing Jungle". ISSA '13, July 15-20/2013, Lugano, Switzerland.

[9] Shahriar, H.: "Security Vulnerabilities and Mitigation Techniques of Web Applications". SIN'13, November 26-28/2013, Aksaray, Turkey.

[10] Grossman, J.: "How does your website security stack up against peers?" White Hat Report, Summer. 2012.

[12] Josh Pauli: "The Basics of Web Hacking: Tools and Techniques to Attack the Web". Ed. ELSEVIER. 2013. 25 Wymnan Street, Waltham, MA 02451, USA.

[13] Engebretson, P.: "The Basics of Hacking and Penetration Testing". Elsevier 225 Wyman Street, Waltham, MA 02451, USA. 2011.

[14] Hahanov, V.; Hayford, A.; Ahmetoglu, A. H.; Nunmirradovich, J. D.; Abeid, A. M.; Stanley, O.: "Pentesting and vulnerability diagnosis". CADSM. Poyana - Svalyava (Zakrpatia). UKRAINE 19- 23/february 2013

[15] Nagpal, B.; Chauhan, N.; Shing, N y Panesar, A.: "Tool based implemetation of SQL Injection for penetration testing". International Conference on Computing, Communication and Automation (ICCCA). IEEE. 2015.

[16] Xynos, K.; Sutherland, I.; Read, H.; Everitt, E.; Blyth, A. J. C.; "Penetration Testing and Vulnerability Assessments: A Professional Approach". International Cyber Resilience conference. 2010

[17] "Open Source Security Testing Methodology Manual (OSSTMM)". Institute for Security and Open Methodologies (ISECOM). Diciembre 2010. Cataluña. España.

[11] Guía OWASP (Open Web Application Security Project). Vs. 3.

[19] Ciampa, A.; Visaggio, C. A.; y Di Penta, M. "A heuristic-based approach for detecting SQL-injection vulnerabilities in Web applications". SESS. Cape Town, South Africa. Mayo 2/2010.

[18] Nagpal, B; Chauhan, N.; Singh, N.; Panesar, A.: "Tool Based Implementation of SQL Injection for Penetration Testing". International Conference on Computing, Communication and Automation (ICCCA2015). 2015

[20] Yusof, I; Pathan, A. S.; "Preventing Persistent Cross-Site Scripting (XSS) Attack By Applying Pattern Filtering Approach". IEEE. 2014.

[21] Takahashi H.; Yasunaga K.; Mambo M.; Kim K.; Youl Youm H.: "Preventing Abuse of Cookies Stolen by XSS". Eighth Asia Joint Conference on Information Security. IEEE 2013.

Anonimato Incondicional en Sistemas de Voto Electrónico Presencial

Pablo García¹; Silvia Bast¹; Germán Montejano^{1 2}

¹Departamento de Matemática
 Universidad Nacional de La Pampa
 Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
 Tel.: +54-2954-425166– Int. 28
 [pablogarcia, silviabast]@exactas.unlpam.edu.ar

²Departamento de Informática
 Universidad Nacional de San Luis
 Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
 Tel.: +54-2652-424027 – Int. 251
 gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

Resumen

La discusión sobre la viabilidad de la implementación del voto electrónico parece interminable. Por ejemplo, luego de arduas discusiones legislativas, el Congreso de la Nación ha rechazado la propuesta del Poder Ejecutivo Nacional de implementar un sistema de boleta única electrónica. En los últimos días, un importante grupo de expertos informáticos de Universidades Nacionales se manifestaron en contra del voto electrónico¹ e invitan a firmar una adhesión a tal rechazo².

En un momento de la historia en el que por medios digitales se hacen múltiples operaciones que involucran, por ejemplo, grandes cantidades de dinero o riesgo para vidas humanas, parece apresurado afirmar que es imposi-

ble generar un sistema de votación electrónica que resulte apropiado.

Si se analiza en detalle, la característica distintiva de este tipo de sistemas es la necesidad de mantener el anonimato de quienes emiten un sufragio. En otros aspectos, el voto electrónico no se diferencia demasiado de otras aplicaciones cuyo uso se encuentra absolutamente generalizado.

Este documento expone una serie de avances que, en el ámbito de un Proyecto de Investigación que involucra a los autores, se han obtenido en el tema de la privacidad del votante, que es el punto más complejo de resolver en un sistema de E – Voting.

Palabras clave: *E-Voting, Anonimato Incondicional, Secreto Perfecto, One Time Pad, Recuperación de Colisiones, Voto Presencial.*

¹<http://www.cronista.com/economia politica/Expertos-universitarios-lanzaron-una-campana-contra-el-voto-electronico-20161101-0113.html>

²<http://www.dc.uba.ar/solicitada-voto-electronico>

Contexto

Este trabajo se enmarca el Proyecto de Investigación: "Aspectos de Seguridad en Proyectos de Software", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (UNLPam), Resolución N° 488/14 del Consejo Directivo. Tal proyecto surge desde la línea de Investigación "Ingeniería de Software y Defensa Cibernética", presentada en [1], y que a su vez se enmarca en el Proyecto "Ingeniería de Software: Aspectos de alta sensibilidad en el ejercicio de la Profesión del Ingeniero de Software" de la Facultad de Ciencias Físico - Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL) y que incluye acciones de cooperación con la Universidad Federal de Minas Gerais (UFMG, Brasil).

Entre tales acciones deben mencionarse:

- Pablo García realizó una estadía de un año en la Universidad Federal de Minas Gerais (UFMG), aprobando seminarios de posgrado y trabajando en el grupo "Criptografía Teórica y Aplicada", dirigido por Jeroen van de Graaf, PhD. Desde el 1/3/2012 hasta el 15/12/2012.
- Jeroen van de Graaf, PhD., Docente de UFMG, y el Dr. Germán Montejano (UNSL) fueron orientadores del Mg. Pablo García en el desarrollo de su tesis de para obtener el grado de Magister en Ingeniería de Software, defendida en 2013.
- Pablo García realizó una estadía de intercambio y actualización en el laboratorio 4303 de DCC, ICEX, UFMG, a través de una beca CAFPBA. Desde el 18/10/2106 hasta el 18/11/2106.
- Se prevé otra estadía de un mes en 2017 para intercambio de avances entre el grupo de investigación UNLPam – UNSL y el de DCC (UFMG).

1. Introducción

Una de las premisas fundamentales para este grupo de investigación se refiere a la necesidad de otorgar seguridad incondicional a la privacidad del votante. En efecto, muchos esquemas de voto electrónico, (aquellos que se basan en Mix Nets, [2]), han otorgado seguridad incondicional al proceso de votación (que sólo debe ser protegido por las diez horas que dura la elección) y seguridad computacional al anonimato, que debe asegurarse indefinidamente.

Bajo el convencimiento de que esa propuesta es incorrecta, se buscan alternativas que operen de manera exactamente inversa, asegurando el secreto eterno de la opción que realizó un votante, aún sabiendo que esa información podría permanecer indefinidamente disponible en medios digitales para su análisis.

Los avances en ese sentido fueron desarrollándose en múltiples publicaciones desde 2014 hasta la fecha (por ejemplo, [3] y [4]) y fueron finalmente recopilados y revisados para la publicación, como libro, en [5].

En consecuencia, se comenzó a trabajar en un esquema concreto e integral de voto electrónico presencial que atendiera a todas esas consideraciones y que, simultáneamente, otorgara niveles razonables de seguridad computacional a la hora de proteger el acto eleccionario. La primera versión de tal propuesta se denomina "OTP – Vote" y fue presentada en [6]

El esquema inicial, todavía en etapa de diseño final, propone dos condiciones básicas para garantizar la privacidad incondicional:

- Rigurosa separación de la identificación del votante y el acto de votación.
- Almacenamiento del sufragio en posiciones aleatorias (potencialmente diferentes) en un esquema basado en canales paralelos de slots.

Como consecuencia de lo anterior, las colisiones son posibles; por lo tanto, es factible perder votos. La implementación de canales paralelos de slots es una optimización del esquema conocido como Birthday Paradox ([7]), que enuncia lo siguiente:

“En un grupo de 23 personas, la probabilidad de que dos cumplan años el mismo día es superior a $\frac{1}{2}$.”

En la práctica, eso se relaciona con almacenar los sufragios en un único vector unidimensional. La alternativa propuesta implementa un conjunto de arreglos paralelos. Cada sufragio se almacenará una vez en cada canal, en posiciones aleatorias potencialmente distintas. Aplicando el nuevo esquema, un voto sólo se perderá si colisiona en todos los canales. La probabilidad de ese evento puede llevarse a cualquier valor exigido mediante la aplicación de las fórmulas presentadas en [8] y [9].

El esquema de múltiples claves, inspirado en [10], exige que para poder realizar un fraude, la totalidad de las personas que cuentan con una clave individual, deben ser corruptas. En efecto, con que uno sólo de los involucrados muestre una conducta honesta, cualquier intento de fraude será detectado.

Finalmente, OTP – Vote incorpora una técnica de recuperación de votos basada en XOR que disminuye aún más la probabilidad de

pérdida de sufragios. De acuerdo con un alto número de simulaciones, la misma es despreciable si se administran apropiadamente los parámetros del sistema.

Concretamente, por la forma en que opera el modelo, no es posible que se pierda un único voto individual. Y la probabilidad de que se pierdan n votos (con $n > 1$), es que los mismos coincidan en todos los canales implementados. Los primeros análisis en ese sentido se presentan en [11] y muestran que, en las condiciones correctas, la probabilidad de que se pierdan sufragios puede minimizarse tanto como se desee.

2. Líneas de Investigación, Desarrollo e Innovación

OTP – Vote se encuentra en una etapa avanzada de su desarrollo. Si bien el esquema muestra características interesantes, quedan algunas cuestiones pendientes:

- Definir el sistema criptográfico exacto que se utilizará entre las estaciones de votación y los servidores.
- Formalizar los commitments necesarios para generar una comunicación confiable.
- Concretar, en base al punto anterior, un protocolo antifraudes que garantice una operación segura con un nivel aceptable de eficiencia.
- Desarrollar técnicas de Verificabilidad “End to End” que permitan, simultáneamente, la verificación individual de cada votante y la verificación global del resultado del proceso.

La concreción de los cuatro pasos mencionados permitiría construir un sistema de voto electrónico sólido y confiable.

3. Resultados y Objetivos

OTP – Vote es una propuesta para implementación de voto electrónico presencial que se obtiene como resultado de cinco años de investigación. En él se plasman todos los avances individuales que se fueron obteniendo desde 2012:

- Se implementa un sistema de almacenamiento basado en canales paralelos que minimiza de manera significativa la pérdida de sufragios por colisiones y garantiza el anonimato incondicional.
- Se utiliza una técnica basada en múltiples claves de estilo One Time Pad que garantizan secreto perfecto.
- Se aplica un modelo de recuperación de colisiones por XOR que refuerza la seguridad del modelo.

A futuro, se espera implementar todos los puntos mencionados (y los que aún están en análisis) en una aplicación concreta de voto electrónico.

4. Formación de Recursos Humanos

En el marco del presente proyecto se presentan múltiples acciones relacionados con la formación de recursos humanos:

1. Pablo García defendió su tesis para obtener el grado de Magister en Ingeniería de Software de la Universidad Nacional de San Luis, el día 13/11/2013. La tesis se tituló: “Optimización de un Esquema Dining Cryptographers Asíncrono” y fue dirigida por Jeroen van de Graaf, PhD (UFMG) y Dr. Germán Montejano (UNSL).
2. Silvia Bast defendió su tesis para obtener el grado de Magister en Ingeniería de Software de la Universidad Nacional de San Luis, bajo la dirección del Dr. Germán Montejano (UNSL) y del Mg Pablo García (UNLPam) el día 14/12/2016. La tesis se tituló: “Optimización de la Integridad de Datos en Sistemas de E- Voting”.
3. Pablo García está desarrollando su tesis para obtener el grado de “Especialista en Ingeniería de Software”. Su plan de trabajo fue aprobado y se planea su defensa para agosto de 2017. La tesis se titula: “Anonimato en sistemas de Voto Electrónico” y es dirigida por Jeroen van de Graaf, PhD (UFMG) y Dr. Germán Montejano (UNSL).
4. Silvia Bast está desarrollando su tesis para obtener el grado de “Especialista en Ingeniería de software”. Su plan de trabajo fue aprobado y se planea su defensa para agosto de 2017. La tesis se titula: “Sistemas de e-Voting: Integridad de Datos” y es dirigida por el Dr. Germán Montejano (UNSL) y del Mg Pablo García (UNLPam).
5. Estela Marisa Fritz: completó su etapa de capacitación en un tema en el que no era experta. Durante 2017 realizará todos los aportes relacionados con la temática de generadores aleatorios, insumo necesario para los nuevos avances en el proyecto. Los mismos deberían plasmarse en una tesis de posgrado.
6. Silvia Bast y Pablo García completaron el cursado de la totalidad de los créditos para el Doctorado en Ingeniería Informática (FCFMyN – UNSL).

5. Referencias

- [1] **Uzal R., van de Graaf, J., Montejano G., Riesco D., García P.**: “Inicio de la Línea de Investigación: Ingeniería de Software y Defensa Cibernética”. Memorias del XV WICC. Ps. 769 - 773. ISBN: 9789872817961. 18-19/04/2013.
- [2] **Jakobsson M., Juels a., Rivest R.**: ”Making Mix Nets Robust for Electronic Voting by Randomized Partial Checking”. USENIX Security '02, ps. 339-353. 2002.
- [3] **van de Graaf J., Montejano G., García P.**: “Optimización de un Protocolo Non - Interactive Dining Cryptographers”. Congreso Nacional de Ingeniería Informática / Sistemas de Información. CoNaIISI 2013. 21 y 22 de noviembre de 2013. Córdoba, Argentina
- [4] **van de Graaf J., Montejano G., García P., Bast S.**: “Anonimato en Sistemas de Voto Electrónico”. Memorias del XVI Workshop de Investigadores en Ciencias de la Computación 2014 (WICC 2014). Ps. 822 – 826. ISBN: 9789503410844.
- [5] **García, P.**: “Una Optimización para el Protocolo Non - Interactive Dining Cryptographers: una Propuesta Alternativa para Obtener una Implementación Eficiente”. ISBN - 13: 978-3-639-85270-7. ISBN - 10: 3639852702. EAN: 9783639852707. Idioma: Español. Editorial Académica Española Número de páginas: 180. Fecha de publicación: 31/01/2017.
- [6] **Bast S.**: “Optimización de la Integridad de Datos en Sistemas de E-Voting”. Tesis de Maestría defendida en la Universidad Nacional de San Luis. 14 de Diciembre de 2016. San Luis, Argentina.
- [7] **Flajolet P., Gardy D., Thimonier L.**: “Birthday Paradox, Coupon Collectors, Catching Algorithms and Self - Organizing Search”. Discrete Applied Mathematics 39, ps. 207-223. North-Holland. 1992
- [8] **van de Graaf J., Montejano G., García P.**: “Manejo de Colisiones en un Protocolo Non - Interactive Dining Cryptographers”. Anales de las 42° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Páginas 29 a 43. Septiembre 2013.
- [9] **García P., van de Graaf J., Montejano G., Bast S., Testa O.**: “Implementación de Canales Paralelos en un Protocolo Non - Interactive Dining Cryptographers”. 43° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO 2014), Workshop de Seguridad Informática (WSegI 2014).
- [10] **Broadbent A., Tapp A.**: ”Information - Theoretic Security without an Honest Majority”. Computing Research Repository - CORR. vol. abs/0706.2, ps.410-426, 2007.
- [11] **García P., Bast S., Montejano G., Fritz E.**: “Codificación de Sufragios con Detección de Colisiones en NIDC con Canales Paralelos de Slots”. Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2016.

Aplicabilidad de Normas de Gestión de la Configuración a Medianas y Pequeñas Instalaciones TIC de Organismos Públicos

Miguel A. Fernández, Raúl D. Montoya, Gabriel Antoniutti, Diana F. Fernández, Horacio D. Benítez

Universidad Nacional de Entre Ríos – Facultad de Ciencias de la Administración
migfer@fcad.uner.edu.ar, radamonto@hotmail.com.ar, gabrielantoniutti@gmail.com,
fernandezdiana@gmail.com, dhbenitez@gmail.com

Luis Enrique Prudenza

Universidad Tecnológica Nacional – Regional Concepción del Uruguay
lunenro@gmail.com

RESUMEN

Configuración es un término genérico usado para describir un grupo de componentes que funcionan en forma conjunta para dar servicios TIC (Tecnologías de la Información y las Comunicaciones). La Gestión de la Configuración, por su parte, tiene por objetivo controlar los activos y elementos de configuración que forman parte de la infraestructura TIC, asegurando su contribución al valor de los servicios o productos institucionales.

Las normas ITIL (Information Technologies Infrastructure Library)[9] e ISO20000[8], son estándares internacionales que, entre otros temas, dan un encuadre ordenado para tratar de resolver este asunto.

La pregunta que surge es: ¿este modelo de gestión de la configuración, emergente de ITIL e ISO20000, es aplicable en las MyPITIC (Medianas y Pequeñas Instalaciones TIC) de los organismos de la APER (Administración Pública de Entre Ríos) ? Nuestra hipótesis es que no es aplicable por el grado de complejidad que conlleva.

Los resultados de las encuestas y trabajo de campo realizados sobre una muestra representativa de las MYPITIC de la APER que se sintetiza en el presente trabajo permite dar validez a la hipótesis planteada.

PALABRAS CLAVES: ITIL, ISO20000, MyPITIC, APER, Gestión de la Configuración.

CONTEXTO

El trabajo se enmarca en el Proyecto PID N° 7043 de la Universidad Nacional de Entre Ríos.

1. INTRODUCCIÓN

Para que una organización alcance exitosamente su objetivo institucional, la administración eficiente de los datos es un aspecto crítico. Ello surge del aumento de la dependencia hacia la información y de la fiabilidad de los sistemas que la proveen; del incremento de la vulnerabilidad y del amplio espectro de las amenazas; de los costos de las inversiones en sistemas de información (SI); y del potencial que posee la TI para cambiar drásticamente las organizaciones y crear nuevas oportunidades[7].

La configuración es la esencia de la infraestructura TIC. Son muchos los procesos operativos y de decisión que dependen de su correcta gestión. El mantener bajo control y disponer de información confiable y actualizada sobre los elementos específicos de la infraestructura (Elementos de Configuración) y sobre las relaciones entre ellos, contribuye a ejecutar de una manera confiable los procesos relacionados con la misión del organismo y resguardar la integridad de la instalación y de los servicios que se prestan.

Realizar una correcta gestión de la configuración implica mantener registrados todos los datos de los elementos de configuración sustantivos, requeridos para la prestación del servicio. Esos datos pueden ir desde su descripción e interconexión, hasta un nivel de detalle que incluya la categoría, las relaciones, los atributos y los posibles estados en los cuales pueden estar en determinado momento.

En especial, cuando la institución a la que sirve la instalación es un organismo de administración pública, no solo influyen en los criterios de gestión de la configuración los aspectos relacionados con los servicios que debe prestar, sino también aquellos que deben respetarse por el encuadre jurídico de la institución, sometida siempre a regulaciones y controles propios de su naturaleza.

A medida que las tecnologías de la información y de las comunicaciones fueron evolucionando e integrándose, la aparición de novedades tecnológicas se fueron haciendo cada vez más frecuentes, complejas y especializadas. Esto llevó a que las empresas, tanto públicas como privadas, se vieran forzadas a encauzar su

propio devenir tecnológico, de manera de tratar de mantener el proceso bajo control, al menos respecto de sus propios intereses y necesidades.

Son muchos los aspectos de los servicios TIC que deben gestionarse adecuadamente para asegurar su contribución al valor de los servicios o productos institucionales, siendo el de la configuración de las instalaciones tan solo uno. Pero de gran importancia y fuerte incidencia en otros. Existen conjuntos de recomendaciones basadas en las buenas prácticas, de los cuales uno de los más conocidos es ITIL que describe procesos para gestionar cada uno de los aspectos importantes en una instalación de servicios informáticos.

Estos estándares son fruto de la recopilación y ordenamiento de buenas prácticas, producto de relevamientos realizados originalmente en Gran Bretaña en grandes instalaciones TIC de la administración pública y luego enriquecidos con otros aportes referidos en general a grandes organizaciones. Los aspectos ordenados por las recomendaciones ITIL/ISO20000 son abarcativos de la mayoría de las funciones que deben desarrollarse en una instalación TIC, de una manera exhaustiva, detallada y especializada.

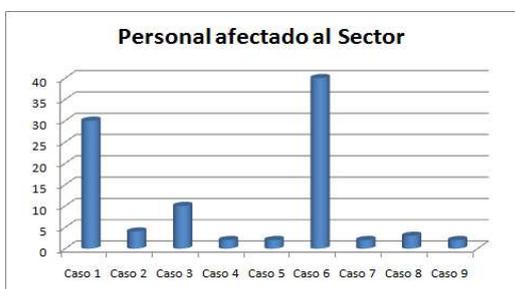
¿Es posible pensar en un modelo de Gestión de la Configuración que siga los lineamientos ITIL/ISO20000, teniendo en cuenta el marco regulatorio de la administración pública provincial, trabajando con el personal TIC propio del organismo, y sentando las bases para un ordenamiento tipo ITIL en las demás funciones, adaptado a la envergadura real y a los recursos habituales de dichas instalaciones?

Para responder a esta pregunta, se ha llevado adelante una tarea de relevamiento de datos sobre una muestra de nueve (9) instalaciones de la provincia – Cuatro (4) municipios, Una (1) instalación del Poder Judicial y Cuatro (4) organismos del Estado Provincial.

A partir de la selección de informantes claves y de una encuesta inicial[4], se realizaron contactos personales, telefónicos, videoconferencia, etc. obteniéndose valiosa información que en sus aspectos mas significativos se detallan.

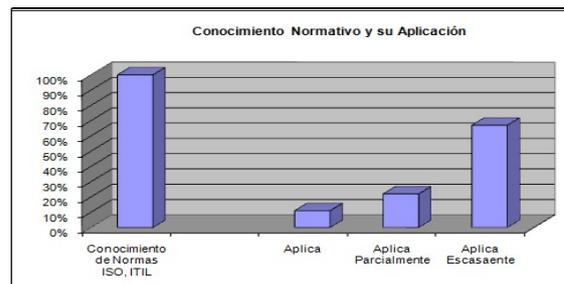
a) Caracterización del Sector TIC:

El 66% de los casos tienen nivel de oficina careciendo de una orgánica formal, con dependencia de niveles jerárquicos superiores. Del total de la muestra el 88% no cuenta con partidas presupuestarias propias.



Casi el 67% de los casos consultados cuenta con una infraestructura de baja intensidad, con poco personal y con presupuestos que dependen de otros niveles de decisión, lo que origina que las inversiones necesarias del sector TIC deben competir con otras necesidades de la organización

b) Nivel de conocimientos de normas ITIL/ISO.



Solamente un caso declara una aplicación suficientemente completa vinculada con la Gestión de la Configuración (entre otras). El 67% de la muestra declara una aplicación escasa y el 22% una aplicación parcial.

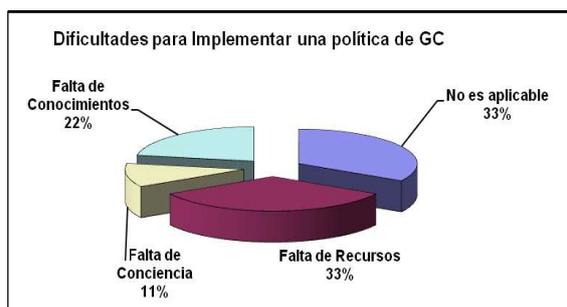
c) Necesidad de Proteger Activos.



Se aprecia un relativo grado de concientización sobre la necesidad de protección de activos. Pero un tercio de la muestra declara no tenerlo o no siempre, lo que guarda correlato con que solo el 43% reconoce la necesidad de invertir preventivamente.



d) Implementación de Políticas de GC.



Finalmente, de las razones esgrimidas sobre la falta de implementación de una adecuada política de GC, los 2/3 de la muestra declara la falta de aplicabilidad por la complejidad y adaptabilidad de las normativas ITIL a su realidad y concurrentemente la falta de recursos.

2. LINEA DE INVESTIGACIÓN Y DESARROLLO.

El proyecto de investigación se centra en el análisis de las normas ITIL e ISO sobre GC y en particular sobre su grado insuficiente de aplicabilidad en las MyPITC. Dada la importancia de gestionar eficientemente los activos de una organización, se pretende realizar un aporte para mejorar la calidad de servicios de la administración pública, desarrollando un modelo aplicable y adaptado al contexto en que desarrollan su actividad.

3. FORMACION DE RECURSOS HUMANOS

Durante el año 2015 y 2016 han recibido el título de postgrado “Especialización en Gestión de la Innovación y Vinculación Tecnológica” en la Universidad Nacional de Entre Ríos (UNER) los integrantes Miguel Fernandez (Dip.Nº 18850), Diana Fernández (Dip.18.801), y Gabriel Antoniutti (Dip.Nº19.950). La integrante Diana Fernández se encuentra cursando la “Maestría en Sistemas de Información” en la UNER, habiendo aprobado los módulos desarrollados durante el 2015 y 2016. Se realizaron varios cursos de capacitación y asistencia a Jornadas.

4. RESULTADOS OBTENIDOS.

Se logró la convocatoria de un número suficientemente representativo para colaborar en el proyecto, destacándose la posterior predisposición a brindar información, incluso en los casos en que se abordaban temas sensibles.

El trabajo se inició con una encuesta con interrogantes considerados básicos[4], para luego continuarse con sucesivos contactos a fin de ampliar, aclarar y profundizar aspectos que permitan arribar a conclusiones lo más cercanas posible a la realidad, y fundamentalmente que permitan validar o no, la hipótesis planteada.

En base al análisis de los datos recopilados, salvo casos puntuales (municipios de ciudades importantes y organizaciones provinciales), se pudo verificar situaciones caracterizadas por:

- Alta precariedad de las instalaciones e infraestructura.
- Insuficiencia de recursos.
- Insuficiente y/o falta de implementación de normas sobre GC.
- Falta de criterios suficientemente sistematizados para implementar recomendaciones de buenas prácticas sobre gestión de la calidad en general.

Como conclusión final se puede afirmar que existe dos razones básicas que imposibilitan la aplicación de las normas sobre gestión de la configuración:

- a) insuficiencia de recursos
- b) alta complejidad de las normas.

En base a dicha conclusión, es posible afirmar que la elaboración de un modelo de configuración que permita definir recomendaciones básicas aplicables a las particulares características de las MyPITIC, significaría un aporte

importante para estas organizaciones, posibilitando con su implementación, una mejor y más eficiente gestión de su infraestructura.

5.RESULTADOS ESPERARADOS.

En la próxima etapa del proyecto se desarrollarán las actividades para lograr:

- Los criterios para elaborar modelos de configuración para las MyPITIC y orientaciones de las actividades que afectan a la configuración hacia el modelo definido.
- Elaborar las recomendaciones relacionadas con los aspectos de la configuración.

6. BIBLIOGRAFIA.

1. ABAD, Luis M.; SANCHEZ, Alejandra P.; GAONA, Juan T; FERNANDEZ, David B.; SANZ, Miguel J.G; ISO/IEC 2000 – Guía completa de aplicación para la gestión de los Servicios de Tecnología de la Información; AENOR (Asociación Española de Normalización y Certificación) Ediciones, año 2009.
2. ABAD GARCIA, María Francisca; *Evaluación de la Calidad de los Sistemas de Información*; Editorial Síntesis; Enero 2005; ASIN B00628F9B0.
3. ARDITA, Julio César. Director de Cybsec S.A. Security System y ex-Hacker. Entrevista personal realizada el día 15 de enero de 2001 en instalaciones de Cybsec S.A. <http://www.cybsec.com>.
4. CASAS ANGUITA, J.R. CAMPOS DONADO J.. *La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos*. Departamento de Planificación y Economía de la Salud. Madrid. España. 2002.
5. ESCALERA IZQUIERDO, Gregorio; GARCIA ROJO, Elena, *Los Sistemas de Información y su Influencia en los*

estándares de calidad; Primera Edición, Editorial Universitas internacional S.L.; Enero 2006; ISBN-13 9788493463809.

6. FARABOLLINI, Gustavo R. *Gobierno Electrónico: una oportunidad para el cambio*, Segundo Congreso de Argentino de Administración Pública, Sociedad y Estado.

<http://www.ag.org.ar/2congreso/Ponencias/Farabollini.pdf>

7. FERNANDEZ SANCHEZ, Carlos Manuel; PIATTINI VELTHUIS, Mario. *Modelo para el Gobierno de las TICs basado en las normas ISO*. AENOR 2012. Impreso en España. ISBN: 978-84- 8143-764-5.

8. ISO/IEC 20000. *Guía de Bolsillo (spanish versión)*. ITSM Library. ISBN 978-90-7721-288-2

9. ITIL Version 3 – SERVICE TRANSITION – Versión Digital. <https://tomjmyth.files.wordpress.com/2015/04/itilv3st-til-v3-service-transition.pdf>

10. ONTI (Oficina Nacional de Tecnologías de la Información) de la Jefatura de Gabinete de Ministros de la Secretaría de Gabinete y Coordinación Administrativa, Presidencia de la Nación Argentina; *Modelo de Política de Seguridad de la Información para Organismos de la Administración Pública*; - Versión 1 – Julio de 2005.

11. PABLO-HEREDERO, Carmen de; LOPEZ HERMOSO, José J.;PIATTINI VELTHUIS, Mario; GARCIA RUBIO, Félix O.; GARCIA RODRIGUEZ, Ignacio; PINO, Francisco; *Calidad de Sistemas de Información*; Segunda versión actualizada; Ediciones Ra-Ma, Enero 2011; ISBN-13 978-8499640709.

12. WALTER, Jorge y PANDO, Diego, “*Planificación Estratégica: Nuevos Desafíos y Enfoques en el Ámbito Público*”, Sociedad Argentina de Análisis Público, Primera Edición, 2014, ISTC A-10-2014-00000002-4.

Aproximación a la Seguridad de las Comunicaciones en Internet de las Cosas

Mg. Jorge Eterovic; Esp. Marcelo Cipriano;

Instituto de Investigación en Ciencia y Tecnología
 Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
 Universidad del Salvador.
 Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{jorge.eterovic; cipriano1.618}@gmail.com

RESUMEN.

Ya entrado el siglo XXI un sinnúmero de dispositivos y objetos almacenan, transmiten y reciben información con escasa o nula intervención de los seres humanos. Las Redes WSN¹ y dispositivos de tipo *RFID*² son ejemplo de ello.

Para el público en general estos sistemas resultan ser “invisibles”. Es decir que se ignora su existencia o se tiene una visión parcial o incompleta de los mismos.

Esta invisibilidad también incluye a las técnicas de protección y seguridad de dichas comunicaciones, transporte y almacenamiento de datos personales, en los sistemas que así lo requieren.

Todos ellos tratan con información que en la mayoría de los casos es de índole personal y por ello surge la imperiosa necesidad que sea tratada de manera segura y confidencial[1,2]. Este es un objetivo a cumplir por empresas y organismos que emplean estas tecnologías.

Este proyecto persigue realizar un estudio y análisis de los protocolos de comunicaciones seguras que podrían ser usados en Internet de las Cosas³. En particular en los aspectos de privacidad y protección de datos personales usando Criptografía Ligera[3].

Palabras Clave:

Internet de las Cosas, Internet of Things Protocolos, Seguridad. RFID.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad Nacional del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad, entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación de la y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto, con una duración de 2 años (2017-2018).

¹ Wireless Sensor Network: Redes Inalámbricas de Sensores.

² Radio Frequency Identification: identificación por radiofrecuencia.

³ Internet of Things: Internet de las Cosas.

1. INTRODUCCIÓN.

Aunque se emplee con más asiduidad día tras día y se tenga la sensación que siempre “ha estado ahí”, la IoT fue dada a conocer al mundo no hace tantos años atrás. Fue en una presentación[4] para la empresa Procter & Gamble (P&G) hace 18 años, más precisamente en 1999.

Kevin Ashton⁴ es miembro fundador del Laboratorio de Investigación Auto-ID Center del MIT⁵ (hoy llamado Auto-ID Labs[5], junto a David Brock, Dr. Daniel Engels, Sanjay Sarma y Sunny Siu y patrocinado por las empresas Procter and Gamble, Gillette, the Uniform Code Council entre otras empresas fabricantes de nivel internacional) presentó las ventajas económicas del uso RFID⁶.

Y de allí al concepto de “Internet de las Cosas” no hay mucho más que un paso. Se trata de la conectividad, mediante Internet, entre objetos para una gran diversidad de objetivos y formas

Este concepto revoluciona al mundo y aún no se ha visto casi nada de lo que vendrá. En el próximo mes de Octubre en Barcelona, España, se realizará la 2da edición del “IoT Solutions World Congress” [6] por ejemplo con usos medicinales, comerciales, científicos, hogareños y hasta militares[7].

La historia se inicia con las conocidas “etiquetas antirrobo” que se adhieren a libros, prendas y demás objetos en librerías y shoppings. Luego aparecieron otros objetos, como las llaves “computadas” de vehículos, los “tags” para abonar peajes y tarjetas para el pago electrónico de pasaje en transporte público (tarjetas SUBE, Monedero, etc.). Pero menos conocidos por su reciente aparición y no tan masiva difusión como son los pasaportes, licencias de conducir, documentos de Identidad y hasta incluso minúsculos chips

subcutáneos⁷, entre otros dispositivos y sistemas a implementar[8].

La consultora internacional McKinsey & Company tiene una visión de la IoT desde un punto de vista más global y no solamente técnico: “Connecting physical objects is creating new business models, improving processes, and can reduce costs and risks”⁸ [9]. Sino que además augura que el impacto de IoT a escala mundial será de u\$s 11,1x10¹¹ (más de once billones de dólares⁹) por año, para el año 2025[10].

Se espera que la IoT ingrese a casi todos los ámbitos de nuestras vidas y cambie la sociedad en la que vivimos. La mayoría de estos dispositivos manejarán información vital y sensible. Y de ahí la imperiosa necesidad de proteger su valiosa carga. Sin embargo no todos estos dispositivos tienen la capacidad para hacerlo por las capacidades limitadas de espacio, energía y recursos. Es por ello que el empleo de la Criptografía Ligera aplicada a la Internet de las Cosas se convierte en una solución [11,12].

Esta investigación se centrará en encontrar la mejor relación costo-beneficio-seguridad para las comunicaciones en Internet de las Cosas mediante el uso de Criptografía Ligera

Incluso las modificaciones de protocolos ya existentes, como el advenimiento de nuevos protocolos específicos exclusivos pensados para aplicaciones de IoT deben somerterse a exhaustivos análisis que permitan dotar de seguridad a esta tecnología. [13-15].

⁴ Tecnólogo y científico británico.

⁵ Massachusetts Institute of Technology: una de las principales universidades de Estados Unidos y del mundo, se enfoca principalmente en investigación, disciplinas científicas y educación tecnológica.

⁶ Radio Frequency Identification: en español identificación por radiofrecuencia.

⁷ Más propio de la Ciencia Ficción que de la realidad. Sin embargo ya es una “vieja” tecnología del año 2004, registrada por la empresa VeryChip que logró la autorización de la FDA, la Administración de Alimentos y Medicamentos de Estados Unidos, para el registro de datos médicos en seres humanos.

⁸ La conexión de objetos físicos está creando nuevos modelos de negocio, mejorando procesos y puede reducir costos y riesgos.

⁹ “11.1 trillions of dollars” según el informe, equivalen a 11.100.000.000.000 de dólares en nuestra forma de expresar cantidades.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

Se realizará un relevamiento, estudio y análisis exhaustivo de los principales protocolos de comunicaciones que podrían ser usados en IoT.

Se realizará un análisis de riesgos para determinar el grado de exposición en los aspectos de privacidad, protección de datos personales y seguridad en las comunicaciones electrónicas.

Se definirán indicadores utilizando las experiencias publicadas en trabajos internacionales para evaluar comportamientos y permitir comparaciones.

Se volcarán los resultados obtenidos en una tabla comparativa sobre el comportamiento de algoritmos usando los protocolos de comunicaciones estudiados.

Finalmente se redactará un informe final con los resultados obtenidos.

3. RESULTADOS OBTENIDOS/ ESPERADOS.

El objetivo de este proyecto es realizar un análisis comparativo, de acuerdo a criterios de aplicabilidad y seguridad, de 3 Algoritmos Criptográficos Livianos para dispositivos RFID de bajo costo.

Se realizará un relevamiento exhaustivo de los principales algoritmos criptográficos ligeros existentes y determinará cuáles se podrían utilizar para dispositivos RFID de bajo costo.

Se definirán indicadores utilizando otras experiencias internacionales para evaluar comportamientos y permitir comparaciones.

Se simulará el funcionamiento de los algoritmos seleccionados y se realizará una tabla comparativa sobre el comportamiento de los algoritmos estudiados.

Finalmente se redactará un informe final y se presentarán en diferentes congresos los resultados obtenidos de esta investigación,

para difusión y conocimiento de la comunidad científica.

4. FORMACIÓN DE RECURSOS HUMANOS.

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas en la Facultad de Ingeniería, el área de la Seguridad Informática, de la Universidad del Salvador.

Dado que este proyecto recién inicia se espera que en breve se sumen a él alumnos de las carreras de Ingeniería en Informática y Licenciatura en Sistemas de Información.

5. BIBLIOGRAFÍA.

[1] Román R., Nájera P., López J. “Los Desafíos De Seguridad En La Internet De Los Objetos” University of Malaga, España. 2010.

[2] Heer, T.; Garcia-Morchon, O.; Hummen, R.; Keoh, S.L.; Kumar, S.S.; Wehrle, K. “Security challenges in the IP-based internet of things”. *Wirel. Pers. Commun.* 61, 527– 542. 2011.

[3] ISO/IEC 29192. Information technology - Security techniques - Lightweight Cryptography. 2012. <https://www.iso.org>.

[4] <http://www.rfidjournal.com/articles/view?4986>. Consultada el 1-3-17.

[5] <https://autoidlabs.org/>. Consultada el 1-3-17.

[6] <http://www.iotsworldcongress.com> Consultada el 1-3-17.

[7] Radio frequency identification ready to deliver, Armed forces communications and electronics association 2005. <http://www.afcea.org>.

[8] <http://www.lanacion.com.ar/1892969-club-tigre-chips-bajo-la-piel-una-tecnologia-de-identificacion-practica-o-invasiva>. Consultada el 1-3-17.

[9] <http://www.mckinsey.com/global-themes/internet-of-things>. Consultada el 1-3-17.

- [10] Manyika, J.; Chui, M.; Bisson, P.; Woetzel, J.; Dobbs, R.; Bughin, J.; The Internet of Things: Mapping the Value Beyond the Hype. Executive Summary. McKinsey Global Institute. 2015.
- [11] Masanobu Katagi; Shiho Moriai, Lightweight Cryptography for the Internet of Things; Sony Corporation; 2016.
- [12] Bhattasali Tapalina. "LICRYPT: Lightweight Cryptography Technique for Securing Smart Objects in Internet of Things Environment". University of Calcutta. 2013.
- [13] Garcia-Morchon, O.; Keoh, S.; Kumar, S.; Hummen, R.; Struik, R. "Security Considerations in the IP-based Internet of Things". IETF Internet Draft draft-garcia-core-security-04; The Internet Engineering Task Force (IETF): Fremont, CA, USA, 2012.
- [14] Cirani S., Ferrari G., Veltri L. "Enforcing Security Mechanisms in the IP-Based Internet of Things: An Algorithmic Overview". Algorithms 2013, 6, 197-226;
- [15] Garcia-Morchon, O.; Keoh, S.; Kumar, S.; Hummen, R.; Struik, R. "Security Considerations in the IP-based Internet of Things". IETF Internet Draft draft-garcia-core-security-04; The Internet Engineering Task Force (IETF): Fremont, CA, USA, 2012.

Arquitectura de Seguridad por Capas en Sistemas Críticos

Oscar Martín Bianchi ^(a,b)
oscarmartinbianchi@gmail.com

Ignacio Martín Gallardo Urbini ^(a,b)
ignaciommgu@gmail.com

German Luis Vila Krause ^(a)
g.vilakrause@gmail.com

Ignacio Arrascaeta ^(a)
ignacioarrascaeta@gmail.com

^(a)CIDESO⁰, DIGID¹- Ejército Argentino

^(b)EST², IESE³ - Ejército Argentino

RESUMEN

El rotundo avance de la tecnología nos presenta continuamente nuevas herramientas y facilidades para el desarrollo de sistemas electrónicos e informáticos, que permiten mejorar y/o complementar los procedimientos existentes en el campo militar. Por este motivo, es esencial identificar los posibles riesgos que estas tecnologías pueden acarrear, para así poder preparar a los sistemas críticos desarrollados en dicho ámbito para posibles contingencias.

La presente investigación plantea la posibilidad de adaptar técnicas de la industria, más específicamente del área de seguridad informática y ciberdefensa, a los sistemas desarrollados por el Centro de Investigación y Desarrollo de Software Operacional (en adelante CIDESO), en el ámbito de los sistemas de propósito crítico^a (tanto en el área de Comando y Control, como en la de apoyo en situación de catástrofe).

Palabras Clave: Seguridad Informática, Ciberdefensa, Comando y Control, Arquitectura, Sistemas Críticos.

CONTEXTO

El Ejército Argentino financia y patrocina la construcción de un sistema de Comando y Control (C2) para sus Grandes Unidades de nivel táctico (Brigadas), así como de sistemas de C2 a nivel unidad de combate, y cuenta con la asignación de presupuesto específico y subsidios asignados por el Programa de Investigación y Desarrollo para la Defensa (PIDDEF) del Ministerio de Defensa de la República Argentina. La finalidad de los sistemas de C2 es dar soporte a los procesos de toma de decisiones que realizan los comandantes y sus equipos de asesores (Estado Mayor), optimizando el flujo de información operativa y decisoria en todos los niveles de la estructura orgánica de las brigadas, integrándose en forma horizontal y vertical.

El CIDESO, posee una amplia experiencia en el desarrollo de aplicaciones militares y sistemas de propósito crítico, tanto para problemas militares operativos – Batalla Virtual (BV), Sistema Integrado Táctico del Ejército Argentino (SITEA) – como para operaciones militares de paz – Simupaz –.

Cómo explotación de esa capacidad, y con la experiencia obtenida tanto en el desarrollo del sistema de C2 del Ejército Argentino (EA), bajo el nombre de proyecto Sistema Integrado Táctico de Comando y Control del Ejército Argentino (SITEA, (BIM) N° 413, PIDDEF

^a Sistemas cuyo fallo atenta contra el cumplimiento de los objetivos previstos para el mismo, y que pone en riesgo vidas humanas o bienes materiales.

¹ CIDESO: Centro de Investigación y Desarrollo de Software

² DIGID: Dirección General de Investigación y Desarrollo

³ EST: Escuela Superior Técnica - Facultad de Ingeniería del Ejército Argentino

⁴ IESE: Instituto de Enseñanza Superior del Ejército - Universidad del Ejército Argentino

037/10), como de BV, se propuso la aplicación del conocimiento logrado para el desarrollo de un sistema que permita la conducción de las operaciones de apoyo de fuego dentro de la red de apoyo de fuego. Así surgió el Sistema de Automático de Tiro de Artillería de Campaña (SATAC).

Ambos sistemas por separado aportaron un gran valor en conocimiento para el desarrollo que se emprendió con el proyecto anteriormente mencionado.

En este contexto, la introducción de la posibilidad de sumar elementos de seguridad resulta absolutamente crucial, dada la criticidad de los proyectos desarrollados.

1. INTRODUCCIÓN

Una de las razones por la que la presente investigación toma especial relevancia es el hecho de que se trata del primer esfuerzo en el diseño de una arquitectura de seguridad para un sistema de propósito crítico por parte del CIDESO, que esté a la altura de las exigencias modernas.

Para esto resultaron claves tanto la pluralidad de conocimientos del equipo conformado por profesionales del ámbito de Sistemas, Defensa y alumnos de distintas universidades de Ingeniería, como el conocimiento del negocio y del entorno del sistema, así como de los posibles factores de riesgo a los que se encontraba expuesto.

Uno de los principales problemas planteados, fue la necesidad de que la solución aportada por la arquitectura a diseñar fuese transparente tanto al medio como a la tecnología a utilizar (1), centrándose más específicamente en la definición de políticas y procedimientos (2).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco del desarrollo del proyecto presentado anteriormente, se pretende lograr un avance significativo, no solo en lo que respecta

a la seguridad intrínseca de los desarrollos llevados a cabo por el CIDESO, sino que además plantee una base de conocimientos y requerimientos mínimos para el desarrollo de los sistemas de ahora en adelante. Por lo tanto, se plantea lograr:

- Incluir dentro de los sistemas de propósito crítico siendo desarrollados por el CIDESO, estrategias de seguridad que atiendan la criticidad de los sistemas.
- Aportar nuevas herramientas que permitan potenciar los atributos de calidad asociados a la seguridad en los sistemas que están siendo desarrollados actualmente en el CIDESO.
- Posibilidad de desarrollar nuevas estrategias de seguridad basándose en el estado del arte de la industria y las experiencias obtenidas por los actores más destacados.

Así mismo, se decidió separar la investigación en dos grandes áreas:

- Seguridad Física: Dada la naturaleza de los sistemas que están siendo desarrollados por el CIDESO, el acceso físico a los equipos resulta una problemática clave a tener en cuenta en cualquier solución propuesta, ya que el mismo puede comprometer seriamente la integridad de los sistemas.
- Seguridad lógica: Esta línea agrupa todas las medidas, procedimientos y decisiones destinados a proteger la información del sistema que impactan directamente sobre la implementación del mismo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Las presentes líneas de investigación tienen por objetivo:

- Facilitar la comprensión de la temática, y el trabajo a realizar, capacitando a los técnicos y artesanos responsables de la implementación mediante tareas de investigación y actualización.

- Obtener una arquitectura de seguridad para el sistema, que sea independiente del medio, y cuya implementación no impacte fuertemente en la programación temporal del proyecto.
- Obtener una arquitectura de seguridad para el sistema, cuya implementación garantice los atributos de seguridad mínimos para un sistema de C2.

Los primeros resultados de la línea de investigación decantaron hacia un diseño de defensa por capas (“Layered Security Approach^b”) (3) (2) (4) (5) dadas las ventajas que ofrece:

- Independencia del medio.
- Independencia de la Tecnología.
- Muy Bajo índice de acoplamiento entre capas.
- Alto índice de transparencia al usuario.

Este diseño implica:

- Definición de Políticas.
- Definición de la Arquitectura.
- Definición de Mecanismos de seguridad.

Entre las políticas determinadas, podemos contar las siguientes:

- Autenticación Multifactor: Combinación del uso de dos o más credenciales distintas, de distinta naturaleza (distinto factor, por ejemplo, una clave alfanumérica y una clave biométrica como la huella digital) (6) (7).
- Encriptado de archivos, disco y dispositivos extraíbles: Encriptado de los medios que intervienen en el funcionamiento del sistema.
- Autenticación para las Comunicaciones: Definición de mecanismos de autenticación para el intercambio de información entre los diferentes usuarios del sistema.
- Cifrado de Comunicaciones: Cifrado de toda la información que se transmite

en el sistema, entre las distintas terminales del mismo.

- Segmentación de los Datos: Segmentar una Base de Datos implica agrupar a los usuarios de la misma según determinados aspectos y características en común, con el fin de acotar el dominio de los datos a los que cada grupo o individuo tiene acceso.
- Seguridad por Control de Contenido: No todos los datos e información que circulan en un sistema tienen el mismo nivel de criticidad, ni requieren que se tomen las mismas medidas de seguridad.

Como objetivo final de la investigación se espera obtener una correcta definición de los mecanismos, así como una separación lógica de las distintas capas de la arquitectura, de modo tal que se pueda implementar cada una de forma independiente de la anterior, pero que a su vez sigan siendo complementarias entre sí (5) (8).

La investigación debería, en última instancia, dar como resultado un prototipo experimental-operacional que atienda a cada una de las políticas suscriptas por la arquitectura, de modo que permita ser evaluado por los organismos correspondientes.

4. FORMACIÓN DE RECURSOS HUMANOS

Una característica distintiva de los sistemas desarrollados por el CIDESO, es la estrecha colaboración con los laboratorios y departamentos de la EST. Esto pone al proyecto en un ámbito privilegiado para la formación de recursos humanos.

Por un lado, el CIDESO tiene amplia experiencia en la formación de recursos humanos en el terreno de la investigación aplicada en sistemas de información de diversa

^bConsiste en la combinación de distintos controles y medidas de seguridad de efecto acumulativo, destinadas a resguardar recursos o información.

índole, incluyendo sistemas de simulación para el adiestramiento, sistemas de información geográfica, sistemas de visualización, sistemas inteligentes, sistemas móviles, sistemas de comunicación de alta complejidad y sistemas de cómputo de alto rendimiento. Por el otro, la EST tiene experiencia en investigación básica en el terreno de la ciencia de la informática, asociado estrechamente al hecho de poseer la carrera de Ingeniería Informática como parte de su oferta académica.

También cabe mencionar que cuenta con expertos informáticos, matemáticos y criptógrafos que dan cuerpo en numerosas ocasiones a las investigaciones de los trabajos finales de carrera que se realizan en los posgrados que brinda la universidad.

Tanto el CIDESO como los laboratorios de la EST, a través del dictado de materias de grado en Ingeniería Informática, aportan recursos humanos a la misma universidad. Es así que investigadores de los laboratorios dan cátedras en la EST y, de manera análoga, alumnos de la escuela aportan sus análisis a los laboratorios a través de trabajos prácticos de laboratorio, prácticas profesionales supervisadas o tesis y tesinas de grado y posgrado.

Para el próximo paso, se pretende continuar con esta interacción fluida entre los centros de investigación y el alumnado, formado profesionales con conocimientos de campo en el terreno de la computación de alto rendimiento y un conocimiento acabado sobre temas criptográficos.

Además, al expandirse el sistema para ser aplicado en cualquier problema que requiera altos niveles de cómputo, se pretende incorporar alumnos y docentes de otras cátedras, de cualquiera de las ingenierías que se dictan en la Facultad.

Así, pues, se formarán recursos humanos de todos los niveles, grado, posgrado o investigadores activos, incorporando más alumnos a los laboratorios y, potencialmente, becarios que se dediquen de modo formal (no

sólo académico) a la profundización de los modelos propuestos.

5. BIBLIOGRAFÍA

1. **Blobel, Bernd and Roger-France, Francis.** *A systematic approach for analysis and design of secure health information systems.* s.l. : International Journal of Medical Informatics, 2001.
2. **Jim, Alves-Foss, Carol, Taylor and Paul, Oman.** *A Multi-layered Approach to Security in High Assurance Systems.* s.l. : University of Idaho .
3. **Baker, Bruce and Scheye, Eric.** *Multi-layered justice and security delivery in post-conflict and fragile states.* s.l. : Conflict, Security & Development, 2007.
4. **Mehmet, Yildiz, Jemal, Abawajy and Tuncay, Ercan.** *A Layered Security Approach for Cloud.* s.l. : 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009.
5. **Shenk, Jerry.** *Layered Security: Why It Works.* s.l. : SANS Institute, 2013.
6. **Abhishek, Kumar, et al.** *A Comprehensive Study on Multifactor Authentication Schemes.* Berlin : Advances in Intelligent Systems and Computing, vol 177, 2013.
7. **on, Guidance.** *Guidance on Multi-factor Authentication.* s.l. : State Services Commission, 2016. 0-478-24466-5.
8. **Gupta, Kapil Kumar, Nath, Baikunth and Kotagiri, Ramamohanarao.** *Layered Approach Using Conditional Random Fields for Intrusion Detection.* Washington : IEEE Transactions on Dependable and Secure Computing, 2010.
9. **Zhiyuan, Shi and Lianfen, Huang.** *Layered security approach in LTE and simulation.* s.l. : 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, 2009., 2009.
10. **Komninos, Nikos, Vergados, Dimitrios D. and Douligeris, Christos.**

Authentication in a layered security approach for mobile ad hoc networks. 2007.

11. **Behl, Akhil.** *Emerging security challenges in cloud computing.* s.l. : World Congress on Information and Communication Technologies, 2011.

12. **Payne, Bryan D., Sailer, Reiner and Cáceres, Ramón.** *A layered approach to simplified access control in virtualized systems.* New York : s.n., 2007.

13. **Banyal, Rohitash Kumar, Jain, Pragma and Jain, Vijendra Kumar.** *Multi-factor Authentication Framework for Cloud Computing.* 2013. 978-0-7695-5155-5.

14. **Multi-factor Authentication.** s.l. : SafeNet, 2013.

Controles y Métricas Asociadas en el Contexto de la Ciberdefensa

Pablo G. Sack, Jorge Etherovic, Jorge S. Ierache
 Facultad de Informática Ciencias de la Comunicación y Técnicas Especiales
 Universidad de Morón
 Cabildo 134 Morón, 5627 2000 int 189
sackpablo@gmail.com, jierache@unimoro.edu.ar

RESUMEN

El presente artículo presenta la línea de investigación aplicada en el contexto de la Ciberdefensa. Se presenta el modelo de un Framework basado en el análisis del estado del arte en materia de controles de seguridad y métricas para asistir a la gestión y diagnóstico de seguridad en el contexto de la Ciberdefensa. Finalmente se comentan los resultados iniciales

Palabras clave: Ciberespacio; Ciberguerra; Ciberdefensa; Seguridad Informática; Métricas; Control de Seguridad

CONTEXTO

El proyecto se inserta en la línea de investigación “Seguridad” de la Facultad de Informática Ciencias de la Comunicación y Técnicas Especiales de la Universidad de Morón, el proyecto se encuentra radicado en el instituto ISIER-UM., y se desarrolla bajo el marco del PID 01-001-16 financiado por UM

1. INTRODUCCIÓN

La constante evolución humana ha llevado a la generación de un nuevo espacio artificial que resulta transversal a los espacios naturales (terrestre, marítimo, aéreo, espacial) en los que la humanidad se desarrolla naturalmente. El nuevo ambiente o espacio denominado ciberespacio está plenamente integrado en las actividades humanas, no reconoce

fronteras físicas ni estados naciones, permite la evolución de las operaciones en términos de interoperabilidad de los sistemas en los distintos ambientes naturales. Desde este nuevo espacio artificial se presenta un campo de batalla, presente en la puerta de cada computadora infectada por un software malicioso (malware) puede ser un elemento de ataque a gran escala a un país/agencia o empresa sin que su dueño esté enterado. Internet no es un lugar seguro, ya que hay personas que buscan delinquir en la red ya sea por diversión, por dinero, por motivos políticos, etc.

Ciberespacio

La palabra ciberespacio surge de la conjunción de la palabra “cibernao” - proveniente del griego que significa “pilotear una nave” y es utilizado comúnmente en el ámbito de las redes -, y espacio dando así la idea de estar piloteando o navegando sobre un mundo virtual. En la actualidad el ciberespacio se le da un significado más amplio al que se lo aglomera en la conjunción de toda la información disponible (digitalmente) junto con el intercambio de la información y las comunidades electrónicas que surgen en base al uso de esa información [1]. El Ciberespacio en el contexto del campo de batalla ha ido creciendo y convirtiéndose en algo más difícil de definir y defender [2]. En una visión particular de ciberespacio por parte de los autores propone como definición

del ciberespacio: al ámbito artificial transversal a los ambientes naturales (terrestre, marítimo, aéreo y espacial) que conforma el espacio virtual de interacción en el que se desarrollan actividades propias de humanos y máquinas relacionadas con la creación, procesamiento, publicación, almacenamiento, modificación y explotación de datos, información y conocimiento digitales, en un contexto distribuido (computación en nube) a través de redes interdependientes e interconectadas globales, públicas, privadas, híbridas, software y firmware de máquinas, cuyo carácter distintivo está dado por el empleo de las tecnologías de información y comunicaciones.

Ciberguerra

La ciberguerra es definida por Richard Clarke como las acciones realizadas por un estado Nación que penetra computadoras o redes de otras naciones con el propósito de causar daño o ruptura de las mismas. Estados Unidos (DoD) Cyber warfare (CyW) — Cualquier acto destinado a obligar a un oponente para cumplir nuestra voluntad nacional, ejecutado contra el software de control de procesos dentro del sistema de un oponente. [3]

Ciberdefensa

La Ciberdefensa es definida por La OTAN como: El desarrollo de la capacidad de prevenir, detectar, defenderse y recuperarse de los ataques cibernéticos [4]. Defensa implica [5] la capacidad de colocarse en el camino de penetración, identificar tal intento, y frustrar a través de la interrupción y suspensión de la tareas. Para este propósito, los sistemas informáticos se utilizan bloquear vías de acceso; limitación de permisos; verificación de identidad; proporcionar cifrado y habilitar

la copia de seguridad y recuperación de desastres. Para Estados Unidos (Comprehensive National Cybersecurity Initiative (CNCI)): La defensa de todo el espectro de amenazas mediante la mejora de las capacidades de contrainteligencia de EEUU y el incremento de la seguridad de las cadenas claves de suministro de información [6].

Controles de Seguridad

En el contexto de la Ciberdefensa para su estudio se propone armar una base sólida y progresiva de controles de seguridad tomando como esqueleto y estructura el documento “The Critical Security Controls for Effective Cyber Defense” [7]: este documento está dividido en veinte controles críticos de seguridad que contienen a su vez un grupo de subcontroles. Dichos subcontroles se categorizan con el fin de implementar controles de manera progresiva y escalonada de la siguiente forma: a) Logros Rápidos: Gran reducción de riesgo, poca inversión financiera y técnica; b) Medidas de Visibilidad y Atribución: Mejoran el proceso, la arquitectura y las capacidades técnicas para monitorear sus redes y sistemas informáticos; c) Mejora de la Configuración de Seguridad de la Información: Estos subcontroles ofrecen reducir el número y magnitud de las vulnerabilidades de seguridad y mejorar las operaciones de los sistemas informáticos en red; d) Subcontroles Avanzados: procedimientos que proveen máxima seguridad pero son difíciles de implementar, por ser más caros o requiere de personal altamente capacitado.

Complementar la base de controles de seguridad con lo planteado en “Strategies to Mitigate Targeted Cyber Intrusions” [8]: El documento está dividido en tres partes y realizado por el Departamento de Defensa de Australia y se compone de la

siguiente manera: a) Mitigation Strategies 2014: contiene una breve introducción y un poster con el resumen de las treinta y cinco estrategias para mitigar ciberataques; b) Mitigation Strategies 2014 Details: Describe cada una de las treinta y cinco estrategias para mitigar ciberataques y sus controles recomendados; c) Information Security Manual 2014 Control: Describe los controles mencionados en el documento anterior. Complementar la base con el documento de “Security and Privacy Controls for Federal Information Systems and Organizations (800-53 Rev.4)” [9] que forma parte de un “ciclo de vida de la seguridad”. Por último se toma como punto de partida y de referencia el documento “Framework for Improving Critical Infrastructure Cybersecurity” de National. [10]. El mismo está compuesto de tres partes principales. a) Núcleo: Está compuesto por un conjunto de actividades de ciberseguridad, resultados deseados y referencias. Se compone de cinco funciones concurrentes y continuas: Identificar, Proteger, Detectar, Responder y Recuperar. A su vez, cada una de estas funciones principales se dividirá en Categorías, Subcategorías y Referencias informativas. b) Niveles de Implementación: Proporciona un contexto sobre cómo una organización ve los riesgos de ciberseguridad y los procesos para gestionar ese riesgo. Se compone de cuatro niveles: Parcial, Riesgo Informado, Repetible y Adaptativo. c) Perfiles: Representa los resultados en base a las necesidades del negocio que una organización ha seleccionado de las categorías y subcategorías del marco de trabajo.

Métricas

Se seleccionaron los siguientes documentos que contienen un gran número de métricas que se aplicarán en la

elaboración del marco de trabajo, siendo los mismos: The CIS Security Metrics [11], “Measurement Frameworks and Metrics for Resilient Networks and Services: Technical report ” [12], “Cyber Resiliency Metrics V 1.0 Rev.1” [13] y “Measurement, Identification and Calculation of Cyber Defense Metrics” [14].

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación se corresponden con las normas y métricas en el marco de la ciberdefensa, se explora el desarrollo de frameworks aplicados al diagnóstico de una unidad. En el marco de las futuras líneas se considera la incorporación de sistemas basados en conocimientos para asistir al diagnóstico en materia de ciberdefensa

3. RESULTADOS OBTENIDOS/ESPERADOS

Modelo de Framework

Sobre la base de controles y el marco de trabajo presentados en la sección anterior, se elaboró el Modelo de Framework [15], [16] que facilita la implementación de controles de manera progresiva y realiza un seguimiento de avance a través de las métricas en cada uno de los controles, ofreciendo documentos de consulta por cada control involucrado y links sugeridos. Para esto utilizamos la división de los controles de seguridad propuestos en Framework for Improving Critical Infrastructure Cybersecurity [10]. Dicho documento plantea una división de los controles de seguridad propuestos en cinco fases. Estas fases aglutinarán controles de seguridad que cumplan con el objetivo de cada fase. Las fases serán: **Identificar** (recursos permitidos de la empresa como Hardware, Software, PCs,

etc.); **Proteger:** (configuraciones de Hardware y software, usuarios, contraseñas, cifrados, etc.); **Detectar:** vulnerabilidades, códigos maliciosos, desactualizaciones, accesos no autorizados, etc.; **Responder:** Aislamiento de dispositivos, administración de Logs, eliminación de dispositivos, etc.; **Recuperar:** tiempo, daño, restauración, generación de backups, etc.

Iteraciones dentro del marco de trabajo

Adicionalmente a la división de los controles de seguridad en cada una de las fases, el marco de trabajo constará de cuatro iteraciones que cruzarán todas las fases del mismo. Estas iteraciones tendrán como objetivo lo descrito en la categorización de los subcontroles comentados en la sección anterior del documento The Critical Security Controls for Effective Cyber Defense [7], logrando así la conformación de una grilla que permitirá ubicar los controles de acuerdo a la fase (Identificar, Proteger, Detectar, Responder, Recuperar) e iteración. Las cuatro iteraciones se clasifican como: **a) Logros Rápidos** (primera iteración); **b) Medidas de Visibilidad y Atribución** (segunda iteración); **Mejora de la Configuración de Seguridad de la Información** (tercera iteración); **d) Subcontroles Avanzados** (cuarta iteración).

Se desarrollo un demostrador del Framework que permitió una evaluación orientada a la primera cuadrícula correspondiente a los controles de seguridad de dos de los documentos previamente seleccionados ([7] y [8]) e incorporando a cada uno de los controles de seguridad dentro del framework propuesto Ejemplo de una interacción se observa en la Fig.1

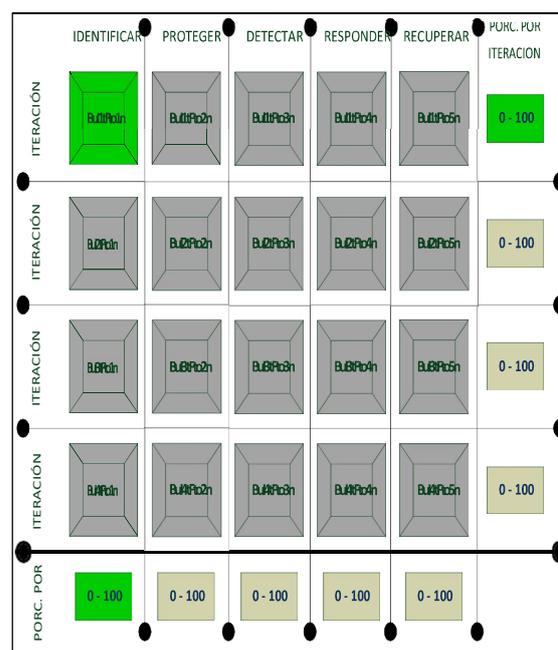


Figura 1. Framework. Iteración 1- Logros Rápidos/ Fase 1 Identificar

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo trajo se conforma de dos investigadores formados un investigador alumno. Se finalizo una tesis de grado en relación a la línea de investigación presentada.

5. BIBLIOGRAFÍA

- [1] David A Umphress, T.C. “El Ciberespacio”: ¿Un Aire y un espacio Nuevo?, Air&Space Power Journal, august 2007.
- [2] Geers, K. A, “Brief Introduction to Cyber Warfare”. Common Defense Quarterly, pp16-17, Spring 2010
- [3] Alford Jr., Lionel D, CYBER WARFARE: PROTECTING MILITARY SYSTEMS., 2000. (Disponible: https://turnitin.com/viewGale.asp?r=53.8117643938158&svr=06&lang=en_us&oid=25740519&key=427186604ab0a08b8bbb71d05c16f6f4)
- [4] C. Czosseck, R. Ottis, K. Ziolkowski, 4TH INTERNATIONAL CONFERENCE ON CYBER CONFLICT (CYCON 2012),

- C. , pp162, Disponible:https://ccdcoe.org/sites/default/files/multimedia/pdf/CyCon_2012_Proceedings_0.pdf)
- [5] Tabansky, L. “Basic Concepts in Cyber Warfare”, Military and Strategic Affairs , vol. 3. N° 1 , pp. 75-92 , May 2011.
- [6] The Comprehensive National Cybersecurity Initiative, (Disponible:<https://www.whitehouse.gov/issues/foreign-policy/cybersecurity/national-initiative>)
- [7] The Critical Security Controls for Effective Cyber Defense. Version 5 (CSC-5) del “Council on CyberSecurity. <https://www.sans.org/media/critical-security-controls/CSC-5.pdf>
- [8] Strategies to Mitigate Targeted Cyber Intrusions” del “Department of Defense – Intelligence and Security of Australian Government.<http://www.asd.gov.au/infosec/mitigationstrategies.htm>
- [9] Security and Privacy Controls for Federal Information Systems and Organizations (800-53 Rev.4) de National Institute of Standards and Technology (NIST).<http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf>
- [10] Framework for Improving Critical Infrastructure Cybersecurity de National Institute of Standards and Technology (NIST).<http://www.nist.gov/cyberframework/upload/cybersecurity-framework-021214-final.pdf>
- [11] The CIS Security Metrics. The Center for Internet Security, november 2010.(Disponible: https://benchmarks.cisecurity.org/tools2/metrics/CIS_Security_Metrics_v1.1.0.pdf)
- [12] Measurement Frameworks and Metrics for Resilient Networks and Services, Technical report ENISA, february 2011.
- [13] Deb Bodeau, Rich Graubart, Len LaPadula, Peter Kertzner, Arnie Rosenthal, Jay Brennan “Cyber Resiliency Metrics V 1.0 Rev.1”., MITRE , april 2012.
- [14] Juan E. Sandoval, Suzanne P. Hassell. “Measurement, Identification And Calculation Of Cyber Defense Metrics”, MILCOM, pp. 2174-2179,october 2010.
- [15] Sack, P , Ierache, J “Controles de seguridad propuesta inicial de un framework en el contexto de la ciberdefensa” XXI CACIC (Junín, 2015) 11p, ISBN: 978-987-3806-05-6
- [16] P. G. Sack and J. S. Ierache, "Initial proposal of a framework in the context of cyberdefense to articulate controls and metrics associated," Computing, Communication and Security (ICCCS), 2015 International Conference on, Pamplermousses, 2015,pp.1-6.doi: 10.1109/CCCS.2015.7374178 URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7374178&isnumber=7374113iee>

Criptografía Post Cuántica

Diego Cordoba³, Miguel Méndez-Garabetti^{1,2}

¹Universidad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería
diego.cordoba@um.edu.ar, miguel.mendez@um.edu.ar

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Universidad de Mendoza, Facultad de Ingeniería, Subsede San Rafael

RESUMEN

La criptografía es la base de cualquier mecanismo de seguridad informática. Se utiliza habitualmente en un login web, en el envío de correos electrónicos, o incluso cuando se produce la sincronización de archivos en la nube, entre otros. Todos los protocolos de comunicación que utilizan SSL/TLS en TCP/IP [1] hacen uso de criptografía asimétrica para autenticación y firma digital. Estos algoritmos se basan en complejos cálculos matemáticos de una sola vía, es decir, son fáciles de realizar, pero muy difíciles de revertir. Si bien los ordenadores actuales no son capaces de romper estos algoritmos en periodos de tiempo aceptables, las computadoras cuánticas, hoy en sus albores de desarrollo, sí podrán hacerlo fácilmente. Es aquí donde surge la necesidad de algoritmos de cifrado que sean resistentes a ataques cuánticos. Estos algoritmos, denominados post cuánticos, si bien están en sus primeras etapas de investigación, resultarán de suma utilidad en un futuro cercano, en el que las técnicas de cifrado asimétrico actuales no puedan brindar la privacidad, autenticación e integridad de los datos en Internet. El presente trabajo de investigación pretende dar luz en este moderno campo de estudio, analizando las implementaciones de software y bibliotecas de programación disponibles en la actualidad.

Palabras clave: criptografía; post cuántico; seguridad informática; criptografía asimétrica

CONTEXTO

El presente trabajo de I+D se desarrolla como proyecto de tesis de posgrado de la Maestría en Teleinformática, Dirección de Posgrado, perteneciente a la Facultad de Ingeniería de la Universidad de Mendoza, (Ciudad, Mendoza). El presente proyecto será presentado como propuesta de tesis.

1. INTRODUCCIÓN

El advenimiento de la computación cuántica, al menos hoy en día en forma teórica más que práctica, resulta una amenaza a los algoritmos de cifrado clásicos debido a que un ordenador cuántico puede procesar mucha más información en menos tiempo, que un ordenador clásico. Esto lleva a pensar que, en el caso de que un atacante provisto de un ordenador cuántico logre interceptar tráfico de red cifrado en forma tradicional, este podría romper dicho cifrado y obtener el contenido de los mensajes originales.

La seguridad de la criptografía moderna se basa en supuestos de complejidad computacional no demostrados, de modo

que en general, las claves se generan mediante operaciones denominadas “de una sola vía”, es decir, operaciones matemáticas muy simples de calcular pero extremadamente difíciles de revertir. Actualmente, por ejemplo, una clave de cifrado puede generarse a partir de la multiplicación de dos números primos grandes. Esta operación es muy simple de calcular con cualquier ordenador clásico. Luego, teniendo el resultado de este producto, un supuesto atacante debería invertir mucho tiempo y procesamiento computacional para encontrar los dos números primos que dieron origen al producto. Hasta hoy, el tiempo necesario es la limitante, ya que las claves de cifrado pueden cambiar periódicamente, por ejemplo, cada hora, de modo que el atacante nunca llegue a tener tiempo suficiente para encontrar la clave y romper el tráfico cifrado que pudo ir acumulando. Si el atacante, por su parte, dispusiera de un ordenador cuántico, encontrar la clave puede ser una tarea muy rápida y permitiría vulnerar casi cualquier tráfico cifrado mediante las técnicas de la criptografía moderna que conocemos.

El algoritmo de Shor[2] fue el primer algoritmo cuántico no trivial que demostró un potencial de crecimiento exponencial de velocidad sobre los algoritmos clásicos. Es un algoritmo cuántico para descomponer en factores un número N en un tiempo $O((\log N)^3)$, y debe su nombre al profesor de matemáticas aplicadas del MIT Peter Shor. Un mensaje cifrado por el algoritmo asimétrico RSA puede ser descifrado descomponiendo en factores la clave pública, que es producto de dos números primos grandes. Los algoritmos clásicos no pueden factorizar la clave pública N de RSA en un tiempo menor a $O((\log N)k)$, para ningún k , por lo que RSA sigue

siendo un algoritmo de cifrado seguro[3][4][5]. Este algoritmo es la base del criptoanálisis cuántico, y ha demostrado ser capaz de romper RSA en un tiempo polinómico [2].

Por otro lado, el algoritmo de Shor, como todos los algoritmos de computación cuántica, da su resultado en forma probabilística con un determinado grado de acierto, por lo que se requieren ejecuciones sucesivas del mismo para aumentar el porcentaje de exactitud del resultado.

En la práctica, un grupo de trabajo de computación cuántica de IBM liderado por Isaac Chuang logró por primera vez factorizar el número 15, el menor número que valida el algoritmo de Shor, en sus factores 3 y 5 mediante una computadora cuántica de 7 qubits. Luego, en marzo del 2016, el mismo Chuang, junto a un grupo de investigadores del MIT, lograron crear un ordenador cuántico de 5 qubits que también pudo correr el algoritmo de Shor para factorizar en número 15 con una confianza del 99%.

Es importante tener en cuenta que, la aparición de las computadoras cuánticas y su posterior comercialización, deja obsoletos muchos de los algoritmos criptográficos que actualmente son seguros; tales como RSA, DSA o ECDSA [3][6][7][8]. No obstante, esto no es razón suficiente para decir que la computación cuántica destruirá la criptografía. Hay muchas clases importantes de criptografía más allá de los algoritmos conocidos:

1. *Criptografía basada en hash (hash-based)*: incluye sistemas criptográficos como las firmas Lamport y el esquema de firmas Merkle [3]. Particularmente estas últimas, creadas en 1970, son de especial importancia como posibles

- sucesores a las firmas digitales RSA y DSA actuales. Perdió importancia en su momento ante las firmas RSA, pero luego tomó mayor renombre al ser un sistema de firmas resistente a los ataques cuánticos.
2. *Criptografía basada en código (code-based)*: aquí el ejemplo clásico es el esquema de firmas McEliece con códigos aleatorios Goppa, y ha sido recomendado por el Post Quantum Cryptography Study Group[3][9] (patrocinado por la unión europea) como candidato para la protección contra ataques cuánticos.
 3. *Criptografía basada en sistemas de ecuaciones multivariable*: aquí se encuentran los sistemas criptográficos como el esquema Rainbow[10] basado en la dificultad de solucionar sistemas de ecuaciones con múltiples variables. Aunque varios sistemas de cifrado basado en ecuaciones multivariables han fracasado, Rainbow podría proporcionar las bases para firmas digitales a prueba de ataques cuánticos.
 4. *Criptografía basada en enrejado (lattice based)*: incluye sistemas criptográficos como el algoritmo de intercambio de claves de aprendizaje con errores, o las firmas de anillos de aprendizaje con errores, los sistemas de cifrado NTRU y GGH, y las firmas digitales NTRU y BLISS. Actualmente se está estudiando la variante Stehle-Steinfeld de NTRU para su estandarización como algoritmo postcuántico válido[11].
 5. *Cifrado simétrico basado en clave secreta de Rijndael*: éste hoy llamado AES, sigue siendo un cifrado seguro contra ataques cuánticos siempre y cuando la longitud de la clave sea mayor, y acorde a las capacidades de cálculo.
- Todos estos algoritmos y técnicas son cuánticamente seguros, y por consecuencia, a ninguno se le ha podido aplicar el algoritmo de Shor.
- En la actualidad casi la totalidad del tráfico web en Internet corre sobre SSL/TLS. El intercambio de datos de autenticación o números de tarjetas de crédito suelen protegerse mediante HTTPS[12]. Debido a esto, surgió la necesidad de comenzar a pensar e implementar algoritmos post cuánticos para proteger este tipo de tráfico en Internet. Con el fin de nuclear el desarrollo de implementaciones prototípicas de algoritmos criptográficos post cuánticos, vio la luz el proyecto Open Quantum Safe (OQS)[13].
- Las ramas de desarrollo de OQS se dividen en dos. Por un lado, el desarrollo de una biblioteca de cifrado post cuántico implementada en lenguaje C, liboqs, y por otro, prototipos de las integraciones de esta biblioteca en implementaciones tradicionales. liboqs es open source licenciado bajo los términos de la licencia del MIT, lo que le permite a cualquier desarrollador disponer del código fuente y utilizar la librería para sus propias implementaciones.
- Una aplicación prototipo importante, propuesta por el proyecto OQS, es un fork de OpenSSL v1.0.2, que hace uso de liboqs para poder cifrar comunicaciones con algoritmos resistente a ataques cuánticos.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación y desarrollo en criptografía post-cuántica son numerosas, principalmente debido a que la criptografía forma parte de prácticamente todos los mecanismos de seguridad de datos en Internet. Entre ellas podemos enumerar las siguientes.

1. Analizar y documentar los algoritmos de cifrado resistentes a ataques cuánticos disponibles en la actualidad.
2. Estudiar las librerías de programación disponibles en la actualidad para escribir aplicaciones criptográficas que utilicen algoritmos de cifrado post-cuánticos. Entre ellas, liboqs, del proyecto OQS, libntruencrypt, del proyecto NTRU Open Source Project, y NFLlib, una librería que implementa fast lattice basada en NTT (Numerical Theoretic Transform).
3. Analizar la viabilidad de proyectos de software que implementen el protocolo SSL, para su uso en servicios en producción. Se pretenden medir parámetros de rendimiento para los algoritmos de cifrado post-cuántico soportados, realizar pruebas de testing y detección de fallos para poder ayudar a la toma de decisiones sobre su implementación en sistemas de producción. En este punto se pretenden estudiar el fork de OpenSSL v1.0.2 del proyecto OQS, y el fork de OpenSSL v1.0.2e de vscrypto, denominado OpenSSL-ringlwe, que implementa de manera práctica el protocolo de intercambio de claves RingLWE.
4. Analizar la estabilidad y viabilidad de StrongSwan para el

establecimiento de túneles VPN con IPSec haciendo uso de algoritmos NTRU.

5. Analizar la estabilidad y viabilidad de CodeCrypt, una implementación abierta GNUPG-like como reemplazo a GNUPG para el cifrado asimétrico.
6. Analizar e implementar sistemas de cifrado para almacenamiento de datos en la nube mediante algoritmos de cifrado resistentes. Aquí se pretenden estudiar EncFS y CryFS[14].

3. RESULTADOS ESPERADOS

Luego de analizar los algoritmos post cuánticos disponibles, las bibliotecas de programación, y las herramientas de software que ya hacen uso de estas bibliotecas, se pretende determinar la viabilidad del uso de estas implementaciones en sistemas en producción.

Como resultados intermedios, también se obtendrán muestras de rendimiento de los algoritmos post cuánticos en el uso de comunicaciones SSL/TLS donde hoy se implementa cifrado asimétrico no resistente a ataques cuánticos, y realizar comparaciones de performance con las herramientas actuales.

Por otro lado, el análisis de las implementaciones de software y bibliotecas de cifrado post cuántico también permitirá detectar errores de funcionamiento, o bugs, y se podrán plantear soluciones de mejora.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de posgrado por parte del Ing. Diego Córdoba, quien es estudiante de la Maestría en

Teleinformática de la Universidad de Mendoza.

5. BIBLIOGRAFÍA

- [1] W. Stallings, Data and computer communication, Prentice Hall, 2006
- [2] P. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Compute, AT&T Research, 1996
- [3] D. J. Bernstein, J. Buchmann, E. Dahme, Post Quantum Cryptograph, Bernstein - Buchmann - Dahme, 2009
- [4] R. Rivest, A. Shamir, L. Adlema, A Method for Obtaining Digital Signatures and Public-Key Cryptosystem, Communications of the AC, 1978
- [5] E. W. Weisstein, RSA-640 Factored, MathWorld Headline New, 200, <http://mathworld.wolfram.com/news/2005-11-08/rsa-640/>
- [6] M. Campagna, L. Che, Quantum Safe Cryptography, An introduction, benefits, enablers and challenge, ETSI White Pape, 2015
- [7] T. Takagi, Post-Quantum Cryptography – 7th International Workshop, PQCrypt, Springe, 2016
- [8] R. A. Perlner, D. A. Cooper, Quantum Resistant Public Key Cryptography: A Surve, National Institute of Standards and Technolog, -
- [9] D. J. Bernstein, T. Lange, C. Peter, Attacking and Defending, the McEliece cryptosyste, PQCrypt, 2008
- [10] P. Oechsli, Making a Faster Cryptanalytical Time-Memory Trade-Of, Advances in Cryptology: Proceedings of CRYPT, 2003
- [11] T. Lange, Initial recommendations of long-term secure post-quantum systems - Horizon 2020 ICT-64562,PQCRYPTO.E, 2015
- [12] EC Council, Network Defense, Fundamentals & Protocol, EC Council Press, 2010
- [13] D. Stebila, M. Mosc, Post-Quantum Key Exchange for the Internet and the Open Quantum Safe Projec, Department of Computing and Software, Mc Master Universit, 2016
- [14] S. Messmer, CryFS: Design and Implementation of a Provably Secure Encrypted Cloud Filesystem, Institute of Theoretical Informatics, Karlsruhe Institute of Technolog, 2015

Diseño y Desarrollo de un Prototipo de Aplicación para la Gestión de las Pericias en Informática Forense Adaptada al Sistema Jurídico Argentino (GEPiF)

Luis Enrique Arellano González, María Elena Darahuge, Carlos Orozco y Agustín Solimine
 Sistemas / Informática / Universidad Argentina John F Kennedy
 Dirección: Bartolomé Mitre 1411 - 1º Piso. C.P. 1037
 Ciudad Autónoma de Buenos Aires
 Teléfono: 05411-5236-1224

e-mails: arellano@kennedy.edu.ar, darahuge@kennedy.edu.ar, corozco@alumnos.kennedy.edu.ar, asolimine@alumnos.kennedy.edu.ar

Resumen

El diseño y desarrollo de un prototipo de aplicación para la gestión de pericias de informática forense adaptado al sistema jurídico argentino, es una inminente necesidad para los peritos informáticos que desempeñan diferentes roles como auxiliares de la justicia (perito oficial, de oficio, de parte, consultores técnicos). El proyecto se centra en la confección de una herramienta que permita guiar a los peritos en informática forense en cada una de las etapas de la elaboración del informe pericial, de manera tal que su aplicación pueda ser realizada de forma homogénea y con un procedimiento básico que luego podrá ser readaptado y retroalimentado acorde a las diversas situaciones que surjan en la gestión de la pericia en informática forense. Se utilizará como marco de referencia y modelo principal la estructura internacionalmente reconocida a mediados del siglo pasado e implementada por el Departamento Scopométrico de la Policía Federal Argentina, ya que permite la normalización, búsqueda, sistematización e intercambio de resultados periciales entre informes procedentes de diferentes áreas del saber, facilitando el apoyo a la decisión judicial obligatoria (sentencia). Asimismo, permite intercambiar datos de manera eficiente, efectiva y eficaz, entre los distintos órganos de investigación judicial, dependientes del poder judicial nacional e internacional.

Palabras clave: informática forense, perito, gestión pericias informático-forenses.

Contexto

El proyecto se encuentra inserto dentro de la investigación y desarrollo de los Proyectos Plurianuales de Investigación (PPI), el cual fue aprobado en octubre de 2015, es financiado por la Universidad Argentina John F Kennedy, su duración es de tres años y con una finalización estimada para el año 2018 y se relaciona con el proyecto Confiabilidad (trazabilidad y responsabilidad) de la cadena de custodia en Informática Forense (CCCIF). El Instituto Argentino de Normalización y Certificación – IRAM- a través del Subcomité Seguridad en Tecnología de la Información colabora con la información relacionada con las normas ISO de informática forense.

1. Introducción

La gestión pericial implica la realización de una serie de etapas que debe efectuar el perito en informática forense en cualquiera de los roles en que se desempeñe, de manera tal que su accionar se encuentre contenido en un entorno sistémico y jurídico que permita efectuar la reconstrucción de las tareas realizadas por parte de cualquier otro perito en informática forense. La estructura del informe pericial que se propone, con sus diferentes etapas, ha sido creada en 1940 por los organismos de seguridad y aplicadas

específicamente al área de las disciplinas criminalísticas [1], [14]. [24]. La informática forense es una disciplina criminalística, por lo tanto requiere la implementación de la estructura del informe pericial adaptado a dicha disciplina. El informe pericial está constituido por los siguientes elementos [7], [8] y [13]: Introducción, Objeto, Elementos ofrecidos, Elementos dubitados, Operaciones realizadas, Conclusiones, Recomendaciones, Anexos. En la actualidad no existe en el país ningún sistema que muestre la cronología y la sucesión de las etapas de un informe pericial de informática forense y que además refleje el entorno jurídico nacional.

Estado actual del conocimiento:

Actualmente los informes periciales de informática forense no se realizan de manera homogénea, ni estructurada, es decir, cada perito, consultor o asesor técnico, confecciona el informe pericial sin tener en cuenta una base sistémica y criminalística [4], [5], [6], [9]. [11] y [12]. La bibliografía actual de referencia respecto de la realización de los informes periciales de informática forense en relación al sistema jurídico argentino, se encuentra descripta a nivel de detalle en los Manuales de Informática Forense I y II [7] y [8]. Las aplicaciones desarrolladas por diferentes países del hemisferio norte (Encase: <https://www.guidancesoftware.com/>) o por los europeos (Deft Association: <http://www.deftlinux.net/>, Digital Forensics Framework, DFF <http://www.digital-forensic.org/>) se limitan solamente a la implementación de herramientas informáticas individuales o conformando un paquete de aplicaciones que únicamente resuelven necesidades de carácter operativo, es decir, que solo se corresponden con el área de operaciones realizadas del informe pericial en informática forense. Las normas internacionales [15], [16], [17], [18], como el RFC [2], relacionados con la informática forense, hacen referencia al sistema jurídico del derecho anglosajón (*common law*) [3], [22] y [23], el cual dista notablemente del sistema jurídico argentino,

que no se basa en la jurisprudencia, sino en la codificación de leyes [26].

Situación problemática: Al efectuar un informe pericial informático forense se deben realizar un conjunto de etapas sistémicas y cronológicas que permitan la reconstrucción de la información por parte de cualquier profesional que así lo requiera. La importancia de registrar las acciones, tareas y documentación de la información en cada etapa es fundamental. Las etapas de la gestión pericial en informática forense son las siguientes:

1. Identificación y registro
2. Autenticación, Duplicación y Resguardo de la prueba
3. Detección, recolección y registro de indicios probatorios
4. Análisis e interpretación de los indicios probatorios. Reconstrucción y / o simulación del incidente
5. Cotejo, correlación y conclusiones, generación de la cadena de custodia e informe pericial informático forense.

Cada una de estas etapas se relaciona con el sistema jurídico argentino y sus códigos de fondo y forma [19] y [21].

La pregunta fundamental se circunscribe a determinar: ¿Cuál es el modelo de informe pericial más adecuado que permita una descripción exhaustiva de las tareas realizadas sobre la prueba indiciaria de informática forense y brinde al mismo tiempo un formulario digital computable que facilite el intercambio de información y resultados desde las diferentes áreas del conocimiento en apoyo a la decisión judicial obligatoria (sentencia)?

Hipótesis: El diseño y desarrollo de un modelo digital de gestión de pericias en informática forense permitirá la homogeneización de la elaboración de los informes periciales de informática forense en el ámbito de la justicia de la República Argentina.

Variable Dependiente: Relación entre el nivel de aplicación de un modelo digitalizado de gestión pericial de informática forense con otras disciplinas criminalísticas. [1], [14], [20] y [24].

Variables Independientes:

2. Relación entre el modelo de gestión pericial propuesto y el entorno judicial involucrado.
3. Grado de integración de la pericia en informática forense con el resto de las disciplinas criminalísticas.

4. Líneas de Investigación, Desarrollo e Innovación

Las principales disciplinas en las que se inscribe el presente anteproyecto son: la criminalística, la informática forense y el derecho. A partir de la criminalística se circunscribe la metodología y los procedimientos del tratamiento de las pruebas o evidencias propias de la informática forense. En relación a la informática forense se consideran las técnicas y métodos específicos requeridos por ésta disciplina al iniciar el proceso de identificación y registro de la prueba y las sucesivas etapas hasta su disposición final. El derecho interviene estableciendo las pautas procesales necesarias para dar validez a los indicios o prueba informático forense desde su etapa inicial de recolección hasta el destino final establecido por el Juez.

El desarrollo de la aplicación **GEPIF** no tiene un precedente en el ámbito nacional. Considerando los altos costos de las aplicaciones de código cerrado comerciales extranjeras existentes en el mercado de la informática forense, tanto a nivel de licencias como en la capacitación, **GEPIF** se realizará utilizando código abierto [10], [26] y licencia GPL, para posteriormente ser difundido en el ámbito nacional. Actualmente el paquete de aplicaciones de desarrollo internacional de código abierto de gestión de pericias como *Autopsy*

(<http://www.sleuthkit.org/autopsy/features.php>) , ofrece un entorno de desarrollo modular, tanto para java como para python (<http://www.sleuthkit.org/sleuthkit/framework.php>). En el proyecto se revisarán los módulos existentes de *Autopsy* como punto de referencia para luego adaptarlos / modificarlos o generar nuevos acorde a las necesidades de los peritos en informática forense, en el marco de la criminalística y del sistema jurídico argentino.

5. Resultados Obtenidos / Esperados

El propósito del presente proyecto se focaliza en: **Determinar** los elementos requeridos para el diseño y desarrollo de un prototipo de aplicación de gestión de pericias de informática forense adecuados al sistema jurídico argentino que permita homogeneizar la elaboración de los informes periciales de informática forense en el ámbito de la República Argentina (en soporte de papel y digital).

Objetivo general: **Diseñar y desarrollar** un prototipo de aplicación para la gestión de las pericias de informática forense adaptado al sistema jurídico argentino.

En relación a los resultados esperados, el diseño de un prototipo de gestión de pericias en informática forense tiene como fin la homogeneización de la labor de los peritos en de informática forense. La aplicación podrá difundirse en el ámbito de la función pública (organismos e instituciones del poder judicial y de las fuerzas de seguridad), como así también en el ámbito académico privado y público de todo el territorio de la nación argentina. El proyecto será presentado en diversos congresos, seminarios y jornadas del área de la criminalística y de la informática forense en diferentes ciudades del país y en las conferencias internacionales de informática forense en la modalidad presencial y virtual. El impacto del proyecto radica principalmente en que en la actualidad no existe una aplicación de gestión de pericias informático forense adaptada al sistema jurídico argentino que

normalice para todo el país la labor de los expertos y/o peritos en informática forense en la gestión de la prueba indiciaria de informática forense, por consiguiente, es una necesidad inminente la formalización e implementación de un modelo común para los peritos en informática forense como auxiliares de la justicia. Al mismo tiempo deja abierta la posibilidad de futuras investigaciones, para comprobar el funcionamiento del modelo propuesto como referente de intercambio de datos entre las distintas disciplinas criminalísticas. A partir del año 2017 se ha concretado la incorporación del equipo de investigación al Subcomité de Seguridad en Tecnología de la Información del Instituto Argentino de Normalización y Certificación – IRAM-. El proyecto GEPIF permitirá generar una infraestructura de procedimientos de gestión de la pericia informático forense que conformarán el inicio de una propuesta de normalización para ser evaluada oportunamente por el IRAM:

6. Formación de Recursos Humanos

El equipo está conformado por un docente de la licenciatura en sistemas de la UAJFK y Director del Curso de Experto en Informática Forense de la UTN – FRA, Especialista en criptografía y seguridad teleinformática, Ingeniero en Sistemas, Abogado, Licenciado en Criminalística y Experto en Informática Forense, actualmente escribiendo la tesis doctoral en filosofía del derecho en la UBA. Una docente Ingeniera en informática, Especialista en Criptografía y Seguridad teleinformática, Magister en dirección estratégica en tecnologías de la información y Experta en Informática Forense, actualmente escribiendo la tesis del doctorado en Psicología Social de la UAJFK. Un abogado que además presentó y aprobó en 2016 el trabajo final de grado de la Licenciatura en Sistemas sobre informática forense y derecho continuando en el equipo de investigación como graduado. Dos alumnos cursantes del cuarto año de la

Licenciatura en sistemas de la UAJFK interesados en el tema de investigación en relación con la seguridad de la información.

7. Bibliografía

1. Albarracín, Roberto. (1969). Manual de Criminalística. Buenos Aires: Editorial Policial.
2. Brezinski, D. Killalea, T. (2002). Request for Comments: 3227 - Category: Best Current Practice. Guidelines for Evidence Collection and Archiving.
3. Broen, Christopher. Computer Evidence Collection & Preservation. (2006). Massachussets, USA: Charles River Media.
4. Darahuge, María Elena; Arellano González, Luis Enrique (2005). “Metodología de la Inspección Ocular en la Informática Forense”, Investigación publicada en el Congreso Virtual Latinoamericano de Psicología Jurídica”.
5. Darahuge, María Elena; Arellano González, Luis Enrique. “La prueba documental informática” (recaudos procesales). Compendio Jurídico ERREIUS-ERREPAR, Septiembre 2010, N° 44.
6. Darahuge, María Elena; Arellano González, Luis Enrique. “La recolección ilegítima de datos (el problema del phishing)”. Compendio Jurídico ERREIUS-ERREPAR, Agosto 2011, N° 54.
7. Darahuge, María Elena; Arellano González, Luis Enrique. (2011). Manual de Informática Forense. Buenos Aires: Errepar.
8. Darahuge, María Elena; Arellano González, Luis Enrique. (2012). Manual de Informática Forense II. Buenos Aires: Errepar.
9. Darahuge, María Elena; Arellano González, Luis Enrique “Prueba documental informática, errores inauditos”. Compendio Jurídico ERREIUS-ERREPAR, Septiembre 2013, N° 77.
10. Darahuge, María Elena; Arellano González, Luis Enrique “Aplicaciones de código abierto en informática forense”. Revista

- Profesional&empresaria, D&G ERREIUS-ERREPAR, Enero 2014, N° 172, Tomo XV.
11. Darahuge, María Elena; Arellano González, Luis Enrique Artículo: “La intervención notarial en la recolección de la prueba informático forense”. Revista Profesional&empresaria, D&G ERREIUS-ERREPAR, Febrero 2014, N° 173, Tomo XV.
 12. Darahuge, María Elena; Arellano González, Luis Enrique “La prueba documental informática (recaudos procesales)”. Revista Profesional&empresaria, D&G ERREIUS-ERREPAR, Abril 2014, N° 175, Tomo XIV.
 13. Darahuge, María Elena; Arellano González, Luis Enrique “La estructura formal del Informe Pericial”. Revista Profesional&empresaria, D&G ERREIUS-ERREPAR, Diciembre 2014, N° 183, Tomo XV.
 14. Guzmán, Carlos. (2000). Manual de Criminalística. Buenos Aires: Ediciones La Rocca.
 15. ISO/IEC 27037:2012 — Information technology — Security techniques — Guidelines for identification, collection, acquisition, and preservation of digital evidence.
 16. ISO/IEC 27041 — Information technology — Security techniques —. Guidelines on assuring suitability and adequacy of incident investigative methods (FDIS).
 17. ISO/IEC 27042 — Information technology — Security techniques — Guidelines for the analysis and interpretation of digital evidence
 18. ISO/IEC 27043 — Information technology — Security techniques — Incident investigation principles and processes (FINAL DRAFT). ISO/IEC 27050 — Information technology — Security techniques —. Electronic discovery
 19. Medina, Graciela; Maíz, Mónica Gabriela. (2008). Derecho procesal civil para peritos. 2da ed, actualizada. Buenos Aires: Rubinzal-Culzoni Editores.
 20. Rosset, Ricardo; Lago, Pedro. (1962). El abc del dactiloscopio. Buenos Aires: Biblioteca policial.
 21. Torres, Sergio Gabriel. (2014). Nulidades en el proceso penal. 6ta.,ed. Buenos Aires: Editorial Ad-hoc.
 22. Smith, Fred Chris; Gurley Bace, Rebecca. (2003). A guide to Forensic Testimony. The art and practice of presenting testimony as an expert technical witness Boston
 23. Vacca, John. (2002). Computer Forensic: Computer Crime Scene Investigation. USA: Ed, Charles River Media, Networking Series.
 24. Zajaczkowski, Raúl Enrique. (2012). Manual de criminalística. Buenos Aires: Ediciones Dosyuna.
 25. Códigos de fondo y forma y leyes relacionadas de la República Argentina
 26. Grupo de Informática Forense: <https://espanol.groups.yahoo.com/neo/groups/informatica-forense/>

Esteganografía Simulada para Análisis de Efectos sobre Portadores Imagen

Mg. Ing. Guillermo Sergio Navas¹, Mg. Ing. Gustavo Rodríguez Medina²

Gabinete de Computación / Fac. de Ingeniería / Univ. Nacional de San Juan ^{1,2}

Av. Libertador Gral. San Martín 1109 (oeste) – San Juan
0264 – 4211700 (Int. 435¹ / 285^{1,2})

snavas@unsj.edu.ar ¹, grodriguez@unsj.edu.ar ²

RESUMEN

La Esteganografía digital hace uso de variadas técnicas para ocultar información en un Portador, con el fin de que ella pase inadvertida para terceros [3]. Generalmente, los objetos portadores utilizados son elementos multimediales (imágenes, audio y video), mientras que el mensaje puede ser de cualquier tipo¹ [2].

La propuesta, expuesta en el presente artículo, se utiliza para implementar una herramienta software que es utilizada para el análisis de efectos esteganográficos sobre portadores imagen tipo BMP [1]. Ella surge como solución a la necesidad de analizar los efectos esteganográficos para poder evaluar los métodos y ponderar características deseables (perceptibilidad visual, detectabilidad y capacidad); teniendo como requisito que los mensajes deben ser de tamaños precisos y adecuados. Pero, disponer de "archivos mensajes" reales, con tamaños específicos para cada ensayo (son cientos) y evaluación, y aplicar esteganografía real, es una tarea excesivamente costosa y complicada.

Por ello, se planteó la *Esteganografía Simulada*, mediante la cual se genera computacionalmente un mensaje ficticio, de tamaño apropiado para cada uno de los ensayos, el mismo se inyecta en el portador, provocándole un efecto

esteganográfico igualmente válido al que si se utilizara un mensaje real, pero el proceso resulta ser mucho más eficiente, rápido y con parámetros controlados por el operador.

La *Esteganografía Simulada*, expuesta en este artículo, es innovadora, no existen otros antecedentes al respecto.

Palabras clave: Esteganografía, Mensajes ficticios, Métodos LSB.

CONTEXTO

En este artículo se expone uno de los temas desarrollados en la tesis "*Exploración de efectos esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos*", de la Maestría en Informática de la Univ. Nacional de la Matanza.

Cabe destacar que ese trabajo de Maestría a servido para la formación de un equipo de investigación en la temática, en la Facultad de Ingeniería de la Universidad Nacional de San Juan, y a partir del cual se han generado otras propuestas de Tesis de Posgrado.

1. INTRODUCCIÓN

La esteganografía trata el proceso de ocultar un objeto software de cualquier tipo (mensaje), en otro, denominado portador, con el fin de ser enviado desde un emisor hacia un receptor de

¹ Se hace referencia a un "archivo mensaje", de tamaño limitado a la capacidad del portador.

manera subrepticia, esto es, se lleva a cabo una transmisión encubierta [2].

Si bien pueden ser utilizados portadores de cualquier tipo, se prefiere los objetos multimediales, aprovechándose la limitación de los sentidos humanos: la vista y el oído no pueden detectar cambios sutiles en presentaciones visuales o de audio [2] [3].

Estos portadores resultan ser los preferidos en aplicaciones esteganográficas como elemento portador, dada la eficacia de las técnicas que aprovechan tales limitaciones humanas. En este tratado se utilizan portadores imagen Bitmap en formato BMP color de 24 bpp.

En la figura 1 se muestra un esquema general de un sistema de esteganografía.

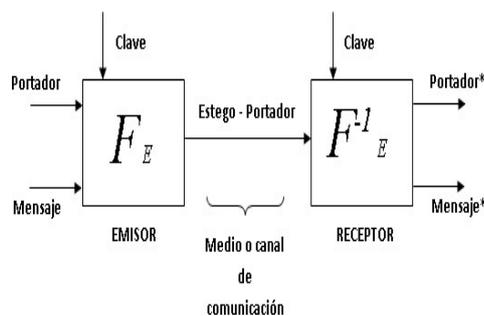


Fig. 1. Esquema de un sistema de Esteganografía.

Exploración de efectos esteganográficos

Existen diversas técnicas o métodos para ocultar un mensaje en un portador imagen. Entre las más difundidas están: la de Sustitución LSB 1 bit y la de inyección al final del portador[6][1]. En el desarrollo del trabajo de Tesis de Maestría [1] se han propuesto las variantes y combinaciones de la técnica de sustitución básica, como por ej. variar la cantidad de bits menos significativos a sustituir en el portador, utilizar diferentes canales de color RGB, alternar el uso entre ellos, usar dispersión en los pixels, etc, además de las posibles combinaciones entre variantes.

El objetivo central del trabajo de Tesis ha sido la exploración de los efectos provocados en portadores BMP por la aplicación de técnicas de sustitución, a fin de encontrar *las mejores* en

cuanto a las características de *perceptibilidad* visual y de *capacidad*, eventualmente también de la característica de *detectabilidad*.

La implementación, y posterior prueba y valoración, de tal cantidad de variantes (más de mil preseleccionadas) puede llegar a ser bastante complicada y laboriosa si se la encara con un desarrollo de esteganografía real. Además, si para el caso que se requieran realizar varias pruebas sobre diversos portadores para determinar la perceptibilidad y la evaluación de la capacidad, esta última característica requiere encontrar un mensaje que se inserte exactamente en el 100% del portador, es decir que "encaje" en su totalidad exacta (byte a byte) en toda la imagen.

Simulación de esteganografía:

Por lo dicho en el punto previo, se ha tenido que pensar en soluciones más eficientes y sencillas para realizar el estudio y exploración. Es así que se ha propuesto y desarrollado, con excelentes resultados, una forma de "provocar" los efectos bajo estudio sobre portadores imagen, pero con niveles de inserción controlados, elegidos por el usuario. De esta forma se puede, por ej., con total facilidad encontrar con certeza el valor de la capacidad que provee cada método o técnica, cualquiera sea la variante utilizada. También generar, en escaso tiempo, efectos de multitud de técnicas sobre portadores testigo elegidos a fin de compararlas, valorarlas y ponderar características de perceptibilidad, incluso analizar detectabilidad. Obviamente en función de las características halladas también se podrá evaluar la bondad de cada método.

Lo anterior se ha logrado implementando por software una suerte de "Simulación de Esteganografía". Esto significa simular en un portador los efectos que puede provocar el ocultamiento de un "mensaje verdadero", pero usando un "mensaje ficticio", que es generado computacionalmente, por tanto factible de elegirlo en un tamaño muy preciso. El operador simplemente elige un método de sustitución a analizar

y el nivel de inserción, y la posterior generación del archivo estego es automática e inmediata.

Esteganografía real Vs. simulada:

La única diferencia entre la esteganografía real y la simulada es que, en la última, el mensaje oculto es ficticio, y por lo tanto no se requiere crearlo, leerlo, cargarlo, procesarlo ni recuperarlo, como sí se debería en esteganografía real. De hecho, para los fines perseguidos, no tiene ninguna relevancia recuperar el mensaje, sólo importa saber que sí se podrá recuperar en caso de una implementación análoga pero real y que los efectos provocados sobre el portador serán los mismos.

La sencillez de la lógica simulada frente a la esteganografía real es notable, dado que el mensaje se genera con lógica simple y en paralelo a la simulación.

La eficiencia de los algoritmos es mucho mayor, ya que no se lee ni se procesan archivos de mensajes, sino que se los genera usando rápidas lógicas booleanas.

La generación de estegoportadores es más simple y requiere muy escaso tiempo, lo cual agiliza notablemente los ensayos para análisis.

No se requiere buscar mensajes de largo adecuado para la prueba de cada portador y cada técnica (son cientos). El nivel de inserción lo elige el operador y el proceso se realiza automáticamente con un "mensaje" de largo exacto. No se requiere la función inversa de recuperación del mensaje, sólo la certeza de que es factible lograrla sin complicaciones fuera de lo común.

Validez de la metodología por simulación:

Desde el punto de vista del diseño y la implementación, las *técnicas de sustitución simulada* y sus lógicas de generación, respetan exactamente los conceptos y reglas de los mismos métodos; es decir que la simulación esteganográfica se realiza de manera idéntica a como si fuera la real, por lo cual, desde ese punto de vista no hay nada que comparar ni comprobar,

simulada o real usan idénticas técnicas. Lo único que cambia es el mensaje. La esteganografía simulada no requiere tener que operar con un mensaje real, evitando procesarlo, conocer su tamaño, la forma de lectura del archivo, acceso al disco, las lógicas de separación por canales, ocultamiento de parámetros de recuperación en el portador, etc.

En cuanto a los efectos provocados sobre el portador, la simulación de esteganografía podría llegar a diferir mínimamente de una aplicación real, la diferencia sería solamente debida a que se utiliza un mensaje ficticio y por lo tanto simulado; por lo cual, para lograr los efectos correctos se debe generar adecuadamente el mensaje, con validez comprobada. Se intenta que los efectos provocados sobre el portador sean lo más parecidos posible a los derivados de una esteganografía con mensaje real. Esto se logra realizando alteraciones, a nivel de bits, de carácter aleatorio con distribución uniforme.

Resulta necesario definir un elemento de medición, denominado *Nivel de Afectación*. Conceptualmente, es la relación porcentual entre la cantidad de Bytes utilizados en el proceso esteganográfico respecto a la cantidad de Bytes que efectivamente sufrieron alguna modificación. La expresión es la siguiente:

$$\text{NivelAfectación} = \frac{\text{BytesUtilizados}}{\text{BytesModificados}} \times 100$$

Se realizaron diversas y numerosas pruebas con LSB 1 real [6], utilizando variados tipos de mensajes u objetos a ocultar, tales como texto ASCII, imágenes BMP, archivos ZIP, ejecutables, etc. Como resultado, el nivel de afectación medio resultó del 50%, con una dispersión media en torno a $\pm 1,5\%$.

La conclusión de las pruebas con esteganografía real es que se alteran los bits LSB de los Bytes del portador en aproximadamente un 50% de los casos. Y alrededor de la mitad de los bytes utilizados no sufre ninguna alteración. Esto provoca finalmente en el portador determinado nivel de ruido.

Respecto a pruebas de validez de mensajes ficticios, se preparó para el análisis una muestra creada con un programa que genera números binarios aleatoriamente (pseudoaleatorios), y que cuenta la cantidad de unos y su promedio general. Se generaron 20 secuencias de 32 números cada una, en la suposición de que se está emulando 20 mensajes distintos de 4 caracteres cada uno. Se hicieron varias corridas del programa obteniéndose varias tablas de números binarios aleatorios.

Como resultado, se obtuvo que los binarios generados tienen una media que tiende al 50% de unos (y ceros) con una desviación bastante escasa, algo menor que la de mensajes reales. Por tanto se prueba que el algoritmo de simulación por generación aleatoria binaria es completamente válido, provocando efectos análogos a los que produce un mensaje real sobre un portador.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación corresponde a la temática *Esteganografía*, la que se enmarca en el área de Seguridad Informática[3]. En este sentido, los autores del presente artículo han elaborado hasta el momento dos tesis de Maestría, como así también la publicación de resultados en diferentes congresos [4][5][6].

Los resultados obtenidos desde el año 2006 a la fecha, aportan estudios y desarrollos innovadores en el área.

3. RESULTADOS OBTENIDOS/ESPERADOS

De las pruebas efectuadas, descriptas anteriormente, se desprende que si se generan los cambios en el bit menos significativo de un portador usando la función de generación aleatoria, los cambios obtenidos son casi iguales a una sustitución LSB 1 bits realizada con un mensaje real [6]. Y en consecuencia se debe obtener en el portador niveles de perturbación general en el canal y perceptibilidad de valores análogos a los conseguidos con mensajes verdaderos o reales.

Varias pruebas numéricas, como las descriptas, y otras mucho más extensas, dieron resultados satisfactorios, que indican que la simulación esteganográfica así realizada arroja resultados correctos, cualquiera sea el método de sustitución.

Cabe destacar que oportunamente se desarrolló un software piloto, al que además de implementar algunas técnicas esteganográficas reales, se le dotó de la capacidad de aplicar esteganografía simulada, para efectuar exploración de efectos. Ese software, cada vez que se hace una simulación, presenta un informe del *Nivel de Afectación* en los bytes del portador utilizado. Si con este software se aplica esteganografía real con la técnica LSB 1 bit, también se obtiene un reporte del *Nivel de afectación* (o de ruido). Por tanto el soft provee medios comparativos para conocer la bondad de la simulación.

En numerosas pruebas realizadas con el software piloto se ha comprobado que en el caso de mensajes ficticios la tendencia del *Nivel de Afectación* es muy marcada hacia el 50%, con una dispersión bastante baja, en general no mayor a $\pm 1\%$, y en caso de mensajes reales es algo dependiente del tipo de mensaje, para un mismo portador, pero también con fuerte tendencia hacia un 50% aunque con algo más de dispersión, en torno a $\pm 1,5\%$.

En cuanto a la evaluación del nivel de perceptibilidad de la simulación, comparado con la esteganografía real, es necesario "*poder ver*" los cambios LSB provocados. Esto se logra también con el software implementado, a través del filtrado de bits (los menos significativos del portador) y ajustes gráficos usando un editor de imágenes. La Fig. 2 muestra un caso, se observa la afectación de píxeles en una imagen totalmente blanca, los píxeles blancos son los no afectados.

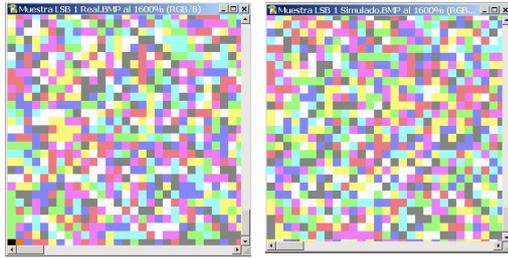


Fig. 2. Afectación de píxeles en un portador, con efecto real y simulado. Zoom 1600.

La validez y bondad de la simulación ha sido comparada numérica y visualmente. Por tanto es completamente válido realizar exploraciones de los efectos esteganográficos utilizando mensajes simulados, con lo cual se simplifica la implementación del software, y los ensayos para análisis resultan ser notablemente más sencillos, rápidos y precisos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación, encabezado por el autor principal, viene trabajando en la temática de Esteganografía desde 2005.

En 2006, defendió el trabajo de Tesis de Maestría titulado “*Exploración de efectos esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos*”.

Con el transcurso del tiempo, el anterior trabajo dio lugar a la propuesta, desarrollo y defensa de otra Tesis de Maestría titulada: “*Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganográficas*”, 2015, posibilitando de esta manera la formación de nuevos recursos humanos en el área.

5. BIBLIOGRAFÍA

- [1] G. Sergio Navas. Exploración de efectos Esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos. Tesis. Argentina. Univ. Nacional de la Matanza – Escuela de Posgrado. 2006.
- [2] Abbas Cheddad, Joan Condell, Kevin Curran, Paul McKeivitt. Digital image steganography: Survey and analysis of current methods

[en línea]. El Sevier (Ed), Journal of Signal Processing, 90(3), 727-752, 2009. [Consultado Junio 2014]

Disponible en:

<http://www.sciencedirect.com/science/article/pii/S0165168409003648>

[3] David Frith, Steganography approaches, options, and implications [en línea]. Ed. Elsevier Ltd., Network Security, 2007. [Consultado marzo 2016.]

Disponible en:

<http://www.sciencedirect.com/science/article/pii/S1353485807700715>

[4] Navas, Sergio; Rodríguez, Gustavo; Eterovic, Jorge; “Aplicación del filtro de Canny a la esteganografía digital”. XVI WICC, ISBN 978-950-34-1084-4. Ushuaia, 2014. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/40706>

[5] Rodríguez, Gustavo; Navas, G. Sergio; Eterovic, Jorge; “Selección óptima de métodos de sustitución en aplicaciones esteganográficas”. XVII WICC, ISBN 978-987-633-134-0. Salta, 2015. Disponible en :

<http://sedici.unlp.edu.ar/handle/10915/45218>

[6] Rodríguez, Gustavo; Navas, G. Sergio; “Esteganografía: Sustitución LSB 1 bit utilizando MatLab”. XVIII WICC. ISBN 978-950-698-377-2. Entre Ríos, 2016. Disponible en : <http://sedici.unlp.edu.ar/handle/10915/52766>.

Estudio Comparativo de Buenas Prácticas para la Recolección de la Evidencia Digital

Nicolás Armilla, Jorge Eterovic, Marisa Panizzi, Luis Torres

Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales,
Universidad de Morón.

Cabildo 134 – CP (1708) – Morón – Prov. de Bs. As. Tel: 5627-2000

nicolasarmilla@hotmail.com; jorge_eterovic@yahoo.com.ar; marisapanizzi@outlook.com;
torreslu@ar.ibm.com

Resumen

Este trabajo de investigación consiste en la comparación de un conjunto de buenas prácticas para la recolección de la evidencia digital. En la República Argentina se evidenciaba la ausencia de un manual, un procedimiento o de un código sobre la recolección de la evidencia digital. Esto conlleva a que una gran cantidad de casos quedasen inconclusos y sin resolución, hasta la creación de la Guía de obtención, preservación y tratamiento de evidencia digital de la Procuración General de la Nación Argentina, en Marzo del año 2016.

Se consideraron buenas prácticas de nivel internacional y nacional, entre ellas la Guía de buenas prácticas para evidencia digital de ACPO (Association of Chief Police Officers), la Guía para las mejores prácticas en el examen forense de tecnología digital de ISFS (Information Security and Forensic Society), la Guía para recolectar y archivar evidencia de RFC 3227 e Investigación en la escena del crimen electrónico del Departamento de Justicia de los Estados Unidos de América y la Guía de obtención, preservación y tratamiento de evidencia digital de la Procuración General de la Nación Argentina.

Se realizará una revisión sistemática para la identificación de los aportes y áreas de

vacancias de las buenas prácticas consideradas.

Palabras clave: Informática forense, perito informático, evidencia digital, buenas prácticas, procedimientos en la informática forense.

Contexto

Este trabajo de investigación se encuentra radicado en el Instituto de Ingeniería de Software Experimental perteneciente a la Facultad de Informática, Ciencias de la Comunicaciones y Técnicas Especiales de la Universidad de Moron. El Instituto articula con las cátedras de tesis de la carrera Licenciatura en Sistemas y con la cátedra de Auditoría y Seguridad de los Sistemas de información.

Introducción

Se han realizado una investigación exploratoria documental respecto a definiciones de informática forense, antecedentes actuales en el ámbito internacional y nacional.

Darahuge et al. definen la Informática Forense como el conjunto multidisciplinario de teorías, técnicas y métodos de análisis, que

brindan soporte conceptual procedimental a la investigación de la prueba indiciaria informática (Darahuge, 2011).

Kovacich define la Informática Forense como la aplicación legal de métodos, protocolos y técnicas para obtener, analizar y preservar evidencia digital relevante a una situación en investigación (Kovacich, 2000).

Gómez define la Informática Forense como aplicación de técnicas científicas y analíticas especializadas a infraestructura tecnológica que permite identificar, preservar, analizar y presentar datos que sean válidos dentro de un proceso legal. O también lo define como una ciencia que busca reproducir científicamente con una metodología estricta de los hechos acontecidos y su correlación para determinar el grado de impacto, y posteriormente establecer en coordinación con otros entes intervinientes, mecanismos tendientes a evitar nuevamente su ocurrencia, que van desde el marco normativo hasta la utilización de mecanismos técnicos (Gómez Luis, 2012).

Listek en el Diario La Nación plantea que el Gobierno quiere normas claras para obtener pruebas digitales en los procesos judiciales (La Nación, 2016).

La Procuración General de la Nación menciona que uno de los temas que puede tocarse desde ahora es el relativo a la evidencia digital, ya que su adecuada obtención, conservación y tratamiento es un elemento clave, entre muchos otros, para asegurar el éxito de las investigaciones (Procuración General de la Nación, 2016).

Luego de revisar los antecedentes en nuestro país, nos planteamos como problema de esta investigación la escasa maduración de procedimientos para la recolección de la

evidencia digital en la informática forense en la República Argentina.

- En los últimos años los peritos informáticos se basaron en procedimientos y buenas prácticas de otros países tales como Canadá, Estados Unidos, Reino Unido y Hong Kong. En la actualidad, a partir del año pasado se cuenta con la nueva resolución de la Procuración General de la Nación (Procuración General de la Nación. Argentina, 2016).

Los procedimientos y buenas prácticas más considerados en nuestro análisis se detallan a continuación:

- Guía de buenas prácticas para evidencia digital. (ACPO, 2012).
- Computación Forense - Parte 2: Mejores Prácticas. (ISFS, 2009).
- Guía para recolectar y archivar evidencia - RFC 3227. (RFC, 2002).
- Investigación en la escena del crimen electrónico. (NIJ, 2001).
- Guía de obtención, preservación y tratamiento de evidencia digital. (Procuración General de la Nación. Argentina, 2016).

Se ha detectado que el inconveniente de basarse en procedimientos y buenas prácticas de otros países presenta diferencia de factores tecnológicos, sociales, culturales y legales respecto a los de nuestro país.

En este estadio de la investigación, se ha realizado una revisión sistemática de los procedimientos y buenas prácticas tanto a nivel internacional como nacional contemplando un conjunto de dimensiones a considerar en el análisis y como resultado se ha obtenido la Tabla 1. Cuadro Comparativo de buenas prácticas y procedimientos a nivel internacional y nacional.

Tabla 1 - Cuadro comparativo de buenas prácticas y procedimientos a nivel internacional y nacional.

	Buenas prácticas y procedimientos a nivel internacional y nacional				
	Internacional				Nacional
	Guía de buenas prácticas para la evidencia digital (ACPO, 2012)	Computación Forense - Parte 2: Mejores Prácticas (ISFS, 2009)	Guía para recolectar y archivar evidencia – RFC3227 (RFC, 2002)	Investigación en la escena del crimen electrónico (NIJ, 2001)	Guía de obtención, preservación y tratamiento de evidencia digital (Procuración General de la Nación, 2016)
Evaluación de Escena		■		■	■
Herramientas y equipamientos		■	■	■	■
Dispositivos electrónicos				■	■
Recolección	■	■	■	■	■
Almacenamiento y transporte	■	■	■	■	■
Análisis	■				■
Reporte	■	■			

Líneas de Investigación,

Desarrollo e Innovación

En el marco de la investigación, según la investigación exploratoria y documental realizada sobre buenas prácticas para la recolección de la evidencia digital en Argentina, se evidencia poca maduración y pruebas realizadas con el novedoso conjunto

de buenas prácticas. También se observa que no existen recomendaciones para poder generar un reporte de forma óptima. Se pretende madurar en el conjunto de buenas prácticas para la recolección de la evidencia digital contemplando otras dimensiones en el procedimiento.

Resultados y Objetivos

El objetivo de este trabajo de investigación consiste en la comparación, propuesta de un procedimiento de un conjunto de buenas prácticas para la recolección de evidencia digital en Argentina, contemplando elementos de los existentes y proponiendo nuevos elementos.

En la actualidad, con la revisión de antecedentes se ha logrado la elaboración del instrumento que permitió analizar cada uno de los procedimientos considerados logrando identificar las áreas de vacancia existentes en cada uno de ellos.

Para la validación del procedimiento a construir, se simulará la recepción de un mail con amenazas, luego será analizado y se sacará una conclusión sobre los hechos.

Formación de Recursos Humanos

El equipo de investigación está compuesto por dos docentes-investigadores, un investigador en formación y un estudiante de la carrera Licenciatura en Sistemas de la Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales, Universidad de Morón.

Dicho trabajo de investigación dará como resultado una tesina de grado de la carrera Licenciatura en Sistemas.

Bibliografía

- Darahuge Maria Elena – Arellano González Luis Enrique, Manual de informática forense 1, Buenos Aires, 2011.
- Darahuge Maria Elena – Arellano González Luis Enrique, Manual de informática forense 2, Buenos Aires, 2012.
- Kovacich Gerald, High-Technology Crime Investigator's Handbook: Working in the

- Global Information Environment, United States of America, 2000.
- Gómez Luis A., La informática forense: una herramienta para combatir la ciberdelincuencia, Buenos Aires, 2012. <http://www.minseg.gob.ar/node/1050>
- Listek Vanesa, El gobierno quiere normas claras para obtener pruebas digitales en los procesos judiciales, Diario La Nación - Argentina, viernes 19 de agosto de 2016. <http://www.lanacion.com.ar/1929918-EL-GOBIERNO-QUIERE-NORMAS-CLARAS-PARA-OBTENER-PRUEBAS-DIGITALES-EN-LOS-PROCESOS-JUDICIALES>
- Procuración General de la Nación, Guía de obtención, preservación y tratamiento de evidencia digital, publicada en la Resolución PGN-0756-2016-001, 31 de marzo de 2016.
- ACPO: Association of Chief Police Officers, Good Practice Guide for Digital Evidence, Reino Unido, 2012.
- ISFS: Information Security and Forensic Society, Computación Forense – Parte 2: Mejores Prácticas, Hong Kong, 2009.
- RFC: Request for Comments, RFC 3227: Guía para recolectar y archivar evidencia, 2002. <http://rfcs.org/pendientes/rfc3227-es.txt>
- NIJ: National Institute of Justice, Electronic Crime Scene Investigation: A Guide for First Responders - Second Edition, Washington, 2001. <https://www.ncjrs.gov/pdffiles1/nij/219941.pdf>
- Piccirilli Dario, La forensia como herramienta en la pericia informática, Buenos Aires, 2013. <http://sistemas.unla.edu.ar/sistemas/redisla/ReLAIS/relais-v1-n6-237-240.pdf>
- Piccirilli Dario. PROTOCOLOS A APLICAR EN LA FORENSIA INFORMÁTICA EN EL MARCO DE LAS NUEVAS TECNOLOGÍAS (PERICIA – FORENSIA y CIBERCRIMEN), La Plata – Prov. Buenos Aires, 2015.
- ENFSI: European Network of Forensic Science Institutes, GUIDELINES FOR BEST PRACTICE IN THE FORENSIC EXAMINATION OF DIGITAL TECHNOLOGY, Europa, 2009. <https://pdf.yt/d/D2rLh6ku8yFjUbt3>
- Acurio Del Pino Santiago, Manual de Manejo de Evidencias Digitales y Entornos Informáticos. Versión 2.0, Ecuador, 2009. http://www.oas.org/juridico/english/cyb_pan_manual

Herramientas de Criptoanálisis

Castro Lechtaler, Antonio^{1,2}; Cipriano, Marcelo¹; García, Edith¹,
Liporace, Julio¹; Maiorano, Ariel¹; Malvacio, Eduardo¹; Tapia, Néstor¹;

¹Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad de Ingeniería. Instituto Universitario del Ejército.

² CISTIC/FCE - Universidad de Buenos Aires.

acastro@est.iue.edu.ar , marcelocipriano@est.iue.edu.ar,
{edithxgarcia; jcliporace; maiorano; edumalvacio; tapianestor87}@gmail.com

RESUMEN.

Este proyecto persigue el estudio, análisis, desarrollo e implementación de técnicas o métodos criptográficos para ser aplicados a determinados generadores de secuencias pseudoaleatorias tipo *Stream Ciphers*¹, en particular a aquellos algoritmos que involucran *LFSR's*², *NLFSR's*³, *CCG*⁴ y *CA*⁵.

Se orientará el estudio a los métodos *Criptoanálisis Diferencial*[1-2], *Lineal*[3], *Algebraico*, *Guess-and-Determine*[4] y una de las últimas técnicas criptológicas denominada *Cube Attack*⁶ [5].

Desarrollar un conjunto de herramientas que permitan analizar algoritmos de cifrado, generadores de secuencias pseudoaleatoria, primitivas criptológicas, protocolos de seguridad de la información, claves secretas de distintos criptosistemas.

Palabras Clave:

Criptología, Criptoanálisis. Stream Ciphers.

¹ Stream cipher: generadores pseudoaleatorios conocidos también como generadores en flujo o cadena.

² Linear Feedback Shift Registers: registros de desplazamiento realimentados linealmente.

³ Non Linear Feedback Shift Registers: registros de desplazamiento realimentados no linealmente.

⁴ Clock Controlled Generators: generadores controlados por reloj.

⁵ Cellular Automata: autómatas celulares.

⁶ Presentado en el congreso EuroCrypt del año 2009 por sus autores: Itai Dinur y Adi Shamir.

CONTEXTO.

El *Grupo de Investigación en Criptología y Seguridad Informática (GICSI)* pertenece al *Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática (CriptoLab)* pertenece a los *Laboratorios de Informática (InforLabs)* de la *Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)*, dependiente de la *Facultad del Ejército, Universidad Nacional de la Defensa (UNDEF)*. El mismo se enmarca en el área de la carrera de grado de *Ingeniería en Informática* y del posgrado en *Criptografía y Seguridad Teleinformática* que se dictan en esta institución.

1. INTRODUCCIÓN.

En la actualidad, el desarrollo de las comunicaciones electrónicas, el uso masivo y generalizado de las computadoras, la transmisión y almacenamiento de grandes flujos de información, hace necesario tomar una serie de medidas para poder protegerlos manteniendo su confidencialidad, autenticidad e integridad.

Es entonces cuando la *Criptología* pasa a ser una exigencia, una necesidad real, donde la falta de protección de los datos privados pasa a ser una amenaza latente.

Los sistemas criptológicos o las primitivas criptográficas creadas para sortear diferentes amenazas deben ser cuidado-

sa y eficientemente desarrolladas y evaluadas. Al momento de realizar el diseño de un criptosistema se deben tener en cuenta todos los ataques que éste puede sufrir. Cada filosofía de diseño que se aplica está respondiendo a un hipotético procedimiento de criptoanálisis y así demostrar su resistencia a él.

Hoy en día no se puede hablar de una única modalidad general de criptoanálisis. Cada algoritmo, cada primitiva, cada protocolo debe ser atacado mediante una técnica adecuada a su estructura.

El criptoanálisis tiene un impacto significativo en el mundo real, puesto que los algoritmos criptológicos, los protocolos y también los tamaños de las claves entre otros, son seleccionados basándose en el estado del arte del criptoanálisis.

En los últimos años los métodos para el diseño de algoritmos seguros han tenido un gran avance e impulso a nivel mundial. Basta recordar, entre otros, los llamados en 1997 del *NIST*⁷ para escoger un nuevo algoritmo como estándar de cifrado llamado AES [6]. El concurso europeo *e-Stream* en 2004, organizado por el *E-CRYPT* [7] del cual superaron todas las pruebas y ataques, 7 algoritmos: 4 para software y 3 para hardware. Y el concurso aún en proceso *CAESAR*⁸ el cual se espera que este año 2017 emita el ganador o un portfolio de los algoritmos que lleguen al final de las etapas del certamen [8].

Aunque bienvenido, este renacimiento mundial por la búsqueda de nuevos algoritmos por sí sólo resulta insuficiente a la hora de establecer parámetros criptográficos seguros.

El objetivo fundamental del criptoanálisis es hallar las vulnerabilidades en uno o varios aspectos de la seguridad de los algoritmos criptológicos, implícita o explícitamente.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

Se ha dado en planificar este proyecto de investigación siguiendo 6 etapas:

1. Estado del Arte del Criptoanálisis de los Stream Ciphers. Mediante el estudio de bibliografía actualizada y oportunamente solicitada, así como la asistencia a Cursos, Congresos y Workshops específicos del área, se profundizará en el estado del arte del *Criptoanálisis de los Stream Ciphers* y los nuevos ataques que se han desarrollado en la comunidad criptológica mundial.
2. Estudio, análisis y selección de los generadores de secuencias cifrantes. A través del estudio de las distintas plataformas y entornos se seleccionarán algoritmos para aplicar las prácticas y métodos criptográficos.
3. Relevamiento de los métodos criptográficos que se analizarán. Ataque por Fuerza Bruta, Criptoanálisis Lineal, Criptoanálisis Diferencial, Ataque por Correlación, Cube Attack, Ataque Algebraico.
4. Estudio de técnicas criptográficas. Determinar el o los métodos adecuados a la estructura del algoritmo estudiado.
5. Implementación de los métodos de criptoanálisis. Desarrollar e implementar las técnicas de criptoanálisis aplicadas a determinados generadores.
6. Análisis de los resultados obtenidos. Evaluar los resultados obtenidos para poder establecer el grado de fortaleza del algoritmo elegido.

⁷ Institución de Estados Unidos, llamada Instituto de Normas y Estandarización (National Institute of Standards and Technology) por sus siglas en inglés.

⁸ CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness.

3. RESULTADOS OBTENIDOS / ESPERADOS.

Como resultado de esta investigación se propone realizar el estudio y análisis para el desarrollo de técnicas y/o herramientas criptoanalíticas que posibiliten la realización del diseño de aplicaciones criptográficas, como así también su evaluación, determinar sus vulnerabilidades o si es posible, quebrarlas.

Los alcances del criptoanálisis podrán ser:

- a- Obtención de la/s clave/s del cifrado.
- b- Hallar patrones estadísticos en la salida del sistema estudiado.
- c- Desarrollar nuevas técnicas criptoanalíticas de acuerdo a las propiedades del sistema estudiado.
- d- Analizar el algoritmo de generación de la/s clave/s y estudiar su vulnerabilidad.

Para ello se perseguirán los objetivos particulares:

- 1- Estudio y análisis de técnicas criptoanalíticas.
- 2- Diseño y desarrollo de herramientas de evaluación, ataque o quiebre de aplicaciones criptográficas.
- 3- Pruebas y testeos de las herramientas desarrolladas sobre algoritmos específicos.

4. FORMACIÓN DE RECURSOS HUMANOS.

Los docentes investigadores de este proyecto se encuentran dictando las asignaturas *Matemática Discreta*, *Paradigmas de Programación I, II y Criptografía y Seguridad Teleinformática*. Desde allí se invita a los alumnos a participar en los proyectos de investigación que se llevan adelante. Es por ello que los alumnos *LEIRAS, F. MIGLIARDI*

A., MONTANARO, L. ROMERO, E. y UVIEDO, G. han demostrado su interés y se han sumado en calidad de colaboradores.

El *Cap. Pérez, P.* integra el equipo de investigación desde el año 2015 y se espera que este año realice su Proyecto Final de Carrera en un tema afín con este proyecto de investigación.

Atendiendo a la responsabilidad ética y social que compete a la actividad científica y tecnológica, el Grupo Integrante de este Proyecto de Investigación, ya sea durante su ejecución o por la aplicación de los resultados obtenidos, desea expresar su compromiso a no realizar cualquier actividad personal o colectiva que pudiera afectar los derechos humanos, o ser causa de un eventual daño al medio ambiente, a los animales y/o a las generaciones futuras.

5. BIBLIOGRAFÍA

[1] Ding C.; *The differential cryptanalysis and design of natural stream ciphers*. In: Anderson R. (eds.) *Fast Software Encryption. FSE 1993. Lecture Notes in Computer Science*, vol. 809. Springer Berlin, Heidelberg.

[2] Wu H., Preneel B. *Differential Cryptanalysis of the Stream Ciphers Py, Py6 and Pypy*. In: Naor M. (eds.) *Advances in Cryptology. EUROCRYPT 2007. Lecture Notes in Computer Science*, vol. 4515. Springer Berlin, Heidelberg. 2007.

[3] Muller F., Peyrin T. *Linear Cryptanalysis of the TSC Family of Stream Ciphers*. In: Roy B. (eds.) *Advances in Cryptology - ASIACRYPT 2007. Lecture Notes in Computer Science*, vol. 3788. Springer, Berlin, Heidelberg. 2005.

[4] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; IEEE Transactions on

Information Theory. Vol. 55 Ed.7º, 2009.

[5] Dinur I., Shamir A. *Cube Attacks on Tweakable Black Box Polynomials*. Advances in Cryptology - EUROCRYPT 2009. Lecture Notes in Computer Science, vol 5479. Springer, Berlin, Heidelberg. 2009.

[6] Daemen, J.; Rijmen, V.; *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer. New York. 2002.

[7] <http://www.ecrypt.eu.org/stream/>
Consultada el 10-3-17.

[8] <https://competitions.cr.yt.to/caesar.html>.
Consultada el 10-3-17.

Introducción al Estudio de Algoritmos Criptográficos Livianos

Mg. Jorge Eterovic; Esp. Marcelo Cipriano;

Instituto de Investigación en Ciencia y Tecnología
Dirección de Investigación Vicerrectorado de Investigación y Desarrollo.
Universidad del Salvador.
Lavalle 1854 – C1051AAB -Ciudad Autónoma de Buenos Aires - Argentina

{jorge.eterovic; cipriano1.618}@gmail.com

RESUMEN

La llamada Internet de las Cosas¹ trata acerca de la conectividad, usando la red Internet, que se realizará entre objetos de diversa naturaleza (tanto en hardware como en software). Los cuales a su vez podrán interconectarse de manera variada y persiguiendo diferentes finalidades. Extendiendo los límites del concepto MtoM² (o también M2M).

Las Redes WSN³ y dispositivos de tipo RFID⁴ de manera invisible se suman al ecosistema en el que los seres humanos conviven. Conforman el andamiaje de la IoT, que promete un cambio de paradigma para la humanidad.

Sólo a modo de ejemplo ya existen zapatillas con sensores GPS y conexión Bluetooth, que reciben órdenes de un teléfono móvil conectado a Google Maps[1] pensadas para ayudar a personas invidentes para guiarlos.

Esta nueva era que se vislumbra con el advenimiento de la IoT conlleva consigo un enorme desafío: proteger la información que procesan los dispositivos, que se mueve por las redes y es almacenada en equipos y reservóeos. Muchos de ellos, tal vez, no

cuenten con los mecanismos de seguridad adecuados.

Volviendo al ejemplo pero aportando un nuevo punto de vista: ¿Qué consecuencias le podrían acarrear al usuario de las zapatillas que terceras partes tengan acceso a la información de su posición en tiempo real, al recorrido o a su destino? ¿Qué podría ocurrir si se hiciesen cambios no autorizados de los mismos?

Éstas y muchas otras preguntas surgen al realizar un análisis del tipo de información que estos dispositivos procesan. Queda expuesto el impacto desfavorable sobre los usuarios si su información no estuviese protegida por técnicas de confidencialidad, entre otras.

Este proyecto persigue realizar un estudio y análisis de algoritmos criptográficos que podrían ser ejecutados en dispositivos con limitados recursos de hardware y software haciendo uso de Criptografía Ligera[2].

Palabras Clave:

Criptografía Ligera, RFID, Internet de las Cosas, Internet of Things.

CONTEXTO

El Vicerrectorado de Investigación y Desarrollo (VRID), perteneciente a la Universidad Nacional del Salvador (USAL), dicta las políticas referidas a la investigación, concibiéndola como un servicio a la comunidad, entendiendo que los nuevos conocimientos son la base de los cambios sociales y productivos. Con el impulso de las propias

¹ Internet of Things: Internet de las Cosas.

² Machine to Machine: máquina a máquina. Se refiere a la comunicación para el intercambio de información entre dos dispositivos distantes o remotos.

³ Wireless Sensor Network: Redes Inalámbricas de Sensores.

⁴ Radio Frequency Identification: identificación por radiofrecuencia.

Unidades Académicas se han venido desarrollando acciones conducentes a concretar proyectos de investigación uni/multidisciplinarios, asociándolos a la docencia de grado y postgrado y vinculando este accionar, para potenciarlo, con otras instituciones académicas del ámbito nacional e internacional.

La Dirección de Investigación, dependiente del VRID, brinda soporte a las distintas Unidades de Investigación de la y a sus investigadores para el desarrollo de Proyectos y Programas de Investigación, nacionales e internacionales, como así también, apoyo y orientación de recursos para la investigación.

A ella pertenece el Instituto de Investigación en Ciencia y Tecnología (RR 576/12) en el cual se enmarca este proyecto, con una duración de 2 años (2017-2018).

1. INTRODUCCIÓN

El cambio social y cultural que ofrece a la humanidad la llamada IoT promete influir sobre algunos aspectos de la sociedad[3]:

- Cuidados médicos.
- Manufactura de productos.
- Uso de la energía.
- Infraestructura urbana.
- Seguridad
- Extracción de recursos naturales.
- Agricultura
- Ventas
- Vehículos

Sin embargo este cambio está sustentado en dispositivos que por sus características físicas están limitados en el uso de determinados recursos, como son, entre otros:

- Espacio
- Consumo de energía
- Memoria
- Capacidad de cómputo

En el último Ericsson Mobility Report del año 2015, el gigante de la telefonía prevee que 28.000.000.000 de teléfonos estarán conectados para el año 2021, más de la mitad de ellos con capacidades de IoT y M2M[4].

Dado que la telefonía móvil incrementará su población actual, se espera que esta demanda de conectividad en aumento sea satisfecha con el advenimiento de la nueva tecnología 5G⁵.

Estos dispositivos móviles intercambiarán información con objetos de la vida cotidiana: desde zapatillas con GPS y Bluetooth (como las del ejemplo anterior), heladeras que hacen la lista de los alimentos que faltan y cajones que se abren/cierran automáticamente[5] esta lista sigue incrementándose a diario.

Estos aparatos y los que vendrán tienen en común que la comunicación que establezcan entre ellos debe estar protegida. Volviendo a las zapatillas, es de esperar que su potencia de cómputo sea muy limitada (por el espacio, por ejemplo) y es válido preguntar ¿qué tanto puede hacer para securizar la información?

La Criptografía Ligera o Liviana estudia algoritmos que por sus propiedades matemáticas pueden ejecutarse en plataformas o dispositivos de recursos limitados, como lo son los que conforman IoT.

Algunos algoritmos son diseñados para ser eficientes en un entorno determinado: hardware o software. Otros en cambio pueden serlo en cualquiera de los dos entornos.

Todas las aplicaciones de la Criptografía tradicional encuentran su par en la Criptografía Liviana. Por ejemplo Block Ciphers⁶ y Stream Ciphers⁷. También existen algoritmos de Clave Pública⁸.

A modo de ejemplo cabe mencionar que recientemente la agencia[6] gubernamental

⁵ 5G: es la llamada 5ta generación de Tecnologías de Telefonía Móvil. Su antecesora, la tecnología 4G aún no del todo difundida en nuestro país.

⁶ Algoritmo de Cifrado por Bloques: algoritmos que dividen el mensaje a cifrar en bloques de n bits y luego proceden al cifrado del bloque.

⁷ Algoritmo de Cifrado en Cadena o Flujo: algoritmos que generan largas secuencias pseudoaleatorias de bits, los cuales uno a uno pueden ser operados con cada bit del mensaje a cifrar.

⁸ Algoritmos que utiliza 2 claves, una de ellas es pública y sirve para cifrar el mensaje. La otra permanece secreta y se usa para descifrar el mensaje. También son llamados Algoritmos Asimétricos, por el uso que se hace de sus claves.

NSA⁹, ha dado a conocer para uso público, dos algoritmos de cifrado por bloques, llamados SIMON y SPECK[7] que por sus características se enmarcan en Criptografía Liviana y están orientados a hardware y software, respectivamente.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Simon y Speck, como así también muchos otros algoritmos deben demostrar su robustez frente a ataques criptoanalíticos¹⁰. Profundizar el estudio de sus propiedades criptológicas, matemáticas, hallar vulnerabilidades, debilidades y la búsqueda de posibles claves débiles son objetivos que se propone esta línea de investigación.

Para ello se realizará un relevamiento, estudio y análisis exhaustivo de los principales algoritmos, que podrían ser usados en IoT, poniendo énfasis en los stream ciphers. Los que por sus características podrían ser empleados por la mayoría de los dispositivos de tipo RFID.

Se definirán indicadores utilizando las experiencias publicadas en trabajos internacionales para evaluar comportamientos y permitir comparaciones entre algoritmos, si ello es posible.

Se volcarán los resultados obtenidos en una tabla comparativa sobre el comportamiento de algoritmos.

Finalmente se redactará un informe final con los resultados obtenidos.

3. RESULTADOS OBTENIDOS/ ESPERADOS

El objetivo de este proyecto es abordar y profundizar en el conocimiento de las

propiedades criptológicas y de seguridad de Algoritmos Criptográficos Livianos que puedan emplearse en Internet de las Cosas[8] u otros dispositivos semejantes, que así lo requieran por sus limitaciones.

Se realizará un relevamiento exhaustivo de los principales algoritmos criptográficos ligeros existentes y determinará cuáles se podrían implementar en esos dispositivos.

Se definirán indicadores utilizando otras experiencias internacionales para evaluar comportamientos y permitir comparaciones.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas en la Facultad de Ingeniería, el área de la Seguridad Informática, de la Universidad del Salvador.

Dado que este proyecto recién inicia se espera que en breve se sumen a él alumnos de las carreras de Ingeniería en Informática y Licenciatura en Sistemas de Información, que se dictan en la Facultad de Ingeniería.

5. BIBLIOGRAFÍA.

- [1] <http://www.lanacion.com.ar/1753934-las-zapatillas-con-gps-dan-un-primer-paso-buscando-nuevos-mercados>. Consultada el 1-3-2017.
- [2] ISO/IEC 29192. Information technology - Security techniques - Lightweight Cryptography. 2012. <https://www.iso.org>.
- [3] Manyika, J.; Chui, M.; Bughin, J.; Dobbs, R.; Bisson, P.; Marrs, A. Disruptive technologies: Advances that will transform life, business, and the global economy. McKinsey Global Institute. 2013.
- [4] https://www.ericsson.com/mx/news/2015-11-17-emr-es_254740126_c. Consultada el 1-3-2017.
- [5] http://tn.com.ar/tecnof5/ces-2016-las-heladeras-del-futuro-conectadas-y-con-multiples-sensores_647274

⁹ National Security Agency: Agencia de Seguridad Nacional. Organismo gubernamental de Estados Unidos.

¹⁰ Criptoanálisis: parte de la Criptología que se encarga de analizar, estudiar y desarrollar ataques para el descubrimiento de los mensajes cifrados o las claves que fueron empleadas.

[6] <http://www.nsa.gov/>. Consultada el 1-3-2017.

[7] <http://eprint.iacr.org/2013/404.pdf>

[8] Masanobu Katagi; Shiho Moriai, Lightweight Cryptography for the Internet of Things; Sony Corporation; 2016.

Longitud de la Clave RSA vs Poder Computacional

Aristides Dasso

SEG / Departamento de Informática / Facultad de Ciencias Físico-Matemáticas y Naturales
 / Universidad Nacional de San Luis
 Ejército de los Andes 950, D5700HHW San Luis, Argentina
 +54 (0) 266 4520300, ext. 2126
 arisdas@unsl.edu.ar

Resumen

Millions Instructions Per Second (MIPS) es una, aunque no la única, de las métricas tradicionales de performance de los procesadores. Por otro lado Bit Length (BL) puede ser considerado como una métrica para medir la fortaleza de un método de encriptación asimétrico.

Dentro del contexto de desarrollo de sistemas y métodos de seguridad, esta investigación tiene como objetivo el concretar un análisis integral de las fortalezas y debilidades de métodos de encriptación asimétricos que permita predecir el nivel de seguridad que dichos métodos presentan hacia el futuro, considerando la longitud de la clave con relación al poder computacional existente.

Proponemos estudiar la seguridad, en un lapso de tiempo dado, de un método de encriptación basado en factorización, tal como RSA [19], estableciendo una relación entre el poder computacional necesario para quebrar una clave y el BL usado en la encriptación. Esta relación permitirá una estimación del lapso de tiempo en que una encriptación con un BL dado será segura frente a posibles ataques.

Palabras clave: Seguridad de Sistemas Informáticos. Evaluación de la Seguridad de Sistemas Informáticos. Métodos asimétricos de encriptación. Factorización. Poder computacional. Esfuerzo computacional de factorización.

Contexto

Este trabajo de investigación se viene llevando a cabo dentro del SEG (Software Engineering Group), en el ámbito de la

Universidad Nacional de San Luis y se encuentra enmarcado dentro de una de las líneas de investigación del Proyecto de Ciencia y Técnica P-031516 “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad” (Director: Daniel Riesco, Co-Director: Roberto Uzal. Acreditado con evaluación externa. Financiamiento: Universidad Nacional de San Luis.

Introducción

Aunque hay varias formas de explicar el mecanismo de cifrado y descifrado de RSA [19], podemos decir que está basado en dos claves: una clave pública $k_{\text{pub}} = (n, e)$, y una clave privada $k_{\text{pr}} = (d, n)$. Donde $d \cdot e \equiv 1 \pmod{(p-1) \cdot (q-1)}$, (e es el inverso multiplicativo de d módulo $(p-1) \cdot (q-1)$)

El proceso de cifrado en RSA consiste en que dada la clave pública $k_{\text{pub}} = (n, e)$, y un texto claro (no cifrado) convertido a un número M , la función de cifrado, que produce el texto cifrado C es:

$$C = e(M) \equiv M^e \pmod{n}$$

donde:

$$\begin{aligned} n &= p \cdot q \\ &\bullet q; (p, q) \text{ son números primos} \\ &\text{y } M, C \in \mathbb{Z}_n \end{aligned} \quad \text{Eq 1}$$

Descifrar, en RSA, consiste en que dada la clave privada $k_{\text{pr}} = (d, n)$ y el texto cifrado C la función de descifrado es:

$$M = d(C) \equiv C^d \pmod{n}$$

donde:

d es un entero grande que es relativamente primo a:

$$(p - 1) \cdot (q - 1), \quad \text{Eq 2}$$

que es lo mismo que

$$\Phi(n) = (p - 1) \cdot (q - 1)$$

o sea, $\text{gcd}(d, (p - 1) \cdot (q - 1)) = 1$

$$\text{y } M, C \in \mathbb{Z}_n$$

Como hemos dicho, la clave (n, e) es pública, es decir conocida por todos, la clave de descifrado d es privada; además, (p, q) también son sólo conocidos por el propietario de la clave de descifrado.

n es también llamado el módulo, y siendo que n es el producto de dos números primos (p y q en Eq 2) la seguridad de RSA está basada en la dificultad del problema de la factorización. Por otro lado, el tema de si el llamado 'RSA problem' es más o menos difícil que factorizar un número grande, está tratado en [1], [2], [5], [6]. En general, los casos informados en la literatura, se trata de la factorización de un número RSA, por ejemplo [8], [9], [10], [22].

Líneas de Investigación, Desarrollo e Innovación

Así, guardar un secreto basado en la dificultad de una computación (factorización en este caso) en un momento dado y pretender que así sea por un período de tiempo dado, entonces, entre otros aspectos, tanto el incremento del poder computacional como el avance del criptoanálisis deben ser considerados. Es decir que, si al tiempo t_0 , un documento es cifrado con una clave de x_0 bits y al tiempo t_1 el cifrado es hecho normalmente con una clave de x_1 bits ($x_1 > x_0$), ¿puede un documento cifrado con x_0 bits ser descifrado al tiempo t_1 ? y, de ser así, ¿cuánto más grande debe ser x_0 ? ¿al menos $x_0 = x_1$?, ¿o mayor?; y ¿cómo estimar, cuando se está en t_0 , el valor de x_0 para que resista, al menos, hasta t_1 ?

Esto es significativo, ya sea que la codificación sea empleada para seguridad en comunicaciones o en el uso de la firma digital de documentos público o privados. Además, es sustancial señalar que, en la mayoría de los casos, un mensaje requiere de un corto lapso de

tiempo para su seguridad, cosa que no es así cuando se usa en la firma digital donde el lapso de tiempo requerido para la seguridad de la firma digital, puede llegar a ser mucho más extenso, tan es así que se puede hablar de décadas o generaciones, por lo que la seguridad de la clave en el tiempo es mucho más crítica.

Dado que la seguridad de RSA está basada en la dificultad de encontrar los factores primos de un número, en este caso de n , si un atacante consigue factorizar n , es decir obtiene p y q , puede conseguir d (ver Eq 2), y por supuesto M , como se muestra en Eq 3.

Paso

$$\begin{array}{l} 1^{\text{er}} \quad \Phi(n) = (p - 1) \\ \quad \quad \quad \cdot (q - 1) \\ 2^{\text{do}} \quad d^{-1} \equiv e \pmod{\Phi(n)} \\ 3^{\text{ro}} \quad M \equiv C^d \pmod{n} \end{array} \quad \text{Eq 3}$$

Hay un gran número de algoritmos de factorización. Actualmente el más empleado es el Number Field Sieve [12]. El esfuerzo computacional requerido para factorizar un número n empleando alguno de los algoritmos más conocidos (NFS, QFS) [11], [15], es (empleando notación L):

$$L[n, u, c] = \exp(c(\ln n)^u (\ln \ln n)^{1-u})$$

$$\text{Donde } c = (64/9)^{1/3} \approx 1.923 \text{ y } u = 1/3 \quad \text{Eq 4}$$

$L[n, u, c]$ es a veces mencionado como $L[n]$

Resultados y Objetivos

En una primera etapa nos encontramos trabajando en un primer modelo de análisis del problema, empleando por un lado el esfuerzo computacional (tal como se muestra en Eq 4) requerido para factorizar un número y por otro lado el crecimiento del poder computacional considerado a partir de la llamada ley de Moore [16], [17], el cual es empleado, entre otras métricas, en la literatura [7], [13], [14], [18], y que permite relacionar el tiempo con el poder computacional y así establecer un parámetro para el tamaño de la clave dado un lapso de tiempo dado; es decir, tratar de 'relacionar', esfuerzo computacional, su incremento en el tiempo con el tamaño (p.e. en bits) de la clave para poder 'predecir' cuánto tiempo durará la seguridad provista por una clave de longitud

dada.

Como parte del trabajo futuro, esperamos, en una etapa siguiente, incluir el costo monetario en la evaluación, así como otros parámetros tales como la importancia del procesamiento en la nube en el poder computacional. En principio y de acuerdo con algunas evaluaciones preliminares que ya hemos realizado, podría decirse que el mismo sigue de cerca lo establecido en Moore.

También nos encontramos trabajando en la generación de un cuestionario para las empresas proveedores de sistemas de seguridad informática, y organizaciones que emplean encriptación, en todos sus niveles, con el objeto de obtener mayor información sobre las características de los sistemas ofrecidos, con el objeto de evaluar las características de seguridad de los mismos con relación a los niveles de encriptación empleados.

Formación de Recursos Humanos

Dentro del SEG (Software Engineering Group), en el ámbito de la Universidad Nacional de San Luis, en el que se enmarca el Proyecto de Ciencia y Técnica P-031516 “Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad”, se han llevado a cabo numerosas tesis de grado y de posgrado. Asimismo, se han desarrollado modelos de evaluación de seguridad informática [3], [4], se ha iniciado una línea en Ciberdefensa [21] y se están realizando otros trabajos en el tema [20], así como tesis de posgrado en seguridad informática.

Referencias

- [1] Aggarwal, Divesh; Maurer, Ueli M. “Breaking RSA Generically is Equivalent to Factoring” *Advances in Cryptology - EUROCRYPT 2009*, 28th Annual International Conference on the Theory, Applications of Cryptographic Techniques. Volume 5479 *Lecture Notes in Computer Science* Editors Antoine Joux. Springer Berlin, Heidelberg. 2009.
- [2] Aggarwal, Divesh; Maurer, Ueli M. “FACTORIZING IS EQUIVALENT TO GENERIC RSA”. *IACR, Cryptology ePrint Archive*, 2008. Retrieved 12/07/2016 <http://eprint.iacr.org/2008/260> SAMEas previous one
- [3] Aristides Dasso, Ana Funes, “Un Modelo para la Evaluación de la Seguridad en Sistemas Informáticos”. *WICC 2016*, XVIII Workshop de Investigadores en Ciencias de la Computación, Concordia, Entre Ríos, Argentina. 14 y 15 de abril, 2016.
- [4] Aristides Dasso, Ana Funes, Germán Montejano, Daniel Riesco, Roberto Uzal, Narayan Debnath; “Model Based Evaluation of Cybersecurity Implementations”. *ITNG 2016*. Las Vegas, Nevada, USA, 11-13 abril 2016. In S. Latifi (ed.), *Information Technology New Generations, Advances in Intelligent Systems and Computing* 448. DOI: 10.1007/978-3-319-32467-8_28. Springer International Publishing, Switzerland 2016.
- [5] Boneh, D; Venkatesan, R. “Breaking RSA may not be equivalent to factoring”. *Proceedings of the International Conference on the Theory, Application of Cryptographic Techniques*. Espoo, Finland, May 31 – June 4, 1998. *Advances in Cryptology — EUROCRYPT'98. Lecture Notes in Computer Science*, Volume 1403. Editor: Kaisa Nyberg. Springer Berlin, Heidelberg 1998.
- [6] Brown, Daniel R. L. “Breaking RSA May Be As Difficult As Factoring”, *IACR, Cryptology ePrint Archive*, 2005. Retrieved 07/12/2016 <http://eprint.iacr.org/2005/380>
- [7] ECRYPT II, European Network of Excellence in Cryptology II, “Yearly Report on Algorithms and Keysizes (2011-2012)”. *ICT-2007-216676*. Retrieved 2/11/2016, <http://www.ecrypt.eu.org/ecrypt2/>
- [8] European Union Agency for Network and Information Security. “Algorithms, Key Size and Parameters Report”. November, 2014
- [9] FactorWorld. “General Purpose Factoring Records.” Retrieved 06/08/2016, <http://www.cryptoworld.com/FactorRecords.html>
- [10] Kleinjung, Thorsten; Aoki, Kazumaro; Franke, Jens; Lenstra, Arjen; Thomé, Emmanuel; Bos, Joppe; Gaudry, Pierrick;

- Kruppa, Alexander; Montgomery, Peter; Osvik; Dag Arne; te Riele, Herman; Timofeev, Andrey; Zimmermann, Paul; "Factorization of a 768-bit RSA modulus". IACR, Cryptology ePrint Archive, Report 2010/006, 2010, Retrieved 21/07/2016, <http://eprint.iacr.org/2010/006>.
- [11] Kostas Bimpikis and Ragesh Jaiswal. "Modern Factoring Algorithms". University of California, San Diego USA. 2005. Retrieved February 2017, <http://www.cs.columbia.edu/~rjaiswal/factoring-survey.pdf>
- [12] Lenstra, A.K., Lenstra, H.W.J. (Eds.). "The Development of the Number Field Sieve". Series: Lecture Notes in Mathematics, Vol. 1554. 1993, VIII, 140 p. Springer-Verlag Berlin Heidelberg,
- [13] Lenstra, Arjen K. "Unbelievable Security Matching AES security using public key systems". Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT '01), Pages 67-86. Springer-Verlag London, UK 2001. Retrieved 2/11/2016, <http://iacr.org/archive/asiacrypt2001/22480067.pdf>
- [14] Lenstra, Arjen K.; Verheul, Eric R.. "Selecting Cryptographic Key Sizes". Journal of Cryptology (2001), 14: 255–293. DOI: 10.1007/s00145-001-0009-4.
- [15] Lenstra, H. W.; Pomerance, Carl (July 1992). "A Rigorous Time Bound for Factoring Integers" (PDF). Journal of the American Mathematical Society. 5 (3): 483–516. doi:10.1090/S0894-0347-1992-1137100-0. MR 1137100.
- [16] Moore, Gordon E. "Cramming more components onto integrated circuits". Electronics, Volume 38, Number 8, April 19, 1965
- [17] Moore, Gordon E. "Progress In Digital Integrated Electronics". Technical Digest 1975. International Electron Devices Meeting, IEEE, 1975, pp. 11-13.
- [18] République Française. Premier ministre. Agence nationale de la sécurité des systèmes d'information. "Référentiel Général de Sécurité". version 2.0. Annexe B1. "Mécanismes cryptographiques. Règles et recommandations concernant le choix et le dimensionnement des mécanismes cryptographiques. Version 2.03 du 21 février 2014
- [19] Rivest, R. L.; Shamir, A.; Adleman, L. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". Communications of the ACM, Volume 21 Issue 2, Feb. 1978. Pages: 120-126 doi>10.1145/359340.359342
- [20] Roberto Uzal, Daniel Riesco, Germán Montejano, Claudio Baieli, Walter Agüero. "Conflictos en el Ciber Espacio entre estados naciones: Potenciales aportes para la eventual actuación de las Naciones Unidas y de la Unión Internacional de las Telecomunicaciones". 8° Simposio Argentino de Informática en el Estado - SIE 2014. Argentina.
- [21] Roberto Uzal, Jeroen van de Graaf, Germán Montejano, Daniel Riesco, Pablo García. Inicio de la Línea de Investigación "Ingeniería de Software y Defensa Cibernética". XV Workshop de Investigadores en Ciencias de la Computación, Paraná, Argentina. 2013.
- [22] Valenta, Luke; Cohny, Shaanan; Liao, Alex; Fried, Joshua; Bodduluri, Satya; Heninger, Nadia. "Factoring as a Service". In <http://eprint.iacr.org/2015/1000>, Retrieved 05/29/2016. Also in Financial Cryptography, Data Security 2016, Twentieth International Conference, February 22–26, 2016. Barbados.

OTP-Vote: Avances en la Generación de un Modelo de Voto Electrónico

Silvia Bast¹; Pablo García¹; Germán Montejano^{1 2}

¹Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-425166– Int. 28
[pablogarcia, silviabast]@exactas.unlpam.edu.ar

²Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950 – (5700) San Luis – San Luis – Argentina
Tel.: +54-2652-424027 – Int. 251
gmonte@unsl.edu.ar – web: <http://www.unsl.edu.ar>

RESUMEN

En la actualidad, gran parte de las actividades humanas se ven mediadas por el uso de las plataformas tecnológicas. Resulta lógico pensar que con el transcurrir del tiempo, aún mayor cantidad de tareas irán migrando a su forma electrónica y permitirán que los usuarios las lleven a cabo de manera remota. Debido a los avances tecnológicos, al crecimiento poblacional y a las distancias geográficas, los sistemas de voto presentan un gran potencial para ser llevados a cabo parcial o totalmente de manera electrónica. Existen, sin embargo, cuestiones de confianza por parte de los ciudadanos en este tipo de sistemas, dado que contienen información sumamente sensible para el elector. La seguridad es entonces un aspecto central que debe reforzarse en tales sistemas. Básicamente debe ser posible mantener de forma anónima la información del voto de cada uno de los electores y debe resultar imposible modificar los votos ya emitidos.

Se plantean entonces problemas relacionados con la confidencialidad y la integridad de los datos

Se exponen en el presente documento los avances obtenidos en la investigación de un modelo que optimiza los aspectos de confidencialidad e integridad en Sistemas de Voto Electrónico.

El modelo propuesto se denomina: OTP-Vote.

Palabras clave: *sistemas de voto electrónico, confidencialidad, integridad, seguridad.*

CONTEXTO

Este trabajo se enmarca en el Proyecto de Investigación: "Aspectos de Seguridad en Proyectos de Software", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (UNLPam) Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de

La Pampa y es dirigido por el Doctor Germán Antonio Montejano y codirigido por el Magister Pablo Marcelo García e incluye a la Mg. Silvia Gabriela Bast y la Profesora Estela Marisa Fritz como investigadoras.

Surge desde la línea de Investigación “Ingeniería de Software y Defensa Cibernética”, presentada en [1], y que a su vez se enmarca en el Proyecto “Ingeniería de Software: Aspectos de alta sensibilidad en el ejercicio de la Profesión de Ingeniero de Software” de la Facultad de Ciencias Físico - Matemáticas y Naturales de la Universidad Nacional de San Luis (UNSL) (<http://www.sel.unsl.edu.ar/pro/proyec/2012/index.html>) y que incluye acciones de cooperación con la Universidad Federal de Minas Gerais (UFMG, Brasil).

Entre tales acciones debe mencionarse que Jeroen van de Graaf, PhD., Docente de UFMG, y el Dr. Germán Montejano (UNSL) fueron orientadores del Mg. Pablo García en el desarrollo de su tesis de maestría titulada “Optimización de un Protocolo Dining Cryptographers Asíncrono”, defendida en 2013. El Mg. García ha realizado estadias de investigación en UFMG en 2013 mientras desarrollaba su tesis y en 2016 en el Laboratorio 4303 en el Departamento da Ciências da computação do Instituto de Ciências Exatas.

1. INTRODUCCIÓN

El voto electrónico puede definirse como *“una forma de votación basada en medios electrónicos que se diferencia del método tradicional por la utilización de tecnologías como hardware, software y procedimientos*

*que permiten automatizar los procesos que comprenden unas elecciones”*¹

En el ámbito del proyecto de investigación se propone una nueva metodología para aplicar a los sistemas de E-Voting, presentada en [2]. El desarrollo de la misma implicó llevar a cabo las siguientes actividades:

- Analizar las características y requerimientos de sistemas de Voto electrónico.
- Analizar las herramientas y modelos que focalizan en la seguridad de los datos.
- Proponer en base a los análisis realizados un modelo que optimice los aspectos de confidencialidad e integridad de los datos ofreciendo anonimato incondicional y seguridad computacional (que puede llevarse al nivel exigible) durante el proceso electoral.
- Aplicar el modelo a casos para validar su funcionamiento.

Los Requerimientos de los Sistemas de Voto Electrónico

Del análisis de las características de estos sistemas ([3], [4] y [5]), surge un grupo de requisitos que deben respetar los mencionados desarrollos. Los mismos pueden resumirse en: Anonimato, Autenticación del votante, Verificabilidad, Simplicidad, Costo, Auditabilidad, Inviolabilidad, Seguridad, No coerción, Robustez.

Entre los requerimientos mencionados, existen algunos que pueden satisfacerse de forma sencilla, pero otros presentan mayor nivel de dificultad. En los sistemas de Voto electrónico es necesario proteger:

- Indefinidamente la privacidad del votante: aún después de

¹ Sitio Observatorio del voto-E en América Latina disponible en <http://www.voto-electronico.org/index.php/definicion/definicion-amplia>

finalizada la elección, dado que en caso de que algún intruso obtenga una copia digital de registros que permitan relacionar el votante con su voto contaría con todo el tiempo para intentar descifrarlo. Las personas desean mantener su privacidad asegurada indefinidamente y existen casos en los que sería de suma gravedad que se conociera por quién votó alguna persona en particular. Por ejemplo, conocer la trayectoria como votante de un candidato actual podría influir en el electorado

- Mientras dure el proceso electoral, la seguridad de los datos: la protección de la información circulante sólo debe soportar el lapso de tiempo que corresponda al proceso de votación.

Aportes Teóricos para la Propuesta de un Modelo de Datos para Sistemas de Voto Electrónico.

El modelo propuesto se basa en:

- 1) El modelo de almacenamiento de Canales Paralelos – Múltiples Canales Dato Único (MCDU, [6], [7], [8], [9] y [10]), que ofrece :
 - Anonimato Incondicional: a través de su característica de aleatoriedad.
 - Uso eficiente del almacenamiento.
 - Disminución de la probabilidad de colisiones, que puede llevarse a cualquier nivel exigible a través del uso de las fórmulas para la configuración de parámetros de la elección.
- 2) El uso de claves One Time Pad (OTP) [11]. OTP es un algoritmo criptográfico que puede crear un texto cifrado del que nadie puede obtener el texto plano y que no puede quebrarse

aún con potencia de cálculo infinito e ilimitada cantidad de tiempo.

Las claves OTP presentan las siguientes características:

- Son aleatorias.
 - Son tan largas como el mensaje mismo.
 - A partir del mismo texto cifrado, aplicando una clave diferente se produce un texto plano distinto.
- 3) El uso de la Redundancia Adecuada en la definición de las dimensiones de los atributos que formarán parte del Voto [12] con el objetivo de:
 - Disminuir considerablemente la posibilidad de detectar cuáles son los valores válidos que se han usado.
 - Disminuir también la probabilidad de que por el efecto de colisiones simples o múltiples, se produzca como resultado otro Identificador de voto válido.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación que se siguen son:

- Desarrollo de un modelo de datos que optimice las características de confidencialidad e integridad de los datos.
- Desarrollo de compromisos y protocolos que permitan asegurar que la información que es intercambiada se mantiene inalterable.
- Desarrollo de una técnica para verificabilidad E2E del modelo propuesto, basada en funciones de Hash.
- Implementación del sistema propuesto.

3. RESULTADOS Y OBJETIVOS

Se desarrolló un Modelo de Datos que ofrece:

- Anonimato Incondicional: aportado por: la aleatoriedad en el almacenamiento de datos del modelo subyacente, esto es, Canales Paralelos MCDU, las claves OTP que se usan para la encriptación y la separación total de los procesos de acreditación y emisión de voto.
- Seguridad Computacional que puede llevarse a cualquier nivel exigible a través de: la aleatoriedad en el almacenamiento de datos provista por el uso de Canales Paralelos MCDU, las claves OTP que se usan para la encriptación, el uso de la redundancia suficiente en los atributos de las tuplas y la configuración de los parámetros de la elección.

Se desarrollaron dos propuestas de recuperación de colisiones adicionales al modelo.

Se validó el modelo a través de su aplicación a diferentes casos generados por un simulador desarrollado ad-hoc.

Se espera avanzar en el refinamiento de protocolos antifraude [13] y en la implementación del modelo.

También se está trabajando para agregar a la propuesta, la característica de Verificabilidad End to End.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco del presente proyecto se presentan los siguientes puntos relacionados con la formación de recursos humanos:

- Pablo García realizó una estadía de un año en la Universidad Federal de Minas Gerais (UFMG), aprobando seminarios de posgrado y trabajando en el grupo “Criptografía Teórica y Aplicada”, dirigido por Jeroen van de Graaf, PhD.

- Pablo García defendió su tesis para obtener el grado de Magister en Ingeniería de Software de la Universidad Nacional de San Luis, bajo la dirección de Jeroen van de Graaf, PhD (UFMG) y Dr. Germán Montejano (UNSL). La tesis se tituló: “Optimización de un Esquema Dining Cryptographers Asíncrono” y recibió la calificación de Sobresaliente.

- Silvia Bast defendió su tesis para obtener el grado de Magister en Ingeniería de Software de la Universidad Nacional de San Luis, bajo la dirección del Dr. Germán Montejano (UNSL) y del Mg Pablo García (UNLPam). La tesis se tituló: “Optimización de la integridad de datos en Sistemas de E-Voting” y recibió la calificación de Sobresaliente.

- Pablo García está desarrollando su tesis para obtener el grado de “Especialista en Ingeniería de Software”. Su plan de trabajo fue aprobado y se planea su defensa para Agosto de 2017. La tesis se titula: “Anonimato en Sistemas de Voto Electrónico” y es dirigida por Jeroen van de Graaf, PhD (UFMG) y Dr. Germán Montejano (UNSL).

- Silvia Bast está desarrollando su tesis para obtener el grado de “Especialista en Ingeniería de Software”. Su plan de trabajo fue aprobado y se planea su defensa para Agosto de 2017. La tesis se titula: “Sistemas de E-Voting: Integridad de Datos” y es dirigida por el Dr. Germán Montejano (UNSL) y el Mg. Pablo García (UNLPam).

- Silvia Bast y Pablo García completaron el cursado del Doctorado en Ingeniería Informática en la Facultad de Ciencias Físico Matemáticas y Naturales de la

Universidad Nacional de San Luis
(UNSL).

5. BIBLIOGRAFÍA

- [1] **UZAL R., VAN DE GRAAF J., MONTEJANO G., RIESCO D., GARCÍA P.:** “Inicio de la Línea de Investigación: Ingeniería de Software y Defensa Cibernética”. Memorias del XV WICC. Ps 769-773. ISBN: 9789872817961. 2013. <http://sedici.unlp.edu.ar/handle/10915/27537>
- [2] **BAST S.:** “Optimización de la Integridad de Datos en Sistemas de E-Voting”. Tesis de Maestría defendida en la Universidad Nacional de San Luis. 14 de Diciembre de 2016. San Luis, Argentina.
- [3] **EPSTEIN J.:** “Electronic Voting”, Cyber Defense Agency LLC.
- [4] **KAZI M., ALAM R., TAMURA S.:** Electronic Voting - Scopes and Limitations IEEE/OSA/IAPR International Conference on Infonnatics, Electronics & Vision.
- [5] **PRINCE A.:** Consideraciones, aportes y experiencias para el Voto electrónico en Argentina. 2005.
- [6] **VAN DE GRAAF J., MONTEJANO G., GARCÍA P.:** “Manejo de Colisiones en un Protocolo Non Interactive Dining Cryptographers”. Anales de las 42^o Jornadas Argentinas de Informática e Investigación Operativa (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Páginas 29 a 43. Septiembre 2013. Disponible en: <http://42jaiio.sadio.org.ar/proceedings/Simposios/Trabajos/WSegI/03.pdf>.
- [7] **GARCÍA P., VAN DE GRAAF, MONTEJANO G., RIESCO D., DEBNATH N., BAST S.:** “Storage Optimization for Non-Interactive Dining Cryptographers (NIDC)”. The International Conference on Information Technology: New Generations. 2015. Las Vegas, Nevada, USA. Disponible en: <http://ieeexplore.ieee.org/document/7113449/>.
- [8] **GARCÍA P., BAST S., FRITZ E., MONTEJANO G., RIESCO D., DEBNATH N.:** “A Systematic Method for Choosing Optimal Parameters for Storage Parallel Channels of Slots”. IEEE International Conference on Industrial Technology (ICIT 2016). 14 - 17 March 2016 / Taiwan, Taipei. Disp. en: <http://ieeexplore.ieee.org/document/7475019/>.
- [9] **GARCÍA P., MONTEJANO G., BAST S., FRITZ, E.:** “Loss of Votes in NIDC Applying Storage in Parallel Channels”. Congreso Argentino de Ciencias de la Computación, CACIC 2016. San Luis, 3 al 7 de octubre de 2016. Universidad Nacional de San Luis (UNSL). Se obtiene distinción como MEJOR EXPOSITOR del Workshop de Seguridad Informática. Seleccionado para ser publicado en el libro de los mejores artículos de CACIC 2016.
- [10] **GARCÍA P., VAN DE GRAAF J., HEVIA A., VIOLA A.:** “Beating the Birthday Paradox in Dining Cryptographer Networks”. The third International Conference on Cryptology and Information Security in Latin America, *Latincrypt 2014*. September 17-19, 2014. Florianopolis, Brasil. Progress in Cryptology an Information Security in Latin America. Revised Selected Papers Lecture. Springer (2014). Diego F. Aranha, Alfred Menezes Eds. ISBN: 978-3-319-16294-2. ISSN: 0302-9743. ITEM:9783319162942. Disp en: http://rd.springer.com/chapter/10.1007/978-3-319-16295-9_10.
- [11] **PAAR C., PELZL J.:** “The One-Time Pad” - Chapert 2: Stream Ciphers in Understanding Cryptography Springer Berlin Heidelberg New York. ISBN 978-3-642-04100-6 e-ISBN 978-3-642-04101-3. DOI 10.1007/978-3-642-04101-3
- [12] **GARCÍA P., BAST S., MONTEJANO G., FRITZ E.:** “Codificación de Sufragios con Detección de Colisiones en NIDC con Canales Paralelos de Slots”. Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2016.
- [13] **VAN DE GRAAF J., MONTEJANO G., GARCÍA P.:** “Optimización de un Protocolo Non-Interactive Dining Cryptographers”. Congreso Nacional de Ingeniería Informática / Sistemas de Información. CoNaIISI 2013. 21 y 22 de noviembre de 2013. Córdoba, Argentina. Disponible en: <http://conaiisi.unsl.edu.ar/2013/25-483-1-DR.pdf>.

Propuesta de un Modelo de Proceso para Resolver Vulnerabilidades de Seguridad en Infraestructura Utilizando Herramientas de Computación Cognitiva

Darío Propato, Jorge Eterovic, Marisa Panizzi, Luis Torres
Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales,
Universidad de Morón.

Cabildo 134 – CP (1708) – Morón – Prov. de Bs. As. Tel: 5627-2000
dariopropato@gmail.com; jorge_eterovic@yahoo.com.ar; marisapanizzi@outlook.com;
torreslu@ar.ibm.com

Resumen

El objetivo de este trabajo de investigación consiste en el desarrollo de un modelo de proceso que permita resolver vulnerabilidades de Seguridad Informática utilizando herramientas de la computación cognitiva. Este servirá como guía a los profesionales de informática que se desempeñen en áreas de seguridad de la información. Los sistemas cognitivos pueden aprender de sus experiencias, encontrar correlaciones, crear hipótesis y recordar los resultados y aprender de ellos. Esta capacidad de los sistemas cognitivos tendrá una influencia crucial en la toma de decisiones por parte de múltiples interesados y la solución a construir tomara como base las ventajas de este tipo de sistemas. Se realizó una revisión sistemática de un par de modelos de procesos de software como por ejemplo MoProSoft y Métrica versión 3, para considerar de ellos los elementos propuestos y ser contemplados en la propuesta de solución de esta investigación. Para llevar a cabo la validación del modelo propuesto se utilizará un caso real de una organización.

Palabras clave: Computación cognitiva, Modelo de Proceso, Seguridad en Infraestructura, Vulnerabilidades, Tiempo de Detección.

Contexto

Este trabajo de investigación se encuentra radicado en el Instituto de Ingeniería de Software Experimental perteneciente a la Facultad de Informática, Ciencias de la Comunicaciones y Técnicas Especiales de la Universidad de Morón. El Instituto articula con las cátedras de tesis de la carrera Licenciatura en Sistemas y con la cátedra de Auditoria y Seguridad de los Sistemas de información.

Introducción

Ha sido necesario llevar a cabo una investigación exploratoria documental de antecedentes de la computación cognitiva, y aspectos de vulnerabilidades y modelos de procesos de software.

Según Kelly, los sistemas cognitivos son probabilísticos, significa que son diseñados para adaptarse y dar sentido a la complejidad e impredecibilidad de la información. Interpretan la información, la organizan y ofrecen una explicación de su significado junto con la justificación de sus conclusiones. No ofrecen una respuesta definitiva (John E. Kelly, 2015).

Desde el punto de vista de Briggs, la analítica cognitiva está inspirada en como el cerebro humano procesa información, saca conclusiones y codifica instintos y la experiencia de aprendizaje. En lugar de aprender sobre reglas predefinidas y consultas estructuradas para encontrar respuestas, el análisis cognitivo se basa en tecnología para generar hipótesis de una amplia variedad de información relevante y conexiones (Briggs, 2014).

Entre los desarrollos recientes, basados en computación cognitiva y que actualmente se encuentra funcional es la herramienta Watson de IBM, que es un sistema de inteligencia artificial con la capacidad de responder a preguntas formuladas en lenguaje natural, desarrollado por IBM. Es parte del proyecto del equipo de investigación DeepQA para la generación de hipótesis, la recopilación de pruebas masivas, el análisis y la calificación (DeepQA Project, 2011).

Para la construcción del modelo se tomaron como base modelos de procesos de software y metodologías que actualmente se encuentran validados y permiten una adaptabilidad al proyecto planteado. Algunos de los seleccionados se detallan a continuación:

- Métrica versión 3, es una metodología de planificación, desarrollo y mantenimiento de sistemas de información para la sistematización de actividades del ciclo de vida de los proyectos de software en el ámbito de la administración pública de España. Está basada en el modelo de procesos del ciclo de vida de desarrollo ISO/IEC 12207 así como en la norma ISO/IEC 15504 SPICE (Ministerio de Hacienda y Obras Públicas del Gobierno de España, 2001).

- MoProSoft (Modelo de Procesos para la Industria del Software), es un modelo para la mejora y evaluación de los procesos de desarrollo y mantenimiento de sistemas y productos de software. Está dirigido a las organizaciones dedicadas al desarrollo y mantenimiento de software (NORMA NMX-I-059-NYCE-2005)

Se consideró la revisión de la ISO/IEC 27002, la cual proporciona recomendaciones de las mejores prácticas en la gestión de la seguridad de la información a todos los interesados y responsables en iniciar, implantar o mantener sistemas de gestión de la seguridad de la información. La seguridad de la información se define en el estándar como la preservación de la confidencialidad (asegurando que sólo quienes estén autorizados pueden acceder a la información), integridad (asegurando que la información y sus métodos de proceso son exactos y completos) y disponibilidad (asegurando que los usuarios autorizados tienen acceso a la información y a sus activos asociados cuando lo requieran) (ISO/IEC 27002, 2013)

El problema que motiva la elaboración de este trabajo es la existencia de vulnerabilidades en la seguridad de los sistemas de información y los elevados tiempos de detección. El objetivo es solucionar las vulnerabilidades con el uso de computación cognitiva y, a partir del mismo, generar un modelo específico, dado que los modelos propuestos no soportan o no son adecuados para el proceso de seguridad de la información. Es decir, solo son aplicables para el ciclo de vida del software. La computación cognitiva provee un desafío en cuanto a su utilización e implementación por el escaso conocimiento que se tiene de la misma. Todo esto ocasiona que los profesionales de sistemas involucrados en la seguridad de la información no cuenten con

un modelo específico para desarrollar la implementación de una herramienta de computación cognitiva dentro de una organización para solucionar vulnerabilidades de seguridad.

Entre los conceptos de vulnerabilidad informática revisados se puede mencionar el propuesto por Cencini, en el cual la define vulnerabilidad como una falla se convierte en una vulnerabilidad si el comportamiento expuesto es tal que puede ser explotado para permitir el acceso no autorizado, elevación de privilegios o denegación de servicio (Cencini, 2005).

La ISO 27005 define vulnerabilidad como: una debilidad de un activo o grupo de activos que pueden ser explotados por una o más amenazas (ISO 27005, 2008).

La Agencia Europea de Seguridad de las Redes y de la Información (ENISA) conceptualiza vulnerabilidad como la existencia de una debilidad, diseño o error de implementación que puede conducir a un evento inesperado e indeseable que compromete la seguridad del sistema informático, la red, la aplicación o el protocolo involucrado (ENISA, 2016)

Entre las principales vulnerabilidades informáticas analizadas, se seleccionaron aquellas de mayor importancia en los informes anuales de seguridad:

- Malware
- Phishing
- Amenaza interna

Según Moir, el término malware es muy utilizado por profesionales de la informática para referirse a una variedad de software hostil, intrusivo o molesto” (Moir, 2003).

El análisis de malwares de Cisco encontró que la mayoría (91,3%) utilizan los DNS (Sistema de Nombre de Dominios) para ganar control y acceso a comandos, filtrar datos y redirigir tráfico.” (Cisco Annual Security Report, 2016).

Phishing, “es el intento de obtener información confidencial como nombres de usuario, contraseñas y detalles de tarjetas de crédito (e indirectamente dinero), a menudo por razones maliciosas, disfrazándose como una entidad confiable en una comunicación electrónica” (Ramzan, 2010).

Una Amenaza interna (en Inglés, Insider threat) según el FBI: es una amenaza malintencionada a una organización que proviene de personas dentro de la organización, como empleados, contratistas o asociados al negocio, que tienen información privilegiada sobre las prácticas de seguridad de la organización, los datos y los sistemas informáticos. La amenaza puede incluir el fraude, el robo de información confidencial o comercialmente valiosa, el robo de la propiedad intelectual o el sabotaje de los sistemas informáticos (FBI, 2016).

Una de las amenazas con mayor incidencia son los ataques a las redes informáticas, los mismos tienen variantes y diferentes objetivos, como:

- Los ataques de denegación de servicio, según RFC4732 es un ataque a un sistema de computadoras o red que causa que un servicio o recurso sea inaccesible a los usuarios legítimos (RFC4732, 2006).
- Ataques de fuerza bruta, según la Electronic Frontier Foundation: es un ataque criptoanalítico que puede ser utilizado para descifrar y encriptar datos. Este tipo de ataque es utilizado

cuando no es posible obtener ventajas de otras debilidades en un sistema de encriptación (Electronic Frontier Foundation, 1998)

Las vulnerabilidades mencionadas afectan en la mayoría de los casos a los dispositivos pertenecientes a la infraestructura de sistemas de la compañía. El volumen de información siempre creciente eleva el tiempo de detección (TTD), es decir, “la ventana de tiempo entre la primera observación de un archivo que ha pasado por todas las tecnologías de seguridad para llegar a un punto final y la detección de una amenaza asociada con ese archivo. Se estima en la industria actual un tiempo inaceptable de 100 a 200 días.”(Cisco Annual Security Report, 2016).

Resultados y Objetivos

En esta etapa del proyecto se ha llevado a cabo una investigación exploratoria y documental de los tres ejes sobre el cual se basa la solución propuesta: computación cognitiva, vulnerabilidades de infraestructura y modelos de procesos.

Se ha comenzado a bosquejar las fases/actividades/tareas que contemplaran el modelo de proceso de proponer.

También se está analizando cual será el caso testigo para la validación del modelo propuesto.

Formación de Recursos Humanos

Actualmente el equipo de investigación está integrado por un estudiante de la carrera Licenciatura en Sistemas, por dos docentes-investigadores y un docente que inicia su proceso de formación como investigador pertenecientes a la Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales, Universidad de Morón.

Bibliografía

- Kelly, John E., (2015). Computing, cognition and the future of knowing, IBM.
- Briggs, Bill. (2014). Tech Trends 2014 Inspiring Disruption. Consulta realizada en <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/technology/deloitte-uk-tech-trends-2014.pdf>
- IBM Research (2011). DeepQA Project. Consulta realizada en <http://www.research.ibm.com/deepqa/deepqa.shtml>
- Ministerio de Hacienda y Administraciones Públicas del Gobierno de España, (2001). MÉTRICA v3. Consulta realizada en http://administracionelectronica.gob.es/pae/Home/pae_Documentacion/pae_Metodolog/pae_Metrica_v3.html
- Asociación Mexicana para la Calidad en Ingeniería de Software, (2005). MoProSoft. NMX-059/02-NYCE-2005.
- Hurwitz, Judith, (2015). Cognitive Computing and Big Data Analytics: Implementing Big Data Machine Learning Solutions. ISBN: 1118896629.
- ISO/IEC 27002:2013, (2007). Information technology Security techniques. Consulta realizada en <http://www.iso27001security.com/html/27002.html>.
- Cencini, Andrew, (2005). Software Vulnerabilities: Full-, Responsible, and Non-Disclosure. Consulta realizada en https://courses.cs.washington.edu/courses/csep590/05au/whitepaper_turnin/software_vulnerabilities_by_cencini_yu_chan.pdf
- ISO/IEC FIDIS 27005:2008, (2008). Information technology -- Security techniques-Information security risk management Consulta realizada en <http://www.iso27001security.com/html/27005.html>

- ENISA (2016) ,Vulnerabilities and exploits.
Consulta realizada en
<https://www.enisa.europa.eu/topics/national-csirt-network/glossary/vulnerabilities-and-exploits>
- Moir, Robert, (2003). Defining Malware. Microsoft Security MVP. Consulta realizada en
<https://technet.microsoft.com/en-us/library/dd632948.aspx>
- CISCO, (2016). Annual Security Report. Consulta realizada en <http://mkto.cisco.com/rs/564-WHV-323/images/cisco-asr-2016.pdf>
- Ramzan, Zulfikar, (2010). Phishing attacks and countermeasures. Handbook of Information and Communication Security. ISBN 9783642041174.
- FBI, (2016).The Insider Threat: An Introduction to Detecting and Deterring an Insider Spy, ,
https://www.fbi.gov/file-repository/insider_threat_brochure.pdf.
- IETF Trust, (2006). Internet Denial-of-Service Considerations (RFC4732). Consulta realizada en <https://tools.ietf.org/html/rfc4732>
- Electronic Frontier Foundation, (1998) . Cracking DES – Secrets of Encryption Research, Wiretap Politics & Chip Design. ISBN 1- 56592-520-3.
- Symantec. (2016). Internet Security Threat Report. Consulta realizada en <https://www.symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf>

Red de Comunicación y Seguridad Inalambrica: Implementación del Programa “RedCoFi, Wifi al alcance de todos” en la FI UNLZ

Servetto Diego,¹ Pascal Guadalupe¹, Lobo Mirassón Unelén¹

¹Facultad de Ingeniería - Universidad Nacional de Lomas de Zamora Instituto
de Investigaciones de Tecnología y Educación (IIT&E)
Camino de Cintura y Avda. Juan XXIII –Llavallol
Teléfono: 011 – 4282-7880

diegoservetto@gmail.com; guadapascal@gmail.com; uneleen@gmail.com.

Resumen

En la actualidad, a medida que aumenta el volumen de información disponible en la web, aumentan las tecnologías para la comunicación y el procesamiento de la información. Y en un mundo, donde la tecnología es parte de la vida diaria, las organizaciones deben trabajar en estrategias específicas para penetrar los canales de comunicación.

El presente artículo presenta una planificación estructural del acondicionamiento de la red WI FI, garantizando conectividad y seguridad en Internet a toda la comunidad. Asegurando legitimidad de la comunicación, consistencia en los datos, disponibilidad de red e identificación de los nodos conectados de manera confidencial para garantizar el libre acceso a Internet. Para finalizar, se muestran el análisis de los resultados de las encuestas realizadas usuarios de la con el objetivo de censar frecuencia y zonas de uso, como así también la valoración al servicio.

Contexto

En la actualidad, existe un amplio volumen de información disponible en la web, las cuales requieren cada vez mayor velocidad y

calidad en la transmisión de datos. Así mismo, existe una amplia variedad de tecnologías para la comunicación, dando lugar a generar retos importantes para la gestión Informática.

El Área de Sistemas Informáticos del Instituto de Investigaciones en Tecnología y Educación (IIT&E) de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (FIUNLZ) desarrolla estrategias orientadas a mejorar los medios de comunicación, con el objetivo de satisfacer las necesidades de la comunidad que pretende acceder a los servicios on-line. En este marco, se desarrolla el programa “RedCoFi: Wifi al alcance de todos.”

Introducción

En tiempos actuales, la comunicación es un factor fundamental en la sociedad y con el paso del tiempo se ha acentuado el uso de nuevas tecnologías de información y comunicación (TIC). *La comunicación y el entretenimiento, los celulares inteligentes, la computadora y la tecnología en general se han vuelto parte de la vida diaria (Gutierrez, Lopez, Ramirez, 2015, ISSN: 2007-531).*

Tal es así que las instituciones implementan acciones (en el caso de las instituciones

académica, esas acciones se ven materializadas en la creación de un Campus Virtual) que permitan penetrar los canales de comunicación. Sin embargo, la llegada a la Universidad de los llamados nativos digitales, genera nuevos desafíos que se relacionan con la inmediatez y las redes sociales. (Campoli, Minnaard, Morrongiello, Pascal, 2013).

Todo esto, implica una adecuación y planificación estructural que garantice conectividad. Por ello, *hoy en día, al hablar de redes inalámbricas no se habla sólo de comunicación, si no también se deben estudiar grandes características técnicas y operativas, tales como configuraciones básicas de seguridad, prestaciones, alcance, posición, ubicación, etc. (Mendoza, Vidal, Almanza, 2017).*

Como se ha mencionado, *una de las características más importantes de las tecnologías de comunicaciones es la seguridad, principalmente en redes inalámbricas. Por lo tanto se deben desarrollar mecanismos específicos para proteger datos, validar usuarios, etc. (Aransay, 2009).*

Estos mecanismos están explícitos en el estándar IEEE 802.11, el cual define las normas de funcionamiento en una red local inalámbrica (WLAN). A lo largo del desarrollo de los grupos de trabajo en torno al estándar IEEE 802.11, se ha detectado diversas amenazas, entre las que se caracterizan: el análisis de tráfico, la eliminación de mensajes e interceptación, secuestros de sesión (Session Hijacking), Man-in-the-Middle (MitM), denegación de servicio, etc. (Mitchell, 2005).

Han sido exhaustos los desarrollos necesarios para alcanzar, en la actualidad, la respuesta que demandaban los usuarios de

redes inalámbricas. Es en ese momento cuando surge WAP (WI FI Protected Access), un subconjunto de la norma IEEE802.11i, con características como: distribución dinámica de claves, utilización robusta del vector inicialización, y el protocolo TKIP (Temporal Key integrity Protocol) encargado de asegurar integridad de datos.

Frente a este escenario, la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora se ve en el desafío de planificar e implementar una red WI FI siguiendo los protocolos necesarios para garantizar seguridad y conectividad entre todos los usuarios.

En el año 2016, la unidad académica contaba con con numerosos puntos de acceso inalámbricos (WiFi) para poder conectarse a Internet. Sin embargo, dichos puntos de acceso estaban en su mayoría bajo contraseña, haciéndolos exclusivos para el plantel Docente y Administrativo, lo que imposibilitaba que los alumnos de la misma puedan acceder libremente a dicho servicio.

Así mismo, ofrecía una gran cantidad de puntos de accesos libres, pero sin la capacidad de soportar la masividad de usuarios conectados en simultáneo, implicando que los alumnos desistan del servicio por su escasa o nula conectividad.

Líneas de Investigación, Desarrollo e Innovación

La presente línea de investigación, desarrollo y innovación consiste primeramente en analizar y estudiar el desempeño de la Red, re diseñar su topología, adquirir equipamiento específico que garantice la conectividad masiva de usuarios, estudiar, desarrollar e implementar nuevas estrategias de seguridad, proveer un nuevo

canal de comunicación, y evaluar constantemente el uso y la satisfacción del servicio por parte de los usuarios.

Para llevar adelante este plan, en la FIUNLZ se han desarrollado tareas como:

- Analizar y Estudiar la Red.
- Realizar pruebas de stress.
- Aplicar reglas de control para el Tráfico de la red.
 - Re diseñar la topología.
 - Reducir saltos en la Red.
 - Llevar la velocidad de transmisión a GigaE.
- Diseñar una red exclusiva para la red Wifi Publica.
 - Adquirir equipamiento específico. Switch L2, Access Point, Servidor para Controlador de Red.
 - Analizar resultados de encuestas realizadas.
 - Detectar zonas de influencias más conflictivas de la Facultad, es decir, donde más se requiere este servicio.
 - Analizar detalladamente los comentarios que han dejado en las encuestas para llevar adelante la estrategia más adecuada.

Resultados y Objetivos

Hasta el momento se implementó equipamiento Access Point para proveer la red pública WIFI en dos de las tres alas de la Planta Alta que tiene el edificio. Y en esta línea, el equipo de trabajo dispone de dos grupos de resultados: el primero corresponde a las encuestas realizadas entre los meses Noviembre-Diciembre del 2016 y otro de las estadísticas actuales de la red Wifi pública.

La encuesta estuvo destinada a alumnos, docentes y no docentes de la FIUNLZ y tenía como objetivo censar si habían notado mejora

de conexión en los últimos 3 meses, si estaban conforme con el servicio y por qué zona del edificio de la unidad académica frecuentaba mayormente, con el objetivo de realizar un mapeo de zonas críticas y trabajar sobre esos sectores.

Los datos obtenidos a partir de las encuestas durante el año pasado, fueron muy importantes, ya que permiten conocer el grado de satisfacción que la comunidad tenía sobre el acceso libre a internet, conocer a través de qué dispositivos se conectaban, por qué lugar del edificio frecuentaban y también se permitió un campo libre de expresión para dejar comentarios sobre la conectividad. El resumen de estos resultados se muestran las Figuras 1, 2, 3 y 4. Éstos fueron considerados el principal input del programa **“RedCoFi: Wifi al alcance de todos.”**

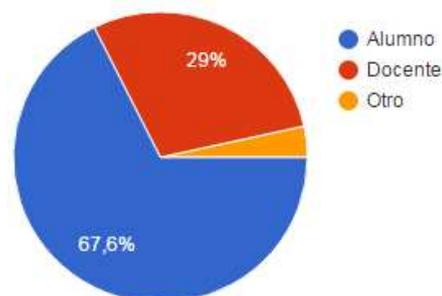


Figura 1: Resultados de la pregunta ¿Cuál es tu desempeño en la Facultad?

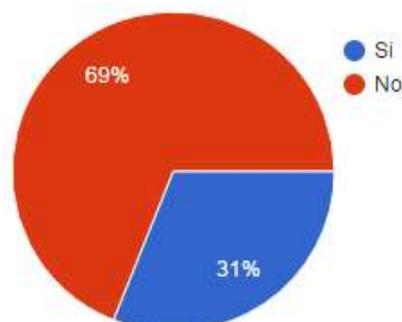


Figura 2: Resultados de la pregunta ¿Has notado mejora en la conexión a Internet en los últimos 3 meses?



Figura 3: Resultados de la pregunta ¿Por qué zona de la Facultad es la que mayor tiempo frecuentas? Se ponen algunos sitios de referencia.

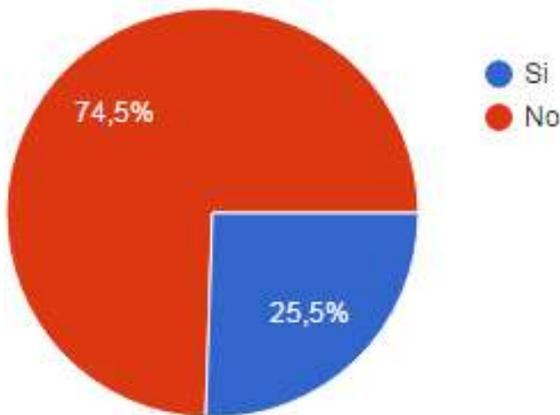


Figura 4: Resultados de la pregunta ¿Estás conforme actualmente con la conectividad a Internet?

Se puede observar en la figura 3, que el 35,2% de las personas encuestadas circulan frecuentemente por la Planta Baja, en el ala de las aulas 2/3/4, Gabinete 2/3/4, Sala de Profesores. Y otro 20,7% de personas frecuenta por la Planta Alta, en el ala de Aula Magna, Aulas 5a/5b/6, Gabinete 9/10. De estos resultados, también se puede observar que a fines del 2016 el 74,5% de las personas censadas no estaba conforme con el servicio de Internet.

Con este conjunto de datos, en el mes de Enero 2017, se ubicó estratégicamente los puntos de acceso para proveer internet wifi público y lograr el mayor alcance posible. En

la figura 5 se observa la distribución propuesta.

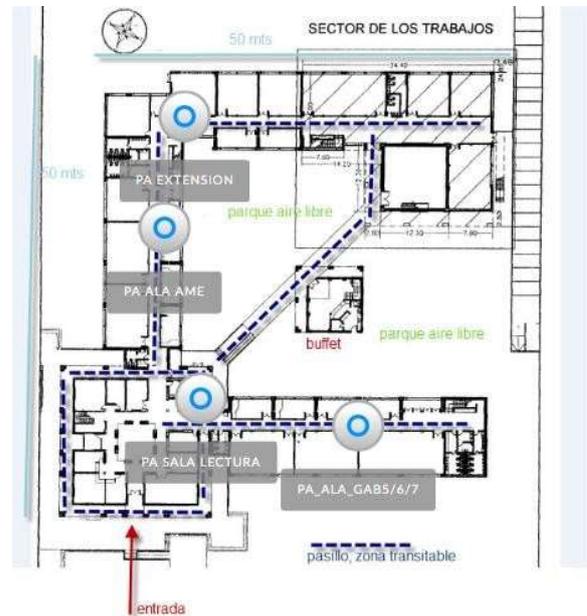


Figura 5: Distribución de los Access point dentro del Edificio.

En las figuras se aprecia claramente que los Access Point que mayor uso tiene son aquellos que están ubicados en Planta Alta, Ala Gab/5/6/7 y en PA Ala A, los cuales fueron ubicados estratégicamente en función de los resultados obtenidos de las encuestas.

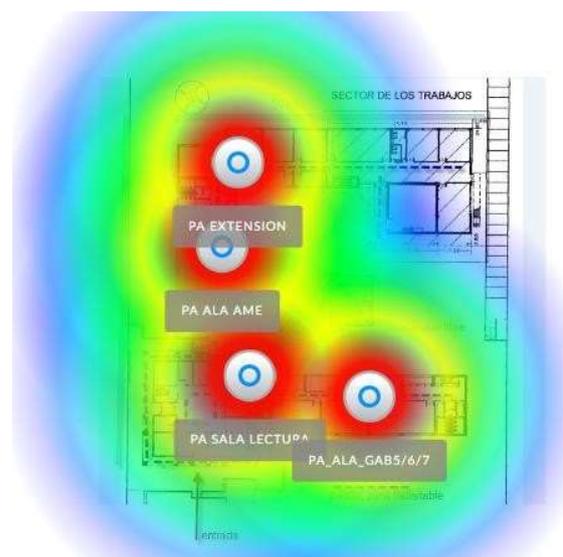


Figura 6: Cobertura de los Access Point.

Finalmente en la figura 7 se puede observar como a través de nuestro controlador de la red Wifi Pública, obtenemos información sobre el consumo de la red y corroboramos la información obtenida en las encuestas realizadas.

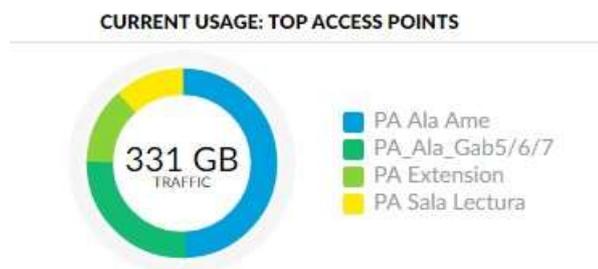


Figura 7: Tráfico de datos según Access point.

Siguiendo en esta línea de mejoramiento; el Programa **“RedCoFi: Wifi al alcance de todos”**, del Área de Sistemas Informáticos del IIT&E, prevé un plan de actividades para el corriente año, el cual tiene como objetivo extender este servicio a toda la Unidad Académica, garantizando conectividad y seguridad a Internet a toda la comunidad. Asegurando legitimidad de la comunicación, consistencia en los datos, disponibilidad de red e identificación de los nodos conectados de manera confidencial.

Formación de Recursos Humanos

La presente línea se integra al programa de investigación, liderado por el Dr. Ing. Oscar Pascal, a través del programa “Enseñanza de la Ingeniería. Desarrollo y Evaluación de Modelos Estrategias y Tecnologías para Mejorar los indicadores Académicos y la Eficiencia Organizacional”, bajo la dirección del Esp. Marcelo Estayno. Cabe destacar que el programa integra cuatro proyectos y en ellos se desempeñan docentes investigadores, personal técnico, personal de apoyo y seis becarios CIN; todos contribuyendo a la multidisciplinariedad.

Referencias

Aransay, A. S. (2009). Seguridad en Wi-Fi.

Campoli, O., Minnaard, C., Morrongiello, N., & Pascal, G. (2013). Impacto de las Redes Sociales en ámbitos académicos universitarios. Indicadores de aceptabilidad en docentes y estudiantes de ingeniería. Exposición presentada en el Congreso Argentino de Ingeniería Industrial (COINI) realizado en la ciudad de San Rafael, Mendoza, Argentina durante el.

IEEE Standards Association (2017), URL <http://www.standards.ieee.org/>

Graciela Martínez Gutierrez, Edgar Josué García Lopez, Jessica García Ramirez. Universidad Iberoamericana León, 2015, ISSN: 2007-531.

Mendoza, C. M. H., Vidal, L. M. R., & Almanza, M. A. (2017). Análisis de seguridad en redes inalámbricas de las MiPyME y propuesta de mejora. *Revista Iberoamericana de Producción Académica y Gestión Educativa*

Mitchell, C. H. J. C. (2005). Security Analysis and Improvements for IEEE 802.11 i. In *The 12th Annual Network and Distributed System Security Symposium (NDSS'05) Stanford University, Stanford* (pp. 90-110)

Wireless Fidelity Alliance (2017), URL <http://www.wi-fi.org>

Temporal Key Integrity Protocol (TKIP), URL http://en.wikipedia.org/wiki/Temporal_Key_Integrity_Protocol

Seguridad en Entornos BPM: Firma Digital y Gestión de Clave

Patricia Bazán , Paula Venosa, Nicolas Macia, Ivan Grcevic
LINTI, Facultad de Informática, UNLP

1. Resumen

Los sistemas de gestión de procesos de negocio (en inglés BPMS - *Business Process Management Systems*) se están estableciendo como las soluciones de IT adoptadas por las organizaciones actuales que buscan transformar su visión funcional clásica por un enfoque basado en procesos de negocio. Un proceso representa las tareas que la organización debe realizar para producir sus productos, el orden de ejecución de las mismas y las personas responsables de realizarlas.

En este sentido los BPMS, vistos como sistemas de información, deben incorporar mecanismos para garantizar los atributos de seguridad de autenticidad, confidencialidad, integridad, no repudio y disponibilidad.

La línea de investigación presentada propone incorporar conceptos de firma digital a un BPMS y a su vez, aplicar BPM a la gestión de seguridad, en particular al proceso de gestión de claves y su aplicación para la firma digital.

Palabras clave: Firma digital – BPMS – Seguridad — SGSI

2. Contexto

En el LINTI (Laboratorio de Investigación en Nuevas Tecnologías) [1] un grupo de docentes investigadores se dedican a estudiar temas relacionados a seguridad y privacidad en redes, en particular, lo referente a certificados digitales, su gestión y aplicación. Por otra parte, otro grupo aborda la línea de trabajo vinculada a metodología BPM, en particular, sus

soluciones tecnológicas.

En esta propuesta se entrelazan ambas líneas de trabajo desde un punto vista de aplicación práctica incorporando firma digital en procesos de negocio que, por su criticidad, demanden altos requisitos de autenticidad, no repudio e integridad.

3. Introducción

Los procesos de negocios como concepto clásico modelan y ejecutan tareas repetitivas y estructuradas de una organización a fin de producir los productos (o servicios) ofrecidos por la misma. La gestión de procesos de negocio se basa en la idea de que cada producto es el resultado de un conjunto de actividades que se realizan a fin de obtener dicho producto [5]. Por este motivo, la correcta y eficiente gestión de los procesos de negocio es un aspecto importante para la productividad de toda organización, ya que permite identificar las tareas que la misma debe realizar para producir sus productos, el orden de ejecución de las mismas y las personas responsables de realizarlas.

Los BPMS constituyen la herramienta tecnológica para alcanzar este objetivo y que puede enriquecerse incorporando los conceptos de firma digital [4], que garantiza que los datos han llegado al sistema tal como fueron ingresados (integridad), que el usuario que completó la tarea es realmente quien dice ser (autenticidad) y que no pueda evadir responsabilidades sobre la tarea completada (no repudio).

De esta manera no sólo se aplica la firma digital para proteger documentos como mensajes, facturas, transacciones

comerciales, notificaciones, decretos; sino que cualquier tarea completada, desde ingresar datos de una venta hasta aprobar un presupuesto, podría estar digitalmente firmada.

Como ejemplo, consideremos una compañía de seguros que ha incorporado la gestión por procesos. Los productores tratan cada caso de siniestro y los supervisores deberán aprobar ciertas solicitudes creadas por los productores. A su vez, los supervisores pueden derivar ciertos casos a los abogados especialistas y peritos. Estos últimos evaluarán los casos críticos y aprobarán ítems de las solicitudes de cobertura ante el siniestro. Este tipo de proceso tiene distintos niveles de seguridad para cada tarea. Los productores tienen requerimientos de seguridad de **autenticidad** e **integridad**, mientras que para los supervisores podemos agregar el requerimiento de **no repudio**. Si un supervisor dice no haber aprobado una solicitud, puede alegar la posibilidad de un fraude producido por un atacante externo y eludir responsabilidades sobre sus acciones. Similar situación se plantea para las tareas realizadas por los abogados especialistas y peritos.

Los BPMS actuales incluyen sólo un mecanismo de autenticación a través del inicio de sesión web con nombre de usuario y contraseña, que permite garantizar la autenticidad del usuario que inicia sesión. Esto cubre un aspecto muy recortado de los servicios de seguridad enunciados.

3.1- Seguridad en BPMS

La propuesta de incorporar seguridad en BPMSs se aborda desde dos enfoques: 1- Incorporar firma digital en las tareas para autenticidad, no repudio e integridad, 2- definir un proceso específico para la gestión administrativa de claves y certificados asociados.

3.1.1- Firmado de tareas

Hoy en día muchas aplicaciones

incorporan las firmas digitales para garantizar autenticidad y no repudio del autor o emisor, e integridad de los datos firmados. Si bien un sistema BPM podría incorporar la funcionalidad de firmar documentos o mensajes, como lo hacen las aplicaciones de correo electrónico o los editores de pdfs, ésto no sería una verdadera integración de firma digital con BPM, sino una integración más con un servicio dentro de un proceso de BPM común y corriente.

La parte central de una integración de firma digital para enriquecer la seguridad de una organización que implementa BPM, es la firma de tareas.

En el caso de la compañía de seguros presentado en la sección anterior, con el uso de una conexión segura a través de HTTPS, el usuario del BPMS puede confiar en la autenticidad del servidor y en la confidencialidad e integridad de los datos que el servidor le hace llegar.

En esta relación de confianza, supongamos que un **supervisor A** está realizando una tarea y el BPMS informa que un determinado **usuario B** ha realizado otra tarea importante previamente. El protocolo HTTPS garantiza que la información no fue modificada durante la transmisión por la red. Ahora bien: ¿Puede el servidor estar seguro de que el **usuario B** realizó esa tarea?

La posibilidad de **firmar tareas** con una **clave privada** incrementa en forma significativa la seguridad brindada en torno a la **autenticidad y no repudio** del usuario y sus acciones [7].

Una vez que incorporamos un mecanismo de estas características, podemos decir que el servidor tendrá una certeza mucho mayor de la autenticidad del usuario. En ese caso el **supervisor A** puede confiar realmente, porque no sólo sabe que lo que le llega del servidor es certero y auténtico, también confía en que lo que previamente

llegó de otro cliente -el **usuario B**- al servidor era certero y auténtico.

En este sentido se define un subproceso abstracto para cada tarea que requiera firma digital y que denominaremos **Firma de Tarea**, cuyos participantes: el actor que ejecuta la tarea, la aplicación cliente que se ejecuta en el navegador por HTTPS, y el BPMS. Éste último verificará la firma basándose en la clave pública del usuario en cuestión.

Este proceso abstracto es un modelo en BPMN que luego se implementa con funcionalidades específicas del BPMS Bonita BPM [9].

3.1.2 Proceso de gestión de claves

La incorporación de firma digital a cualquier sistema de información conlleva realizar un conjunto de pasos administrativos para la gestión de claves y certificados que puede abordarse como un proceso adicional, no específicamente vinculado al negocio, pero que es compartido por todos los sistemas de información que incluyan firma digital. Estos nuevos procesos formarán parte del Sistema de Gestión de Seguridad de la Información (SGSI) [3].

Las principales actividades relacionadas con la gestión de claves dentro de una

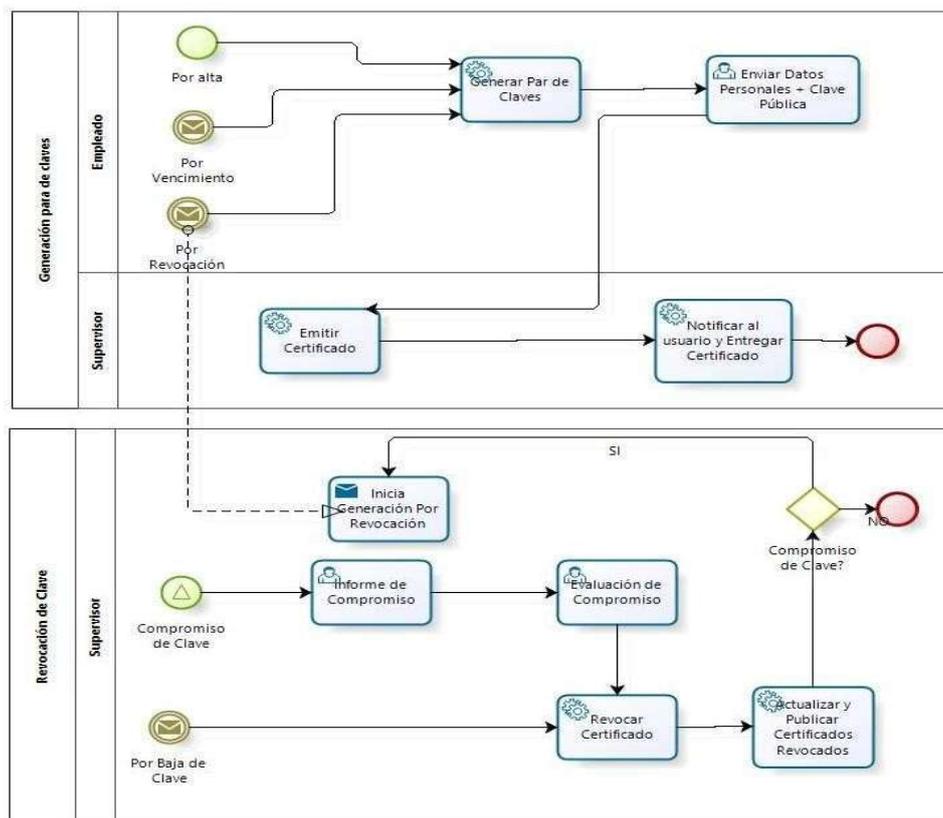


Figura 1 – Procesos de Generación y Revocación de Claves [11]

infraestructura de firma digital son: 1- la generación y distribución de claves, 2- la revocación debido a claves comprometidas o a baja de usuarios. Estas actividades en parte pueden modelarse como un subproceso dentro de un proceso de negocio de Contratación de Personal o bien de Desvinculación del mismo.

Estos subprocesos son **Generación de par de claves** y **Revocación de claves**.

El subproceso de **Generación de par de claves** como se muestra en la Figura 1 tiene tres eventos de inicio posibles: 1- Por alta, se inicia el proceso como un subproceso de Contratación e Integración de Personal, 2- Por expiración, cuando una clave generada previamente alcanza su fecha de vencimiento y es necesario gestionar al usuario un nuevo par de claves y 3- por revocación de la clave de un usuario debido a un compromiso de su clave anterior.

El subproceso de **Revocación de clave** como se muestra en la Figura 1 tiene dos eventos de inicio posibles: 1- compromiso de la clave utilizada y 2- por baja de un empleado. El evento de compromiso se dispara cuando cualquier empleado informa sus sospechas de que su clave o la de otro empleado ha sido comprometida. Luego de evaluar si corresponde revocar la clave, se procede al subproceso de **Generación de par de claves**.

4. Líneas de Investigación, Desarrollo e Innovación

La incorporación de mecanismos para garantizar los atributos de seguridad en la ejecución del proceso de negocio, conlleva a mejorar las soluciones basadas en gestión por procesos, equiparándolas con cualquier sistema de información seguro.

Por otra parte, la gestión de procesos de negocio constituye una buena metodología para instrumentar infraestructuras de gestión de claves públicas y privadas.

Para alcanzar este objetivo, se trazan puntos en común entre las líneas de investigación vinculadas a modelado y desarrollo de sistemas de información basados en procesos de negocio con la línea de seguridad y privacidad en redes, ambas enmarcadas dentro del LINTI (Laboratorio de Investigación en Nuevas Tecnologías Informáticas) de la UNLP.

5. Resultados y Objetivos

El objetivo de este trabajo consiste en realizar una investigación en los campos de BPM y la firma digital, centrada en el análisis de la integración de estas dos tecnologías, con el fin de que la firma digital enriquezca los componentes de BPM y a su vez los conceptos de BPM se puedan aplicar al sistema de gestión de seguridad, en particular al proceso de gestión de claves y su aplicación para la firma digital [8].

Entre los principales resultados obtenidos en la incorporación de firmas digitales en un entorno BPM cabe resaltar el aporte brindado a los BPMS incorporando el uso de firma digital de manera nativa.

Por otra parte, la distribución segura de claves, así como la detección de claves comprometidas y su necesidad de revocación, resultan aspectos críticos para las organizaciones que de hecho se ven reguladas por normas internacionales [3][6] y cuya estructuración por procesos le brinda un mayor formalismo y agilidad en el uso.

6. Formación de Recursos Humanos

BPM y la mejora continua de procesos de

negocio aplicada a los ambientes de ejecución, es una línea de trabajo que ha formado docentes e investigadores en torno a la solución de problemas reales. Por su parte la línea de seguridad y privacidad en redes, particularmente en temas relacionados a criptografía y sus aplicaciones, reúne el interés de docentes y alumnos que se traduce en la realización de varias tesis, 3 entre el año 2016 y el 2017, la continuidad desde el año 2007 del proyecto PKIGrid UNLP [2], infraestructura de clave pública para eficiencia y, en consecuencia, la participación de docentes e investigadores del LINTI en la comunidad de TAGPMA[10] en forma activa.

7. Referencias

- [1] <http://www.linti.unlp.edu.ar/>
- [2] www.pkigrd.unlp.edu.ar
- [3] Disterer, G. (2013). ISO/IEC 27000, 27001 and 27002 for information security management.
- Figura 1 - Procesos de Gestión de pares de claves*
- [4] William Stallings. Cryptography and Network Security Principles and Practices. Prentice Hall. 5th Edition.
- [5] Mathias Weske. Business Process Management. Concepts, Languages, Architectures. Springer Berlin Heidelberg New York. 2007
- [6] Chokhani, S., Ford, W., Sabett, R., Merrill, C., & Wu, S. „RFC 3647: Internet X. 509 Public Key Infrastructure Certificate Policy and Certification Practices Framework “, November 2003. *Obsoletes RFC2527*.
- [7] Jutta Mülle, Silvia von Stackelberg and Klemens Böhm. Modelling and Transforming Security Constraints in

Privacy-Aware Business Processes. IEEE International Conference on Service-Oriented Computing and Applications. 2011

[8] Jens Müller and Klemens Böhm. The Architecture of a Secure Business- Process-Management System in Service- Oriented Environments. Ninth IEEE European Conference on Web Services. 2011.

[9] <http://es.bonitasoft.com/>

[10] <http://tagpma.es.net/>

[11] <http://www.bizagi.com/es>

Seguridad en Redes las Industriales: Clave para la Ciberdefensa de las Infraestructuras Críticas

Jorge Kamlofsky¹, Samira Abdel Masih¹, Hugo Colombo¹, Daniel Veiga¹, Eugenio Costa¹, Claudio Milio¹, Marcelo Semería¹, Pedro Hecht²

¹ CAETI - Universidad Abierta Interamericana Av.
Montes de Oca 725 – Buenos Aires – Argentina
{Jorge.Kamlofsky, Samira.Abdel Masih, Hugo.Colombo, Daniel.Veiga}@uai.edu.ar

² Universidad de Buenos Aires, Facultades de Ciencias Económicas, Ciencias Exactas y Naturales e Ingeniería. Maestría en Seguridad Informática, Buenos Aires, Argentina phecht@dc.uba.ar

Resumen

Los procesos de producción industrial a gran escala se automatizan mediante los sistemas de control industrial. Por su robustez y efectividad, estos sistemas también fueron adoptados en la automatización de las infraestructuras críticas de las naciones: plantas de tratamientos de líquidos, distribución de energía, siderúrgicas y demás. La seguridad se basa en su aislamiento. Los nuevos requerimientos de mayor flexibilidad y eficiencia promueven su conexión con las redes corporativas, dejando expuestas sus vulnerabilidades a gran cantidad de amenazas provenientes de estas últimas.

En este proyecto se estudian las principales vulnerabilidades de estos sistemas, y se analizan y se desarrollan soluciones: algunas basadas en mejoras de procesos y herramientas tradicionales y otras basadas en esquemas compactos de criptografía adaptables a los dispositivos de las redes industriales.

Palabras clave: Seguridad en Redes Industriales, Seguridad en SCADA,

Criptografía compacta, criptografía en PLCs, Infraestructuras críticas, ciberdefensa en redes industriales.

Contexto

Los proyectos radicados en el CAETI¹ se clasifican en cinco líneas de investigación. Este proyecto se enmarca dentro la línea de Seguridad Informática y Telecomunicaciones.

Se pretende obtener conocimiento teórico y desarrollar e implementar soluciones que permitan mejorar la situación de vulnerabilidad de estas redes lo que permitiría otorgar ciberseguridad a las infraestructuras críticas.

Introducción

Los Sistemas de Control Industrial (ICS de sus siglas en inglés) son redes de telemando y telecontrol de procesos compuestos por autómatas industriales llamados PLC (del inglés: Programmable

¹ CAETI: Centro de Altos Estudios en Tecnología informática, dependiente de la Facultad de Tecnología informática de la UAI.

Logic Controller) interconectados entre sí y cada uno de ellos a sensores digitales o analógicos (caudalímetros, sensores de nivel, de temperatura, microswitches, etc.) y/o a actuadores (motores, válvulas, llaves, etc.). Fueron diseñados para supervisar y actuar en los procesos industriales. El aislamiento del proceso de producción, dio por muchos años una sensación de seguridad ilusoria gracias al ocultamiento [1, 2]. En la tecnología industrial, la prioridad siempre fue el proceso, y no la seguridad.

Los ICS son muy robustos, y por ello se los utiliza en sistemas que requieren uso continuo y permanente. Están presentes en plantas de potabilización de agua, producción y distribución de energía, transporte, telecomunicaciones, siderúrgicas, entre otras: están en infraestructuras críticas de naciones.

Los ICS se controlan desde el SCADA. Un SCADA (por sus siglas en inglés: Supervisory Control and Data Acquisition) es software que muestra gráficamente el estado de cada componente de la planta, con la posibilidad de actuar sobre ellos. Se diseñaron para controlar sistemas industriales, conectando PCs con las redes de autómatas industriales; conformando la interfaz hombre máquina.

Originalmente los SCADA se instalaban en salas aisladas, con acceso restringido. Con el tiempo surgió la necesidad de vincularlos a la red corporativa e incluso a internet. Y esta tendencia es inevitable. Su interconexión dejó a los ICS expuestos a amenazas y riesgos provenientes desde el exterior, los que suponen serias consecuencias [3].

Hoy es posible, mediante dispositivos móviles, controlar un ICS, desde cualquier lugar del mundo con cobertura de red móvil [4], suponiendo un escenario ideal para explotar vulnerabilidades e inyectar malware. La tecnología

corporativa y la industrial dejaron a la seguridad entre ambas [5].

Hasta hace pocos años, era impensable que un ICS se pudiera infectar con virus informático. En el año 2010 el sistema SCADA de las plantas de enriquecimiento de uranio de Irán fue atacado por un virus llamado Stuxnet. Esto desconcertó a analistas estratégicos de todo el mundo. La comunidad internacional mostró gran preocupación por la seguridad de las infraestructuras basadas en estas tecnologías [6 – 8], y se encuentra trabajando en soluciones [9 – 12].

En el ámbito de las tecnologías corporativas se tiene experiencia en Seguridad. Las recomendaciones de las normas ISO27000 [13] y NIST SP800-30 revisión 1 [14] y los múltiples desarrollos realizados, ayudan a proteger la seguridad de los activos informáticos. La criptografía es clave para asegurar sistemas informáticos. Es posible dar seguridad criptográfica a dispositivos con baja capacidad de cómputo gracias al desarrollo de algoritmos criptográficos de clave pública basada en estructuras algebraicas de anillos no conmutativos [15-17] la cual a fecha actual es inmune a ataques cuánticos y esquemas simétricos compactos como el presentado en [18].

Este proyecto pretende desarrollar soluciones en las redes industriales: en los SCADA usando avances en la seguridad de redes corporativas, y en la profundidad de la red industrial mediante soluciones criptográficas basadas en álgebra no conmutativa integrándolas con esquemas simétricos compactos.

Líneas de Investigación, Desarrollo e Innovación

El equipo de investigación trabaja en dos ramas: matemático-criptográfica y redes-sistemas.

La rama matemático-criptográfica trabaja estudiando las estructuras de anillos no conmutativos y no asociativos y su posibilidad de aplicarlos criptográficamente como sistema de intercambio seguro de claves. Las variantes generadas se programan y se las pone a prueba en ambientes de simulación controlados.

La rama de redes-sistemas se encuentra estudiando las normas de seguridad en sistemas de información más importantes y los protocolos de comunicaciones intervinientes con la intención de lograr implementar los algoritmos generados. Se estudian las vulnerabilidades más frecuentes en estos sistemas.

Resultados y Objetivos

La rama matemático-criptográfica ha logrado implementar el protocolo presentado en [15] y ha llegado a mejoras en tiempos de ejecución usando cuaterniones [16], en diferentes conjuntos numéricos [17]. También se analiza el funcionamiento del algoritmo de Shor para computación cuántica [19]

La rama de Redes-Sistemas analizó ataques a infraestructuras críticas publicados en [20] y ha propuesto un enfoque para disminuir los efectos de ciberataques [21]. Hoy se encuentra analizando en detalle las normas ISO27000 [13], la NIST [14] y NIST específica de los ICS [22].

El objetivo final del proyecto es el desarrollo de soluciones de Seguridad que puedan implementarse en las redes de los ICS. El problema en cuestión es crítico y se encuentra latente en toda la infraestructura crítica e industrial del mundo.

Se espera lograr transferencia de resultados a la comunidad y a la industria.

La transferencia a la comunidad se logrará mediante cursos de extensión universitaria. La transferencia a la industria consistirá en el patentamiento de métodos y algoritmos que permitirán incrementar la seguridad de las redes industriales.

Formación de Recursos Humanos

El proyecto está dirigido por el Esp. Lic. Jorge Kamlofsky y la Dra. Samira Abdel Masih. Tiene la colaboración especial del Dr. Pedro Hecht. Integran el proyecto los siguientes docentes de la Facultad de Tecnología Informática de la UAI: el PhD. Hugo Colombo, los ingenieros Marcelo Semería, Eugenio Costa, Claudio Milio y el Lic. Daniel Veiga.

El equipo de investigación se completa con alumnos de la Facultad de Tecnología Informática de la UAI: Matías Sliafertas, Juan Manuel Pedernera, Oscar Morales, Pablo Oviedo, Jesica Valente, Christian Martin, Damián Romero. Daniel Sola, Enrique Belaustegui, Facundo Coronel, Federico Romero, Federico Tabarez Rosa, Fernando Ribas, Matías iacobuzio, Sandra Biondini, Joan Mutti Ferreyra, Andrés Perez, Angel Orlauskas, Rodrigo Gomez, Nicolás Mayer, Maximiliano Ríos, Nicolás Carella, Nicolás Salas, Gastón Suarez y Montserrat Patiño.

Los docentes integrantes del proyecto adquieren conocimientos y técnicas de seguridad que pueden complementar los conocimientos por ellos enseñados en las diferentes materias que integran.

Gran parte de los alumnos que se desempeñan como auxiliares de investigación inician experiencias en la investigación científica adquiriendo la correspondiente metodología, en una temática que resulta ser muy atractiva. Otros se encuentran promediando la

carrera, y el conocimiento adquirido se incorporará en sus trabajos finales de carrera. En particular Oscar Morales, Jesica Valente y Pablo Oviedo están finalizando las tesis de final de la carrera Licenciatura en Matemática.

Referencias

- [1] Courtois, N. *The dark side of security by obscurity, and Cloning MiFare Classic Rail and Building Passes Anywhere, Anytime*. IACR Cryptology ePrint. Archivo 2009: 137, 2009.
- [2] Menezes, A., Van Oorschot, P., and Vanstone, S. *Handbook of applied cryptography*. CRC press, 1996.
- [3] Sánchez P. *Sistema de Gestión de la Ciberseguridad Industrial* [En línea]. Universidad de Oviedo, (2013). [Consulta: 11/02/15]. Disponible en: <<http://dspace.sheol.uniovi.es/dspace/bitstream/10651/17741/1/TFM%20-%20PABLO%20SANCHEZ.pdf>>.
- [4] Opto 22, *Press Release: Updates groov to Easily Connect Modbus/TCP Devices with Smartphones and Tablets* [En línea], (2015). Disponible en: <http://www.modbus.org/member_docs/OPTO22-Jan2015.pdf> [Consulta: 14/08/2015].
- [5] Carrasco Navarro, O. y Villalón Puerta, A. *Una visión global de la ciberseguridad de los sistemas de control*. Revista SIC: ciberseguridad, seguridad de la información y privacidad 106, (2013), pp. 52-55.
- [6] Veramendi, R. *Ataques a la Seguridad Informática y Telecomunicaciones en el Contexto Internacional*. Revista del Instituto de Estudios Internacionales IDEI-Bolivia, 45(2), (2012), pp. 4-11.
- [7] Vazquez, S. *Ciberseguridad en Paraguay*. III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, 2015.
- [8] Corvalan, F. *Seguridad de Infraestructuras Críticas: Visión desde la Ciberdefensa*. III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, 2015.
- [9] Blackmer, M. *Cibersecurity for Industrial Control Networks*. III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, 2015.
- [10] Simoes, P., Cruz, T., Proença, J. and Monteiro, E. *Honeypots especializados para Redes de Control Industrial*. VII CIBSI. Panamá, 2013.
- [11] Arias, D. *Seguridad en Redes Industriales*. Trabajo Final, Universidad de Buenos Aires, 2013.
- [12] Paredes, I. *La protección de infraestructuras críticas y ciberseguridad industrial*. Red seguridad: revista especializada en seguridad informática, protección de datos y comunicaciones 62, (2013), pp. 49.
- [13] ISOTools, *ISO 27001* [En línea], (2015). Disponible en: <<https://www.isotools.org/normas/riesgos-y-seguridad/iso-27001>>. [Fecha de consulta: 14 de Agosto de 2015].
- [14] NIST. *Special Publication 800 – 30, revision 1: Information Security*. National Institute of Standards and Technology, U.S. Department of Commerce, [En línea] (2012). Disponible en: <<http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf>> [Consulta: 8/3/2017].
- [15] Hecht J. *Un modelo compacto de criptografía asimétrica empleando anillos no conmutativos*. V CIBSI, Montevideo, 2009.
- [16] Kamlofsky J., Hecht J., Abdel Masih, S. Hidalgo Izzi, O. *A Diffie Hellman compact model over commutative rings using quaternions*. VIII CIBSI, Quito, 2015.
- [17] KAMLOFSKY, Jorge. *Improving a Compact Cipher Based on Non Commutative Rings of Quaternion*. XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), 2016.
- [18] Castro Lechtaler, A., Cipriano, M., García, E., Liporace, J., Maiorano, A., y Malvacio, E.. *Model design for a reduced variant of a Trivium Type Stream Cipher*. Journal of Computer Science & Technology, 14.
- [19] Shor, Peter W. *Algorithms for quantum computation: Discrete logarithms and factoring*. Foundations of Computer Science,

1994 Proceedings., 35th Annual Symposium on. IEEE, 1994.

[20] Security Incidents Organization, *RISI: The Repository of Industrial Incidents* [En línea], (2015). Disponible en: <<http://www.risidata.Com / Database>> [Consulta: 14/08/2015].

[21] Kamlofsky J, Colombo H, Sliafertis M y Pedernera J, *Un Enfoque para Disminuir los Efectos de los Ciber-ataques a las Infraestructuras Criticas*. III CONAISI, Buenos Aires, 2015.

[22] NIST. *Special Publication 800 – 82, revision 2: Guide to Industrial Control Systems (ICS) Security*. National Institute of Standards and Technology, U.S. Department of Commerce, [En línea] (2015). Disponible en: <<http://nvlpubs.nist.gov / nistpubs / Legacy / SP / NIST.SP.800-82r2.pdf>> [Consulta: 8/3/2017].

Seguridad en Servicios Web

Edgardo Bernardis, Hernán Bernardis, Mario Berón, Germán Montejano

Departamento de Informática
Facultad Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis

Ejército de los Andes 950 – San Luis – Argentina
{ebernardis, hbernardis, mberon, gmonte}@unsl.edu.ar

Resumen

Con el auge de internet y las distintas dinámicas de la sociedad actual, ha cambiado en gran medida la forma de interactuar entre las personas y las empresas. Este cambio notable se observa en la forma de intercambiar información entre los distintos actores. Este intercambio se vuelve de particular interés siendo blanco de ataque por parte de todos aquellos actores que quieren obtener información útil y valiosa a sus propios intereses o de terceros. Es aquí donde cobra particular relevancia implementar todo tipo de medidas y acciones tendientes a evitar estos ataques, por tal motivo surge lo que se denomina Seguridad Informática.

En este artículo se describe una línea de investigación cuyo principal objetivo es el desarrollo de métodos, técnicas y estrategias orientadas a incrementar el nivel de seguridad de Servicios Web.

Palabras clave: Aplicaciones Web, Servicios Web, Seguridad Informática, Seguridad de la Información.

Introducción

Con los avances de la tecnología, sobre todo en el ámbito de internet, se vuelve sumamente importante y necesario la protección de todo tipo de información. En la actualidad, es realmente alta la cantidad de delitos que se llevan adelante en contra de

información personal o de empresas. Todo tipo de información es valiosa, ya sea desde simples datos personales hasta sistemas y bases de datos empresariales.

Con el auge de internet, el intercambio de archivos se ha vuelto un punto esencial en la sociedad actual. Los consumidores intercambian información no sólo entre ellos sino también con los vendedores. Para el intercambio de información se utilizan diferentes medios entre los que se pueden mencionar redes sociales, correo electrónico, sistemas punto a punto, pagos online, juegos. Todo esto se fundamenta en la confianza y el correcto funcionamiento del software y del hardware subyacente a dicho proceso.

Por lo antes mencionado es que surge lo que se conoce como *Seguridad Informática* (SI). Existen diversas definiciones de SI, algunas más extensas que otras, pero todas enfocadas en los mismos aspectos comunes. A los fines de este trabajo, se adhiere a la siguiente definición de SI: *Preservación de la confidencialidad, integridad y disponibilidad de la información en el Ciberespacio. A su vez, el Ciberespacio se define como el entorno complejo que resulta de la interacción de las personas, software y servicios en Internet por medio de redes y dispositivos tecnológicos conectados a el, y que no existe en ninguna forma física* [1].

La información es un conjunto organizado de datos, que cambia su enfoque y su estado de conocimiento dependiendo del ámbito en

la que se la utilice. Por ejemplo, si la información se conceptualiza bajo el punto de vista de la ingeniería: *Estudio de las características y estadísticas del lenguaje que permite su análisis desde un enfoque matemático, científico y técnico.*

Desde el punto de vista de una empresa: *Conjunto de datos propios que se gestionan y mensajes que se intercambian personas y/o máquinas dentro de una organización* [2].

La información se ve afectada por muchos factores, motivo por el cual se vuelve importante su seguridad. De aquí que *Seguridad de la Información* es: *una disciplina, cuyo principal objetivo es mantener el conocimiento, datos y sus significados libres de eventos indeseables, tales como el robo, espionaje, daños, amenazas y otros peligros. La Seguridad de la Información incluye todas las acciones tomadas con anticipación, para evitar eventos no deseados* [3].

El objetivo de la SI es obtener un nivel aceptable de seguridad, entendiéndose por aceptable un nivel de protección suficiente para que la mayor parte de potenciales intrusos, interesados en los equipos con información de una organización o persona, fracasen en cualquier intento de ataque contra los mismos. Asimismo, se encarga de establecer los mecanismos para registrar cualquier evento fuera del comportamiento normal y tomar las medidas necesarias para reestablecer las operaciones críticas a la normalidad [4].

El punto o centro de ataque a la seguridad informática se da en una *Vulnerabilidad: debilidad de un activo o control que puede ser explotado por una o más amenazas* [5]. La presencia de una vulnerabilidad no puede causar daño en sí misma, ya que es necesario que exista una amenaza que la aproveche. Una vulnerabilidad que no tiene una amenaza, puede no requerir la aplicación de un control, pero debe ser reconocida, supervisada por si tiene cambios y, en lo posible, eliminada.

Las amenazas surgen a partir de la existencia de vulnerabilidades, es decir que una amenaza sólo puede existir si existe una vulnerabilidad que pueda ser aprovechada,

independientemente de que se comprometa o no la seguridad de un sistema de información. Una amenaza se puede definir como: *Cualquier elemento o acción que es capaz de aprovechar una vulnerabilidad y comprometer la seguridad de un sistema de información* [6]. Las amenazas se pueden clasificar o dividir en dos tipos; las intencionales, en caso de que deliberadamente se intente producir un daño (por ejemplo el robo de información). Las no intencionales, en donde se producen acciones u omisiones de acciones que si bien no buscan explotar una vulnerabilidad, ponen en riesgo los activos de información y pueden producir un daño (por ejemplo las amenazas relacionadas con fenómenos naturales).

Dentro de las diferentes áreas y enfoques que se pueden llevar adelante, en lo que respecta a seguridad informática, existe uno particularmente interesante y de gran crecimiento como lo son las aplicaciones que ejecutan en la web. Estas aplicaciones son de lo más utilizado en la web por parte de los usuarios, debido a su versatilidad, beneficios, fácil utilización, comodidad, etc. Existe una gran variedad y cantidad de las mismas; desarrolladas en diferentes tecnologías y lenguajes de programación. Todas contienen información formal (código fuente) e informal (identificadores, comentarios, documentación, etc.). Es a partir de estos tipos particulares de información que se puede detectar y medir el nivel de comprensibilidad de la misma y a través de esto, aumentar o disminuir su seguridad.

La organización de este artículo se expone a continuación. La sección 2 describe la línea de investigación y desarrollo abordada. La sección 3 presenta los resultados obtenidos hasta el momento, junto con todos aquellos esperados a corto plazo. Finalmente, la sección 4 describe las tareas realizadas por los recursos humanos en formación.

Líneas de Investigación y Desarrollo

En las sub-secciones siguientes se describen de manera concisa la línea de investigación presentada en este trabajo.

Técnicas para Medir el Nivel de Comprensibilidad de las Aplicaciones Web

Una de las formas de analizar si un software es o no seguro, se puede llevar a cabo a través de la estimación de su nivel de comprensibilidad. Entendiendo tal concepto como una medida que establece a priori la facilidad de entender la tarea que un sistema de software lleva a cabo. Es de suponer que mientras más fácil sea de comprender más fácil será de vulnerar. Por lo antes mencionado, se estudian métodos y estrategias que permitan medir la comprensibilidad con alto nivel de precisión.

En este contexto, el grupo de investigación está centrado en la elaboración de estrategias que permitan medir a priori o a posteriori el nivel de comprensibilidad de Servicios Web (SW). Dicho enfoque se lleva a cabo debido al crecimiento, interoperabilidad y ventajas que ofrecen los SW, lo que los vuelve un blanco atractivo por aquellos actores interesados en robar o manipular la información valiosa que pueden contener.

Visualización de Vulnerabilidades y partes Comprensibles

No solo basta con detectar los puntos vulnerables y comprensibles de un sistema de software sino que se deben visualizar de la manera adecuada. Esta tarea es necesaria ya que cuando se analizan sistemas de gran envergadura, la información extraída puede ser muy grande y compleja de entender. Por consiguiente, si no se plantea apropiadamente la forma de visualizarla, es decir, presentada de forma sintetizada y con extrema facilidad de análisis será sumamente complejo examinarla y por lo tanto, aplicar las estrategias tendientes a mejorar e incrementar el nivel de seguridad de los puntos vulnerables será todo un desafío. En este contexto, los estudios se centran en la generación de visualizaciones innovadoras que permitan disminuir la brecha existente entre la información extraída de las aplicaciones y la estructura de conocimiento del programador.

Aproximaciones para Incrementar la Seguridad de las Aplicaciones Web

Como se mencionó anteriormente, toda aplicación web está conformada por distintos tipos de información, tanto formal como informal. Al analizar detalladamente y estimar su nivel de comprensibilidad, permitirá detectar las partes más entendibles, lo que las vuelven más vulnerables para su manipulación y/o robo de información. En este punto, es posible definir estrategias que permitan subsanar las vulnerabilidades y proteger las partes del software que sean susceptibles de ataques.

En este contexto, se toma como base aproximaciones basadas en el análisis del código de las aplicaciones y las transformaciones correspondientes que permitan incrementar el nivel de seguridad.

Existen diversas formas para poder analizar y extraer información de un código fuente. En el caso de esta línea de investigación, se utilizan las técnicas de compilación tradicionales [16] que usan la representación de Árbol de Sintaxis Abstracta decorado para representar el código fuente. De manera simplificada, como primer paso se utiliza un Analizador Lexicográfico (Lexer), que toma el código fuente y lo divide en fragmentos o partes denominadas tokens [7]. Estas partes o tokens son la entrada del Analizador Sintáctico (Parser), el cual tiene dos funcionalidades principales: Verificar si la especificación del Programa no tiene errores sintácticos y realizar acciones semánticas para llevar a cabo actividades tales como: recolección de información específica, control, transformación de código, etc. La salida del parser es un Árbol de Sintaxis Abstracta sobre el cual se pueden aplicar diferentes recorridos para recuperar, efectivamente, la información compleja del sistema. Una vez obtenida la información requerida y de acuerdo al grado de comprensibilidad de las mismas, se realizan las modificaciones y/o transformaciones necesarias, según los métodos elegidos, para mejorar la seguridad de la aplicación al volver menos comprensible dicha información.

Resultados Obtenidos/Esperados

Hasta el momento se han llevado a cabo las siguientes tareas:

- Se desarrolló una métrica que posibilita medir a priori la comprensibilidad de aplicaciones. Esta métrica utiliza un método multicriterio que permite que el ingeniero de software pueda volcar su experiencia en el dominio de la aplicación. Experiencia que beneficiará en la precisión del cálculo del nivel de comprensibilidad.
- También, como parte del proceso de validación, se desarrolló un prototipo que implementa la métrica antes mencionada y en la cual se llevaron a cabo pruebas para validar los resultados. Además, como parte de la herramienta, se agregaron diagramas de barras e iluminación de código fuente para visualizar el nivel de comprensibilidad del mismo. Todas las pruebas se aplicaron sobre Servicios Web especificados en WSDLs (Web Services Description Language), esto es debido a la simplicidad que proveen para la extracción de la información. Es importante notar que todas las ideas probadas para un WSDL [11] pueden ser extendidas y utilizadas con relativa facilidad en aplicaciones que usan otros lenguajes de programación.
- Se generaron, como parte de las tareas mencionadas previamente, diferentes publicaciones en congresos nacionales, internacionales, capítulos de libros y revistas indexadas.
- Se desarrolló un prototipo con diferentes recorridos en el árbol de sintaxis abstracta que permite extraer la información de los identificadores. Dicho proceso se aplicó a WSDLs mediante la utilización de un parser. El lenguaje utilizado por los WSDL es el XML [12] y para este lenguaje existen varios parsers. En el caso

particular de este trabajo se utiliza DOM (Document Object Model), que facilita las distintas estrategias de inspección de información y manipulación de la misma [13]. Así mismo, la herramienta es parte de un proyecto de mayor envergadura, cuyas operaciones sobre el código fuente son más complejas y en las cuales utilizar DOM facilita el trabajo.

- Se estuvo trabajando con métodos simples de ofuscación y encriptación de código para mejorar el nivel de seguridad de la información contenida en los identificadores [14, 15]. Las Técnicas de Comprensibilidad permiten, a través de métricas, medir el nivel o grado de entendimiento de un WSDL [8, 9, 10]. Mientras mayor sea el nivel de comprensibilidad o entendimiento de la información, mayor será su vulnerabilidad. Si las métricas indican que el nivel de comprensibilidad de un WSDL es alto, se lo puede manipular y transformar con métodos adecuados de ofuscación y/o encriptación. Al aplicar dichos procesos la información se volverá menos entendible, por consiguiente su nivel de comprensibilidad será mucho menor y por lo tanto más segura.

Entre los objetivos planteados a corto y largo plazo relacionados a este trabajo se pueden mencionar:

- Extender el nivel de seguridad no solo a la información contenida en los identificadores, sino a todos los componentes que forman una especificación de un WSDL.
- Ampliar y aplicar el prototipo a especificaciones escritas en BPEL debido a que este lenguaje es muy utilizado para la ejecución de procesos de negocios. Afortunadamente dicho lenguaje también utiliza el lenguaje XML, lo cual facilita su procesamiento.

- Estudiar, comprender y ampliar el número de métodos de encriptación y ofuscación de código utilizados.
- Crear un entorno que permita aplicar y utilizar métodos de terceros tendientes a mejorar la seguridad; ajenos a los implementados en la herramienta.

Formación de Recursos Humanos

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de trabajos para optar al grado de Magister en Ingeniería de Software. En el futuro se piensa generar diferentes tesis de licenciatura, maestría y doctorado a partir de los resultados obtenidos en la presente línea de investigación. Todas las actividades están enmarcadas en el proyecto de investigación “Ingeniería de Software: Conceptos, Prácticas y Herramientas para el desarrollo de software de Calidad”.

Bibliografía

- [1] ISO/IEC. Iso/iec 27032:2012 information technology - security techniques - guidelines for cybersecurity.
- [2] Jorge Ramió Aguirre. Libro Electrónico de Seguridad Informática y Criptografía. Universidad Politécnica de Madrid, 2006.
- [3] Jeremy Hilton Yulia Cherdantseva. Understanding information assurance and security. 2013.
- [4] Alejandra Stolk. Técnicas de seguridad informática con software libre, 2013. Parque Tecnológico de Mérida. ESLARED.
- [5] ISO/IEC. Iso/iec 27000:2016 information technology - security techniques - information security management systems - overview and vocabulary, 2016.
- [6] <http://www.seguridadinformatica.unlu.edu.ar/>. UNLU. 2016.
- [7] Mario Berón, Germán Montejano, Daniel Riesco, Pedro Rangel Henriques, Narayan Debnath. SIP: A Simple Tool for Inspecting and Evaluating WSDL Specifications. 10th International Conference on Information Technology: New Generations. 2013.
- [8] Hernán Bernardis, Edgardo Bernardis, Mario Berón, Daniel Riesco, Pedro Rangel Henriques, Maria Joao Pereira. Cálculo de Métricas para Medir el Grado de Entendimiento de una Descripción WSDL. WICC. 2016.
- [9] Mario M. Berón, Hernán Bernardis, Enrique A. Miranda, Daniel E. Riesco, Maria João Pereira, Pedro Rangel Henriques. "WSDLUD: a Metric to Measure the Understanding Degree of WSDL Descriptions". Proceedings of the 2015 Symposium on Languages, Applications and Technologies, SLATE'15. Madrid, España 2015.
- [10] Bernardis, Hernán; Berón Mario; Bernardis, Edgardo; Riesco, Daniel; Henriques, Pedro. “Extracción de información y cálculo de métricas en WSDL 1.1 y 2.0”. II Congreso Nacional de Ingeniería Informática / Sistemas de información (CoNaIISI). Argentina. 2014.
- [11] WSDL Specification for W3C <https://www.w3.org/TR/wsd1>.
- [12] Extensible Markup Language (XML) 1.0 (Fifth Edition). <https://www.w3.org/TR/REC-xml/>.
- [13] Parser DOM specification for W3C. <https://www.w3.org/DOM>.
- [14] Cappaert, J. Code obfuscation techniques for software protection. Katholieke Universiteit Leuven. 2012.
- [15] Stallings W. Cryptography and Network Security Principles and Practice. Fifth Edition.
- [16] A. V. Aho, R. Sethi, and J. D. Ullman. "Compilers Principles, Techniques and Tools". Addison-Wesley, 1986.

Utilización de Watermarking para Seguridad en la Nube: el Caso de las Imágenes Médicas

Laura Mónica Vargas^{1,2}, María Alejandra Di Gionantonio³

laura.monica.vargas@unc.edu.ar, ing.alejandradg@gmail.com

¹ Laboratorio de Redes y Comunicaciones de Datos, Departamento de Computación, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

² Laboratorio de Procesamiento de Señales, Departamento de Matemática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

³ Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Córdoba, Universidad Tecnológica Nacional

I. RESUMEN

En los últimos años, acompañando el rápido avance de las telecomunicaciones se ha desarrollado la telemedicina por medio de la cual los médicos pueden transferir y compartir los datos digitales de los pacientes en forma remota para determinar un diagnóstico definitivo. Por otra parte, actualmente la tendencia es llevar la información médica que se almacenaba en el propio centro de salud a la nube siendo esencial en estos casos proteger los datos médicos intercambiados. En las plataformas de Cloud Computing, la seguridad es todavía un problema importante a resolver. Se debe garantizar que las imágenes médicas se puedan compartir en forma segura preservándolas de cualquier intento de distorsión, como así también proporcionar privacidad en las cadenas de datos de las historias clínicas de salud o Electronic Health Records (EHR). Una alternativa de solución es la inserción de marcas de agua en las imágenes médicas, técnica conocida como *watermarking*, cuya aplicación en imágenes digitales empezó hace unas décadas. En este trabajo se propone que acompañe a las clásicas técnicas criptográficas para perfeccionar la seguridad en la nube.

Palabras clave: seguridad informática, watermarking, imágenes digitales, cloud computing, telemedicina.

II. CONTEXTO

El presente trabajo se realiza en los siguientes ámbitos: Laboratorio de Redes y Comunicaciones de Datos perteneciente al Departamento de

Computación y en el Laboratorio de Procesamiento de Señales del Departamento de Matemática, ambos en la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de Córdoba, así como en el Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Córdoba de la Universidad Tecnológica Nacional.

La temática de watermarking ha estado presente en los Proyectos aprobados por SeCyT-UNC de ID, todos ellos Categoría A: **Reducción de ruido, compresión y seguridad en la transmisión de señales 1-D y 2-D**, Código 05M/160, años 2010-2011, **Restauración, reconocimiento de patrones y transmisión segura de imágenes**, Código 05/M232, años 2012-2013 y **Restauración, reconocimiento de patrones y transmisión segura de imágenes**, Código 05M/296, años 2014-2015, de todos los cuales la primera autora ha sido **SubDirectora**.

La primera autora realizó su tesis de doctorado en la FCEfyN-UNC aplicando métodos de Inteligencia Artificial a algoritmos de marcado y la segunda autora logró el grado de Especialista en Sistemas de Información de la FRC-UNC con una tesis sobre la utilización de marcas de agua para autenticación de imágenes médicas.

La primera autora ha dirigido tesis para alcanzar el grado de Ingeniero en Computación de la FCEfyN consistentes en diseño e implementación de marcas de agua durante los años 2012-2013. También ha participado de congresos y jornadas presentando algoritmos de implementación de marcas de agua (INMAT 2008, Congreso

Internacional de Ingeniería 2010, MACI 2013 y 2015, ACCN 2015 entre otros), publicado en IEEE y en Int. Journal of Advances in Processing Images Techniques. Ambas autoras han publicado un artículo sobre el tema en la Revista de la FCEFYN-UNC en 2016.

III. INTRODUCCIÓN

En la etapa actual, la investigación se centra en el análisis del problema de seguridad de la información contenida en las imágenes médicas cuando estas son alojadas en un framework de Cloud Computing.

Cloud Computing es un mecanismo que creció en los últimos años, basado en la Web que permite escalar y virtualizar recursos de TI que son proporcionados como servicios a través de la red. Características inherentes y esenciales que deben ser provistas por las aplicaciones de cloud computing son: servicio bajo demanda, acceso ubicuo, escalabilidad, elasticidad, independencia del usuario respecto al mantenimiento y pago por uso, siendo la seguridad todavía un desafío [1] [2].

En las últimas décadas, antes del desarrollo de la nube, se logró un progreso significativo en el uso de tecnologías de comunicación para almacenar y distribuir datos médicos bajo formatos digitales [3]. Al comienzo de los años 80 apareció el PACS (Picture Archiving and Communication System/ Sistema de Comunicación y Almacenamiento de Imágenes). Inicialmente, estos se desarrollaron para cubrir necesidades específicas de los centros médicos, tales como adquisición de datos y visión de estos en estaciones de trabajo de poca capacidad. En un principio solo se asociaban a radiología, pero luego abarcaron todo tipo de imágenes médicas [4]. Rápidamente se extendieron en los centros médicos, pero presentaban problemas de interoperabilidad ya que eran desarrollados en forma independiente por distintos proveedores que seguían sus propias reglas. Surgió entonces la Norma DICOM (Digital Imaging and Communications in Medicine), desarrollada por el ACR (American College of Radiology) en conjunto con NEMA (National Electrical Manufacturers Association) [5]. Tras varios intentos se aprobó en 1993 y desde entonces sufre una actualización constante. Esta norma define el acceso a la web, la estructura de

intercambio de datos, las capas de comunicación y los comandos para manejo de imágenes médicas que deben respetar todos los fabricantes para obtener interoperabilidad. Su aceptación hizo que se “dicomizaran” los PACs. Actualmente los fabricantes de equipos para imágenes médicas, siguiendo indicaciones de la norma, los acompañan de un CS (Conformance Statement) que asegura que cumplen con la misma. Esta norma permite el acceso remoto a archivos en formato dicom (extensión dcm) utilizando los ya clásicos protocolos TCP/IP, el Protocolo HTTP (Hypertext Transfer Protocol) o HTTPS (Hypertext Transfer Protocol Secure). Si bien aseguró interoperabilidad entre los distintos sistemas y demostró cierta flexibilidad en entornos que manejan imágenes médicas, no hizo un aporte significativo a la seguridad, ni al acceso de datos por fuera de instituciones médicas.

El empleo de las imágenes no es seguro cuando los datos, médicos o de otro origen, circulan libremente por redes abiertas como Internet, expuestos a que los mismos sean alterados o mal utilizados. Esto sucede cuando se utilizan servicios de teleconsulta o telediagnóstico, los que se están difundiendo por todo el mundo con el aporte de las Tecnologías de Información y Comunicación.

En un principio, la norma DICOM no presentó ninguna disposición respecto a seguridad ya que no era una preocupación en el momento de su generación. La parte 15, agregada posteriormente, actualizada por última vez en 2016, pide el uso de protocolos de comunicación seguros, encapsulamiento en formatos seguros, encriptación de los datos y empleo de firmas digitales. Sin embargo, la firma digital puede ser separada del archivo por lo que son varios los autores que señalaron la necesidad de agregar otro nivel de seguridad, incluso antes de que el almacenamiento de la información médica se hiciera en la nube [6]. Almacenar en la nube es una buena alternativa ya que permite a los centros médicos desentenderse del hardware y software usado en los sistemas de archivos, pero conlleva un mayor riesgo de violación de autenticidad e integridad en los registros del paciente [7]. En la actualidad, los sistemas de información no son centralizados, sino distribuidos y, por lo tanto, el control también debe

estar distribuido. La protección primitiva, que controlaba el acceso mediante claves, no es suficiente. Por otro lado, la confidencialidad de los datos es una necesidad ética en el campo de la salud que puede lograrse por medio de la encriptación clásica [8], herramienta que, sin embargo, no resulta suficiente para solucionar todos los problemas de protección de datos digitales.

En 1996, se dictó en EEUU, la HIPAA (Health Insurance Portability and Accountability Act) que indica qué requisitos se deben cumplir para las transacciones de datos de salud con el objetivo de que los datos médicos se almacenen y se puedan recuperar a largo plazo, evitando abusos y fraude. Microsoft en 2007 y Google en 2008 ofrecieron portales Health a los usuarios que querían que sus EHRs estuvieran disponibles para sus servicios de salud y para ellos mismos. En 2010, IBM y Aetna en conjunto anunciaron un nuevo uso de la plataforma de cloud computing de IBM diseñada para ayudar a los profesionales de la salud a acceder rápidamente a la información del paciente: registros médicos, recetas, y datos de laboratorio recolectados de múltiples fuentes para crear un registro detallado del mismo. Se estima que en el año 2020 el 80% de los datos se habrá mudado a la nube. El uso de estas plataformas y otras permite a los centros médicos desentenderse de problemas técnicos (actualización y mantenimiento de software y hardware), económicos y legales relacionados al manejo de datos lo que le conviene más allá de los riesgos que corre. Entre los inconvenientes se encuentra la latencia, la dificultad para tener el servicio disponible todo el tiempo, y la seguridad [9]. Se debe tener especialmente en cuenta que los datos almacenados en la nube son vulnerables a ataques internos. La identidad y ubicación de intermediarios y de los proveedores de servicio está disimulada, oculta, por la nube.

En forma paralela a estos avances en telecomunicaciones, se empezó a desarrollar en la década del 90, el watermarking o marcado de productos multimedia, imágenes, videos, audio, gráficos, etc., como forma de protección de propiedad intelectual [10]. Consiste en embeber bits en el archivo, sea imagen, video o audio, de forma visible (audible) o invisible (no audible). Estos bits extra constituyen la marca y en las primeras

implementaciones permitían identificar al propietario, utilizándose posteriormente para alcanzar otros propósitos como detección de adulteraciones, aseguramiento de integridad e incorporación de metadatos. Algunos métodos permiten en el caso de haberse producido una alteración, determinar en qué sector del archivo sucedió. Así, en estos días, la marca de agua aparece como un medio eficiente para asegurar integridad y verificar autenticidad. Es usual que la marca sea la firma digital del archivo o metadatos encriptados siguiendo técnicas clásicas. Mientras que con la encriptación se espera que no se entienda lo que se ve, con el watermarking hay datos ocultos que no se pueden ver.

El watermarking puede ser reversible o irreversible. El reversible permite recuperar el archivo original. Para lograrlo es preciso haber guardado cierta información. En la Figura 1 se presenta un esquema básico de watermarking reversible, que es el que nos interesa ya que permite recuperar la imagen original algo necesario en imágenes legales como lo son las médicas.

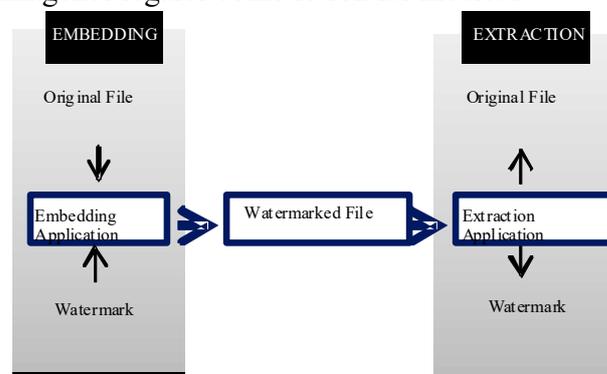


Figura 1. Esquema de Marcado Reversible

En cuanto a los algoritmos de embebido se han realizado numerosos desarrollos, en el dominio espacial (modificación del bit menos significativo o LSB) y en el de la frecuencia, en particular. En este último dominio son de especial interés los desarrollos efectuados usando la transformada wavelet y la transformada discreta de coseno (DCT). Entre los algoritmos clásicos para imágenes se cuentan el de Tian [11] de expansión de la diferencia entre pares de píxeles y el de Ni [12] de corrimiento del histograma. Siempre es necesaria la validación teniendo en cuenta que para distintos tipos de imágenes pueden ser convenientes distintos

algoritmos, no existiendo uno de utilidad universal.

La evaluación de las marcas de agua reversibles se hace considerando el tiempo de ejecución del algoritmo, la capacidad y el error medio cuadrático MSE o relación señal/ruido PSNR para cuantificar imperceptibilidad, la que también debe ser objeto de una evaluación subjetiva por parte de expertos. Estos parámetros responden a las siguientes expresiones:

$$\text{Capacidad}(bpp) = \frac{N_{bm}}{\text{Cantidad total pixeles imagen}}$$

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N (I_w(i,j) - I(i,j))^2}{MN}$$

$$PSNR(dB) = 10 \log_{10} \left[\frac{I_{pico}^2}{MSE} \right]$$

donde N_{bm} indica la cantidad de bits de la marca, M la cantidad de filas y N la cantidad de columnas de la imagen, $I(i,j)$ el valor de la intensidad en la posición del píxel fila i , columna j e $I_w(i,j)$ indica el valor de la intensidad en la posición i,j luego del marcado.

La importancia del watermarking en imágenes médicas se destaca desde hace dos décadas [13], hasta la actualidad [14]. Es de particular interés que se embeban los datos del paciente en sus imágenes médicas personales. Las amenazas, como la destrucción de sistemas de software y violación en los accesos, están surgiendo con frecuencia en la plataforma de la nube, por lo que se hace absolutamente necesario tomar medidas para contrarrestarlas. Se recomienda en telemedicina, la combinación de watermarking con técnicas de criptografía clásicas [13-17]. La encriptación puede impedir problemas en los nodos intermedios, pero no en los puntos finales que deben poder descryptar los datos y si el proveedor del servicio confía, a su vez, en otros proveedores entonces los datos del usuario pueden ser leídos por muchas entidades en la nube. Lo que se precisa para incrementar la confianza en la nube siguiendo esta línea de razonamiento es algún mecanismo que pueda detectar y castigar cualquier problema relativo a la confidencialidad. El usuario final debe confiar en la entidad que administra los EHR (sea Google Health, Microsoft Health Vault u otro proveedor de nube). El administrador de EHR debe

tener los medios para detectar y castigar las violaciones a la confidencialidad que se hubieran producido. Las nuevas técnicas deben tratar especialmente de mitigar los riesgos de que la información sufra ataques internos en la nube [18].

La propuesta actual de investigación consiste en diseñar e implementar un mecanismo que combine watermarking y encriptación para lograr asegurar integridad y autenticidad de datos médicos. Se propone utilizar un sistema de archivos distribuidos Hadoop, y MapReduce, modelo de programación para el manejo de grandes bases de datos [19].

IV. OBJETIVOS Y LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los objetivos de este proyecto de investigación los podemos dividir en:

A. Objetivo General

-Implementación y evaluación de técnicas de watermarking junto con las criptográficas tradicionales, para otorgar seguridad a las imágenes digitales médicas almacenadas en la nube.

B. Objetivos Específicos

- Analizar el estado del arte de cloud computing para servicio de almacenamiento de imágenes médicas.
- Analizar las soluciones implementadas para obtener seguridad en el almacenamiento de imágenes médicas en la nube.
- Utilizar infraestructuras como Hadoop para simular las aplicaciones.
- Difundir los resultados obtenidos para realimentar el proceso de desarrollo de los algoritmos de watermarking.
- Formar recursos humanos específicos en las áreas objeto de estudio.

V. MATERIALES Y MÉTODOS

En nuestra investigación proponemos hacer frente al problema de la seguridad de los datos contenidos en imágenes médicas alojadas en Cloud Computing, utilizando la técnica de marca de agua en el EHR. Luego se envía la imagen con marca de agua al proveedor de la Nube. En particular se trabajará con imágenes médicas en formato dicom y

con marcas reversibles indetectables.

Es común que una imagen médica sea diagnosticada antes de que la misma sea almacenada en un almacenamiento a largo plazo, de este modo la parte significativa de la imagen, conocida como ROI (Region of Interest), es determinada en ese momento. El embebido de información extra se hará fuera de esta zona.

VI. FORMACIÓN DE RECURSOS HUMANOS

La segunda autora está realizando su tesis de Maestría en Sistemas de información en la FRC-UTN con el tema “Análisis comparativo de múltiples plataformas de Health Cloud Computing para hosting de imágenes médicas con marcas de agua”.

Se dirigirán trabajos finales sobre la temática abiertos a estudiantes de Ingeniería en Sistemas de Información e Ingeniería en Computación.

REFERENCIAS

- [1] Youssef *et al.* “Toward a Unified Ontology of Cloud Computing”. Grid Computing Environment Workshop, IEEE, 2008.
- [2] Jadeja and Modi. “Cloud Computing – Concept, Architecture and Challenges”. International Conference on Computing, Electronics and Electrical Technologies, IEEE, 2012.
- [3] Acharya *et al.* “Compact Storage of medical Images with patient Information”. IEEE Transactions on Information Technology in Biomedicine, vol.5, pp. 320-323, 2001.
- [4] Bharath. “Introductory Medical Imaging”. Ed. John Enderle, University of Connecticut, 2009.
- [5] Pianykh. “Digital Imaging and Communications in Medicine (DICOM). A Practical Introduction and Survival Guide”. 1st Ed., Ed Springer, 2008.
- [6] Ahmed and Abdullah. “Telemedicine in a Cloud - A Review”. IEEE Symposium on Computers and Informatics, 2011.
- [7] Ahuja, Mani and Zambrano. "A Survey of the State of Cloud Computing in Healthcare", Network and Communications Technologies, Vol. 1 N°2. Canadian Center of Science and Education, 2012.
- [8] Stallings. "Cryptography and Network Security", Ed. Prentice Hall, 4th Ed., 2005.
- [9] Zhang and Liu. "Security Models and Requirements for Healthcare Application Clouds". 3rd International Conference on Cloud Computing", IEEE, 2010.
- [10] Cox *et al.* "Digital Watermarking". Ed. Morgan Kauffmann. 2002.
- [11] Tian. "Reversible Data Embedding using a Difference Expansion". IEEE Transactions on Circuits Systems and Video Technology, vol. 13, no. 8, pp. 890-896, 2003
- [12] Ni *et al.* "Reversible data Hiding". Proceedings of the 2003 International Symposium on Circuits and Systems, vol. 2, pp. 912-915, 2003
- [13] Coatrieux *et al.* “Relevance of Watermarking in Medical Imaging”. Proceedings of the IEEE EMBS Conf. on Information Technology Applications in Biomedicine, Arlington, USA, pp 250-255, 2000.
- [14] Aminzou *et al.* "Towards a Secure Access to Patient Data in Cloud Computing Environments". Security Days (JNS3), IEEE, 2013.
- [15] Elgamal, Hikal & Abou-Chadi. “Secure Medical Images Sharing over Cloud Computing environment”. International Journal of Advanced Computer Science and Applications (IJACSA), vol. 4, N°5, 2013.
- [16] Bouslimi and Coatrieux. "A Joint/Encryption Watermarking System for Verifying the Reliability of Medical Images". Medical Data Privacy Handbook pp. 493-526, Springer, 2015.
- [17] Al-Haj, Hussein and Abandah. "Combining Cryptography and Digital Watermarking for Secured Transmission of Medical Images". Second International Conference on Information Management (ICIM), IEEE; 2016.
- [18] Garkotti *et al.* "Detection of Insider attacks in Cloud based e-healthcare". International Conference on Information Technology, IEEE, 2014.
- [19] Lee *et al.* "Implementation of MapReduce-based Image Conversion Module in Cloud Computing Environment". Proc. of Int. Conference on Advances in Computing, Control and Telecommunication Technologies, 2011.

Verificación del Hablante Mediante Dispositivos Móviles en Entornos Ruidosos

Graciela Etchart, Silvia Ruiz, Ernesto Miranda, Juan José Aguirre, Mauro Herlein, Carlos Alvez,

Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

getchart@fcad.uner.edu.ar, sruiz@fcad.uner.edu.ar, emiranda@fcad.uner.edu.ar, juaagu@fcad.uner.edu.ar, herlein.mauro@gmail.com, caralv@fcad.uner.edu.ar

Resumen

El interés por las aplicaciones biométricas ha crecido considerablemente en los últimos años debido a que han surgido nuevos servicios que necesitan formas confiables para autenticar a sus usuarios. En este trabajo se abordará la biometría de voz, a través de dispositivos móviles, para la verificación de la identidad del hablante. Actualmente, los dispositivos móviles proporcionan una plataforma conveniente para la captura de voz y su transferencia para la verificación remota. Sin embargo, debido a limitaciones de almacenamiento y de capacidad computacional de estos dispositivos se requiere un adecuado conjunto de parámetros para llevar a cabo la autenticación. Además, por las características del entorno donde se realizan la mayoría de las transacciones de verificación, es necesario trabajar en el fortalecimiento de sistemas de reconocimiento cuando la voz es capturada en entornos ruidosos. El objetivo general de este trabajo es realizar un análisis comparativo de diferentes enfoques para contrarrestar desajustes entre la fase de entrenamiento y prueba en la verificación del hablante a través de dispositivos móviles, analizando su adecuación principalmente en entornos ruidosos.

Palabras clave: biometría de voz, verificación, dispositivos móviles, seguridad.

Contexto

Este trabajo se da en el marco del Proyecto PID 07/G044 “*Gestión de datos biométricos en base de datos objeto - relacionales*”, que da continuidad al Proyecto PID 07/G035 “*Identificación de personas mediante Sistemas Biométricos. Estudio de factibilidad y su implementación en organismos estatales*” [1][2][3].

El mismo constituye un futuro trabajo de tesis de la Maestría en Sistemas de Información (MSI) dictada en la Facultad de Ciencias de la Administración perteneciente a la Universidad Nacional de Entre Ríos (UNER). El objetivo general es realizar un análisis comparativo de diferentes enfoques para contrarrestar desajustes entre la fase de entrenamiento y prueba en la verificación del hablante a través de dispositivos móviles, analizando su adecuación principalmente en entornos ruidosos.

Introducción

El interés por las aplicaciones biométricas ha crecido considerablemente en los

últimos años debido a que han surgido nuevos servicios que necesitan formas confiables para autenticar a sus usuarios. Actualmente, los usuarios deben recordar decenas de contraseñas y números de identificación personal, debido a que resulta poco seguro utilizar la misma contraseña para autenticar su identidad en cada servicio al que se pretende acceder. Además, las contraseñas pueden ser olvidadas o robadas por diferentes medios. La biometría ofrece mecanismos para resolver estos problemas.

En este trabajo se abordará la biometría de voz para la verificación de la identidad del hablante.

La verificación mediante el rasgo biométrico de la voz no es un nuevo campo de investigación. Los primeros intentos de construir sistemas de verificación del hablante se hicieron a principios de la década de 1950 [4, 5].

En biometría, verificar un hablante consiste en autenticar una identidad reclamada mediante el análisis de una muestra de voz del reclamante. Es un problema de detección binaria donde el reclamante debe ser clasificado como el verdadero o como un impostor. De esta manera, pueden ocurrir dos tipos de error: el falso rechazo de un usuario genuino o la falsa aceptación de un impostor [5].

En los servicios de acceso remoto, como ser a través de dispositivos móviles, la verificación de identidad utilizando la voz del usuario presenta numerosas ventajas, debido a la seguridad que ofrece la biometría con respecto a métodos tradicionales de autenticación.

Además, el número de teléfonos móviles ha superado el número de personas en el mundo. Los usuarios se muestran más dependientes de sus teléfonos inteligentes. Su mayor uso se debe a las aplicaciones móviles en creciente aumento. Varias aplicaciones requieren la verificación del usuario para acceder a los servicios.

En este contexto, los dispositivos móviles proporcionan una plataforma conveniente para la captura de voz y la transferencia a distancia para la verificación remota. Sin embargo, debido a limitaciones de almacenamiento y de capacidad computacional de estos dispositivos se requiere un adecuado conjunto de parámetros para llevar a cabo el proceso de autenticación. Además, por las características del entorno donde se realizan la mayoría de las transacciones de verificación, es necesario trabajar en el fortalecimiento de sistemas de reconocimiento cuando la voz es capturada en entornos ruidosos. Estos aspectos serán tratados en este trabajo.

La investigación continua en este campo ha estado en curso durante los últimos veinticinco años [4, 5, 6, 7]. Estudios recientes han encontrado que el desajuste de entrenamiento (enrolamiento) y prueba (verificación) afectan significativamente el desempeño de la verificación del hablante [8, 9, 10].

Las aplicaciones in situ de verificación del hablante, como control de acceso, control de fronteras, domótica, pueden ofrecer mucha libertad en la elección del equipo de captura de datos, el diseño de interfaces de usuario, medios para controlar el comportamiento del usuario y el entorno de fondo. Esto facilita la recolección de muestras de voz de alta calidad bajo condiciones controladas para permitir una verificación exacta del hablante. Sin embargo, con los usuarios físicamente presentes existen otras modalidades biométricas que podrían utilizarse, como huella digital, iris y rostro [11].

Por otro lado, con las aplicaciones basadas en transmisiones remotas, como el caso de los dispositivos móviles, la verificación del hablante presenta ventajas sobre otras modalidades biométricas. Debido a que en cada dispositivo se dispone de un micrófono y un canal de transmisión de voz, mientras que el uso de

otros rasgos biométricos puede requerir instrumentos de captura de datos no disponibles o de costosa incorporación. Sin embargo, existe poca posibilidad de influir en los instrumentos telefónicos que se utilizan en los sistemas de verificación de hablantes, lo que da lugar a diferentes características de señal de entrada entre los usuarios y posiblemente incluso entre capturas de un mismo usuario. También se tienen poca posibilidad de controlar a los usuarios o el entorno acústico. Por lo tanto, en comparación con las aplicaciones in situ, las aplicaciones remotas de la verificación de hablantes generalmente necesitan contrarrestar una gama más amplia de variabilidad entre sesiones que no esté directamente relacionada con las características de voz de los hablantes.

Líneas de Investigación y Desarrollo

Una de las líneas más importantes, consideradas en este trabajo, es la biometría de voz para la verificación de la identidad del hablante, utilizada en diversas aplicaciones o sistemas relacionados con la seguridad.

Entre los principales desafíos que debe enfrentar la verificación del hablante en muchas situaciones prácticas se encuentran [7, 12]:

- Deben desarrollarse métodos eficaces para combatir el desajuste de entrenamiento y prueba que está presente debido a las condiciones acústicas adversas que reducen la precisión de la verificación del hablante.
- Debe reducirse la cantidad de datos necesarios para diseñar sistemas de verificación de hablantes de última generación, ya que es difícil recopilar y registrar gran cantidad de datos de voz.
- Debe adecuarse la duración del discurso requerido para entrenar

modelos de hablantes durante la etapa de entrenamiento así como la duración de los enunciados de prueba o test.

La compensación del canal es un enfoque que se utiliza para reducir la falta de coincidencia entre el entrenamiento y las pruebas. Esta compensación se produce en los diferentes niveles, como el dominio de la característica, el dominio del modelo y el dominio de la puntuación. En el dominio de la característica, la supresión adaptable del ruido, la substracción media cepstral (CMS), el filtrado de RASTA y el warping de la característica se utilizan para compensar la variabilidad del canal. Mientras que para compensar la falta de concordancia en el dominio del modelo se utilizan los enfoques JFA, JFASVM e i-vector. Por otra parte, en el dominio de la puntuación, para compensar la variabilidad de la sesión se utilizan los métodos de normalización, tales como normalización de prueba (norma T), simétrica (norma S) y de prueba cero (norma ZT)[13].

Objetivos

El objetivo general de este trabajo es realizar un análisis comparativo de diferentes enfoques para contrarrestar desajustes entre la fase de entrenamiento y prueba en la verificación del hablante a través de dispositivos móviles, analizando su adecuación principalmente en entornos ruidosos.

Con este propósito se explorarán diferentes modelos y métodos estudiando el impacto de las duraciones de entrenamiento y prueba, el efecto del ruido de fondo, la longitud de la sesión y los canales en el sistema de verificación del hablante.

Para llevar a cabo las pruebas experimentales, se realizará una investigación exploratoria de diferentes herramientas open source utilizadas para el reconocimiento biométrico de personas

mediante voz. Las pruebas se efectuarán sobre una base de datos propia.

Formación de Recursos Humanos

La estructura del equipo de investigación del PID 07/G035 está formada por el Director, Co-directora y cuatro integrantes docentes. El Director del proyecto dirige la tesis de MSI de la Facultad de Ciencias de la Administración de la UNER de la Co-directora y tres integrantes, que se encuentran realizando sus trabajos de investigación en el área del proyecto.

Además, el proyecto cuenta con un becario de Iniciación en la Investigación y un integrante alumno de la Licenciatura en Sistemas, cuyas tareas están relacionadas con la captura, registración y almacenamiento de datos biométricos.

Referencias

1. Etchart G., Luna L., Leal C., Benedetto M., Alvez C., "Sistemas de reconocimiento biométricos, importancia del uso de estándares en entes estatales". CGIV - XIII Workshop de Investigadores en Ciencias de la Computación (WICC 2011), 5 y 6 de Mayo de 2011. Universidad Nacional de Rosario. Rosario – Argentina. pp 339-343.
2. Etchart G., Alvez C., Benedetto M., "Gestión de Datos Biométricos en Bases de Datos Objeto-Relacionales". XV Workshop de Investigadores en Ciencias de la Computación (WICC 2013). Paraná Entre Ríos. pp. 97-101.
3. Ruiz S., Etchart G., Alvez C., Miranda E., Benedetto M., Aguirre J., "Representación e interoperabilidad de imágenes biométricas". XVII Workshop de Investigadores en Ciencias de la Computación (WICC 2015). Salta.
4. Furui S., "50 years of progress in speech and speaker recognition research", ECTI Transaction on Computer and Information Technology, Vol. 1, 2005.
5. Campbell J., "Speaker recognition: A tutorial", Proceedings of the IEEE, Vol. 85, N° 9, 1997.
6. Furui S., "An overview of speaker recognition technology", Kluwer international series in Engineering and Computer science, pp. 31-56, 1996.
7. Naufal A., Phaklen E., Badlishah A., Naseer S., "Speaker Recognition System: Vulnerable and Challenges", International Journal of Engineering and Technology (IJET), Vol. 5, N° 4, 2013.
8. Chen Y., Hong Q., Chen X., Zhang C., "Real-time speaker verification based on GMM-UBM for PDA" 5° IEEE International Symposium on Embedded Computing, 2008.
9. Kinnunen T., Li H., "An overview of text-independent speaker recognition: From features to supervectors", Speech Communication, Vol. 52, N° 1, pp. 12-40, 2010.
10. Ding P., He L., Yan X., Zhao R., Hao J., "Robust technologies towards automatic speech recognition in car noise environments", 8° International Conference on Signal Processing, Vol. 1, 2006.
11. Beranek B., "Voice biometrics: success stories, success factors and what's next", Biometric Technology Today, pp. 9-11, 2013.
12. Ming J., Hazen T. J., Glass J. R., Reynolds D. A., "Robust Speaker Recognition in Noisy Conditions", IEEE Transactions on Audio, Speech and Language Processing, Vol. 15, N° 5, pp. 1711-1723, 2007.
13. Kanagasundaram A., "Speaker Verification using I-vector Features" (Tesis doctoral). Queensland University of Technology, Speech and Audio Research Laboratory, Science and Engineering Faculty, 2014.

**Tecnología Informática
Aplicada en Educación**

Análisis y Detección Temprana de Deserción Estudiantil en UTN FRLP

ISTVAN Romina Mariel; FALCO Mariana; ANTONINI Sergio Andrés

Universidad Tecnológica Nacional, Facultad Regional La Plata

Laboratorio de Ingeniería en Sistemas de Información, LINES

Av. 60 s/n° esquina 124, CP 1900, La Plata, Buenos Aires, Argentina.

ristvan@frlp.utn.edu.ar; mfalco@frlp.utn.edu.ar; antonini@frlp.utn.edu.ar

RESUMEN

La deserción es entendida como aquella situación a la que se enfrenta un estudiante cuando aspira y no logra concluir su proyecto educativo. Esto constituye una de las problemáticas de mayor preocupación en el sistema universitario, debido al aumento de su ocurrencia en los últimos años.

En Argentina, se estima que en las universidades nacionales sólo el 12% de los estudiantes que ingresan se gradúa y si bien no hay datos oficiales para las instituciones privadas, se estima que un 30% concluye con éxito su carrera.

Por este motivo, se plantea el diseño y desarrollo de un sistema informático para la detección y caracterización temprana de posibles desertores, como una propuesta metodológica que brinde un modelo explicativo de los principales factores causales de la deserción y su impacto en la institución. Asimismo se espera que este sistema sirva a tutores, docentes, equipo interdisciplinario y autoridades académicas como fundamento a partir del cual generar estrategias de intervención.

Palabras clave: Deserción estudiantil, educación superior, permanencia académica, retención.

CONTEXTO

La Universidad Tecnológica Nacional, Facultad Regional La Plata reúne los datos académicos en una base de datos centralizada.

Si se observa la relación aspirante/ingresante en el periodo 2011-2015 y considerando los índices acumulados totales, es factible determinar la cantidad de aspirantes que han aprobado el curso de ingreso. De 4989, número total de alumnos aspirantes en los años analizados, sólo 3473 estudiantes logran aprobar el curso de ingreso. Lo cual representa que sólo el 30% de los aspirantes, no llega a ser ingresante. Si se analiza la relación aspirante/egresado e ingresante/egresado, se obtiene que sólo el 10.48% de los aspirantes se gradúa y sólo el 15.05% de los ingresantes se gradúa.

El sistema propuesto constituye el primer estudio sistemático de la deserción estudiantil en la UTN, FRLP, focalizando en la detección temprana de posibles desertores.

1. INTRODUCCIÓN

De acuerdo a Braxton [1] se puede realizar el estudio de la deserción mediante cinco enfoques distintos, el cual varía dependiendo del énfasis que se le otorgue a cada variable involucrada:

1.1. Enfoques psicológicos

Estos enfoques señalan que los rasgos de la personalidad son los que diferencian a los estudiantes que terminan sus estudios de aquellos que no lo logran. El primer modelo de este enfoque es el de Fishbein y Ajzen [2], el cual resalta cómo las intenciones de una persona son el resultado de sus creencias, de forma que las mismas influyen sobre sus actitudes y la llevan a manifestar un

comportamiento. Es así como la intención de una persona es una función de su actitud hacia la conducta y de las normas subjetivas acerca de ella. Este modelo fue ampliado por Attinasi [3], el cual incorporó las percepciones y el análisis que realizan los alumnos en su vida universitaria después de su ingreso como influencias para la persistencia o deserción.

En la década de 1990, Ethington [4] incorporó la elección, la perseverancia y el desempeño a los modelos anteriores.

1.2. Enfoques sociológicos

Los modelos sociológicos toman como base la mirada psicológica y suman y enfatizan la influencia de factores externos al individuo en el estudio de la deserción. El modelo de Spady [5] es uno de los más citados, el cual toma la teoría del suicidio de Durkheim y plantea que el suicidio es el resultado de la ruptura del individuo con el sistema social por su imposibilidad de integrarse a la sociedad. Y lo aplica a la educación superior, de forma que la deserción es el resultado de la falta de integración de los estudiantes en el entorno de educación superior. Señala además que el medio familiar expone a los estudiantes a influencias, expectativas y demandas, las que a su vez afectan su nivel de integración social en la universidad.

1.3. Enfoques económicos

Según las investigaciones desarrolladas por Cabrera [6] se pueden distinguir dos modelos. El enfoque de costo-beneficio, postula que los estudiantes optan por permanecer en la universidad cuando perciben que los beneficios económicos y sociales son mayores a los que conseguirían en comparación con otras actividades, como ser un trabajador. El enfoque que considera la Focalización de subsidios, examina la capacidad o no del alumno y/o sus familiares de costear los estudios universitarios, y analiza los subsidios otorgados a grupos de alumnos con estas limitaciones. Este modelo

busca privilegiar el impacto efectivo de los beneficios estudiantiles por sobre la deserción, dejando de lado las percepciones acerca de la adecuación de dichos beneficios o el grado de ajuste de éstos a los costos de los estudios [7].

1.4. Enfoques organizacionales

Este modelo de análisis pone la mirada en las características y servicios que ofrece la institución, analiza cómo diferentes factores, como la calidad de los docentes y las experiencias de aprendizaje en el aula, afectan los índices de deserción [8]. En este enfoque es altamente relevante la experiencia de aprender en forma activa por parte de los estudiantes en las aulas.

1.5. Enfoques de interacción

Su mayor exponente es Tinto. Este autor expande el modelo de Spady [5], incorporando la teoría del intercambio, la cual expone que los sujetos evitan conductas que implican un costo y buscan recompensas en relaciones, interacciones y estados emocionales.

Según Tinto [9] esto aplicado al ámbito académico, explica el proceso de permanencia en la educación superior como una función del grado de ajuste entre el estudiante y la institución, adquirido a partir de las experiencias académicas y sociales (integración). En la década de 1980, se estudia cómo aspectos demográficos e institucionales sumados a la interacción, la motivación, la aspiración, la personalidad y valores inciden sobre la deserción. El referente más importante de esta línea de investigación es Lenning [10]. Estudios avanzados a partir del año 2000, enfatizan la falta de actitud y compromiso de los estudiantes, el bajo acercamiento al mercado laboral en los tres primeros años y la falta de capacitación docente, entre otras [11].

A modelo de resumen, se puede observar que el abandono estudiantil, constituye un evento educativo explicado a lo largo del tiempo de

diferentes maneras y desde varios puntos de vista. En general, la revisión de estudios recientes, revela una oscilación entre dos macro tendencias explicativas: una que ubica las causas del abandono en el sujeto (genetista-individualista) y la otra que las sitúa en el entorno escolar y social (interacción-socio-estructuralista) [9].

La variabilidad de los índices de deserción entre países y dentro de ellos entre las instituciones públicas y privadas, evidencia un problema común con causas diferentes que hace necesario un estudio que encuentre relaciones entre el sistema educativo, las políticas educativas, el acceso a la educación y las expectativas de formación de los jóvenes, entre otros factores, para cada sociedad y para cada institución en particular.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El proyecto tiene como objetivo generar un sistema informático para la detección y caracterización temprana de posibles desertores, como una propuesta metodológica que brinde un modelo explicativo de los principales factores causales de la deserción y su impacto en la totalidad de las carreras dictadas en la institución.

Tiene como objetivos específicos:

1. Consolidar la primera base de datos unificada para el estudio sistemático de la deserción, la cual recopile información de diversos medios: base de datos académicas, encuestas, asistencias, sistemas de tutores y de análisis de redes sociales.
2. Brindar un modelo explicativo de los principales factores causales de la deserción y su impacto en la institución.
3. Caracterizar tempranamente a posibles desertores a fin de tomar las acciones adecuadas para su retención.

2.1. Modelo

El sistema propuesto se alimenta de las siguientes fuentes de datos:

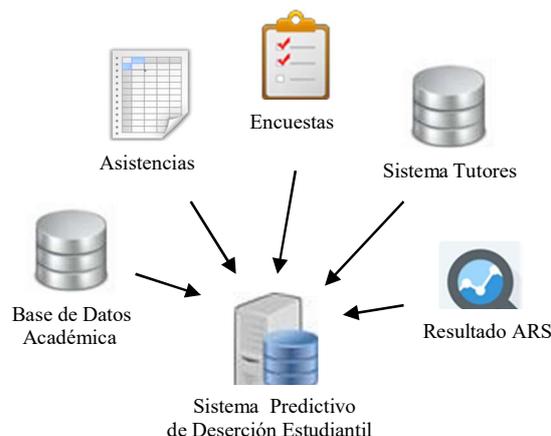


Imagen 1. Modelo de datos.

- Base de Datos Académica: la cual contiene información personal y académica de los alumnos. En ella se contempla: las notas de parciales, finales, año de ingreso y toda otra información derivada de los mismos.
 - Asistencias: las cuales comenzarán a ser registradas a fin de poder identificar tempranamente las ausencias consecutivas a una y/o varias materias por parte de los alumnos.
 - Encuestas: las mismas tendrán objetivos diferentes en cuanto al tipo de información que recaban, dependiendo de la instancia en la que se encuentren los alumnos encuestados. Para los ingresantes, intenta obtener información acerca de la situación general, el grado de satisfacción respecto al curso de ingreso, aspectos vocacionales, sociales, de integración y desarraigo; profesores y actividades académicas junto con el nivel de exigencia de actividades; la situación económica de los ingresantes y las posibilidades de continuación de dicho período.
- Las destinadas a los alumnos intermedios y avanzados, focalizan en la actualización de datos personales, situación laboral y familiar,

disponibilidad de tiempo para el estudio y dificultades encontradas.

- Sistema de Tutores: se encuentra en etapa de análisis y desarrollo el sistema para tutores del Programa PACENI, que agiliza la tarea de los mismos en sus acciones de apoyo para la mejora del rendimiento de los alumnos ingresantes de la carrera.

Se planea que el programa de tutorías sea ampliado para todas las carreras que se dictan en la facultad, y para todos los períodos de la carrera, especialmente en las etapas más vulnerables de abandono.

- Resultados ARS: el Análisis de Redes Sociales, aplicado al ámbito educativo permite mapear el conjunto de interacciones dentro de un aula y así determinar el estatus sociométrico de cada alumno, detectando con anticipación posibles riesgos de deserción.

Es factible su aplicación a través de un tests sociométricos consistentes en la observación o mediante aulas virtuales en las cuales se registran los intercambios y vinculaciones entre los alumnos. De esta manera, es posible ayudar al docente a evaluar cualitativamente y en forma dinámica, los distintos lazos internos del alumnado: estructura interna de los grupos, organización, cohesión, miembros más significativos, alumnos aislados, rechazados y dominantes, entre otros casos.

De este modo, se facilita la determinación de aquellos alumnos que no se encuentren socialmente y/o académicamente integrados al grupo de clase y de esa manera, tengan más posibilidades de desertar.

3. RESULTADOS OBTENIDOS / ESPERADOS

Con la integración de los datos provenientes de las distintas fuentes, se espera que el nuevo sistema predictivo de deserción estudiantil, detecte tempranamente a posibles desertores, registre causas y acciones llevadas a cabo en cada caso en particular, con el consiguiente

seguimiento de los mismos. Mantendrá un registro histórico de causas reales de deserción que retroalimentarán al sistema.

Se espera que este sistema sirva a tutores, docentes, equipo interdisciplinario y autoridades académicas como fundamento a partir del cual generar estrategias de retención. Las partes que se beneficiarán por este proyecto son:

- ✓ Alumnos: que serán, junto a la facultad, los principales destinatarios de esta investigación.
- ✓ Docentes: que podrán conocer las características más vulnerables de sus alumnos y así poder orientar sus clases en función de las mismas.
- ✓ Tutores, Equipo Interdisciplinario y autoridades académicas de la facultad: como fundamento a partir del cual generar estrategias de retención.
- ✓ Empresas: que demandan personal capacitado de las distintas ingenierías, para puestos profesionales.
- ✓ Otras universidades e instituciones educativas: que contarán con una nueva metodología de análisis para el estudio de esta problemática ajustada a su situación.
- ✓ Estado: principal inversor en educación superior.

4. FORMACIÓN DE RECURSOS HUMANOS

La propuesta de investigación presenta como parte de sus objetivos la consolidación de un grupo de I+D, que bajo la tutoría de los directores y coordinadores, y mediante su formación y capacitación en áreas específicas, brinde la posibilidad a sus integrantes de participar en forma activa, con aportes sustanciales en investigación y desarrollo, especialmente en la temática propuesta.

Presenta actualmente la siguiente estructura: director (1), docentes investigadores (2), investigador de apoyo (1), tesis de postgrado

(1), tesistas de grado (2), becarios de investigación (3).

5. BIBLIOGRAFÍA

- [1] J.M. Braxton, A.S. Sullivan, R.M. Johnson (1997). Appraising Tinto's theory of college student departure. En *Journal Higher Education*, Volumen 12, pgs 107-164. New York, Ed: Agathon Press
- [2] M. Fishbein y I. Ajzen (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Ed. Addison-Wesley.
- [3] L. Attinasi (1986). Getting in: Mexican American Students' perceptions of their college going behavior with implications for their freshman year persistence in the University. En *ASH E 1986 Annual Meeting Paper*, ERIC N° 268 869. San Antonio, Texas.
- [4] C. A. Ethington (1990). A psychological model of student persistence. *Research in Higher Education*, 31, (3), 266- 269.
- [5] W. Spady (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1 (1) 64-85).
- [6] A. Cabrera, M. Castañeda, A. Nora, D. Hengstler (1992). The converge between two theories of college persistence. *Journal of Higher Education*, 63, pg. 143-164
- [7] E. Himmel (2002). Modelos de análisis de la deserción estudiantil en la educación superior. *Revista Calidad*.
- [8] J. Berger, J. Milem (2000). Organizational Behavior in higher Education and student outcomes. In: J. Smart. *Higher Education: Handbook of theory and research*. Vol 15: 268-338
- [9] V. Tinto (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125
- [10] O.T. Lenning (1982). Variable - selection and measurement concerns. *New Directions for Institutional Research*. Wiley Subscription Services, Inc., A Wiley Company, 35-53.
- [11] A. Vélez, D. F. López Jiménez (2004). *Estrategias para vencerla deserción universitaria*.
- [12] "Uso de Tecnologías de la Información para detectar posibles deserciones universitarias". M.F. Haderne. VII Congreso de Tecnología en Educación y Educación en Tecnología, Red de Universidades con Carreras en Informática (RedUNCI), 11 y 12 de Junio de 2012; Junín, Buenos Aires, Argentina.
- [13] Programa de Tutorías Para Primer Año: Una experiencia con la intervención de monitores alumnos y las tecnologías como soporte comunicacional. M.A. Odetti, A. I. Haefeli. 2° Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2014) 13 y 14 de Noviembre de 2014, UNSL, Facultad de Ciencias Físico, Matemáticas y Naturales; San Luis, Argentina.

Análisis y Evaluación de Desempeños Escritos en Entornos Virtuales de Aprendizaje: Categorías Discursivas y Rúbrica

Néstor Blanco, Nora Cuello, Paula Penco

Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora

Instituto Tecnológico para la Calidad (ITC)

Camino de Cintura y Avda. Juan XXIII – Lomas de Zamora

Teléfonos: 011-4282 7880 (interno 102)

e- mails: nestorhugoblanco@gmail.com ; norapatriciacuello@gmail.com ;
paucepenco@gmail.com

RESUMEN

Las Tecnologías de la Información y la Comunicación (TIC) se han ido insertando cada vez más en el ámbito educativo, dando lugar a entornos virtuales de aprendizaje. En estos espacios el docente y el alumno interactúan mediante distintas herramientas de comunicación, las cuales se desarrollan a través de desempeños escritos, esperando el logro de la construcción de conocimientos significativa y colaborativamente.

Una herramienta de análisis de los desempeños escritos son las categorías discursivas, las cuales nos permiten estudiar cómo se construye el conocimiento en estos entornos virtuales. En el presente trabajo se recuperan las categorías de análisis del discurso enunciadas por De Pedro Puente (2006) y se crean nuevas categorías para considerar los contenidos escritos que no son abordados por el autor de referencia.

En segundo lugar, se diseña una rúbrica como método de evaluación, análisis y posterior valoración, de la construcción colectiva del conocimiento mediante herramientas de comunicación virtual, como es el foro de debate, y se muestra una experiencia en el área humanística que posee el objetivo de desarrollar competencias lingüísticas, fundamentales para que los futuros profesionales sepan transmitir sus conocimientos, ideas y proyectos.

Palabras clave: categorías de análisis del discurso, herramientas de comunicación virtual, rúbrica.

CONTEXTO

Desde el año 2005, la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (FI-UNLZ) ha desarrollado diversas actividades relacionadas con las TIC en el aula, como complemento de la enseñanza presencial, apuntando a generar, en los estudiantes, un aprendizaje más activo, autónomo, construyendo el conocimiento en forma cooperativa y colaborativa, junto a la acción tutorial del docente como guía, capacitador y motivador para llevar adelante las actividades que se propongan en esos espacios virtuales.

El Instituto Tecnológico para la Calidad, dentro del cual se enmarca el presente trabajo, realiza investigación aplicada para el diseño y desarrollo de sistemas organizacionales de calidad relativos a las funciones directivas o responsables de instrumentar políticas, gerenciar procesos y mejorar las competencias humanas y organizacionales de instituciones del sector público, social y privado.

En cuanto a su visión, pretende desarrollar liderazgo en los procesos de diseño y transferencia de tecnologías apropiadas para la mejora de organizaciones e instituciones, vinculando la universidad, la empresa y las organizaciones sociales, mediante un compromiso de acción por la calidad, el compromiso social y el desarrollo del capital humano.

1. INTRODUCCIÓN

Este trabajo, enmarcado en la experiencia desarrollada en una asignatura del área de las humanísticas dentro de la Escuela Tecnológica Ingeniero Giúdice, perteneciente a la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (Fi-UNLZ), persigue, como objetivo general, fomentar la articulación entre el nivel medio y el superior.

Esta línea de investigación está íntimamente vinculada a la importancia de formar estudiantes con las competencias y el uso de herramientas que se determinan dentro del ámbito de la educación superior y el medio laboral. La Facultad de Ingeniería, en la cual estos estudiantes podrían seguir sus estudios superiores, apunta a formar en el marco de esas competencias y herramientas, algunas de las cuales persiguen incorporar las TIC en la formación académica, lo cual justifica que posea la mayoría de sus materias con aulas virtuales, como complemento de la enseñanza presencial, para el desarrollo de diversas actividades, en pos de que los estudiantes logren esas esperadas competencias.

Incorporando el computador a la educación, la cual pasa a estar mediada por TIC, se apunta a fomentar y lograr, en el estudiante, competencias como ser la autonomía, el trabajo en equipo, la toma de decisiones, la creatividad, la reflexión y autorreflexión, la empatía, bajo un rol del docente como guía y planificador, y un rol activo del estudiante, quien debe construir su propio aprendizaje.

Pérez Gutiérrez y Florido Bacallao afirman que la incorporación de tecnologías al ámbito educativo permite potenciar los modelos de educación a distancia y crear nuevas propuestas con fines de desarrollo profesional y formación permanente. [1]

En esta línea, los autores Vázquez-Cano, Martín-Monje y Fernández-Álvarez exponen que “el uso de las tecnologías de la información y la comunicación (TIC) se ha generalizado en los últimos años en la educación a varios niveles, y docentes y estudiantes usan un mayor número de recursos tecnológicos en las aulas, tanto para la preparación de materiales como en la instrucción y en la producción de tareas y

actividades. El ordenador se ha convertido en una herramienta primordial en la enseñanza, y la evaluación es cada vez más partícipe de esta nueva corriente”. [2]

En cuanto a la enseñanza superior, se aprecia “una tendencia hacia procesos evaluativos más orientados al proceso de enseñanza-aprendizaje, con el estudiante como protagonista, con un esfuerzo continuado por lograr una retroalimentación eficaz que suponga para éstos posibilidad de cambio o mejora (Boud y Falchikov, 2007; Fernández March, 2010). Esta evolución supone una visión diferente a la de la enseñanza tradicional, disociando claramente evaluación y calificación (Fernández March, 2010; Ibarra y Rodríguez Gómez, 2010)”. [2]

Siguiendo a los mencionados autores, “la evaluación es [...] decisiva por ser precisamente el motor del aprendizaje y una condición necesaria para mejorar la enseñanza (Boud y Falchikov, 2007). Dentro de las tres modalidades en que se suele dividir la evaluación (inicial, formativa y sumativa) resulta de gran interés la segunda, puesto que es la que se centra en el seguimiento del estudiante y permite realizar cambios ad hoc para optimizar el proceso de enseñanza-aprendizaje”. [2]

Por lo dicho anteriormente, estos autores afirman que la aplicación de matrices de evaluación o rúbricas apuntan a una evaluación formativa de los trabajos del estudiante, ayudando tanto en la preparación del trabajo, al conocer de antemano los objetivos, como también en la corrección del mismo por parte del docente, que se realiza de forma más objetiva, completa y eficaz.

Por su parte, Hawes B. dice que “...las rúbricas permiten apreciar o evaluar competencias académicas tales como habilidad para criticar, habilidad para producir trabajos académicos, para sintetizar y aplicar conceptos y principios adquiridos recientemente”. [3]

Así, podemos entender las rúbricas como métodos que “resultan bastante eficaces en la observación de indicadores de la productividad y destrezas lingüísticas; especialmente, en aspectos como la producción oral y escrita de

la lengua, y en procesos gramaticales y léxicos”. [2]

Por otro lado, en cuanto a las ventajas asociadas al uso de las rúbricas, ponen el foco en las expectativas y objetivos que el docente espera que los estudiantes logren, especificando los criterios con los cuales va a medir el progreso del estudiante.

Dado que nuestro interés está puesto en estudiar la construcción de conocimientos a partir del desempeño escrito por medios virtuales, como es el foro de debate, “dentro de las cuatro destrezas lingüísticas tradicionalmente consideradas -comprensión y expresión oral, comprensión y expresión escrita- es ésta última la que mayor interés ha suscitado en la utilización de rúbricas y e-rúbricas para su evaluación (Al-Jarf, 2011; Spence, 2010; Wilson, 2006). [2]

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se diseña una rúbrica como método de evaluación, análisis y posterior valoración, de las destrezas lingüísticas en la construcción colectiva del conocimiento en un entorno virtual de aprendizaje, siendo estas destrezas una de las competencias que pretende desarrollar la enseñanza media en articulación con el nivel superior, además del uso de TIC.

El mencionado foro de debate es una herramienta que brinda la Plataforma Virtual E-ducativa que posee la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (UNLZ) y, por ende, su escuela pre-universitaria donde se desarrolla la experiencia de este trabajo. Se indica a los estudiantes que, a partir de distintos interrogantes en torno a un tema propuesto por el docente, deberán debatir teniendo en cuenta los criterios de evaluación del docente. Además, dado que esta experiencia de foro es la primera o una de las primeras que los estudiantes desarrollan en estos medio virtuales, se les provee un material diseñado por el docente con las distintas formas de intervenir en un foro, haciendo hincapié, de manera general, en las categorías del discurso con las cuales se analizará uno de los criterios a evaluar, llamado “Tipo y calidad de las intervenciones”.

Retomando a Hawes B., “la construcción de una rúbrica depende del enfoque que se tenga de la misma. [...] en el enfoque sobre estructuras de desempeño sobre un continuum de logro [...] lo primero es concebir cuál es el desempeño aceptable sobre la tarea que se pretende evaluar [...], siempre habrá un punto en el cual el logro de la tarea es considera ‘aceptable’ y bajo el cual se la considera ‘inaceptable’. [...] A partir de esta definición o toma de posición en torno a qué es lo aceptable, se construyen los restantes componentes de la escala, hacia lo inaceptable y hacia lo superior”. [3]

El diseño de nuestra rúbrica toma el segundo enfoque mencionado, a partir del cual se ha utilizado una escala de 5 puntos, designados por una diferencia de 0,5 puntos cada uno, donde 0,5 es el puntaje o nivel mínimo y 5 es el puntaje o nivel máximo.

Esta escala está íntimamente relacionada con los cinco criterios que se pretende evaluar en el desempeño de los estudiantes, conocidos por ellos antes de llevar a cabo la actividad, ligados a las capacidades y tareas que se pretende que el estudiante logre, que son:

1. Contenido general de las intervenciones
2. Relación de las intervenciones propias con las aportadas por otros participantes
3. Tipo y calidad de las intervenciones (Categorías de análisis del discurso)
4. Uso del lenguaje
5. Pautas de forma (tiempo y espacio)

De todos los criterios mencionados cabe destacar que para definir el “Tipo y calidad de las intervenciones” se tomaron las categorías de análisis del discurso del autor De Pedro Puente, X. (2006), en su ponencia “Cómo evitar el ‘café para todos’ al evaluar trabajos en grupo, y de paso, estimular el aprendizaje reflexivo: resultados preliminares en el marco del proyecto A Wiki Forum”, que son:

- * Aspectos organizativos AO
- * Peticiones de ayudas PS
- ** Ayudas a compañeros AC
- ** Nueva información NI
- ** Reflexión personalizada y argumentada RP
- *** Hipótesis nuevas HN
- *** Preguntas elaboradas y nuevas vías para avanzar PE/NV

- *** Síntesis/ elaboración de información SI
- * Respuestas sencillas RS
- * Definición de conceptos DC
- ** Réplica REP
- * Referencias de ampliación REF
- * Opinión de personal OP
- ** Refuerzo de información RI
- * Facilitación F
- * Agradecimientos A

Los asteriscos delante de cada categoría indican, siguiendo al autor referenciado, indican el nivel de contribución que aportan. Así, cuantos más asteriscos hay delante de la “Categoría / Tipo de contribución” mayor importancia se da a la contribución, en cuanto a la interacción que provoca en pos de la construcción de conocimientos.

Bajo el análisis realizado en el foro, nos encontramos con intervenciones que aportan a la construcción del conocimiento, pero sin encuadrar en ninguna de las categorías propuestas por el autor referenciado anteriormente. Por este motivo, se crean algunas categorías más, con su correspondiente nivel de contribución a la construcción de conocimientos, que se detallan a continuación:

- ** Propuestas de trabajo PT
- * Menció n M
- ** Postura propia fundamentada PPF
- * Emoticones EM
- * Comentarios personales CP
- * Motivación MO
- * Recordatorios REC

El análisis realizado admite que cada categoría pueda aparecer una sola vez por intervención, es decir, cada intervención puede ser categorizada por más de una categoría, pero no por categorías repetidas.

La experiencia del presente trabajo se basa en el análisis del desarrollo escrito de un foro de debate con 38 estudiantes que debían intervenir, como mínimo, dos veces. Pero cada participación debía tener una prudencial distancia una de la otra, con el objetivo de que el estudiante lleve a cabo la lectura de las participaciones previas y realice una contribución que marque una interrelación entre los aportes previos y los nuevos.

Es así que, para llevar adelante la evaluación de los desempeños escritos mencionados se diseña una rúbrica a partir de los criterios de evaluación tenidos en cuenta por el docente, como capacidades o subcompetencias, los cuales se dieron a conocer a los estudiantes previamente.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el caso del “Contenido general de las intervenciones” notamos que el desarrollo escrito refleja, mayormente (en casi la mitad del alumnado), un discurso acorde al tema, alternando información nueva e información replicada y, en segundo lugar, un discurso analítico con aportes de nueva información, demostrando lectura y conocimientos previos. En cuanto a la “Relación de las intervenciones propias con las aportadas por otros participantes”, se destaca notablemente (más de la mitad del alumnado) el nivel más bajo, reflejándose un desempeño sin valor en este criterio, donde los participantes emiten intervenciones nuevas sin demostrar lectura ni relación con las intervenciones anteriores, es decir, con un carácter individualista.

En relación al “Tipo y calidad de las intervenciones”, siendo éste uno de los criterios más complejos para analizar, se puede ver que los niveles que destacan son el medio y superior, con la misma cantidad (en casi la mitad del alumnado), demostrando intervenciones que exponen aspectos organizativos acerca de los contenidos que se deben presentar; ayudas a participantes; réplica de lo aportado por otro/s participante/s; recordatorio acerca de una petición, interrogación o cuestión que ha quedado sin debatir; nueva información; refuerzo de información; opiniones favorables o críticas respecto de lo aportado por otro/s participante/s; facilitación identificando y/o corrigiendo errores o incomprensiones presentadas por otro/s participantes/s; enunciados que transmiten motivación a otro/s participante/s. Como también desarrollan hipótesis nuevas; preguntas elaboradas y nuevas vías para avanzar; síntesis/elaboración de información; reflexión personalizada y

argumentada de lo debatido; agradecimiento respecto de lo aportado por otro/s participante/s, y/o agradecimiento por ser leído por los demás participantes; postura propia fundamentada. En estos últimos desempeños se demuestra lectura, comprensión, análisis y nuevas propuestas/temas de debate.

Pasando al “Uso del lenguaje”, la mitad del alumnado obtiene el nivel medio, lo cual significa que utilizan un lenguaje mayormente correcto, pero presentan algunas faltas de tildes, signos de puntuación y/o de tipeo.

Por último, el criterio a evaluar “Pautas de forma (tiempo y espacio)” es el que presenta más variación en el desempeño de los estudiantes. Casi la mitad del alumnado logra un nivel regular, ingresando una sola vez al foro de debate, sea en el inicio o tardíamente, cumpliendo o no con la extensión permitida para cada intervención, o superándola brevemente, y en una cantidad muy reducida desempeños excelentes, donde se cumplen las pautas en relación al tiempo, ingresando y participando en el debate en sus distintos períodos, de inicio, desarrollo y finalización, cumpliendo con la extensión pedida en los distintos aportes, demostrando firmeza en las ideas incluso sin necesidad de un discurso extenso. Además, en estos últimos casos existe una notoria separación entre cada una de las intervenciones propias, demostrando lectura de las demás intervenciones anteriores.

Por otro lado, si bien la construcción de conocimientos está acentuada en la figura del estudiante, quien debe realizar esta construcción en interrelación con sus pares, el docente es un tutor que guía en este proceso, abriendo nuevos interrogantes, conduciendo el debate, enunciando recordatorios y generando motivación a lo largo de todo el debate, de acuerdo a cómo se va desarrollando éste.

En función de la experiencia se recomienda que el docente, notando un desempeño individualista en el debate, además de recordar las pautas y modos de intervenir, también debe recuperar contribuciones de distintos participantes y fomentar la opinión, favorable o crítica, como también una reflexión o postura propia acerca de lo aportado por otros estudiantes. De esta manera tratar de lograr una

construcción de conocimientos más colaborativa, y no sólo validar aportes de informaciones nuevas o reflexiones sobre los propios aportes.

Con respecto al uso de la rúbrica, como método de evaluación formativa, atendiendo los distintos criterios o capacidades que el docente pretende que el estudiante logre, con su previo conocimiento de las mismas, es una herramienta útil y clara a la hora de realizar dicha evaluación.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo está liderado por el Mg. Néstor Blanco, director del Instituto Tecnológico para la Calidad (ITC), a cargo del área de calidad educativa y acreditación de las carreras de grado y posgrado ante CONEAU.

Lidera proyectos de investigación vinculados a las cuestiones antes mencionadas, junto a la función de director de una tesis de maestría en curso en la misma línea investigativa, habiendo una tesis de maestría ya aprobada.

5. BIBLIOGRAFÍA

- [1] PÉREZ GUTIÉRREZ, A., FLORIDO BACALLAO, R. (2003). Internet: un recurso educativo. Revista Etic@net, España, n.2., disponible online en <http://www.ugr.es/~sevimeco/revistaeticanet/Numero2/Articulos/Intrecedu.pdf>
- [2] VÁZQUEZ-CANO, E., MARTÍN-MONJE, E., FERNÁNDEZ-ÁLVAREZ, M. (2014). El rol de las e-rúbricas en la evaluación de materiales digitales para la enseñanza de lenguas en entornos virtuales de aprendizaje, Revista de docencia universitaria (REDU), EEUU, v.12, n.1, p.135-157, disponible online en <http://red-u.net/redu/index.php/REDU/article/view/756>
- [3] HAWES B., G. (2004). Evaluación: estándares y rúbricas, Proyecto Mecsup TAL101, Universidad de Talca, Chile, disponible online en <http://www.freewebs.com/gustavohawes/Educacion%20Superior/2004EstandaresRubricas.pdf>

Aportes de las TIC desde el Enfoque Ontosemiótico: Análisis de los Errores en la Modelización de Problemas de Probabilidad y Estadística en la FI-UNLZ

Facundo Frende¹, Zulma Torres¹, Guadalupe Pascal¹, Claudia Minnaard¹

¹ Instituto de Investigaciones de Tecnología y Educación (IIT&E) - Centro Asociado CIC Facultad de Ingeniería - Universidad Nacional de Lomas de Zamora

E-correo: facundo.frende@hotmail.com; torreszulm@hotmail.com;
gpascal@ingenieria.unlz.edu.ar, cminnaard@ingenieria.unlz.edu.ar

Resumen

La línea de investigación busca indagar la relación entre los errores en las producciones de los alumnos de la cátedra de Probabilidad y Estadística de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora.

Al analizar los vestigios digitales de la cursada 2016 en el aula virtual de la cátedra, en el marco de un proceso gradual y sostenido de la casa de estudios para incorporar el uso de TIC en la enseñanza, se propone aportar elementos para el diseño de un modelo que describa la tipología de los errores.

Así como ajustar el sistema de prácticas, los objetos de estudio y la aplicación de tecnologías en educación, desde un enfoque Ontosemiótico, que permitan un exitoso desempeño académico del alumno.

Palabras claves: Probabilidad y Estadística; errores; TIC en la enseñanza, Enfoque Ontosemiótico

Contexto

Desde el año 2005, la Facultad de Ingeniería de la Universidad de Lomas de Zamora, viene desarrollando un proceso gradual y sostenido para incorporar el uso de TIC en la enseñanza. Esto no solo implica el empleo de plataformas virtuales para cada una de las cátedras dictadas, sino también la aplicación de diversos Softwares que son usados por los alumnos durante el dictado de las materias.

El presente trabajo se encuadra en la temática y los objetivos del Proyecto de Investigación: "TIC: Modelos y Objetos de Enseñanza, Su Aplicación en Carreras Tecnológicas del NiveSuperior", en el marco del esquema de

trabajo del Instituto de Investigaciones de Tecnología y Educación (IIT&E) de la mencionada casa de estudios.

Contribuyen en las etapas del mencionado proyecto: becarios del Programa Becas de Estímulo a las Vocaciones Científicas 2016 y docentes/tutores en la modalidad Blended Learning, de la cátedra de Probabilidad y Estadística.

Particularmente en la cátedra de Probabilidad y Estadística, el uso del programa Minitab les permite a los alumnos aplicar un software especializado en la materia para facilitar la resolución de los problemas y, a su vez, permite generar bases de datos con las producciones de dichos alumnos, a través de los vestigios digitales en el aula virtual de la cátedra.

Abordando la información proporcionada en éstas bases de datos, se propone realizar análisis estadísticos para indagar sobre la relación entre los errores cometidos por los alumnos de la cátedra a través de la frecuencia de ocurrencia del error y el tipo del problema propuesto.

Introducción

Abordar las investigaciones en matemática educativa, significa considerar las relaciones entre el sistema de prácticas y la configuración de objetos y procesos (Godino, Batanero y Font, 2007; Font, Godino y Gallardo, 2012). El Enfoque Ontosemiótico (EOS) proporciona herramientas de análisis que permiten indagar en estas relaciones. Es importante destacar que el problema epistémico cognitivo no puede desvincularse del ontológico. (Godino, 2012)

A partir de la observación de las producciones de los alumnos, es posible observar que aquellos con muy buen desempeño académico cometen errores. Esto se debe a que la producción del error no es necesariamente falta de estudio, sino la aplicación de un “esquema cognitivo inadecuado”. (Socas, 1997)

Los errores tienen su génesis en el mismo proceso de aprendizaje y estos a su vez se conectan formando redes verdaderamente complejas, actuando como obstáculos que se translucen en la práctica en respuestas erróneas. (Del Puerto, Minnaard, Seminara, 2006)

Las producciones de los alumnos (exámenes parciales y trabajos prácticos) permiten indagar la aparición del error vinculándolo con el sistema de prácticas y la configuración de objetos y procesos.

Radillo Enríquez y Huerta Varela (2007) consideran que “algunos obstáculos o dificultades que encara un estudiante en el manejo del lenguaje matemático son fuente de errores en la solución de problemas. La naturaleza del obstáculo puede explorarse mediante el análisis de los errores cometidos”. En el ámbito de la educación matemática los errores aparecen permanentemente en las producciones de los alumnos: las dificultades de distinta naturaleza que se generan en el proceso de aprendizaje se conectan y refuerzan en redes complejas que obstaculizan el aprendizaje, y estos obstáculos se manifiestan en la práctica en forma de respuestas equivocadas. (Del Puerto, Minnaard & Seminara, 2006) Bachelard (1988) (citado por Rico, 1995) plantea el concepto de “obstáculo epistemológico”, considerando que:

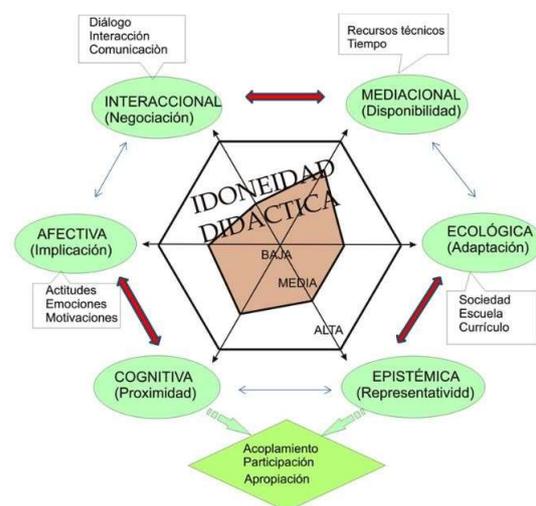
“Cuando se investigan las condiciones psicológicas del progreso de la ciencia hay que plantear el conocimiento científico en términos de obstáculos...el conocimiento de lo real es una luz que siempre proyecta alguna sombra, jamás es inmediata y plena. Al volver sobre un pasado de errores se encuentra la verdad. En efecto, se conoce en contra de un conocimiento anterior, destruyendo conocimientos mal adquiridos o superando aquello que, en el espíritu mismo, obstaculiza”.

Según Socas (1997), “el error debe ser considerado como la presencia en el alumno de un esquema cognitivo inadecuado y no sólo la consecuencia de una falta específica de conocimiento o una distracción”. Brousseau, Davis y Werner (1986) (citados por Rico, 1995), señalan, en el mismo sentido, que “los errores son el resultado de un procedimiento sistemático imperfecto que el alumno utiliza de modo consistente y con confianza”.

El Enfoque Ontosemiótico (EOS) al vincular el sistema de prácticas con la configuración de los objetos y procesos permite evaluar si existe una relación entre la aparición del error y un sistema de prácticas inadecuado. (Pochulu, 2005, Godino, 2012).

Godino (2013) introduce la “noción de idoneidad didáctica” como herramienta dentro del EOS vinculando Idoneidad epistémica, Idoneidad cognitiva, Idoneidad interaccional, Idoneidad mediacional, Idoneidad afectiva e Idoneidad ecológica tal como se observa en el Gráfico 1.

Gráfico 1: Componentes y criterios básicos de idoneidad didáctica



Godino (2013) describe cada una de las idoneidades que llevan a la idoneidad didáctica. La idoneidad epistémica hace referencia a la representatividad de los significados; la cognitiva expresa la proximidad que se encuentran los contenidos con respecto al desarrollo potencial de los alumnos; la interaccional hace referencia a aquellas acciones docentes que permiten

eliminar los obstáculos epistemológicos; la mediacional hace referencia a la disponibilidad de los medios y recursos necesarios en el proceso de enseñanza aprendizaje; la afectiva tiene que ver con el interés y la motivación del alumno y por último la ecológica indica el grado en que el proceso educativo se enlaza en la institución de referencia.

Asimismo, Minnaard (2014) indaga sobre las competencias que en el área de matemática traen los alumnos ingresantes a las carreras de ingeniería. Los exámenes, surgidos de la aplicación del Test Diagnóstico, de alumnos ingresantes a las carreras de ingeniería de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora, fueron el objeto de estudio de la investigación.

Tal como se observa en el Tabla 1 al relacionar la ocurrencia del error con los tipos de problemas propuestos se consideran dos tipologías.

Tabla 1: Relación Tipo de problemas

	TIPO 1	TIPO 2
	Ocurrencia del error: baja	Ocurrencia del error: alta
Tipos de Problemas	<ul style="list-style-type: none"> - No tienen enunciado textual (solamente) - Tienen gráficos. - No es necesario aplicar fórmulas. - Aplican definiciones. - Aplican propiedades. - No se realizan operaciones numéricas. 	<ul style="list-style-type: none"> - Tienen enunciado textual (solamente) - No aplican definiciones. - No aplican propiedades. - No tienen expresiones algebraicas. - Realizan operaciones numéricas. - Aplican fórmulas.

El Análisis Factorial realizado para determinar estas tipologías, permite corroborar una de las hipótesis de trabajo de la investigación realizada, ya que la ocurrencia de los errores que cometen los alumnos en el Test Diagnóstico está relacionada con el tipo de problema que se le presenta para resolver.

Ahora bien, Samaja (2001) citando a Kant, afirma que:

“Nosotros sólo podemos pensar aquello que es pensable con los formatos básicos de nuestro pensamiento. El mundo, la realidad podrán ser lo que sea, pero cualquier cosa que sea en cuanto a sus contenidos, lo que logremos pensar será el resultado de tales contenidos, en la forma en que nuestro intelecto puede considerarlos”.

Tomemos un ejemplo, el símbolo \bar{x} para el común de la gente no tiene significado alguno. Para algún alumno con conocimientos de matemática puede interpretar la x como una incógnita, pero si ese alumno tiene conocimientos de estadística sabrá que simboliza la media aritmética, es decir el promedio aritmético de los valores de la variable considerada.

Líneas de Investigación, Desarrollo e Innovación

El objetivo principal de la investigación consiste en indagar la relación entre los errores cometidos por alumnos de la cátedra de Probabilidad y Estadística de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora en el transcurso del ciclo 2016 a través de la frecuencia de ocurrencia del error y el tipo de problema propuesto.

Las hipótesis de trabajo son:

- Existen patrones de errores propios relacionados con los contenidos de la materia Probabilidad y Estadística.
- Existen factores clave que favorecerían la ocurrencia del error con mayor frecuencia.
- La ocurrencia de error que cometen los alumnos en Probabilidad y Estadística está relacionado con el sistema de prácticas.
- El tipo de error que cometen los alumnos en Probabilidad y Estadística está relacionado con el tipo de objeto semiótico con el que debe trabajar.

Para corroborar o refutar las hipótesis planteadas se considera el portafolio de actividades realizadas por los alumnos de la cátedra de Probabilidad y Estadística de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora: Exámenes parciales y Trabajos prácticos del Primer y Segundo Cuatrimestre de 2016.

Resultados y Objetivos

En la etapa inicial de la investigación, se indagó sobre las relaciones entre las prácticas docentes, los objetos matemáticos involucrados y el tipo de error cometido.

Los resultados y avances en esta instancia fueron presentados recientemente en congresos de divulgación científica, bajo los títulos:

- Entorno Virtual de Aprendizaje en la Formación Estadística del Ingeniero: Nuevas experiencias pedagógicas en la Cátedra Probabilidad y Estadística de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora. (2016)
- Dinámica de las interacciones en el aula virtual de Probabilidad y Estadística. (2017)

El trabajo tiene un importante componente de análisis estadístico, focalizado en el estudio de la información localizada en la plataforma y sus portafolios virtuales. De donde se espera obtener:

- El relevamiento de los errores a partir de un instrumento adecuado.
- Un modelo, que probado, describa la tipología de los errores.

Los objetivos específicos que se han definido son:

- Indagar los errores en las producciones de los alumnos en Probabilidad y Estadística.
- Releva y sistematiza los errores buscando patrones de regularidad.
- Aportar elementos para el diseño de un modelo teniendo en cuenta que los alumnos de la Facultad de Ingeniería de la UNLZ cursan en la modalidad Blended Learning y utilizan las funciones estadísticas básicas y avanzadas de Minitab.

Formación de Recursos Humanos

De la conformación podemos señalar que el equipo de trabajo cuenta con aportes del perfil profesional de cada uno de los miembros que asumimos esta investigación:

Los docentes de la cátedra abordada, brindan al equipo de investigación sus experiencias en

la enseñanza con modalidad Blended Learning, en Metodología de la Investigación, y en el abordaje del Enfoque Ontosemiótico.

Asimismo, se han formado dos becarios: uno en el marco del Programa Estímulo a las Vocaciones Científicas y el otro en el LOMASCYT, aplicando técnicas estadísticas para el análisis de datos siendo acompañados por directores del proyecto marco.

Referencias

Batanero, C., Vera, O. D., & Díaz, C. (2012). Dificultades de estudiantes de Psicología en la comprensión del contraste de hipótesis. *Números. Revista de Didáctica de las Matemáticas*, 80, 91-101.

Brousseau, G., Davis, R. B., & Werner, T. (1986). Observing students at work. In *Perspectives on mathematics education* (pp. 205-241). Springer Netherlands.

Comoglio, M.; Minnaard, C.; Iravedra, C. & Morrongiello, N. (2012). La integración de TIC a la enseñanza de la Ingeniería- Estudio comparativo de su impacto en el rendimiento académico. 1° Congreso Argentino de Ingeniería (CADI 2012) y VII Congreso Argentino de Enseñanza de la Ingeniería (CAEDI 2012), Mar del Plata, Argentina.

Seminara, S. A., Del Puerto, S. M., & Minnaard, C. L. (2006). Análisis de los errores: una valiosa fuente de información acerca del aprendizaje de las Matemáticas. *Revista Iberoamericana de educación*, 38(4), 7.

Godino, J. D. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en Didáctica de la Matemática. *Investigación en Educación Matemática XVI*, 49-68.

Godino, J. D., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM*, 39(1-2), 127-135.

Godino, J. D. (2013) Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. *Cuadernos de Investigación y Formación en Educación Matemática*. 2013. Año 8. Número 11. pp. 111-132. Costa Rica.

Minnaard, C., & Minnaard, V. (2011) Materiales multimediales en el nivel Superior. II Congreso Iberoamericano de Educación y Sociedad (CIEDUC 2011).

Minnaard, C. (2016). Análisis de los errores en matemática de los alumnos ingresantes a las carreras de Ingeniería: el Test Diagnóstico en la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora. Revista Iberoamericana de Producción Académica y Gestión Educativa.

Minnaard, C., Pascal, G., Torres, Z., & Frende, F. (2017). Entorno Virtual de Aprendizaje en la Formación Estadística del Ingeniero: Nuevas experiencias pedagógicas en la Cátedra Probabilidad y Estadística de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora. Revista Electrónica sobre Tecnología, Educación y Sociedad, 4(7).

Minnaard, C., Pascal, G., & Torres, Z. (2017) Dinámica de las interacciones en el aula virtual de Probabilidad y Estadística. IX Congreso Iberoamericano de Educación y Sociedad (CIEDUC 2017)

Pochulu, M. (2005). Análisis y categorización de errores en el aprendizaje de la matemática en alumnos que ingresan a la universidad. Revista Iberoamericana de Educación, 35(4), 1-14.

Rico, L. (1995). Errores y dificultades en el aprendizaje de las matemáticas.

Samaja, J. (2007). Epistemología y metodología: elementos para una teoría de la investigación científica. Eudeba.

Samaja, J. (2001) Ontología para Investigadores. Las categorías puras del intelecto en Imanuel Kant. Revista Perspectivas Metodológicas. Ediciones de la UNLA. Año 1. N° 1 pp. 11 - 42. Buenos Aires.

Robayna, M. M. S. (1997). Dificultades, obstáculos y errores en el aprendizaje de las matemáticas en la Educación Secundaria. In La educación matemática en la enseñanza secundaria (pp. 125-154). HORSORI.

Construyendo un Curso Masivo en Línea sobre Accesibilidad Web. Planteos y Desafíos de los Nuevos Entornos de Aprendizaje

Javier Díaz, Alejandra Schiavoni, Paola Amadeo

LINTI - Laboratorio de Investigación en Nuevas Tecnologías Informáticas.

Facultad de Informática. Universidad Nacional de La Plata

Calle 50 esq. 120, 2do Piso. Tel: +54 221 4223528

jdiaz@unlp.edu.ar, ales@info.unlp.edu.ar, pamadeo@linti.unlp.edu.ar

Resumen

El presente artículo presenta las líneas de investigación que se vienen llevando a cabo en el Laboratorio de Investigación en Nuevas Tecnologías Informática de la Facultad de Informática de la UNLP, en relación con la utilización e integración de diferentes plataformas de software libre y la generación de contenido educativo abierto. El aprendizaje abierto es una concepción de educación que tiene por objetivo eliminar todas las barreras del aprendizaje y proporcionar al estudiante un sistema de educación y entrenamiento centrado en sus necesidades específicas y localizado en múltiples áreas de conocimiento. Esto tiene relación con los REA, Recursos Educativos Abiertos (en inglés OER – Open Educational Resources), que se refiere a cualquier recurso educativo que esté plenamente disponible para ser usado por educadores y estudiantes. El concepto de REA se asocia a otro denominado OpenCourseware – OCW que representa Materiales de Cursos Abiertos, una publicación digital libre y abierta de materiales educativos de nivel universitario. A su vez, los MOOCs – Massive Open Online Courses son cursos online libres sin límite de participación, que incluye interacción, retroalimentación, y evaluaciones.

Las líneas a seguir en esta etapa

comprenden la realización de un relevamiento de iniciativas de OCW y MOOCs analizando herramientas de acceso libre. El objetivo es armar MOOCs sobre temas que se estudian en el laboratorio y que resultan de gran interés. Uno de estos temas es el de Accesibilidad, sobre el cual se viene dictando un curso a distancia desde hace más de cinco años con un gran número de inscriptos.

Palabras clave: Recursos educativos abiertos, OER, OCW, MOOC, accesibilidad

Contexto

El proyecto descrito en este artículo se desarrolla en el Laboratorio de Investigación en Nuevas Tecnologías Informáticas, LINTI de la Facultad de Informática de la UNLP y está enmarcado en el proyecto 11-F020 “Internet del futuro: Ciudades digitales inclusivas, innovadoras y sustentables, IoT, ciberseguridad y espacios de aprendizaje del futuro”, acreditado en el marco del Programa de Incentivos, bajo la dirección del Lic. Javier Díaz.

El tema de accesibilidad se viene estudiando desde hace varios años, a lo largo de los cuales se desarrollan

diferentes proyectos de extensión acreditados por la Universidad Nacional de La Plata. Los proyectos de los últimos dos años son “Por una Web Inclusiva” [1] y “Trabajando por una Web Accesible” [2].

A partir de las plataformas Moodle instaladas en la Facultad de Informática de la UNLP como complemento de las clases presenciales y en cursos completamente no presenciales y semipresenciales desde el año 2003, es relevante llevar adelante distintas experiencias en el diseño e implementación de MOOCs, incluso utilizando Moodle como plataforma de base. García Aretio plantea una serie de interrogantes a partir de numerosas investigaciones consultadas referidas a las posibilidades de los MOOCs para la formación profesional y continua, para abordar las problemáticas de discapacidad y género por ejemplo [3].

Introducción

Los REA, Recursos Educativos Abiertos, han emergido como un concepto con gran potencial para apoyar la transformación educativa. Aunque su valor educativo reside en la idea de usar recursos como método integral de comunicación de planes de estudios en cursos educativos (es decir, aprendizaje basado en recursos), su poder de transformación radica en la facilidad con la que esos recursos, una vez digitalizados, pueden ser compartidos a través de la Internet. La característica distintiva de estos recursos es la licencia de uso, que tiene como objetivo facilitar su reutilización y potencial adaptación, sin tener que solicitar autorización. Los REA ayudan a proporcionar apoyo al aprendizaje abierto, que es en sí una concepción de educación que tiene por

objetivos eliminar todas las barreras del aprendizaje y permitir que el proceso de aprendizaje se centre en el estudiante, y en sus necesidades específicas.

El desarrollo de OpenCourseWare se basó en sus comienzos en la idea de recursos abiertos, donde la atención se concentra en el desarrollo e intercambio de materiales libremente disponibles, autónomos para cursos y enseñanza online. Una buena parte del esfuerzo en este sentido ha sido realizado por el Consorcio OCW [4] [5]. El consorcio cuenta con miembros en todo el mundo y actualmente hay más de 2500 cursos abiertos disponibles en más de 200 universidades.

Actualmente, hay varias iniciativas OCW en las universidades, que se encuentran en etapas avanzadas de desarrollo. El Instituto Tecnológico de Massachusetts (MIT) probablemente tiene el proyecto institucional OCW más conocido y es responsable de involucrar a muchas universidades de todo el mundo en el movimiento REA. El OCW del MIT [6] tiene a disposición en Internet 1900 cursos gratuitos con fines no comerciales. En el MIT se muestra claramente que este proceso de intercambio de materiales ha dado lugar a un aumento significativo del uso compartido de contenidos dentro de su propia institución.

Otra fuente institucional conocida de REA para la educación superior es OpenLearn [7]. La Universidad Abierta es una de las universidades de educación a distancia más exitosas del mundo. Si bien, el conjunto de iniciativas es extenso, mencionamos estas dos a modo de ejemplos.

Después que el proyecto de OCW se consolidó, comenzó a surgir la necesidad de incorporarle herramientas para que los cursos resulten más interactivos, dinámicos y con mayor repercusión social. También, se comenzó a considerar

la incorporación de evaluaciones. Las actualizaciones planteadas dieron lugar a una nueva herramienta educativa: MOOC - Massive Open Online Courses. Mientras los sitios OCW fueron desarrollados básicamente por instituciones universitarias, los proveedores de MOOC son en su mayoría empresas, tales como Coursera, Udacity, Miríada X, entre otras, y que tienden a usar algún tipo de licencia. La plataforma Miríada X surgió como una plataforma para el almacenamiento de contenido de acceso libre, licenciado con Creative Commons, sin embargo no existen herramientas internas para etiquetar los recursos bajo esta licencia.

Dentro de las herramientas de código abierto, se encuentran OpenMOOC y edX, que comparten muchas funcionalidades, pero difieren en su metodología de desarrollo y en su arquitectura.

Los MOOCs representan una herramienta fantástica para enseñar y participar en un espacio colaborativo.

Líneas de Investigación, Desarrollo e Innovación

Las tareas de investigación, desarrollo e innovación que se están llevando a cabo comprenden varias líneas de trabajo relacionadas a e-learning y cuestiones de interoperabilidad e integración de plataformas utilizadas en este campo. Además, se continúa con la línea de generación de recursos educativos abiertos que ofrecen un gran potencial para apoyar la transformación educativa. Asimismo, se aplican los trabajos de investigación que se vienen realizando sobre las normas de accesibilidad y su aplicación en el desarrollo de sistemas y sitios Web.

Continuando con las líneas planteadas en WICC 2015 [8] y WICC 2016 [9], en esta etapa el foco de estudio se relaciona con los pasos a seguir para la creación de cursos masivos MOOCs sobre temas que se investigan y sobre los cuales se viene trabajando en el LINTI. Se investigarán a fondo las plataformas existentes, a partir de las iniciativas que fueron surgiendo a lo largo de los años relativas a OCW y MOOC.

Se estudiarán estrategias para la planificación y diseño de un MOOC sobre Accesibilidad, definiendo la planificación y organización del contenido. El tema de Accesibilidad representa un tema de estudio e investigación en el LINTI desde hace varios años.

La accesibilidad en la Web permite el uso de herramientas de adaptación, el acceso desde diferente hardware y aspectos tecnológicos, uso con o sin teclado, con o sin mouse, etc., uso de distintos sistemas operativos, acepta distintas versiones de navegadores, apunta al diseño universal, incentiva buenas prácticas de programación, mejora la calidad de uso –usabilidad- del sitio y facilita construcción de sitios móviles, entre otras cuestiones.

El curso de accesibilidad Web en modalidad no presencial implementado en Moodle, cursos.linti.unlp.edu.ar, ya va por la IX edición, desde el año 2013. Inicialmente pensado para desarrolladores Web, actualmente se encuentra abierto a toda la comunidad interesada en la temática con un módulo especial para los desarrolladores.

El curso cuenta con 4 unidades temáticas, entregas de ejercicios, participación en foros de debate sobre diferentes problemáticas y la realización de un trabajo que aporte en forma significativa a la accesibilidad Web. Dependiendo de la formación del

participante puede ser un video de concientización, la evaluación de la accesibilidad Web de un sitio público según estándares internacionales, herramientas automáticas y heurísticas, entre otros y desarrollo Web. Ya cuenta con tutores formados y colaboradores que se suman año a año para participar de las nuevas ediciones.

A partir de esta experiencia es interesante evaluar distintas metodologías y herramientas para transformar este curso en un MOOC, analizando y evaluando también nuevas plataformas de base, nuevos medios de comunicación, contenidos, formatos, y demás cuestiones planteadas en distintos artículos y MOOCs como el presentado por la Universidad Carlos III de Madrid [10] o los pasos a seguir para la creación de MOOCs planteado por la Universidad Autónoma de Barcelona [11]. El trabajo interdisciplinario con diseñadores visuales, expertos en comunicación, contenidos y educadores permitirá generar una experiencia y un conocimiento enriquecedor, extensible a otras temáticas y líneas de investigación del LINTI.

Resultados y Objetivos

Según las líneas de trabajo descriptas, se plantean los siguientes objetivos:

- Relevamiento de iniciativas OpenCourseWare - OCW.
- Evaluación de plataformas open source y plataformas abiertas para implementar MOOCs .
- Analizar la creación de un MOOC desde distintas perspectivas docente, institucional y técnica, a partir del estudio de distintas experiencias.
- Analizar, evaluar e implementar una metodología para la creación

de MOOCs.

- Implementar un MOOC de accesibilidad Web de acuerdo a la metodología planteada, a partir de una planificación acorde y utilizando una de las herramientas o iniciativas evaluadas.
- A partir de la planificación estipulada, analizar los resultados obtenidos en cada etapa, aprendiendo de los errores en un ciclo de mejora continua.
- Realizar estudios comparativos entre las dos experiencias que permitan aportar nuevos conocimientos significativos a los nuevos entornos de aprendizaje y enseñanza que se generan con estas herramientas.
- Medir el grado de aceptación del curso por parte de los participantes y realizar análisis de datos considerando distintas variables como formación, edad, género, procedencia, participación y rendimiento, entre otros.

Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por dos profesoras de amplia trayectoria en el campo de la investigación, dos alumnos de la carrera de Licenciatura en Informática de la Facultad de Informática.

Además, se considera integrar a la Prof. Ivana Harari, que realiza trabajos de investigación sobre el tema de accesibilidad Web desde hace varios años.

La participación en eventos de la especialidad, ha permitido seguir estableciendo canales de comunicación con otros investigadores que trabajan en las mismas áreas.

Referencias

[1] Dictamen Proyectos de Extensión UNLP 2014.

http://unlp.edu.ar/uploads/docs/dictamen_proyectos_ext_2014.pdf

[2] Dictamen Proyectos de Extensión UNLP 2015

http://unlp.edu.ar/uploads/docs/dictamen_proyectos_extension_2015.2015.pdf

[3] MOOC en la Universidad

<http://aretio.blogspot.com.ar/2015/11/la-filosofia-educativa-de-los-mooc-y-la.html>

[4] OpenCourseWare - OCW Consortium

<http://ocwconsortium.org>

[5] Guía Básica de Recursos Educativos Abiertos (REA), Organización de las Naciones Unidas, para la Educación, la Ciencia y la Cultura, 2015.

[6] MIT OpenCourseWare

<http://ocw.mit.edu>

[7] OpenLearn, Free Learning from the Open University.

<http://openlearn.open.ac.uk>

[8] J. Díaz, A. Schiavoni, P. Amadeo, A. Osorio, F. Pietroboni, M. Pagano. “Integrando un Repositorio Digital con un Sistema de Gestión de Bibliotecas a través de OAI-PMH”. WICC 2015, XVII

Workshop de Investigadores en Ciencias de la Computación, 1era Edición, ISBN: 978-987-633-134-0, Salta, 16-17 Abril, 2015.

[9] J. Díaz, A. Schiavoni, P. Amadeo, D. Ray. “Construyendo un Repositorio Digital Accesible”. WICC 2016, XVIII Workshop de Investigadores en Ciencias de la Computación, 1era Edición, ISBN: 978-950-698-377-2, Concordia, 14-15 Abril, 2016.

[10] edX edge, “Cómo crear un MOOC de éxito con Open edX”.

https://edge.edx.org/courses/course-v1:UC3Mx+UC3Mx.1x+2016_17_T2/about

[11] Universidad Autónoma de Barcelona. Crear y Planificar un MOOC.

<http://www.uab.cat/web/estudiar/mooc/planificar-y-disenar-un-mooc/definir-objetivos-y-destinatarios-1345668281345.html>

Desarrollo de Competencias en Carreras de Ingeniería con Apoyo en Tecnología de Simulación: Impacto en la Satisfacción y la Calidad Percibida por los Alumnos

Comoglio Marta, Rolon Hugo, Moret Pablo, Lacanna Oscar

Facultad de Ingeniería - Universidad Nacional de Lomas de Zamora Instituto de Investigaciones de Tecnología y Educación (IIT&E)
Camino de Cintura y Avda. Juan XXIII –Llavallol
Teléfono: 011 – 4282-7880

institutoiite@gmail.com; mcomoglio@gmail.com; hrolon@gmail.com ;
pablomore@yahoo.com.ar; oscarlacanna@gmail.com.

Resumen

El Consejo Federal Decanos de Ingeniería (CONFEDI), en el año 2006 elabora un documento en el que se definen las competencias genéricas a desarrollar durante la carrera de ingeniería (CONFEDI, 2006)

La Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora, a través del Instituto de Investigaciones en Tecnología e Investigaciones (IIT&E) desarrolla experiencias orientadas a acompañar una metodología didáctica que organice el proceso de enseñanza y aprendizaje mediante la elaboración y resolución de problemas basados en proyectos que tengan en cuenta la adquisición de aquellas. Uno de los ámbitos de trabajo sistemático, es el Laboratorio de Simulación, al cual concurren alumnos de distintas asignaturas y en el que se les brinda la posibilidad de modelizar, analizar y visualizar mecanismos, piezas, conjuntos como así también procesos. Se trabaja con un método de enseñanza, apoyado en guías de trabajo autónomo elaborada por los docentes de las cátedras y el equipo técnico del laboratorio,

actividades éstas, que requieren tanto para su definición como resolución un trabajo proactivo y planificado (Romero López, M^a y Crisol Moya, 2012)

Los resultados parciales que se presentan surgen de indagar la percepción de los alumnos respecto de la adquisición de competencias tecnológicas y actitudinales a partir de la implementación de trabajos prácticos de simulación y su nivel de satisfacción frente a la estrategia didácticas innovadoras que incorporen tecnología de simulación al proceso de enseñanza.

Palabras clave: Enseñanza de la Ingeniería. Software de simulación, Desarrollo de competencias.

Contexto

En los últimos años el CONFEDI, viene trabajando en la identificación de competencias genéricas de egreso, vinculadas con los perfiles de ingenieros que requiere la sociedad para su desarrollo sustentable.

Asistimos a un cambio de paradigma en materia de educación superior; donde la sociedad exige no sólo formación profesional, es decir el “saber”, sino

también, la adquisición de ciertas competencias profesionales vinculadas al “saber hacer”, cambio que se vio claramente reflejado, en la Declaración de Bolonia del año 1999.

El antiguo paradigma, se sostenía en un esquema de transferencia de conocimientos, postura que fue perdiendo consenso, frente a la visión de que los egresados universitarios, son profesionales de quienes se espera adquieran durante su formación una serie de competencias que les permitan fundamentalmente hacer y ser. Por lo tanto hay consenso en cuanto que el ingeniero no sólo debe saber, sino también saber hacer. El saber hacer no surge de la mera adquisición de conocimientos sino que es el resultado de la puesta en funciones de una compleja estructura de conocimientos, habilidades, y destrezas, que requiere ser reconocida expresamente en el proceso de aprendizaje para que la propuesta pedagógica incluya las actividades que permitan su desarrollo (CONFEDI, 2006). Por su parte la Asociación Iberoamericana de Entidades de Enseñanza de la Ingeniería (ASIBEI) adopta como propias las competencias genéricas de egreso acordadas por CONFEDI, dando lugar a la Declaración de Valparaíso sobre competencias genéricas de Egreso del Ingeniero Iberoamericano (ASIBEI, 2013)

Introducción

En los últimos años, las estrategias de enseñanza para la adquisición de competencias han ido cobrando una importancia cada vez mayor, y ha venido a convertir el aprender a aprender en una de las metas fundamentales de cualquier proyecto educativo (Pozo y Monereo, 1999, citado por Torres Salas, 2010).

La búsqueda de nuevas estrategias llevó a detectar trabajos realizados que buscan enfatizar en la importancia de la simulación industrial para resolver problemas de optimización y también casos puntuales donde se enfatizaba su importancia para casos de estudio de gestión logística, gestión de almacenes y SCM, algunos de estos trabajos se enfocan en el ámbito académico particularmente para acercar herramientas de simulación a los estudiantes (Contreras Castañeda, 2014; González Maya, 2009)

Por su parte, Moran Moguel (2011) muestra los nuevos enfoques en la enseñanza de la ingeniería



El Laboratorio de Simulación¹ desarrolla actividades bajo esta perspectiva teórica, y a través del denominado aprendizaje basado en proyectos colaborativos (ABPC), se orienta a que los docentes, con asistencia técnica del laboratorio, desarrollen en sus

¹ Laboratorio de Simulación integrado al Instituto de Investigaciones en Tecnología y Educación. Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. CIC. FI UNLZ

respectivos espacios curriculares casos de estudio, los que son posteriormente como actividad práctica de la asignatura. El aprendizaje de los alumnos se realiza por descubrimiento, por lo que resulta activo, constructivo y significativo (García, A. y Rodríguez, A., 2008); el responsable del laboratorio orienta y dinamiza la realización de tareas que implica poner en juego habilidades cognitivas tales como la exploración del problema desde diversas perspectivas, la búsqueda de nueva información, y la reflexión sobre el conocimiento generado

Líneas de Investigación, Desarrollo e Innovación

Se utilizan estrategias de investigación cuantitativas a través de la administración de encuestas, al finalizar la actividad de laboratorio..

Los resultados que se presentan, corresponden a alumnos de las carreras de ingeniería mecánica e ingeniería mecánica con orientación mecatrónica, cursantes de la asignatura Seguridad e Higiene en el Trabajo, durante el primer cuatrimestre del año 2016. Para el desarrollo de la experiencia se utilizó el software: Delmia Quest y Delmia Ergonomics.

Se busca mostrar el vínculo existente entre el uso de tecnologías informáticas y de cómputo para el estudio de casos que los estudiantes pueden hallar en la vida real, para comprender cómo el modelado de situaciones manejando este tipo de herramientas puede permitir el aprendizaje en la carrera y dar lugar al conocimiento de dichas herramientas para su uso en la vida profesional.

Resultados y Objetivos

El objetivo de la experiencia, consistió en explorar la percepción de los alumnos

y su nivel de satisfacción ante la propuesta de innovación educativa. El caso de estudio que se presentó, trabajó la adquisición de dos competencias tecnológicas: a) resolución de problemas de ingeniería y b) utilizar de manera efectiva herramientas de ingeniería (en este caso software de simulación); asimismo, dos competencias sociales, políticas y actitudinales: c) desempeñarse de manera efectiva en equipos de trabajo y d) desarrollar trabajo autónomo

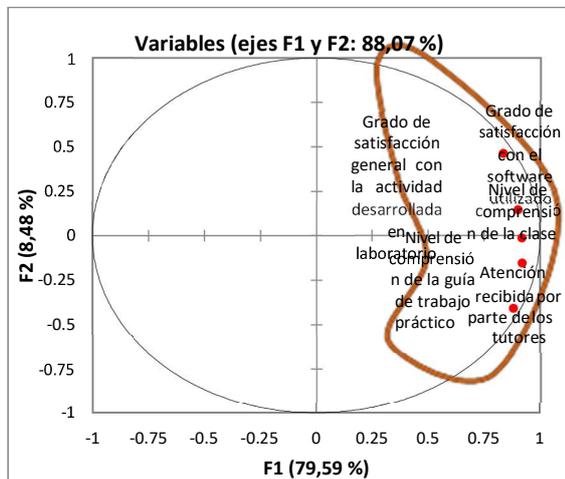
Las variables utilizadas para medir el nivel de satisfacción durante el desarrollo del trabajo práctico fueron:

- Nivel de comprensión de la clase
- Grado de satisfacción con el software utilizado
- Atención recibida por parte de los tutores
- Nivel de comprensión de la guía de trabajo práctico
- Grado de satisfacción general con la actividad desarrollada en laboratorio

A las respuestas obtenidas se le aplicó la prueba estadística, índice de correlación Pearson, a fin de medir el grado de asociación entre las respuestas. Si bien en todos los casos las respuestas mostraron un alto nivel de asociación, las variables que resultaron con mayor nivel fueron:

Variables		Índice de correlación Pearson
Nivel de comprensión de la clase	Nivel de comprensión de la guía de trabajo práctico	0,845
Atención recibida por parte de los tutores	Grado de satisfacción general con la actividad desarrollada en laboratorio	0,848

A partir de dichos resultados se obtuvo la tipología de alumnos que se presenta a continuación:



La encuesta aplicada, además preveía un espacio para comentarios. A continuación se presentan algunos de los realizados:

- Es un software muy interesante y sería bueno que se dicten clases para especializarse en el uso de todas sus herramientas.
- Excelente clase.
- Excelente software de simulación de procesos con fines de mejora continua y reducción de costos.
- Excelente la explicación.
- Mucha información junta en poco tiempo. El software es muy interesante. Un ejemplo más corto sería mejor.
- 10 puntos.

Formación de Recursos Humanos

El presente trabajo se integra al programa de investigación, liderado por el Dr. Ing. Oscar Pascal, a través del programa “Las TIC y su contribución al proceso de

enseñanza y aprendizaje en carreras de Ingeniería: Evaluación de Experiencias en la Universidad Nacional de Lomas de Zamora”(2014-2016). En la actualidad ha derivado en el siguiente programa “Enseñanza de la Ingeniería. Desarrollo y Evaluación de Modelos Estrategias y Tecnologías para Mejorar los indicadores Académicos y la Eficiencia Organizacional”, el que se integra a través de las siguientes líneas y proyectos:

Línea 1: Enseñanza y Competencias de Ingreso y Egreso: 1.-Procesos de Articulación de Competencias de Ingreso a carreras científico tecnológicas: Estrategias de Enseñanza y Evaluación, con soporte en modelos flexibles integrados a ámbitos virtuales y 2.-Innovaciones en los Procesos de Enseñanza en Carreras Científico Tecnológicas: Impacto en la percepción de la calidad, en los Indicadores Académicos y en las competencias de Egreso. Línea 2: Desarrollo e Innovación Tecnológica: aplicación en la gestión interna y externa de las Universidades: 3.- Sistemas de Apoyo a las Decisiones (DSS) para la productividad: aplicación en la Gestión Académica en Universidades y 4.- Vigilancia e Inteligencia Estratégica (VeIE) para la competitividad: aplicación en las actividades de Educación, Investigación y Transferencia en Universidad, los que cuentan con aval institucional y participación en el Programa de Incentivos de la Secretaría de Políticas Universitarias. En el ámbito del programa participan, 22 docentes investigadores, de los cuales, 6 son doctorandos², 6 becarios CIN y 3 alumnos avanzados de la carrera de Ingeniería, de los cuales 2 Becas CIN y un alumno pasante se encuentran afectados concretamente al proyecto en el

² Doctorado de Ingeniería FIUNLZ

que se inscriben las actividades que se presentan.

Referencias

Asociación Iberoamericana de Instituciones de Enseñanza de la Ingeniería (ASIBEI). Declaración de Valparaíso sobre competencias genéricas de egreso del Ingeniero Iberoamericano. Asamblea General de ASIBEI, Valparaíso, Chile, 12 de noviembre de 2013,

Consejo Federal de Decanos de Ingeniería (CONFEDI) “3er. Taller sobre Desarrollo de Competencias en la Enseñanza de la Ingeniería Argentina” – Experiencia Piloto en las Terminales de Ingeniería, Civil, Electrónica, Industrial, Mecánica y Química. Villa Carlos Paz, 14 y 15 de agosto 2006 3er. Informe Agosto 2006

Contreras Castañeda, E. & Silva Rodríguez, J. (2014). Logística Inversa Usando Simulación en la Recolección de envases de Plaguicidas: Estado del Arte. Revista Ingeniería Industrial, 33-50, 18.

Declaración de Bolonia Declaración conjunta de los Ministros Europeos de Educación Bolonia, 19 de Junio de 1999

García, A. y Rodríguez, A. (2008). Las guías de trabajo autónomo en la universidad. En A. Rodríguez, A. Fuentes, M. J. Caurcel Cara y A. Ramos García (Coords.), Didáctica en el Espacio Europeo de Educación Superior. Guías de trabajo autónomo (pp. 95- 116). Madrid: EOS Universitaria

González Maya, L. & Rodríguez Gómez, M. (2009). Juegos y Ejercicios prácticos para las materias del área de gestión de la producción y logística en ingeniería de Producción. Departamento de Ingeniería de Producción, Escuela de Ingeniería, Universidad EAFIT, Medellín.

Moran Moguel, C. (2011) Estrategias de Incorporación del Aprendizaje Basado en Proyectos en las Instituciones de Educación Superior en Ingeniería. (En

http://www.ai.org.mx/ai/images/sitio/edo_delarte/2011/3._estrategia_de_incorporacion_del_aprendizaje_basado_en_proyectos_en_las_ies_en_ingenieria.pdf)

[Consultado: 15/08/2016]

Pozo, J. I. & Monereo, C. (Eds.). (1999). El aprendizaje estratégico. Enseñar a aprender desde el currículo. Madrid, España: Santillana/Aula XXI (

Romero López, M^a y Crisol Moya E. Las guías de aprendizaje autónomo como herramienta didáctica de apoyo a la docencia. En Escuela Abierta: Revista de Investigación Educativa del CES Cardenal Spínola CEU. 2012, 15, 9-31

Desarrollo de Recursos TIC y Métodos Computacionales Aplicados

**Adair Martins, Carina Fracchia, Claudia Allan, Susana Parra,
Renzo García, Estefania Zurbrigk, Natalia Baeza, Lorena Robles, Daniel Benilla,
Ana Alonso de Armiño, Roberto Laurent**

Departamento de Computación Aplicada / Facultad de Informática
Departamento de Electrotecnia / Facultad de Ingeniería
Universidad Nacional del Comahue

Dirección: Buenos Aires 1400, 8300 – Neuquén
Teléfono: 0299 - 4490300 int. 429

e-mails: {adair.martins, carina.fracchia, claudia.allan, susana.parra}@fi.uncoma.edu. ar,
{rgarcia.inf, baeza.natalia, anacarolinaalonsode}@gmail.com,
{dymbe2, dopanga, zureste}@hotmail.com, {rlaurent@yahoo.com.ar}

Resumen

Se presentan los avances que se están realizando en las líneas de investigación “Uso y desarrollo de recursos TIC”, “Estudio y análisis de la tecnología de Realidad Aumentada (RA)” y “Estudio y desarrollo de Objetos de Aprendizaje (OAs)”. Particularmente se desarrolló la herramienta computacional ECDIA como apoyo a la enseñanza de programación en cursos iniciales. Se han realizado experiencias utilizando la metodología de RA en la enseñanza de temas de ciencias naturales y matemática en escuelas del nivel primario de la provincia de Neuquén. También se han diseñado y implementado nuevos OAs para la enseñanza de métodos computacionales en asignaturas de las carreras que se dictan en la Facultad de Informática de la Universidad Nacional del Comahue.

Palabras clave: Recursos TIC, Métodos Computacionales, Programación, Realidad Aumentada, Objetos de Aprendizaje, Educación.

Contexto

Las líneas de investigación se enmarcan en el Proyecto de Investigación “Simulación y Métodos Computacionales en Ciencias y Educación, Facultad de Informática (FAIF), Universidad Nacional del Comahue (UNCo). Está integrado por docentes y estudiantes de la carrera de Licenciatura en Ciencias de la Computación de la FAIF y docentes de la Facultad de Ingeniería, UNCo, y de la Universidad Católica de Brasilia (UCB), Brasil.

Introducción

El aprendizaje de la programación es una tarea difícil, especialmente para los primeros años de estudio de estas disciplinas. Los estudiantes principiantes necesitan una motivación adecuada para aprender a programar en forma exitosa y efectiva. Las tendencias de la enseñanza de programación, en lo que se refiere a enfoques didácticos y metodológicos, y a la utilización de herramientas tecnológicas siguen evolucionando. La base de la programación utilizando la computadora está en el manejo de los algoritmos y en su implementación en un lenguaje especificado. Para la resolución

de un problema de cualquier tipo es necesario seguir una serie de pasos que lleven a su entendimiento y a su solución, es decir, es necesaria la utilización de un algoritmo por medio del cual se resuelva dicho problema. El diseño de algoritmos en computación requiere creatividad y conocimientos sobre técnicas de programación. Las TIC (Tecnologías de la Información y la Comunicación) han experimentado un crecimiento importante generando nuevos paradigmas de interacción, nuevas formas de organización entre personas, modificando diversas prácticas en la vida cotidiana para alcanzar un objetivo común. Una propuesta que viene creciendo en los últimos tiempos es la del aprendizaje colaborativo. La misma busca promover el aprendizaje a través del esfuerzo colaborativo entre principiantes en una determinada tarea, y que interactúan con la ayuda de las tecnologías existentes para construir conocimientos, sumando esfuerzos y competencias. Los alumnos pueden trabajar en forma colaborativa o individual en la solución de un problema, esta estrategia de aprendizaje les permite fortalecer su desarrollo, fomentar el diálogo, ayudándolos en la comprensión y en la resolución de problemas por medio de la computadora [1,2].

En base a lo mencionado, se planteó el desarrollo de la herramienta denominada ECDIA (Entorno Colaborativo para el Diseño e Implementación de Algoritmos) que contempla la edición individual o colaborativa de algoritmos. Se realizó la reingeniería del software PSeInt, adicionando opciones de colaboración. PSeInt es un intérprete de pseudocódigo, su nombre viene de PSeudoInterprete, donde PSe se refiere a PSeudocódigo e Int de Intérprete. Es un software libre y gratuito que se distribuye bajo licencia GPL [3]. Fue utilizada en el desarrollo de ECDIA la arquitectura groupware mediante la cual se describe la herramienta desde tres puntos de vista: estático, funcional y dinámico. Se identificaron para el desarrollo cuatro componentes: las personas, las formas de interacción grupal, las estrategias de colaboración y el ambiente compartido, y

para el sistema distribuido se utilizó el modelo Cliente-Servidor. Para la implementación de la herramienta se eligió el framework Groupkit, producto de software de entorno genérico para aplicaciones colaborativas [4,5].

Se continúa avanzando en la línea “Estudio y análisis de la tecnología de Realidad Aumentada”. La metodología RA combina información real con virtual. El usuario idealmente percibe un escenario mixto donde en algunos casos es difícil o casi imposible distinguir entre la información real y la generada por la computadora. Está siendo muy utilizada principalmente en el ámbito educativo, permitiendo complementar los materiales didácticos con modelos virtuales que estimulen la percepción y ayuden a la comprensión de los conceptos [6-8].

Se realizaron actividades en 8 escuelas primarias donde se ha permitido mostrar el beneficio del uso de estos recursos en el aula, y que han emergido gracias al avance de las TIC. Esto dio origen al proyecto de extensión “Realidad Aumentada: recurso educativo para la enseñanza en la Educación Primaria” conformado por docentes y estudiantes de la Facultad de Informática de la UNCo y docentes del nivel primario de diferentes escuelas públicas y de gestión privada de Neuquén.

Se comenzó trabajando con tres instituciones educativas de nivel primario, y a raíz de la difusión realizada en medios impresos como el diario “Río Negro” y el diario “La mañana de Neuquén”, radiales y televisivos nos llevaron a sumar cinco instituciones más, dos de ellas de la ciudad de Cipolletti, provincia de Río Negro. También se puede mencionar la invitación recibida para participar con un stand en el evento “Neuquén Explora” desarrollado en el espacio DUAM en octubre de 2016. La participación en este evento además de permitirnos observar el interés que despierta esta tecnología en niños y adultos, nos hizo posible el contacto con otras

instituciones que manifestaron su interés en participar en este tipo de actividades con el grupo [9].

En la línea de investigación “Estudios y desarrollo de Objetos de Aprendizajes” se continúa estudiando nuevas formas de comunicación e interacción con la computadora a través del diseño de OAs. Se pueden definir los OAs como: “recursos didácticos e interactivos en formato digital, desarrollados con el propósito de ser reutilizados en diversos contextos educativos que respondan a la misma necesidad instruccional, siendo ésta su principal característica, todo esto con el objetivo de propiciar el aprendizaje” [10-13].

Particularmente se diseñaron e implementaron OAs para la interpretación geométrica del método numérico de Newton Raphson y del método de los Trapecios.

El método de Newton Raphson es utilizado para calcular en forma aproximada la raíz de una ecuación no lineal. Consiste básicamente en partir de una aproximación inicial de la raíz que se desea calcular y mediante la evaluación de la función en este punto se determina el punto por el cual se traza la tangente geométrica a la función. En su intersección con el eje x de las abscisas se encuentra una mejor aproximación a la raíz. Este proceso se repite hasta que se cumpla con una tolerancia deseada.

El método de los Trapecios es un método aproximado para el cálculo de la integral definida de una función $f(x)$ entre los extremos inferior a y el extremo superior b . En la práctica muchas veces no se puede obtener en forma exacta el valor de la integral. Este método se basa en subdividir el intervalo de integración a y b en n partes, y realizar la suma de las áreas elementales de los trapecios formados. Esta suma proporciona el valor aproximado buscado de la integral definida [14,15].

Existen diversas herramientas de autor que facilitan el proceso de construcción de objetos

de aprendizaje como eXe Learning, Reload, Hot Potatoes, GeoGebra, etc. Se utilizó el software GeoGebra para la implementación de los OAs. GeoGebra es un software libre de matemática dinámica utilizado en los distintos niveles de educación, disponible en múltiples plataformas [16]. Permite el trazado dinámico de construcciones geométricas, su representación gráfica, el tratamiento algebraico y el cálculo de funciones reales de variable real, sus derivadas, sus integrales, etc. Los objetos libres de GeoGebra que se utilizaron para la implementación de los OAs fueron la función analizada, la aproximación inicial y la tolerancia deseada para el método de Newton Raphson y para el cálculo aproximado de la integral por el método de los trapecios se utilizó el objeto de GeoGebra, casilla de entrada para el ingreso de la función y los límites del intervalo de integración. La cantidad de subdivisiones del intervalo n se representó utilizando el objeto de GeoGebra deslizador asociando al parámetro n [17].

Los proyectos creados con GeoGebra pueden ser exportados en diversos formatos como páginas web, hojas dinámicas e imágenes. Los objetos exportados se pueden publicar directamente en GeoGebraTube o en una página web y pueden luego ser incluidos en ambientes educativos virtuales como Moodle.

Líneas de Investigación, Desarrollo e Innovación

En forma sintética se presentan las siguientes líneas:

- Uso y desarrollo de recursos TIC
- Estudio y análisis de los aspectos tecnológicos contemplados en el trabajo con Realidad Aumentada
- Estudio y desarrollo de Objetos de Aprendizajes y sistemas para la creación de repositorios

Resultados y Objetivos

Se ha finalizado el desarrollo de la herramienta computacional ECDIA (Entorno Colaborativo para el Diseño e Implementación de Algoritmos) que contempla la edición individual o colaborativa de algoritmos. En una primera instancia podrá ser utilizada como herramienta de apoyo a la enseñanza de la programación e integrada a los procesos de enseñanza y aprendizaje colaborativos en los cursos iniciales de programación de la FAIF. En su desarrollo se realizó la reingeniería del software PSeInt, adicionando opciones de colaboración. Es trabajo dio origen a una tesis de grado de una integrante del grupo.

En el proyecto “Realidad Aumentada: recurso educativo para la enseñanza en la Educación Primaria” mencionado anteriormente se desarrollaron tres juegos didácticos, donde se diseñaron distintos tipos de tableros para trabajar los contenidos de la provincia de Neuquén como: flora, fauna, símbolos provinciales, entre otros y de contenidos de matemática como: operaciones básicas, medidas, resolución de problemas, etc. En los mismos se trabajó con tecnología de Realidad Aumentada, mediante el uso de códigos QR “Quick Response”. Se desarrollaron tableros utilizando las aplicaciones gráficas GIMP e Inkscape. Se utilizaron para los juegos dado, reloj de arena, y también en algunos casos reloj digital. Se elaboró un reglamento y diferentes tarjetas con preguntas de texto, múltiple choice y de opción verdadero/falso, etc.

En el año 2016 se aprobó el proyecto de extensión “Tecnología Realidad Aumentada Aplicada al Ámbito educativo”, y se continúa trabajando con las escuelas mencionadas anteriormente con las pruebas de estos recursos en experiencias concretas. El objetivo es que nos permitan analizar su potencial en el ámbito educativo, teniendo como meta lograr una mejor comprensión de los temas que se enseñan, además de motivar a los estudiantes en el aprendizaje. La realización de experiencias en escuelas,

sumado a los talleres y curso dictados permitió el trabajo con más de 300 niños y 80 docentes. El objetivo que se persigue es motivar que las instituciones educativas participantes puedan introducir estas nuevas tecnologías para el desarrollo de sus currículas. Desde la Facultad de Informática se ha observado la necesidad de vincular las instituciones de nivel primario y medio con la Universidad, compartir de cerca las deficiencias y dificultades que atraviesan los sistemas educativos en nuestra región, y promover actividades de formación docente que contribuyan al afianzamiento y fortalecimiento del desarrollo de la educación.

En el estudio de los OAs se continúa avanzando en el diseño de objetos de aprendizaje dinámicos. Se han desarrollado dos OAs interactivos utilizando la herramienta GeoGebra con el propósito de ayudar en la comprensión de los conceptos teóricos y en la interpretación geométrica de distintos métodos numéricos. Particularmente se trabajó con el método de Newton Raphson para la resolución de ecuaciones no lineales y el método de los Trapecios para la resolución de integrales definidas. El objetivo de la implementación de los OAs es permitir la visualización por parte de los estudiantes del comportamiento de los métodos gráficamente para lograr una correcta interpretación geométrica de los mismos. Fueron utilizados como complemento de los recursos teóricos y prácticos en la materia “Métodos Computacionales para el Cálculo” de la carrera de Licenciatura en Ciencias de la Computación y en la materia Cálculo Numérico del Profesorado en Matemática de la Facultad de Economía y Administración.

Se lograron OAs interactivos, que permitieron a los mismos la modificación de parámetros para observar el comportamiento de los distintos métodos. Con los OAs se pudo observar un gran incremento del interés y de la motivación de los estudiantes, lo que permitió que tengan una mejor comprensión

de los conceptos teóricos e interpretación geométrica de los distintos métodos.

Formación de Recursos Humanos

Una integrante alumna aprobó su tesis de grado en noviembre de 2016 obteniendo la máxima calificación. Otra integrante finalizó los cursos del doctorado en: “Enseñanza de las Ciencias Exactas y Naturales” en la Facultad de Ingeniería (FI), UNCo y se encuentra en etapa de tesis. Otra docente finalizó los cursos de la maestría “Enseñanza de las Ciencias Exactas y Naturales”, FI, UNCo y se encuentra en etapa de tesis. Actualmente otra integrante está realizando el último curso de la maestría en Tecnología Informática Aplicada a la Educación en la Universidad Nacional de La Plata.

Referencias

- [1] L.M. Serrano Cámara, M., Paredes Velasco Alcover C.M, Avaluation of Studens' Motivation in computer-Supported Collaborative Learning of Programming Concepts, *Computers in Human Behavior*, pp. 499-508, 2014.
- [2] Olson, G. M., & Olson, J. S. *Groupware and Computer Supported Cooperative Work*, 2002.
- [3] PseInt.sourceforge.net
- [4] Roseman, M., & Greenberg, S. *GroupKit: A groupware toolkit for building real-time conferencing applications*. In *Proceedings of the ACM conference on Computer-supported cooperative work*, pp. 43-50, ACM, 1992.
- [5] Sommerville, I, *Software Engineering*. 9^a ed., Addison-Wesley Iberoamericana, 2011.
- [6] A. Martins, C. Fracchia, C. Allan, S. Parra, R. García, E. Zurbrigk, N. Baeza, L. Robles, D. Benilla, A. Alonso de Armiño, R. Laurent. *Computación Aplicada: Búsqueda y Desarrollo de Nuevas Estrategias de Enseñanza y Recursos TIC*. Wicc 2016. XVIII Workshop de Investigadores en Ciencias de la Computación, Concordia, 2016.
- [7] J. Cubillo Arribas, S. Martín Gutiérrez, M. Castro Gil, A. Colmenar Santos. *Recursos digitales autónomos mediante realidad aumentada*. RIED. *Revista Iberoamericana de Educación a Distancia*, vol.17, N°2, pp. 241-274, 2014.
- [8] C. Fracchia, A. Alonso de Armiño, A. Martins, *Realidad aumentada aplicada a la enseñanza de Ciencias Naturales*. *Revista: TE & ET*; N°. 16. ISSN: 1850-9959. P. 7-15, 2015.
- [9] Balance 2016- Neuquén INNOVA y Neuquén EXPLORA.
<https://mail.google.com/mail/u/0/#inbox/1580bf7bf81340ee?projector=1>
- [10] Polsani, P., *Use and abuse o reusable learnig journal of digital information*, 2003.
<https://journals.tdl.org/jodi/index.php/jodi/article/view/89/88>.
- [11] L. Weitzel, P. Quaresma, R. Britos, R. Pimentel, *Recuperación de Objetos de Aprendizaje Accesibles*, X Conferência Latino- Americana de Objetos e Tecnologías de Aprendizagem (LACLO 2015), Maceió, 2015.
- [12] C. Allan, S. Parra, A. Martins, *Objetos de Aprendizaje para la Interpretación Geométrica de Métodos Numéricos: Uso de GeoGebra*. XI Congreso sobre Tecnología em Educación & Educación en Tecnología (TE & ET). Buenos Aires, 2016
- [13] C. Allan, S. Parra, A. Martins. *Una Experiencia en la Enseñanza de la Matemática con Objetos de Aprendizaje*. X Congreso sobre Tecnología en Educación & Educación en Tecnología. Corrientes, 2015.
- [14] R. L. Burden, J. D. Faires, *Análisis Numérico*, Cengage Learning, 2011.
- [15] S. C. Chapra, R. P. Canale, *Métodos Numéricos para Ingenieros con Programas de Aplicación*, Mc Graw Hill, 2005.
- [16] <http://geogebra.org>
- [17] Y. Hernández Bieliukas, A. Silva Sprock, *Una experiencia en el desarrollo de objetos de aprendizaje como apoyo a los ambientes virtuales de enseñanza y aprendizaje: integrando el conocimiento entre disciplinas*, 1° Jornadas Internacionales de Educación a Distancia, ISBN 978-980-402-063-6, Universidad de Zulia, 2011.

Desgranamiento Temprano en Materias de Primer Año en las Carreras de la Facultad de Ingeniería de la UNJu: Evaluación de la Influencia de Factores Cognitivos

Roberto D. Lamas, Adelina García Verónica
M. Torres, C. Marcelo Pérez Ibarra

Facultad de Ingeniería / Universidad Nacional de Jujuy
Ítalo Palanca 10, +54-388-4221587

rdlamas@fi.unju.edu.ar, agarcia@fi.unju.edu.ar, vtorres@fi.unju.edu.ar, cmperezi@fi.unju.edu.ar

Resumen

En las últimas dos décadas, la mejora en los índices de rendimiento académico y graduación es un tema relevante en las instituciones de educación superior debido a las acreditaciones que lleva a cabo la CONEAU. En particular, las carreras de ingeniería presentan un reducido porcentaje de graduados en el tiempo de duración de las mismas y una tasa de egresados sobre ingresantes muy baja. Esto se relaciona al importante número de estudiantes que no logran avanzar en la carrera universitaria o que deciden abandonar sus estudios. Cuatro son los factores que se asocian a la deserción estudiantil: individuales, socio-económicos, institucionales y académicos. En la Universidad Nacional de Jujuy, los primeros tres factores están cubiertos por programas institucionales que intentan mitigar los problemas derivados de ellos. Sin embargo, respecto al factor académico aún son incipientes los esfuerzos por integrar y potenciar el uso de las nuevas tecnologías en la implementación de diferentes estilos de aprendizaje. Por tanto, el presente proyecto se enfoca en el estudio, integración y evaluación de innovaciones educativas apoyadas en TIC que pueden incorporarse para mejorar el rendimiento de los alumnos y así contribuir a la

disminución de los índices de desgranamiento y deserción.

Palabras clave: TIC, innovación educativa, desgranamiento, estrategias metodológicas, *b-learning*

Contexto

La línea de investigación presentada se encuentra inserta en el proyecto *Desgranamiento temprano en materias de primer año en las carreras de la Facultad de Ingeniería de la UNJu: Evaluación de la influencia de factores cognitivos*, ejecutado a partir de 2016 por un grupo de docentes investigadores de la Facultad de Ingeniería de la Universidad Nacional de Jujuy (UNJu). El proyecto se encuentra acreditado y financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales de la UNJu.

Introducción

En América Latina, la mejora en los índices de rendimiento académico y graduación se ha incorporado como tema de alta relevancia en la agenda de políticas públicas e institucionales [Fiegehen, 2005]. En parte ello se debe a la agudización de ciertos problemas derivados de la masificación de la enseñanza y de la escasez de recursos

financieros para el sector educativo. El acceso a la educación universitaria de un público estudiantil cada vez más heterogéneo (en términos de su perfil socioeconómico, educativo y en aspiraciones académicas y laborales) demanda que las universidades exploren nuevos caminos pedagógicos e institucionales para lograr que estos jóvenes se gradúen, adquiriendo además los conocimientos y habilidades necesarias para desenvolverse con éxito en su campo académico y profesional. La formación de profesionales y científicos de calidad es actualmente un desafío de alta complejidad en el contexto de organizaciones de gran tamaño y modesto presupuesto. La masificación de la educación superior y las restricciones presupuestarias contribuyeron entonces al diseño de políticas tendientes a promover la elevación de la calidad y la eficiencia organizacional [García de Fanelli, 2014]. En Argentina, la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU) lleva a cabo el proceso de acreditación de las carreras de grado de las universidades nacionales a fin de valorar su situación actual y consecuentemente implementar políticas de mejora. Por ejemplo, en el caso de las carreras de ingeniería, se observó que sólo el 13% de los estudiantes (inscritos entre 1988 y 1998) se graduaron según la duración de la carrera prevista en el plan de estudios (6 años en promedio), mientras que la mitad de los restantes lo hizo en 7 años y el último grupo en un plazo aún mayor. Esto se reflejó en una tasa de egresados sobre ingresantes muy baja: 17% en el año 2000 y 12% en el 2001 [Gutiérrez, 2005]. Estos valores se explican al identificar un importante volumen de estudiantes que no logran avanzar en la carrera universitaria (cursando y recursando materias) o que deciden no continuar sus estudios. El

término desgranamiento hace referencia a los estudiantes que no logran regularizar algunas materias de su cohorte y al año siguiente se presentan como recursantes de estas materias. En tanto que el término deserción se aplica a aquellos estudiantes que abandonan definitivamente la vida universitaria [Dalfaro *et al.*, 2011].

De acuerdo a datos del Ministerio de Educación de la Nación, durante el primer año de estudios universitarios ocurre el mayor número de deserciones (en 2009 alcanzó 60% para todas las universidades nacionales). Según María Elena Duarte existen demasiadas diferencias entre el nivel secundario y el universitario y es por ello que los alumnos que aspiran a ingresar a éste se encuentran con una transición muy desafiante y complicada [Duarte, 2013]. Esto se debe a cuestiones tales como carencia de conocimientos sobre técnicas y estrategias de estudio, falta de organización respecto al material y el tiempo, falta de planificación del esfuerzo requerido para cada asignatura según el nivel de dificultad [Parrino, 2005], falta de adaptación a un nivel de exigencia elevado (en relación al nivel educativo previo), dudas respecto a la vocación, una interacción no personalizada con docentes y compañeros, entre otras. Sin embargo, en años superiores de la carrera universitaria la deserción está más relacionada a cuestiones institucionales: falta de integración entre materias, falta de flexibilidad entre horarios de clase y trabajo, poca relación entre teoría y práctica profesional, falta de contactos con el medio empresarial o profesional, entre otros [Pierini, 2013].

El rendimiento académico (número de materias aprobadas por los alumnos que se reinscriben en un año determinado) constituye un importante indicador que permite valorar el desempeño académico de los estudiantes y de la institución

misma. En el año 2010, el 25,7% de los jóvenes que se reinscribieron en las universidades nacionales había aprobado ninguna o una materia el año previo. En algunas universidades este guarismo era aún superior, llegando al extremo de representar la mitad de los estudiantes reinscriptos [SPU, 2013].

En vista de esta realidad surgen diferentes proyectos y/o programas enfocados en los factores o determinantes que impactan en el rendimiento académico. El ICFES (Instituto Colombiano para el Fomento de la Educación Superior) define 4 factores a considerar: individuales, socio-económicos, institucionales y académicos. Los factores individuales se refieren a características (edad, sexo, estado civil), situaciones (problemas familiares, salud) y sentimientos del estudiante (expectativas). Los factores socio-económicos comprenden el estrato social del estudiante, situación laboral, dependencia económica, carga familiar, etc. Los factores institucionales están asociados a la institución, sus normativas, ayudas y becas de estudio, entorno del estudiante y recursos de la institución. Los factores académicos corresponden a la orientación profesional, tipo de colegio secundario, métodos de estudio, carga académica, entre otros [ICFES, 2002] y [Castaño *et al.*, 2009].

A fin de mejorar el rendimiento académico es preciso fortalecer los aspectos contemplados en estos 4 factores a través de políticas, estrategias y acciones institucionales. En la Universidad Nacional de Jujuy, los factores individuales, socio-económicos e institucionales están cubiertos por programas de becas nacionales (bicentenario, TIC, progresar) y de la propia universidad, cursos de nivelación para ingresantes, un sistema de tutorías, la articulación con escuelas secundarias, una plataforma virtual de aprendizaje y un

programa de discapacidad (UNJuProDis). Sin embargo respecto al factor académico, vinculado estrechamente al proceso de enseñanza aprendizaje protagonizado por docentes y estudiantes, aún son incipientes los esfuerzos por integrar y potenciar el uso de las nuevas tecnologías en la implementación de diferentes estilos de aprendizaje. Por tanto, el presente proyecto se enfoca en el estudio, integración y evaluación de innovaciones educativas apoyadas en TIC que el docente puede incorporar para mejorar el rendimiento de los alumnos y consecuentemente contribuir a la disminución de los índices de desgranamiento y deserción.

Líneas de Investigación, Desarrollo e Innovación

En Argentina el ingreso a la universidad pública es libre, irrestricto y gratuito, siendo el único requisito haber finalizado el nivel de educación media. Este sistema explica en parte la cantidad de ingresantes y la tasa bruta de escolarización universitaria del orden del 45%. Esta alta tasa de escolarización tiene su contracara con una alta tasa de deserción especialmente en los primeros años de las carreras, siendo la retención promedio del sistema universitario del orden del 50% en primer año. La deserción en años posteriores ubica que la graduación final de las carreras de grado es del orden del 20% de los ingresantes a cada carrera.

En la última década, el estudio de los factores que inciden sobre el abandono, la graduación y el rendimiento de los estudiantes universitarios adquirió un nuevo interés, enfocándose no sólo en la institución (diseño de políticas) sino también en sistemas de enseñanza-aprendizaje más flexibles y accesibles

apoyados por las tecnologías de la información y la comunicación (TIC).

La incorporación de nuevos materiales, nuevos comportamientos y prácticas de enseñanza y nuevas creencias y concepciones son cambios que están relacionados con los procesos de innovación en cuanto a mejoras en los procesos de enseñanza-aprendizaje [Fullen & Stiegelbauer, 1991]. La organización de procesos de enseñanza- aprendizaje en entornos virtuales puede verse como un proceso de innovación pedagógica basado en la creación de las condiciones para desarrollar la capacidad de aprender y adaptarse tanto de las instituciones educativas como de los protagonistas del proceso educativo [Morin y Seurat, 1998].

Respecto a estas nuevas incorporaciones se debe considerar que no se inventan nuevas metodologías, sino que la utilización de las TIC en educación supone nuevas perspectivas respecto a una enseñanza mejor y apoyada en entornos *online*, cuyas estrategias son estrategias habituales en la enseñanza presencial, pero ahora son simplemente adaptadas y redescubiertas en su formato *online* [Masson, 1998].

Los cambios sustanciales que las TIC permiten introducir en el sistema educativo mediante la adecuada combinación de elementos tecnológicos, pedagógicos y organizativos constituyen una herramienta esencial para superar las deficiencias que adolecen los sistemas convencionales (presenciales o a distancia) [Salinas, 2004]. Así, la problemática del desgranamiento presente en los primeros años de las carreras universitarias puede abordarse desde una perspectiva de innovación educativa que proporcione los recursos para la captación, permanencia y graduación de los alumnos del nivel universitario.

Resultados y Objetivos

El proyecto, cuya ejecución comprende los periodos 2016-2017 tiene como objetivo general:

- Elaborar una propuesta de intervención educativa virtual a efectos de contribuir al mejoramiento del índice de retención en materias de primer año.

Asimismo persigue los siguientes objetivos específicos:

- Clasificar y evaluar “enfoques o estrategias de aprendizaje” implementables mediante recursos de educación virtual.

- Clasificar y evaluar herramientas tecnológicas aplicadas al proceso de enseñanza-aprendizaje.

- Diseñar un espacio virtual adecuado para el desarrollo de educación mixta (*b-learning*)

- Implementar y evaluar el espacio virtual.

Durante el primer año del proyecto se recopiló y clasificó un conjunto de herramientas software aplicables a las distintas fases del ciclo de enseñanza aprendizaje, se investigaron diferentes enfoques pedagógicos innovadores y se realizaron experiencias con alumnos para evaluar las herramientas y enfoques pedagógicos seleccionados. Para el segundo año se prevé diseñar un espacio virtual en función de los resultados obtenidos el año anterior, evaluar las posibles variantes de la propuesta de diseño pedagógico-tecnológico y finalmente comparar los resultados obtenidos a fin de formular una flexible pero robusta propuesta de educación virtual.

Formación de Recursos Humanos

El equipo de trabajo dedicado a la presente línea de investigación está integrado por 4 docentes.

Estudios de postgrado.

- Especialidad finalizada: 2
- Trabajo final de especialidad: 1
- Especialidad en curso: 1

Referencias

Castaño, E., Gallón, S., Gómez, K., & Vásquez, J. (2009). Deserción estudiantil universitaria: una aplicación de modelos de duración. *Lecturas de economía*, 60(60), 39-65.

Fiegehen, L. (2005). Repitencia y deserción universitaria en América Latina. Presentado en el Seminario de Educación Superior de América Latina y el Caribe. Capítulo, 11.

Dalfaro, N. A., Maurel, M. D. C., & Sandobal Verón, V. C. (2011). El blended learning y las tutorías: Herramientas para afrontar el desgranamiento. In *Primera Conferencia Latinoamericana sobre el Abandono en la Educación Superior*. (IClabes). Managua, Nicaragua. ISBN (pp. 978-84).

Duarte, M. E. (2013). "Tutorías para ingresantes: experiencias en la UNC". 1ª Ed. Córdoba. Universidad Nacional de Córdoba.

Fullen, M., & Stiegelbauer, S. (1991). *The new meaning of educational change*. Ontario Institute for Studies in Education, Toronto, Canada.

García de Fanelli, A. M. (2014) "Rendimiento académico y abandono universitario: Modelos, resultados y alcances de la producción académica en la Argentina". *Revista Argentina de Educación Superior* (ISSN 1852-8171) Año 6, nro 8. Pp 9-38.

Gutiérrez, R. (2005). *Panorama General de las Carreras de Ingeniería de la Argentina*. Buenos Aires: Programa de Calidad Educativa, Secretaría de Política Universitarias, Ministerio de Educación http://www.me.gov.ar/spu/guia_tematica/CALIDAD/calidad_promei.html

Instituto Colombiano para el Fomento de la Educación Superior –ICFES–, 2000. "Educación superior en la década 1990-1999: resumen estadístico" y "Estudio De La Deserción Estudiantil En La Educación Superior En Colombia. Documento sobre Estado del Arte". www.icfes.gov.co

Mason, R. (1998): *Models of On Line Courses*. *ALN Magazine* 2(2) http://www.aln.org/alnweb/magazine/vol_2_issue2/masonfinal.html

Morin, J. & Seurat, R. (1998): *Gestión de los Recursos Tecnológicos*. Cotec, Madrid

Parrino, M. del C. (2005). "Aristas de la Problemática e la Deserción Universitaria". V Coloquio Internacional sobre Gestión Universitaria de América del Sur. Mar del Plata, Argentina.

Pierini, S. (2013) *La Problemática de la deserción universitaria en los años superiores de la carrea de Ingeniería Industrial de la Universidad Nacional de Rosario* (Tesina). Profesorado Universitario para educación Secundaria y Superior. Universidad Nacional de Rosario. Rosario-Argentina.

Salinas, J. (2004). *La integración de las TIC en las instituciones de educación superior como proyectos de innovación educativa, comunicación presentada al I Congreso de Educación Mediada con Tecnologías "La Innovación Pedagógica con el uso de las TIC"*, Barranquilla, Colombia.

SPU (2013). *Anuario 2010 de Estadísticas Universitarias*. Recuperado de <http://portales.educacion.gov.ar/spu/investigacion-y-estadisticas/anuarios/>

DetECCIÓN AUTOMÁTICA DE CONFLICTOS GRUPALES EN ENTORNOS DE APRENDIZAJE COLABORATIVO

María Cecilia Colman¹, Pablo Santana-Mansilla^{1,2} y Rosanna Costaguta¹

¹Instituto de Investigación en Informática y Sistemas de Información (IISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Avda. Belgrano (S) 1912 – Santiago del Estero CP 4200
chechukolman@gmail.com; psantana@unse.edu.ar; rosanna@unse.edu.ar

Resumen

Estudios realizados han comprobado que los entornos de Aprendizaje Colaborativo Soportado por Computadora (ACSC) pueden traer aparejados beneficios educativos, sociales y motivacionales para los grupos de estudiantes. Sin embargo, la tecnología por sí misma no garantiza la creación de un ambiente donde la colaboración pueda ocurrir. Para colaborar efectivamente los estudiantes necesitan (entre otras cosas) de un e-tutor (docente) que coordine las discusiones, promueva la participación, y ayude a enfrentar los problemas que surjan durante la dinámica de trabajo grupal. Para coordinar efectivamente a los grupos de estudiantes en ACSC es preciso que los e-tutores analicen las interacciones que se produzcan entre los alumnos pero, esta tarea requiere una considerable cantidad de tiempo y esfuerzo.

En este artículo se presenta una línea de investigación que propone incorporar técnicas de minería de textos y agentes de software a una herramienta de ACSC existente, con el propósito de monitorear las interacciones, identificar conflictos grupales, e informar al docente sobre la presencia de dichos conflictos para propiciar su oportuna intervención.

Palabras clave: *Aprendizaje colaborativo soportado por computadora, conflictos grupales, e-tutor, minería de texto, agente de software.*

Contexto

Este trabajo se desarrolla en el marco del proyecto “Mejorando escenarios de aprendizaje colaborativo soportado por computadora”, acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT-UNSE) para el período 2017-2020. El proyecto es una continuación de la línea de investigación Sistemas Adaptativos Inteligentes, llevada a cabo entre 2005-2009 por el proyecto 23/C062, “Herramientas conceptuales, metodológicas y técnicas de la Informática Teórica y Aplicada”, continuada entre 2009-2010 por el proyecto P09/C002

,"Personalización en Sistemas de Enseñanza Virtual", entre 2010-2011, por el proyecto 23/C089 "Fundamentos Conceptuales y Soportes Tecnológicos de la Informática Educativa", y entre 2012-2016 por el proyecto 23/C097, “Sistemas de información web basados en agentes para promover el Aprendizaje Colaborativo Soportado por computadoras (ACSC)”. Todos los proyectos citados fueron acreditados y financiados por SECyT-UNSE.

1. Introducción

Los ambientes de ACSC suelen traer aparejados beneficios educativos, sociales y motivacionales para los grupos de estudiantes. En el aspecto educativo, los estudiantes que trabajan en entornos de ACSC logran altos niveles de aprendizaje, toman decisiones de

alta calidad, y entregan informes más completos en comparación con estudiantes que trabajan en aulas presenciales [14,15]. En lo que concierne a lo social, los estudiantes que utilizan ambientes de ACSC no sólo participan de manera equitativa sino que también se involucran en discusiones más complejas, amplias, y cognitivamente desafiantes [14]. En cuanto a la motivación, los estudiantes que colaboran online suelen manifestar altos niveles de satisfacción [14] [15].

El éxito de las experiencias de ACSC depende, entre otros factores, de las habilidades con las que cuentan los e-tutores, puesto que son los encargados de coordinar y mediar las discusiones grupales [4] [5] [6]. Sin embargo, la coordinación de interacciones grupales no es una tarea sencilla ya que el análisis de interacciones en búsqueda de conflictos grupales demanda mucho tiempo y esfuerzo por parte de los e-tutores. Se sabe que esta situación se agrava cuando se necesita coordinar varios grupos y cuando se usan herramientas de comunicación sincrónica [7] [8]. En el primer caso la complejidad radica en la monitorización simultánea de las diferentes dinámicas de colaboración de los grupos, y en el segundo caso, se debe a que en las discusiones sincrónicas los patrones de comunicación son más dinámicos, el marco temporal es significativamente corto, y la coordinación debe realizarse en tiempo real [9].

Por otro lado, algunas investigaciones demuestran que el análisis automático en base a interfaces estructuradas o semiestructuradas (oraciones de apertura, interfaces basadas en menú, interfaces basadas en diagramas, etc.) afecta negativamente el proceso de colaboración grupal [10] [11] [12].

Dado el esfuerzo requerido para realizar el análisis manual y las deficiencias asociadas con el análisis automático de interacciones estructuradas o semiestructuradas; esta línea de investigación busca determinar si la minería de textos combinada con agentes de software permite tanto detectar conflictos grupales en entornos de ACSC basados en interfaces de texto libre, como alertar a los e-tutores sobre la

existencia de esos conflictos para propiciar su oportuna intervención.

2. Líneas de Investigación, Desarrollo e innovación

Para asistir efectivamente a los grupos de estudiantes es preciso que los docentes puedan diagnosticar correctamente las actividades de los estudiantes o grupos. El diagnóstico precede a la intervención del docente y le permite decidir cuál sería la acción apropiado en una situación dada [19].

En principio los sistemas de ACSC facilitarían el diagnóstico de la colaboración ya que mantienen un registro completo de las actividades e interacciones de los docentes y estudiantes. Sin embargo, el análisis manual del registro de actividades e interacciones en un entorno colaborativo online no es tarea sencilla. La complejidad del análisis manual se agrava a medida que crece la cantidad de estudiantes y mensajes intercambiados porque es sumamente difícil seguir el hilo de las discusiones, las ideas, y las contribuciones de cada estudiante [4] [7] [10] [11] [20].

La inspección manual de las actividades colaborativas online es difícil no solamente por el tiempo y esfuerzo que demandan sino también, debido a que con frecuencia el registro de actividades y eventos que mantienen los entornos de ACSC suele ser de un nivel de abstracción muy bajo como para permitirle a los e-tutores comprender lo que verdaderamente sucede durante las actividades de aprendizaje [10] [21].

Las interfaces semiestructuradas son bastante populares en ambientes de ACSC debido a su facilidad de uso y a que reducen la cantidad de procesamiento de lenguaje natural necesario para comprender la colaboración [22]. Pero, existen algunas investigaciones que demuestran que las interfaces semiestructuradas socavan el proceso de colaboración porque restringen los tipos de actos comunicativos, provocan que la comunicación sea lenta [10] [11] y crean stress relacional [12].

Debido al esfuerzo requerido para realizar el análisis manual y las deficiencias asociadas

con el análisis automático de interacciones estructuradas o semiestructuradas, sería deseable que las herramientas de ACSC sean capaces de detectar automáticamente los conflictos grupales e informar a los docentes para su oportuna intervención. Tales herramientas tendrían que ser capaces de hacer su tarea sin afectar negativamente la dinámica de trabajo de los estudiantes o de los e-tutores, ni representar un alto costo o sobrecarga de trabajo [13]. La combinación de técnicas de minería de textos con agentes de software podría permitir que las herramientas de ACSC cumplan con las condiciones antes mencionadas. Aquí nace la pregunta de investigación que se busca responder con este trabajo: ¿Cómo se puede combinar minería de texto con la tecnología de agentes de software para alertar a los e-tutores de la presencia de conflictos grupales en herramientas de ACSC basados en interfaces de texto libre?

Para responder esta pregunta se comenzará utilizando algoritmos de clasificación de minería de textos. Los clasificadores automáticos de conflictos grupales se construirán inductivamente en base a las características de logs de interacciones grupales clasificadas a mano por personas expertas. Para obtener el conocimiento experto necesario para construir los clasificadores se recopilarán logs de interacciones grupales en entornos de ACSC donde e-tutores guiaron a grupos de estudiantes universitarios durante el desarrollo de sus tareas.

A continuación, sobre los mensajes publicados por los estudiantes se aplicará la técnica de análisis de contenido [1]. Mediante esta técnica los expertos, un psicopedagogo y un docente con experiencia en ACSC, indicarán en cada unidad de análisis (mensaje, párrafo, u oración) los tipos de conductas IPA (Interaction Process Analysis) que reconozcan. IPA es un método propuesto por Bales [2] que codifica las conductas grupales en doce categorías específicas: muestra solidaridad, muestra relajamiento, muestra acuerdo, da sugerencia, da opinión, da información, pide información, pide opinión, pide sugerencia, muestra desacuerdo, muestra tensión y muestra antagonismo. A partir de esta categorización

de las conductas grupales, el método IPA permite reconocer seis tipos de problemas o conflictos por los que puede atravesar evolutivamente un grupo: comunicación, evaluación, control, decisión, reducción de tensiones y reintegración. Estos seis tipos de problemas se manifiestan mediante cantidades inapropiadas de interacciones de determinados pares de conductas. Se considera que una conducta se manifiesta inapropiadamente cuando la cantidad de interacciones registradas para la categoría cae por debajo del límite inferior o por encima del límite superior definido por Bales. Estos límites se expresan como porcentajes calculables sobre el total de interacciones registradas.

Una vez aplicado el método de análisis de contenido sobre el conjunto de interacciones grupales, se procederá a construir los clasificadores automáticos de conductas IPA siguiendo los lineamientos de la metodología CRISP-DM [3]. Este paso requiere que el conjunto de interacciones grupales sea dividido en un conjunto de entrenamiento y otro de prueba [16]. Cada clasificador puede considerarse como una función que, analizando las características de las interacciones grupales clasificadas a mano por personas expertas (conjunto de entrenamiento), determina de manera inductiva las condiciones que deben cumplir interacciones nuevas (no utilizadas durante la construcción de un clasificador) para ser asignadas a un determinado tipo de conducta IPA. El clasificador automático puede adoptar la estructura de un árbol de decisión, reglas de decisión, red neuronal, etc., dependiendo del algoritmo utilizado para su construcción [16] [17]. Por su parte, la coincidencia entre las conductas IPA reconocidas por los clasificadores en el conjunto de prueba y el reconocimiento realizado por los expertos mediante el análisis de contenido, permitirá determinar el grado de efectividad de la minería de textos en la detección automática de conductas IPA.

Considerando que los productos de software de minería de texto no aplican sus algoritmos de descubrimiento de conocimiento a colecciones de documentos no estructurados,

será necesario recurrir a lo que se conoce como operaciones de pre procesamiento [17] [18]. Las operaciones de pre procesamiento (Tokenización, Lemmatization, Parsing, Remoción de stopwords, etc.) son responsables de transformar datos no estructurados almacenados en colecciones de documentos en un formato intermedio estructurado más explícitamente [17].

Los clasificadores automáticos construidos serán utilizados por un agente de software, el Agente de Grupo, quien se encargará mediante esos clasificadores de diagnosticar la interacción entre estudiantes y reconocer conflictos grupales. Detectado un conflicto, el Agente de Grupo le notificará tal situación a otro agente de software, el Agente Docente. Cuando esto ocurra, el Agente Docente alertará al e-tutor y le sugerirá acciones a realizar para resolver el conflicto detectado.

Finalmente, los clasificadores automáticos y los agentes de software serán implementados en una herramienta asincrónica de e-learning colaborativa para realizar experiencias con e-tutores y alumnos reales que permitan validar su funcionamiento.

Es importante mencionar que la presente propuesta es el primer intento de automatizar la detección de conflictos IPA cuando las interacciones se producen mediante herramientas de comunicación basadas en texto libre.

3. Resultados esperados y Objetivos

Los objetivos generales fijados en esta línea de investigación son:

- Contribuir al éxito de las sesiones de ACSC.
- Contribuir a la ampliación del área de aplicación de la minería de textos.
- Favorecer la labor de los e-tutores de ACSC en la detección de conflictos grupales.
- Propiciar la oportuna intervención de los e-tutores en sesiones de ACSC.

Como objetivos específicos se fijaron los siguientes:

- Desarrollar un sistema multiagente capaz tanto de detectar la ocurrencia de conflictos en la dinámica de trabajo de un grupo de estudiantes en un entorno de ACSC, como de alertar al e-tutor de tales ocurrencias.
- Determinar el grado de efectividad de diversos algoritmos de clasificación de minería de textos en la identificación automática de conductas IPA.
- Identificar las operaciones de pre procesamiento de minería de textos que permiten lograr el mayor grado de efectividad en la identificación de conductas IPA.

Como resultado final de esta investigación se espera obtener agentes de software que puedan incorporarse en una herramienta de colaboración asincrónica existente para apoyar la labor de los e-tutores de ACSC en la detección de los conflictos grupales IPA.

4. Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto por un docente investigador formado, un investigador en formación y un estudiante. El investigador en formación es un becario CONICET que está desarrollando su tesis para obtener el título de Doctor en Ciencias de la Computación. El estudiante está desarrollando su trabajo final para obtener el título de Licenciado en Sistemas de Información. Ambos realizan sus actividades en el marco del proyecto de investigación "Mejorando escenarios de aprendizaje colaborativo soportado por computadora".

5. Referencias

- [1] Krippendorff Klaus (2004): *Content analysis: an introduction to its methodology*, 2da edición. USA, SAGE Publications.
- [2] Bales, R.F., (1950): A Set of Categories for the Analysis of Small Group Interaction. *American Sociological Review*, 15 (2), 257-263.
- [3] Chapman P., Clinton J., Kerber R., Khabaza T., Reinartz T., Shearer C., y Wirth R. (2001): *CRISP-*

DM 1.0 Step-by-step data mining guide. USA, SPSS Inc.

- [4] Borges M. A. F. y Baranauskas M. C. (2003): CollabSS: a Tool to Help the Facilitator in Promoting Collaboration among Learners. *Educational Technology & Society*, 6(1), 64-69.
- [5] Kukulska-Hulme A. (2004): Do Online Collaborative Groups Need Leaders? En: *Online Collaborative Learning: Theory and Practice*. Information Science Publishing, pp. 262-280.
- [6] Santana-Mansilla P., Costaguta R., y Missio D. (2013): Clasificación de habilidades de e-tutores en aprendizaje colaborativo soportado por computadora. *Revista Electrónica Iberoamericana de Educación en Ciencias y Tecnología*, 4 (2), 1-36.
- [7] Dönmez, P. et al. (2005): Supporting CSCL with automatic corpus analysis technology. *Proceedings of th2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*. Taipei, Taiwan, International Society of the Learning Sciences, pp. 125-134.
- [8] Trausan-Matu, S., Dascalu, M. & Rebedea, T., (2014): PolyCAFe-automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning*, 9 (2), 127-156.
- [9] Shwarts-Asher, D., Ahituv, N., & Etzion, D. (2009): Computer-Mediated Group Interaction Processes. *2009 International Conference on Complex, Intelligent and Software Intensive Systems*. Japan, IEEE, pp. 255–262.
- [10] Chen, W. & Wasson, B., (2005): Intelligent Agents Supporting Distributed Collaborative Learning. En: *Designing Distributed Learning Environments with Intelligent Software Agents*. IGI Global, pp.33–66.
- [11] Constantino-González, M. de L.A., Suthers, D.D. & Escamilla De Los Santos, J.D., (2003): Coaching Web-based Collaborative Learning based on Problem Solution Differences and Participation. *International Journal of Artificial Intelligence in Education*, 13(2-4), 263–299.
- [12] Olivares, O.J., (2007): Collaborative vs. Cooperative Learning: The Instructor's Role in Computer Supported Collaborative Learning. In K. L. Orvis & A. L. R. Lassiter, eds. *Computer-Supported Collaborative Learning: Best Practices and Principles for Instructors*. Information Science Publishing, pp. 20–39.
- [13] Santana Mansilla, P., Costaguta, R., & Missio, D. (2013): Aplicación de algoritmos de clasificación de minería de textos para el reconocimiento de habilidades de E-tutores colaborativos. *Inteligencia Artificial*, 17(53), 57–67.
- [14] Janssen, J. et al., (2007): Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers and Education*, 49(4), 1037–1065.
- [15] Phielix, C., Prins, F.J. & Kirschner, P. a., (2010): Group awareness of social and cognitive behavior in a CSCL environment. *Proceedings of the 9th International Conference of the Learning Sciences - Volume I*. Chicago, University of Illinois at Chicago, pp.230–237.
- [16] Sebastiani F. (2002): Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, 34(1), 1–47.
- [17] Feldman R. y Sanger J. (2007): *The text mining handbook. Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.
- [18] Weiss S. M., Indurkha N., Zhang T., y Damerau F. J. (2005): *Text Mining Predictive Methods for Analyzing Unstructured Information*. USA, Springer.
- [19] Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014): Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers and Education*, 79, 28–39.
- [20] Rosé, C. et al., (2008): Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237–271.
- [21] Harrer, A., Hever, R. & Ziebarth, S. (2007): Empowering researchers to detect interaction patterns in e-collaboration. *Frontiers in Artificial Intelligence and Applications*, 158, 503–510.
- [22] Soller A., Martínez M. A., Jermann P., y Muehlenbrock M. (2005): From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. *International Journal of Artificial Intelligence in Education*, 15 (4).

Detección de Perfiles de Rendimiento Académico en la Universidad Nacional del Este de Paraguay

David L. la Red Martínez¹, Gabriela Bobadilla²

¹Facultad Regional Resistencia / Universidad Tecnológica Nacional
French 414, (3500) Resistencia, Argentina, +54-379-4638194
laredmartinez@gigared.com

²Facultad Politécnica / Universidad Nacional del Este
Campus Km 8 Acaray, Calle Universidad Nacional del Este y Rca. del Paraguay, Ciudad del Este, Paraguay, 59561-577-261

Resumen

El problema de la deserción temprana en las Universidades se ve agravado cada día más por diferentes causas. La situación académica de pobre rendimiento se ve afectada por múltiples factores que conforman un escenario complejo de análisis. En general no existe un solo aspecto que determine el fracaso de los alumnos, por el contrario, es la interacción de varias circunstancias la que provoca el abandono de las carreras universitarias. Claramente, la situación socio-económica es clave a la hora del análisis y no puede dejarse de lado cuando se intenta estudiar el fenómeno de la deserción. Es importante, por tanto, estudiar y determinar cuáles son las variables que inciden en el rendimiento académico a fin de poder establecer estrategias de acción pedagógicas que permitan mejorar dicho rendimiento. Claramente, la interacción de dichas variables es un problema complejo de abordar con técnicas de análisis tradicionales. Por esto se propone realizar un análisis de los datos utilizando herramientas computacionales y algoritmos adecuados para tal fin.

Este proyecto tiene como propósito intentar establecer los perfiles de rendimiento académico de los alumnos de la Universidad Nacional del Este (UNE) de Ciudad del Este – Paraguay a fin de poder describir la situación de deserción e intentar explicar las causas que motivan el abandono de las carreras que se dictan allí. Para ello, se utilizarán técnicas de minería sobre los datos académicos y socio-económicos de los alumnos, aplicando algoritmos de búsqueda de conocimiento en grandes volúmenes de información. Entre esos algoritmos de Minería de Datos (del inglés Data Mining o DM) se destaca la utilización de redes neuronales, algoritmos genéticos, predicción dinámica, agentes inteligentes, clustering, reglas de

asociación, árboles de decisión, análisis de correlación, análisis semántico, análisis de regresión, entre otros.

Si bien el proyecto se llevará a cabo analizando carreras de la UNE, ha sido desarrollado en el ámbito de la Red de Cooperación Interuniversitaria en TICs del Mercosur (ReCITic) integrada por la Universidad Tecnológica Nacional, Facultad Regional Resistencia (UTN-FRRe, Resistencia, Chaco, Argentina), la Universidad Gastón Dachary (UGD, Posadas, Misiones, Argentina), la Universidade Estadual do Oeste do Paraná (UNIOESTE, Foz do Iguaçu, Brasil) y la Universidad Nacional del Este (UNE, Ciudad del Este, Paraguay). En el proyecto trabajarán docentes investigadores de las mencionadas universidades, lo que lo hace de carácter interdisciplinario e internacional.

Palabras clave: rendimiento académico; almacenes de datos; minería de datos; modelos predictivos.

Contexto

La Universidad Nacional del Este fue creada en el año 1993 por Ley de la Nación Paraguaya N° 250/93. Actualmente cuenta con las Facultades de: Ciencias Agrarias; Filosofía; Politécnica; Ciencias Económicas; Derecho y Ciencias Sociales y la de Ciencias de la Salud. Cuenta con tres Escuelas Superiores y con filiales en varias localidades.

Cuenta con más de 10.000 alumnos distribuidos en las distintas unidades académicas y escuelas. La UNE enfrenta el reto expuesto por (Montero Rojas et al., 2007) de la pro-

blemática de las Instituciones públicas de educación superior, que es el de mejorar su calidad académica con recursos cada vez más escasos, y a la vez, hacer frente a las demandas de los nuevos contextos sociales y económicos de una sociedad globalizada.

En el sondeo exploratorio realizado en la FPUNE (Bobadilla y La Red Martínez, 2015), se muestra la acumulación de alumnos en los primeros semestres de las carreras tecnológicas y el bajo porcentaje de egresados en función al año de cohorte, que llevan a la situación de agotar recursos que debieran estar disponibles para atender la demanda natural generada por los alumnos provenientes de la educación media. A fin de atender el problema de la calidad académica y la escasez de recursos, este proyecto propone identificar las variables que inciden en el rendimiento académico de los alumnos de la UNE. Esto permitirá determinar perfiles de alumnos exitosos (los que promocionan en su cohorte), como así también los perfiles de alumnos que no lo logran (los que quedan desfasados a su cohorte o desertan). Una vez determinados los perfiles de alumnos, se podrán plantear acciones tendientes a evitar potenciales fracasos académicos. Se considera oportuno utilizar las técnicas de Almacenes de Datos (Data Warehouse: DW) y Minería de Datos (Data Mining: DM), para la obtención de los perfiles.

El proyecto “Estudio del rendimiento académico y detección temprana de perfiles de alumnos en la Universidad Nacional del Este de Paraguay, aplicando técnicas de minería de datos” fue presentado en la convocatoria 2015 del CONACYT (Consejo Nacional de Ciencia y Técnica de Paraguay) y fue aprobado por Resolución N° 310 del 22/07/16, comenzando su ejecución en el año 2017.

Introducción

La Universidad enfrenta actualmente el desafío de mejorar su calidad académica enfocándose no sólo en el sistema de enseñanza-aprendizaje, sino contemplando otras variables, como la sistematización de procesos de evaluación permanentes (Briand et. al., 1999). Entre estas variables, se destaca el

estudio del perfil de rendimiento académico de los estudiantes.

Se define al rendimiento académico como la productividad del sujeto, matizado por sus actividades, rasgos y la percepción más o menos correcta de los cometidos asignados (Maletic et. al., 2002).

Generalmente al evaluar el rendimiento académico, se analizan en mayor o menor medida los elementos que influyen en el desempeño como ser, entre otros, factores socioeconómicos, la amplitud de programas de estudio, las metodologías de enseñanza, conocimientos previos del alumno (Marcus, 2003).

Se ha demostrado con varios estudios que el factor más relacionado con la calidad educativa es el propio alumno como co-productor, medido a través del nivel socioeconómico del hogar de donde proviene (Maradona & Calderón, 2007) y se ha evidenciado que la productividad del estudiante es mayor para las mujeres, para los estudiantes de menor edad y para quienes provienen de hogares con padres más educados (Porto & Di Gresia, 2003).

También se ha mostrado el contraste que hay entre las personas que trabajan y estudian y las que solamente estudian, encontrándose que no existen diferencias en el rendimiento académico de los dos conjuntos (Reyes, 2004).

El problema de encontrar buenos predictores del rendimiento futuro de manera que se reduzca el fracaso académico en los programas de postgrado ha recibido una especial atención en EE. UU. (Wilson & Hardgrave, 1995), habiéndose encontrado que las técnicas de clasificación como el análisis discriminante o la regresión logística son más adecuadas que la regresión lineal múltiple a la hora de predecir el éxito/fracaso académico.

La diversidad de estudios sobre el rendimiento académico muestra que no existe una manera única para evaluarlo. Por ello, la determinación de grupos o clases de alumnos es un elemento a tener en cuenta para establecer las causas de los problemas relacionados al desempeño de los mismos. Más aún, los problemas pueden variar dependiendo el contexto regional y la realidad social donde está

inserto el alumno. Es decir, no existen herramientas que se puedan aplicar a todos los ámbitos y los resultados tampoco pueden ser extensibles para explicar todas las situaciones posibles. Esto denota claramente la necesidad de determinar perfiles en las instituciones educativas específicas adaptando las herramientas a cada situación particular.

Surge, entonces, la necesidad de implementar un mecanismo que permita determinar las características propias del estudiante, analizando la existencia de relaciones y patrones de comportamiento estudiantiles que posibilite la definición clara de perfiles de alumnos. Para ello una alternativa es utilizar técnicas de minería de datos para el modelado descriptivo (La Red Martínez et. al., 2014, 2015).

A su vez, el modelado predictivo puede usarse para analizar una base de datos y determinar ciertas características esenciales acerca del conjunto de datos que permitan predecir el comportamiento de alguna variable (Connolly & Begg, 2005).

El presente proyecto intenta establecer las causas del bajo rendimiento académico de los alumnos de la Universidad Nacional del Este del Paraguay (UNE), utilizando técnicas de Minería de Datos. Para ello se propone detectar patrones de deserción estudiantil a partir de los datos socio económicos, académicos, actitudinales e institucionales de los estudiantes universitarios, haciendo uso de técnicas de DM. Inicialmente se utilizarán los datos de los estudiantes que ingresaron en los años anteriores, con el fin de hacerles un seguimiento hasta la actualidad, determinando las características de los casos particulares. Con estos datos se construirá un repositorio de datos que se pre-procesará y se transformará con el fin de obtener un conjunto de datos limpios y listos para aplicarle las técnicas de minería de datos.

Para la etapa de análisis se utilizarán algoritmos de DM a fin de poder construir modelos descriptivos que expliquen las circunstancias que llevan al bajo rendimiento académico que frecuentemente termina en deserción de los alumnos o a un rendimiento académico bueno o muy bueno que generalmente los lleva a concluir sus estudios. Los modelos descriptivos se utilizarán para explicar la

situación de rendimiento académico de los alumnos.

Los resultados serán analizados, evaluados e interpretados para determinar la validez del conocimiento obtenido. De esta manera se requerirá una metodología de trabajo sólida con un alto componente de interdisciplinariedad que brinde resultados consistentes y útiles para la toma de decisiones directivas.

Se estima que este proyecto tendrá un alto impacto en el ámbito académico debido a que se pretende mejorar la situación de bajo rendimiento académico y el desgranamiento en las carreras de la UNE. Además, se considera que, en el ámbito científico-tecnológico, también se obtendrán resultados significativos ya que se pretende generar un modelo sólido de análisis de datos que pueda ser extensible a otros campos de aplicación y contribuir a la formación de recursos humanos en el área de investigación del proyecto.

Líneas de Investigación, Desarrollo e Innovación

En el marco del proyecto “Determinación de perfiles de estudiantes y de rendimiento académico mediante la utilización de minería de datos”, desarrollado en la UTN-FRRe entre los años 2013 y 2015, se trabajó en la identificación de las variables que explican el desigual rendimiento académico por parte de los estudiantes de la citada asignatura, lográndose modelos descriptivos del rendimiento académico.

En el contexto del proyecto “Diseño de un modelo predictivo de rendimiento académico mediante la utilización de minería de datos”, iniciado en el año 2016 en la UTN-FRRe, se buscará desarrollar modelos predictivos de rendimiento académico.

En la tesis de maestría “Estudio del Rendimiento Académico de Estudiantes de Análisis de Sistemas, Ingeniería Eléctrica e Ingeniería de Sistemas, de la Facultad Politécnica de la Universidad Nacional del Este, Paraguay”, realizada por la Ing. Bobadilla con la dirección del Dr. La Red Martínez, presentada y aprobada en el año 2015, se trabajó en la misma línea de investigación desarrollada en los proyectos mencionados precedentemente.

Esta misma línea de investigación se tiene previsto desarrollar y profundizar en el proyecto aprobado por la CONACYT para la UNE.

La minería de datos ha sido utilizada para la búsqueda de perfiles de alumnos con el propósito de identificar potenciales éxitos o fracasos académicos, teniendo en cuenta su rendimiento académico, situación demográfica y socio económico (La Red Martínez et al., 2012), (La Red Martínez et al., 2014).

En una universidad del sudeste de los Estados Unidos, se pudo determinar qué alumnos tienen tendencia a tomar cursos online o no, mediante el uso de DM teniendo en cuenta el perfil de los estudiantes por medio de las inscripciones a cursos, rendimiento escolar y situación demográfica (Chong Ho Yu et al., 2012).

En (Kabakcheva et al., 2011) se explica cómo utilizaron DM para predecir el rendimiento de los alumnos de la Universidad de Bulgaria, basados en datos personales y características pre-universitarias.

En (Brijesh Kumar Baradwaj & Saurabh Pal, 2012) utilizan DM para evaluar el rendimiento de los estudiantes universitarios utilizando árboles de decisión para la clasificación de datos, pudiendo extraer de los mismos el rendimiento del examen final del semestre, lo cual ayudó a reconocer de manera temprana estudiantes que necesitarán apoyo pedagógico.

En (Bobadilla & La Red Martínez, 2015) se utilizó DM para la identificación de variables características de los perfiles de alumnos con distinto rendimiento académico, habiéndose encontrado que el grado educacional de los padres, la actitud general hacia el estudio, el género, la utilización de las Tecnologías de la Información y Comunicación y los promedios generales de sus calificaciones del segundo y primer semestre inciden en la situación final de las calificaciones (predominantemente), en el rendimiento académico de los alumnos de las carreras tecnológicas de la FPUNE.

Resultados y Objetivos

Resultados

El presente proyecto intenta identificar las variables que influyen en el rendimiento académico, de los alumnos de las carreras de la Universidad Nacional del Este del Paraguay (UNE), utilizando técnicas de DM. Para ello se propone detectar patrones de rendimiento estudiantil a partir de los datos socio económicos, académicos, actitudinales e institucionales de los estudiantes universitarios. Para la identificación de los perfiles se utilizarán algoritmos de DM a fin de poder construir modelos descriptivos que expliquen las circunstancias del rendimiento académico que frecuentemente termina en deserción de los alumnos o un rendimiento académico bueno o muy bueno que generalmente los lleva a concluir sus estudios. Los resultados serán analizados, evaluados e interpretados para determinar la validez del conocimiento obtenido. Se estima que este proyecto tendrá un alto impacto en el ámbito de la calidad académica debido a que se pretende mejorar la situación de bajo rendimiento académico y el desgranamiento en las carreras de la UNE. Además, se busca contribuir a la formación de recursos humanos en la línea de investigación en Tecnologías de Datos y Gestión del Conocimiento de la Facultad Politécnica de la Universidad Nacional del Este.

Objetivo general del proyecto

Obtener el conocimiento necesario para desarrollar un modelo predictivo para determinar rendimientos académicos potencialmente problemáticos, a partir de la implementación de técnicas de DM.

Objetivos específicos del proyecto

- Utilizar técnicas de DM para obtener conocimiento referido a las variables que predominantemente influyen en el rendimiento académico.
- Analizar y observar el comportamiento del modelo bajo condiciones típicas de rendimiento académico, obteniendo el conocimiento necesario para depurar el modelo predictivo.
- Validar el modelo mediante el seguimiento del rendimiento académico de estudiantes-tipo seleccionados, obteniendo un modelo predictivo confiable.
- Utilizar el modelo predictivo de rendimiento académico desarrollado para predecir

la probabilidad de que un estudiante abandone el cursado de una asignatura, dadas sus características socioeconómicas y académicas.

Formación de Recursos Humanos

El equipo de trabajo está integrado por un Doctor, una Magister y varios maestrandos. Actualmente se está trabajando en la definición de planes de tesis de maestría con temáticas afines a la del proyecto.

Referencias

- Bobadilla, G., La Red Martínez, D. L. (2015) Estudio del rendimiento académico y determinación temprana de perfiles de alumnos en la Facultad Politécnica de la Universidad Nacional del Este de Paraguay; N° 10; FPUNE Scientific; pp. 43-48; ISSN N° 2222-2286; Ciudad del Este, Paraguay.
- Briand, L. C., Daly, J., Wüst, J. (1999) A unified framework for coupling measurement in object oriented systems, IEEE Transactions on Software Engineering, 25, 1.
- Brijesh Kumar Baradwaj, Saurabh Pal (2011) Mining Educational Data to Analyze Student's Performance. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6.
- Chong Ho Yu, Digangi, S., Kay Jannasch-Pennell, A., Kaprolet, C. (2012) Profiling Students Who Take Online Courses Using Data Mining Methods. Online Journal of Distance Learning Administration, Volume XI, Number II, Summer 2008, University of West Georgia, Distance Education Center, U.S.A.
- Connolly, T., Begg, C. (2005) Sistemas de bases de datos. Un enfoque práctico para diseño, implementación y gestión. 4ta Edición. Ed. Pearson Addison Wesley.
- Kabakcheva, D., Stefanova, K., Kisimov, V. (2011) Analyzing University Data for Determining Student Profiles and Predicting Performance. Conference: Proceedings of the 4th International Conference on Educational Data Mining, Eindhoven, The Netherlands, July 6-8.
- La Red Martínez, D. L.; Acosta J. C., Uribe V. E., Rambo A.; (2012) Academic Performance: An Approach From Data Mining; V. 10 N° 1; Journal of Systemics, Cybernetics and Informatics; pp. 66-72; ISSN N° 1690-4524; U.S.A.
- La Red Martínez, D. L.; Karanik, M.; Giovannini, M. Y Pinto, N. (2015) Academic Performance Profiles: A Descriptive Model Based on Data Mining; Volume 11 – N° 9; European Scientific Journal (ESJ); pp. 17-38; ISSN N° 1857-7881; University Ss "Cyril and Methodius" Skopje, Macedonia.
- La Red Martínez, D. L.; Podestá, C. E. (2014) Contributions from Data Mining to Study Academic Performance of Students of a Tertiary Institute; Volume 02 – N° 9; American Journal of Educational Research; pp. 713-726; ISSN N° 2327-6126; U.S.A.
- Maletic, J. I., Collard, M. L., y Marcus, A. (2002) Source Code Files as Structured Documents, in Proceedings 10th IEEE International Workshop on Program Comprehension (IWPC'02), Paris, France, pp. 289-292.
- Maradona, G. y Calderón, M. I. (2007) Una aplicación del enfoque de la función de producción en educación. Revista de Economía y Estadística, Universidad Nacional de Córdoba, XLII. Argentina.
- Marcus, A. (2003) Semantic Driven Program Analysis. Kent State University, Kent, OH, USA, Doctoral Thesis.
- Porto, A. y Di Gresia, L. (2003) Características y rendimiento de estudiantes universitarios. El caso de la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata. Documentos de Trabajo, Universidad Nacional de La Plata.
- Reyes R, S. L. (2004) El Bajo Rendimiento Académico de los Estudiantes Universitarios. Una Aproximación a sus Causas. Revista Theoretikos. Año VI, N° 18, enero-junio, El Salvador.
- Wilson, R. L.; Hardgrave, B. C. (1995) Predicting graduate student success in an MBA program: Regression versus classification". Educational and Psychological Measurement, 55, 186-195. USA.

Diseño de Plataforma Remota para Prácticas de Laboratorio

Héctor Hugo Mazzeo hmgvm@yahoo.com; José A. Rapallini josrap@gmail.com;
 Marcelo Zabaljauregui mzabaljauregui@gmail.com; Omar E. Rodríguez rodriguezomarlp@gmail.com
 Universidad Tecnológica Nacional – Facultad Regional La Plata
 La Plata, Buenos Aires, Argentina

RESUMEN

En muchas ocasiones, el contenido de las clases prácticas en asignaturas de carreras como Ingeniería se ve limitado por los recursos disponibles en los laboratorios. Estas limitaciones, que habitualmente tienen que ver con las capacidades de los recursos experimentales -generalmente escasos- pueden ser remediadas en gran medida mediante la utilización de Internet y realidad virtual aplicados al desarrollo de prácticas de laboratorio.

Las aplicaciones existentes en el mercado para realizar actividades de laboratorio en forma virtual y remota suelen ser costosas o difíciles de utilizar y en la mayoría de los casos no se adaptan a todas las necesidades requeridas en la enseñanza universitaria.

Surgió entonces la idea de formular un PID con el objetivo de desarrollar un laboratorio de ensayo de aplicaciones de hardware-software sobre el cual pudiera experimentarse con sistemas reales a distancia, sin necesidad de concurrir físicamente al mismo, facilitando el acceso de los estudiantes a las labores relacionadas con la formación práctica con libertad de horarios y tiempos. El resultado fue el desarrollo de un laboratorio virtual y remoto con software libre, adaptable a la realización de prácticas de distinta índole y con diferentes tipos de usuarios.

Palabras clave: laboratorio, remoto, virtual, codiseño, microcontrolador, embebido.

CONTEXTO

Este proyecto está enmarcado dentro del PID EIIFILP0003619TC, Laboratorios Virtuales y Remotos, homologado en el año 2015. Participa de su ejecución el grupo de Codiseño hardware/software para Aplicacio-

nes de Tiempo Real (CODAPLI), dedicado fundamentalmente a la investigación y desarrollo de proyectos relacionados con los sistemas de tiempo real: sistemas centralizados y distribuidos basados en computadoras personales, control y adquisición de datos, sistemas embebidos basados en placas con microcontrolador, etc.

El proyecto se desarrolla dentro del ámbito del Departamento de Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional La Plata, institución que financia íntegramente el proyecto facilitando asimismo las instalaciones y equipamiento del laboratorio.

1. INTRODUCCIÓN

A través de este proyecto se buscó brindar a los usuarios la posibilidad de realizar prácticas en entornos controlados, simulando ambientes de trabajo en donde se pueda interactuar en forma remota con los dispositivos o sistemas físicos, químicos, electrónicos, etc.

Para cada caso particular, sólo es necesario modificar el software y hardware propio de la aplicación a automatizar, adaptar sus entradas y salidas y reconfigurar algunos parámetros del sistema para acondicionarlo y vincularlo apropiadamente con el sistema principal.

Uno de los objetivos que se planteó en el proyecto fue que el usuario del laboratorio disponga de una plataforma remota con la cual poner en práctica los conocimientos adquiridos en las materias que cursa, sin la necesidad de recurrir a la adquisición de material extra curricular.

Para ello se implementó una red de computadoras con conexión a Internet, cada una de las cuales dispone de los siguientes elementos:

- Una cámara web y una placa microcontrolada con entradas y salidas analógicas y digitales conectada a la computadora local mediante un puerto USB. Esta placa permite implementar un sistema embebido con escaso o nulo hardware adicional.

- Hardware y software específicos para el desarrollo de la práctica de laboratorio. Aquí es donde se debe analizar cada práctica en particular y definir cuáles son las necesidades en cada caso: sensores y/o actuadores, interfaces o drivers de potencia para comandar los dispositivos electromecánicos, etc. En ciertos casos puede ser necesario cambiar directamente la placa microcontrolada por otra de mayores prestaciones en cuanto a número de entradas-salidas, capacidad de memoria, potencia del microcontrolador, etc.

- La aplicación de software principal desarrollada por el grupo de investigación y encargada de automatizar las tareas de la práctica, gestionar las conexiones de datos, turnos y horarios de utilización, etc.

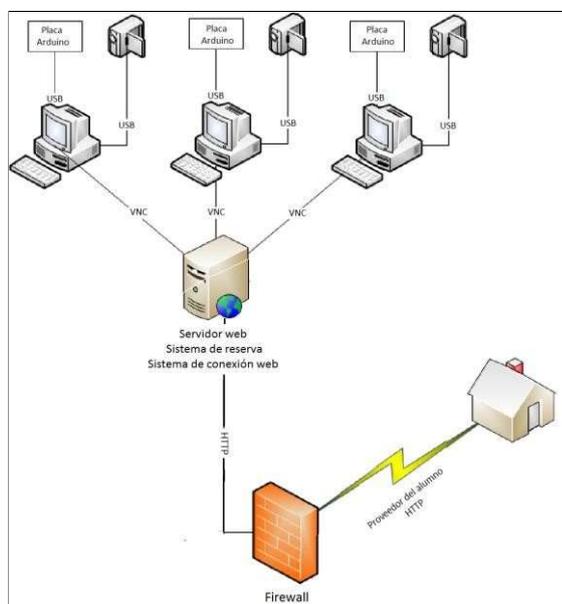


Fig. 1. Diagrama esquemático del laboratorio remoto

En los entornos educativos a distancia, resulta esencial disponer de un mecanismo que permita analizar, gestionar y compartir los recursos entre los usuarios. Cuando se trabaja en un entorno colaborativo con experimentos

reales ubicados en un laboratorio teleoperado, resulta obligatorio la implementación de un "gestor de acceso" mediante reservas. Es por ello que todas las máquinas se conectan a un servidor que se encarga de gestionar las reservas de los laboratorios y permite a los usuarios acceder al mismo por medio de un navegador web con soporte Java instalado (fig. 1).

El proyecto se dividió en tres etapas:

1) Puesta en funcionamiento de las PC locales con las cámaras IP y las placas conectadas y configuradas adecuadamente como también del servidor que permite conectarse a las mismas desde la web.

Para tal fin se instalaron y configuraron las siguientes aplicaciones:

- Para las máquinas locales: un servidor VNC para la visualización de la experimentación remota y la aplicación propia para el desarrollo de la práctica de laboratorio (habrá que implementar en cada caso hardware y software de control, sensores, actuadores, etc.).

- Para el gestor de acceso: un servidor Apache con *Bootstrap*. Se utilizaron distintas tecnologías: HTML, PHP, JQuery y LDAP. La conjunción de las mismas permitió diseñar un portal ágil para la gestión de los laboratorios, tanto para los usuarios como para los administradores de la plataforma. Además el uso combinado de Tomcat y Guacamole provee el acceso remoto a través de HTML 5.

2) Diseño de un sistema de reservas para gestionar los laboratorios y adecuación del sistema de acceso web para que pueda convivir con el sistema de reservas.

Ya que el sistema permite seleccionar distintos laboratorios para realizar, el usuario deberá ingresar primero en el laboratorio correspondiente y luego ir a la opción de reserva. Una vez allí podrá seleccionar el día y horario de realización de la práctica, siempre y cuando estén disponibles.

3) Apertura de los protocolos necesarios a Internet para que el laboratorio pueda ser accedido desde cualquier lugar por el usuario

remoto. Se cuenta con dos tipos de accesos: uno dedicado a la reserva y activación del laboratorio y otro para el acceso de los administradores a fin de visualizar las reservas hechas y realizar estadísticas del uso del mismo o algún deslogueo forzado en el caso de que surja algún problema o situación imprevista.

El procedimiento para la reserva de laboratorios es el siguiente:

a) Ingresar a la URL del proyecto e ir a la solapa “Laboratorios” (fig. 2). Luego en “Reserva de laboratorios” elegir el laboratorio deseado.



Fig. 2. Pantalla de ingreso al laboratorio remoto

b) Seleccionar la fecha y clicar en el botón “Buscar”. Se mostrarán las horas disponibles para seleccionar. Elegir un horario y clicar en el botón “Reservar” (fig. 3).



Fig. 3. Ingreso de datos para la confirmación de reserva

c) Verificar los datos que aparecen en la pantalla e ingresar el e-mail (que será el

usuario necesario para loguearse) y la contraseña (fig. 4).

Fig. 4. Ingreso de datos para la confirmación de reserva

d) El usuario recibirá un email de confirmación de la reserva con el formato de la fig. 5. El laboratorio estará disponible durante 1 hora.

e) Finalmente, ingresando al link que fue recibido por email aparecerá la pantalla de acceso (fig. 6).

Ya se está en condiciones de operar en el laboratorio remoto. Dentro del menú Aplicaciones/ATR se encontrarán los distintos programas de la cátedra, entre ellos “Cheese” que le permitirá abrir la cámara que apunta al proyecto físico instalado en ese laboratorio (fig. 7).



Fig. 5. Confirmación de reserva recibida por e-mail



Fig. 6. Pantalla de acceso al laboratorio remoto

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

En disciplinas técnicas y científicas la realización de experimentos con plantas reales es fundamental para consolidar los conceptos adquiridos en las clases teóricas. Sin embargo, debido a diferentes razones, los laboratorios reales no siempre están disponibles, lo cual impone restricciones en el aprendizaje. Afortunadamente, las nuevas tecnologías basadas en Internet pueden ser utilizadas para mejorar la accesibilidad a los experimentos.

Por consiguiente, trasladando este entorno práctico a la enseñanza a distancia, el elemento necesario para abordar la realización de prácticas sobre diversas materias de Ingeniería en Sistemas es la existencia de un sistema de apoyo a la enseñanza consistente en un laboratorio virtual y de telepresencia accesible a través de una red interna o Internet que permita al alumno practicar de una forma lo más similar posible a como si estuviese en las dependencias del laboratorio, dándole la posibilidad de manejar las simulaciones o interactuar con los sistemas reales.

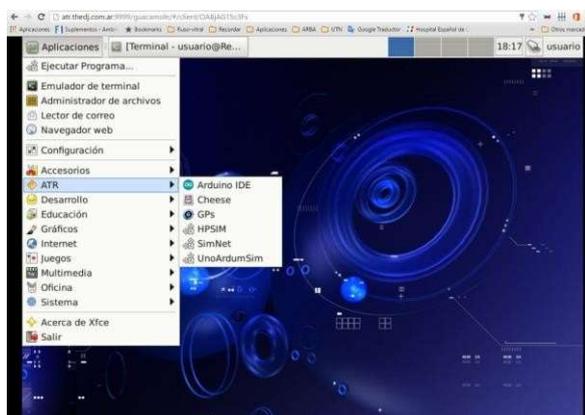


Fig. 7. Aplicaciones disponibles ya dentro del entorno

Así se conseguirán simultáneamente dos objetivos didácticos: (a) realizar prácticas relacionadas con la asignatura ampliando la disponibilidad de los laboratorios y (b) formar a los alumnos en el uso de las TICs.

El desarrollo de este proyecto requirió investigar y trabajar sobre 3 ejes básicos: desarrollo de aplicaciones de tiempo real mediante sistemas embebidos, manejo de protocolos de comunicación sobre redes TCP/IP y configuración de sistemas operativos y aplicaciones de computación virtual en red.

Las posibilidades de aplicación son variadas y cubren diversas áreas tanto en el ámbito de la educación como de la industria, pudiendo mencionarse a manera de ejemplo:

- aplicaciones industriales en desarrollos de sistemas de control de supervisión y adquisición de datos con variables de proceso a través de dispositivos sensores y actuadores de distinto tipo;
- robótica: teleoperación de brazos robóticos para manejo de sustancias peligrosas, control de posicionamiento de antenas de comunicaciones o paneles solares para generación de energía eléctrica, etc.;
- domótica o control de “edificios inteligentes”: sistemas para mejoramiento del confort y ahorro de energía en edificios, seguridad y control en tiempo real de propiedades, controles de riego para el ahorro de agua, etc.

3. RESULTADOS OBTENIDOS/ESPERADOS

Actualmente el sistema está siendo evaluado en forma local, es decir desde la red interna de la Facultad. Las pruebas preliminares que se han hecho para acceder remotamente fueron satisfactorias pero con un delay no tolerable para la visualización de video en tiempo real. Se están gestionando los permisos para obtener las direcciones IP estáticas que permitan acceder vía Internet utilizando un enlace de mayor velocidad. En la fig. 8 puede visualizarse en forma remota la imagen de una placa electrónica sobre la cual se realizan prácticas de laboratorio con

circuitos digitales discretos (multiplexores, contadores, registros, sumadores, etc.).



Fig. 8. Ejemplo mostrando una aplicación operada remotamente y visualizada a través de la webcam

Resta aún continuar con el mejoramiento y actualización del software y la creación de nuevas prácticas de laboratorio que se adapten a la temática de las distintas materias que deseen utilizarlo. Asimismo, falta desarrollar la parte de virtualización de laboratorios, lo que permitiría también mejorar la utilización en situaciones de bajo ancho de banda en las conexiones a Internet.

Se prevé realizar una evaluación del funcionamiento del sistema a través de estadísticas, encuestas de opiniones, etc., lo que contribuirá a encontrar fallas o problemas que requieran nuevas soluciones o mejoras en distintos aspectos del proyecto.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo que participa actualmente en el desarrollo de éste y otros proyectos de investigación relacionados a la temática de los sistemas de tiempo real está conformado por dos Ingenieros en Electrónica (director y codirector del grupo) y dos Ingenieros en Sistemas que también se desempeñan como docentes en el área de Sistemas de Información de la UTN-FRLP. El grupo se completa con cuatro becarios, tres de los cuales son alumnos y un graduado de esta misma facultad.

Uno de los autores de este trabajo (Ing. Omar E. Rodríguez) se encuentra realizando su tesis de postgrado titulada “Laboratorios remotos para facilitar el aprendizaje con experimentos de hardware y software para aplicaciones de tiempo real”.

Colaboran también en forma ad honorem alumnos del último año de la carrera de Ingeniería en Sistemas, quienes durante el año pasado realizaron cuatro tesinas de grado que cubren aspectos relacionados con los sistemas de tiempo real: "Domótica para el hogar", "Inmótica para aula de clases", "Sistema automatizado con caudalímetro para ahorro de agua", "Lector de perfiles de piezas mecánicas".

5. BIBLIOGRAFÍA

- [1] Computers & Education. Revista de Elsevier dedicada principalmente a aspectos pedagógicos del uso del computador y las comunicaciones en la enseñanza, incluyendo la enseñanza a distancia.
- [2] Dormido, S.; Sánchez, J.; Morilla, F. “Laboratorios virtuales y remotos para la práctica a distancia de la automática”, Dpto. de Informática y Automática, UNED, Avda. Senda del Rey s/n, 28040 Madrid.
- [3] Dormido, S. (2004), “Control learning: Present and future”, Annual Reviews in Control, vol. 28, pp. 115-136
- [4] Domínguez, M., Fuertes, J.J., Reguera P. (2009), “Desarrollo de un laboratorio remoto para la formación vía Internet en Automática”.
- [5] F. Torres, F. Candelas, S. Puente, F. Ortiz, J. Pomares, P. Gil, M. Baquero, A. Belmonte, “Laboratorios virtuales remotos para el aprendizaje práctico de asignaturas de ingeniería”, Departamento de Física, Ingeniería de Sistemas y Teoría de la Señal. Universidad de Alicante.
- [6] Isidro Calvo, Ekaitz Zulueta, Unai Gangoiti, José Manuel López, “Laboratorios remotos y virtuales en enseñanzas técnicas y científicas”, Dpto. Ingeniería de Sistemas y Automática Escuela Universitaria de Ingeniería de Vitoria-Gasteiz UPV/EHU.
- [7] IEEE Transactions on Education. Métodos y tecnologías de educación, materiales para la enseñanza, y programas de desarrollo educacionales y profesionales sobre disciplinas de ingeniería eléctrica.

Diseño y Aplicación de Estrategias para la Enseñanza Inicial de la Programación

Depetris Beatriz, Feierherd Guillermo, Pendenti Horacio, Aguil Mallea Daniel, Tejero Carlos Germán, Prisching Guillermo, Fierro Ariel, Aguilar Santiago, Domínguez Juan, Mamani Jonatan

Instituto de Desarrollo Económico e Innovación
 Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
 Hipolito Irigoyen 880 - Ushuaia - Tierra del Fuego
 {bdepetris, gfeierherd, hpendenti, daguilmallea, ctejero, gprisching}@untdf.edu.ar
 arielalejandrofierro@gmail.com, santeex@gmail.com, juani.dz@hotmail.com,
 jonush88@gmail.com

RESUMEN

La enseñanza y el aprendizaje inicial de la programación presentan importantes desafíos para los docentes y los alumnos de las carreras que requieren incorporarlos. Lejos de disminuir, estos han ido aumentando, como consecuencia de la necesidad de incorporar, a los conceptos tradicionales de programación, los que requiere la programación concurrente y paralela. [1]

La problemática no reside en la incapacidad para resolver problemas propiamente dicha sino en el escaso desarrollo del pensamiento computacional y de los procesos de abstracción que ellos requieren. [2].

El objetivo de esta investigación es mejorar algunas estrategias ya utilizadas y promover nuevas propuestas didácticas, que permitan afrontar dichos desafíos, buscando mejorar el desempeño académico de los alumnos ingresantes a las carreras de sistemas de la UNTDF.

Se prevé que la investigación propuesta habilite a encontrar nuevas dimensiones de análisis acerca de las prácticas docentes que aporten

conocimiento significativo al campo de la didáctica de la programación. Además, podrían encontrarse propuestas adecuadas para la introducción de estos temas en los últimos años de la currícula en el nivel secundario, no con el objetivo de que todos los alumnos se conviertan en futuros informáticos, sino que desarrollen su capacidad para resolver problemas.

Palabras clave:

Didáctica de la Programación, Robótica Educativa.

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral en trámite).

El financiamiento requerido es suministrado por la UNTDF.

El proyecto se desarrolla formalmente desde el 01/03/2017 hasta el 28/02/2019.

INTRODUCCIÓN

La enseñanza y el aprendizaje de la programación de computadoras representa una de las principales dificultades que enfrentan los alumnos que deciden estudiar carreras de la disciplina informática [3]. Esto se evidencia, al menos en las universidades argentinas, en elevados índices de fracaso en las materias iniciales que tratan estos temas.

Ahondando en las causas del fracaso, se ha concluido que las mismas no residen en la dificultad de los alumnos para traducir la solución de un problema a las sentencias propias de un lenguaje de programación, sino que, por el contrario, las mismas son mucho más profundas y tienen que ver con la falta de metodología, hábito y capacidad para resolver problemas. A su vez, se nota un marcado déficit de la capacidad de abstracción, la que, según Piaget, debería desarrollarse durante la adolescencia [4]. Por otra parte, es probable que, anticipándose a estas dificultades, muchos alumnos desistan de ingresar a este tipo de carreras luego de haber finalizado sus estudios secundarios.

No obstante, en la sociedad del conocimiento en la que estamos inmersos, las habilidades para resolver problemas son imprescindibles para desempeñarse en forma adecuada en cualquier ámbito. Las motivaciones del proyecto son, entonces, reducir los índices de deserción de los primeros años de las carreras

universitarias de sistemas y, en la medida que el proyecto pueda alcanzar alumnos del nivel medio, contribuir a aumentar la matrícula en las mismas e incrementar las capacidades de estos para resolver problemas.

En el convencimiento de que las dificultades que se enfrentan tienen como causa el déficit en la capacidad de abstracción, se busca subsanarlo recurriendo a estrategias que permiten que los alumnos evolucionen hacia el pensamiento abstracto partiendo de experiencias concretas. [5]

En ese camino, el equipo de trabajo ha desarrollado y viene utilizando con cierto éxito herramientas que permiten visualizar la ejecución de los algoritmos. A partir de desarrollos realizados en el Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la UNLP, se evolucionaron los mismos, incorporando funcionalidades que permiten utilizarlos para la enseñanza inicial de la programación concurrente y paralela [4].

Por otra parte, en los últimos años muchas instituciones han recurrido al empleo de robots con fines educativos. El uso de robots como medio didáctico es una alternativa significativa como motivación para materias introductorias y avanzadas en carreras de Informática e Ingeniería Electrónica, y sirve como eje de integración horizontal y vertical de contenidos curriculares. [6] [7]

La robótica móvil es uno de los campos de investigación que mayor auge está teniendo como investigación aplicada. En este proyecto se recurrirá al uso de robots didácticos de tecnología sencilla y al alcance de los presupuestos de una escuela, centro educativo o universidad de recursos limitados.

El proyecto incluye la adaptación de un robot (hardware) y del soporte informático (software) que permita su manejo por usuarios sin experiencia o formación específica. Mediante técnicas de enseñanza experimental el alumno inicial aprende nociones formales de algoritmo, secuencia y ejecución de autómatas. Para el alumno avanzado de ingeniería y de las carreras de informática el robot proporciona la posibilidad de experimentar en forma práctica con conceptos avanzados. El robot se telecomanda desde una computadora personal y puede incorporar un sistema de sensado de posición y de velocidad basado en la captura de imágenes provenientes de una cámara digital de bajo costo (tipo webcam). El alumno proporciona instrucciones en un lenguaje que inicialmente se ha diseñado de bajo nivel, aunque se contempla el desarrollo de una gramática libre de contexto de alto nivel en desarrollos futuros. Uno de los principales objetivos de este proyecto es formar desarrolladores con la capacidad y experiencia para poder definir nuevos sistemas, productos y mercados que utilicen esta tecnología. Existe en el país un gran potencial para desarrollos en robótica, tanto en el contexto industrial como en aplicaciones para la gestión del medio ambiente, por lo que se espera un alto grado de transferencia tecnológica como producto colateral de la formación que puede darse a través de estas experiencias teóricas y prácticas.

Por otra parte, un proyecto de estas características, dado el atractivo social que genera el uso de la tecnología y el muy bajo costo de la misma, es ideal para ser promocionado en colegios y otras instituciones educativas, buscando fomentar el interés en los estudios universitarios. Ello conseguirá atraer más

y mejores alumnos a las carreras técnicas de la universidad.

A su vez, el uso de esta tecnología como herramienta de estudio en las materias de las carreras de informática, motivará a los alumnos de las mismas para participar en las asignaturas y así disminuir el índice de deserción, minimizar el tiempo de finalizado de la carrera y extender las experiencias más allá de lo meramente formal. El desarrollo es lo suficientemente genérico como para que puedan participar alumnos de diferentes carreras, y diferentes niveles dentro de las mismas, favoreciendo la aplicación de los contenidos de cada materia en la solución de problemas concretos. Esto permite abordar un tema de gran importancia para la formación profesional, como es la integración curricular horizontal y vertical de los contenidos de varias asignaturas, lo cual hace posible que los alumnos comprendan el alcance y utilidad de los diferentes bloques curriculares y mejoren conceptualmente su articulación.

Varios han sido los grupos de investigación que han estudiado la utilización de la robótica desde la perspectiva docente, destacando los beneficios en los estudiantes en cuanto al desarrollo de habilidades tales como creatividad, trabajo en equipo, autoaprendizaje e investigación y como facilitador de aprendizaje contenidos teóricos en la disciplina de programación.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- a) La evolución de productos educativos ya desarrollados en proyectos anteriores, que integren la concurrencia y el paralelismo, poniendo énfasis en la visualización de la ejecución de los algoritmos.
- b) La inclusión de robótica educativa y su integración a los productos mencionados en a), tanto en materias iniciales de programación como en talleres de iniciación a la programación.

- prioritarias por el Ministerio de Educación de la Nación.
4. incorporar a modo de juego estrategias que ayuden a potenciar el desarrollo lógico matemático a través de la robótica educativa.
 5. desarrollar en los estudiantes, desde edades tempranas, la capacidad analítica y de resolución de problemas.
 6. generar entornos ricos para la adquisición de habilidades generales y científicas con objetos reales, que permitan resolver problemas por medio del uso de un pensamiento estructurado lógico y formal empleando robots adecuados en cada caso.

RESULTADOS OBTENIDOS / ESPERADOS

El proyecto se ha iniciado el 1º de marzo de 2017 por lo que a la fecha no pueden enumerarse resultados obtenidos. El grupo de investigación ha comenzado a trabajar en la 1ra. actividad del proyecto que consiste en el relevamiento de experiencias similares en el país y en el exterior a fin de actualizar el conocimiento del tema.

El objetivo general del proyecto es incrementar la matrícula de la carrera de la disciplina, y favorecer el desarrollo de estrategias para la resolución de problemas en los alumnos de los primeros años, buscando disminuir la deserción en esa etapa de la vida universitaria.

Algunos de los objetivos específicos son:

1. mejorar el rendimiento académico de los alumnos de los primeros años de las carreras de informática en la UNTDF.
2. disminuir la deserción en esa etapa.
3. despertar tempranamente vocaciones para el estudio de carreras de informática, declaradas

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está formado por docentes - investigadores, de los cuales dos (2) son Especialistas en Docencia Universitaria, cuatro (4) son Licenciados en Informática, uno (1) es Ingeniero en Sistemas de Computación y tres (3) son alumnos avanzados de la carrera Licenciatura en Sistemas.

Los Licenciados Pendentí y Aguil Mallea están desarrollando su tesis de Maestría en Ingeniería de Software en la Facultad de Informática de la UNLP. En el caso del Lic. Aguil Mallea lo hace en temas vinculados al proyecto. en temas afines con el proyecto.

El alumno Mamaní está desarrollando su tesis de grado, bajo la dirección de los docentes Depetris y Feierherd sobre temas afines al proyecto.

Uno de los beneficios del proyecto es la consolidación de parte del grupo de I/D y la formación de recursos humanos, en

especial introducción a temas de I/D de los tres alumnos que integran el proyecto. Se destaca que varios miembros del proyecto han trabajado en proyectos de investigación en temas afines a este proyecto como son: “Experiencias de intervenciones docentes en espacios virtuales”; “Evaluación y desarrollo de herramientas multimediales para análisis de competencias y aplicación de una metodología didáctica para mejorar el aprendizaje inicial en Informática”.

REFERENCIAS

- [1] Depetris B., Feierherd G. y otros. Experiencias con Da Vinci Concurrente en la enseñanza inicial de la programación y la programación concurrente. En Proceedings del VIII Congreso de Tecnología en Educación y Educación en Tecnología. Disponible en <http://hdl.handle.net/10915/27581>. Accedido el 3 de marzo de 2017.
- [2] Berrocoso J., Fernández Sánchez M., Garrido Arroyo M. Pensamiento Computacional y las nuevas ecologías del aprendizaje. Red - Revista de Educación a Distancia pag. 5 a 18. Setiembre de 2015. Disponible en http://www.um.es/ead/red/46/valverde_et_al.pdf. Accedido el 3 de marzo de 2017.
- [3] Blake, J.D. Language considerations in the first year CS curriculum. J. Comput. Sci. Coll. 26, 6 (2011), 124–129.
- [4] Depetris B., Feierherd G., Aguil Mallea D., Tejero G. A multiplatform interpreter to introduce structured and concurrent programming. Computer Science & Technology Series. RedUNCI; ISBN 978-987- 1985- 20-3; Año 2013
- [5] Cátedra Conceptos de Algoritmos, Datos y Programas, UNLP Por qué “pensar algoritmos” es tan importante en Informática. Revistas → Bit & Byte → Año 02 | Número 04- 2016 Disponible en: <http://sedici.unlp.edu.ar/bitstream/handle/10915/57362/>. Accedido el 3 de marzo de 2017.
- [6] Pásztor A., Pap-Szigeti R., Török E. “Mobile Robots in Teaching Programming for IT Engineers and its Effects”. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, 2013.
- [7] Correa C., Ferreira Szpiniak A. Laboratorios Teórico-Prácticos en Robótica Educativa. XI Congreso de Educación en Tecnología y Tecnología en Educación. 2016. Disponible en http://sedici.unlp.edu.ar/bitstream/handle/10915/53812/Documento_completo.pdf-PDFA.pdf?sequence=1. Accedido el 4 de marzo de 2017

Dispositivos Móviles como Soporte para el Aprendizaje Colaborativo de Programación en el Nivel Universitario Inicial (resultados)

Reyes, C.J., Massé Palermo, M.L., Espinoza, C., Vargas, C., Ramírez, J.,
Trenti, J.E.,

Consejo de Investigación de la Universidad Nacional de Salta (CIUNSa)/ Departamento de
Informática/ Facultad de Ciencias Exactas / Universidad Nacional de Salta (UNSa)

Avenida Bolivia 5150 - Salta - Capital - Argentina A4408FVY Teléfono: 54-0387-425547

e-mails: reyescarina@cidia.unsa.edu.ar; mlmassep@cidia.unsa.edu.ar

Resumen

El objetivo central del proyecto CIUNSa N° 2248 es estudiar la incidencia de estrategias colaborativas para el aprendizaje de Programación en el nivel universitario inicial utilizando tecnologías móviles y plataformas de educación a distancia. Para esto se realizaron estudios de caso en diferentes ámbitos de las carreras de Análisis de Sistemas de Información y Tecnicatura Universitaria en Programación de la Facultad de Ciencias Exactas, UNSa.

La incorporación de TICs en el aula conlleva un necesario proceso de resignificación de la práctica docente, en este caso particular, poniendo énfasis en contextos colaborativos mediados con tecnologías móviles y plataformas de educación a distancia. Se incorporaron dos recursos tecnológicos: una aplicación para dispositivos móviles *Open Diagramar* y un curso montado en la plataforma de educación a distancia Moodle. El diseño e implementación tanto del software *Open Diagramar* como del curso de Moodle fue diseñado y dirigido por el equipo de investigación basándose para ello, principalmente, en los requerimientos de contextos

colaborativos y de software para dispositivos móviles analizados.

En este trabajo se presentan algunos de los resultados obtenidos en el proyecto y los avances alcanzados hasta la fecha.

Palabras clave:

LMS, m-learning, entornos colaborativos, enseñanza de programación, ingreso universitario

Contexto

El Proyecto de Investigación N° 2248 “*Dispositivos móviles como soporte para el aprendizaje colaborativo de Programación en el nivel universitario inicial*” se ejecuta a través del Consejo de Investigación de la Universidad Nacional de Salta (CIUNSa) y es integrado por docentes de primer año de las carreras Licenciatura en Análisis de Sistemas y Tecnicatura Universitaria en Programación de la Universidad Nacional de Salta (UNSa).

Introducción

El proyecto 2248 se fundamenta en cuatro aspectos fundamentales: la Universidad como generadora de nuevos conocimientos, el software libre en las

universidades nacionales y el aprendizaje móvil en contextos colaborativos como estrategias para mejorar los índices de ingreso y permanencia. A continuación describimos brevemente cada uno de estos aspectos.

La Universidad como generadora de nuevos conocimientos y el software libre. La Universidad cumple un rol fundamental en la sociedad en la cual se encuentra inserta, siendo ésta un actor principal en la construcción de nuevos conocimientos que permitan a nuestra sociedad encontrar respuestas a sus necesidades locales. Para esto se requiere que la Universidad se conciba con y para los miembros de la comunidad que la contiene, promoviendo no solo el libre acceso al conocimiento que allí se construye sino propiciando espacios de enseñanza-aprendizaje respetuosos de las trayectorias sociales, culturales y económicas de sus estudiantes[2]. Este rol se potencia en la nueva era de revolución tecnológica y del conocimiento que, como plantean B. Busaniche y D. Saravia, “ponen a la sociedad en las puertas de nuevas y diversas culturas y sociedades, que complementan, potencian o cambian las preexistentes.”[3].

Las ciencias de la computación o informática no se encuentran exentas de ello, todo lo contrario pues “En nuestras sociedades mediatizadas por máquinas, el software se aplica en forma de ley.... El código -software- es ley de una forma muy profunda. El software cuando se ejecuta en automatismos de puntos de transacción, reemplaza al ser humano en la aplicación de la ley. Se ejecuta sin discusión”[3] ¿Cómo pueden las ciencias de la computación contribuir entonces a la construcción de una sociedad libre y democrática en esta nueva era? Principalmente promoviendo el desarrollo de esta Ciencia en el marco de libertades fundamentales que se formalizan en el

uso y desarrollo de Software Libre[4]. En este sentido los valores que una institución educativa requiere ofrecer están estrechamente relacionados con los que promueve el software libre: la libertad de pensamiento y de expresión, la igualdad de oportunidades, el esfuerzo y los beneficios colectivos en lugar de la ganancia individual[5]. De hecho, la libertad debería ser la finalidad central de la educación pues ésta sin libertad se transforma en adoctrinamiento. El Software Libre reafirma el concepto de construcción del conocimiento de forma colectiva y colaborativa, pues es muy difícil pensar que una persona genera su producción sin conocimiento previo, sin intercambio con otros. El nuevo conocimiento es el resultado de la interacción entre los preexistentes y las experiencias actuales.

El modelo de colaboración. Este modelo promueve que los estudiantes junto al docente, definan los objetivos específicos dentro de la temática general de lo que se busca enseñar/aprender, dando lugar a la búsqueda de actividades y estrategias que logren atraer su entusiasmo e interés por lograr las metas propuestas. Es fundamental entonces que la participación del docente se oriente a lograr que todos los integrantes del grupo compartan sus conocimientos, interactúen con respeto, escuchen opiniones diferentes para alcanzar la construcción de un pensamiento crítico y una participación activa, abierta y significativa [6].

El aprendizaje móvil. Es un proceso de enseñanza-aprendizaje mediado por la tecnología móvil. Este paradigma educativo posibilita que pequeños dispositivos disponibles por una gran masa de personas ofrezcan las mismas funcionalidades que una computadora de escritorio o portátil a bajo costo, con

capacidad de interactuar a través de internet o servicios de paquetes de datos.

De una encuesta realizada a los estudiantes de la cátedra Elementos de Programación en el año 2014 se sabe que el 90% tiene acceso a un dispositivo móvil y que en su gran mayoría utilizan el sistema operativo Android. A partir de esta encuesta se planificó el desarrollo de Open Diagramar.

Líneas de Investigación, Desarrollo e Innovación

El proceso de investigación se concreta en cuatro fases. La primera fase, Indagación Preliminar, se centró en la recolección y análisis de datos respecto del uso y accesibilidad a dispositivos móviles tanto por estudiantes como docentes de la cátedra de primer año. La segunda fase, diseño del contexto colaborativo, priorizó la determinación de los requerimientos funcionales que favorecen la construcción de conocimientos individuales y grupales basados en el intercambio de ideas, experiencias y conocimientos previos a través de actividades comunes en un entorno compartido en espacio y tiempo. Las dos primeras fases fueron llevadas a cabo durante el año 2014. La tercera fase, se centró en el desarrollo de Open Diagramar y selección de herramientas colaborativas de Moodle. Esta fase se ejecutó en el año 2015 y parte del 2016. La cuarta y última fase, Relevamiento y Análisis de Resultados, se centra en la utilización del software desarrollado en el marco de diferentes contextos colaborativos que fueron diseñados. Durante el desarrollo de estas actividades y al finalizar cada una de ellas, se aplican instrumentos adecuados para relevar información que permita realizar un análisis de los resultados obtenidos. Esta

fase se implementa durante el año 2016 y 2017.

Resultados y Objetivos

Como actividades de la cuarta fase que se encuentra en ejecución, se llevaron a cabo dos talleres colaborativos. Un primer taller se desarrolló en el marco del redictado de la primer asignatura de ambas carreras, Elementos de Programación, en donde los estudiantes se organizaron en grupos a cargo de un tutor virtual. Se desarrollaron cuatro actividades colaborativas, una por cada guía de Trabajos Prácticos, cada actividad se organizó en seis pasos:

Paso 1. Elección del ejercicio. Se propusieron dos situaciones problemáticas de similar complejidad para elección de los estudiantes. A través de un foro tipo Debate Sencillo los estudiantes interactuaron y acordaron su elección.

Paso 2. Definición de los datos de entrada (DE) y datos de salida (DS). A través de un foro tipo PyR (pregunta y respuesta), definieron los DE. Este tipo de foro permitió una primera instancia de razonamiento individual. Luego se analizan los aportes de todo el grupo y continúan el debate hasta alcanzar un acuerdo. Esta fase requirió de mucha presencia del tutor para guiar el razonamiento y generar aportes significativos.

Paso 3. Elaboración de Casos de Prueba (CP). Los estudiantes trabajaron en la construcción de al menos dos casos de prueba. Los CP se utilizaron en la última fase de validación del algoritmo para comprobar su funcionamiento. Se utilizó un foro de debate sencillo para profundizar el análisis del problema.

Paso 4. Determinación y resolución de sub-problemas. Se basa en el concepto 'divide y vencerás' aplicando una

estrategia de diseño top-down. Los estudiantes trabajaron en un foro de debate sencillo las siguientes actividades:

1) División del problema en sub-problemas. Elaboración de un listado de sub-problemas con la descripción de la/las componentes necesarias para construir una solución. Esta etapa puede requerir de varias instancias hasta llegar a un nivel de abstracción adecuado. El tutor es quien guía el análisis y da el visto bueno cuando se alcanza una partición adecuada; 2) Distribución de los sub-problemas entre los integrantes del grupo. Los miembros del grupo eligen al responsable del armado final del algoritmo. Este rol se rota entre los integrantes en cada uno de los equipos. En el foro cada integrante propuso una solución para el sub-problema que le tocó donde se analizaron, probaron y redefinieron en caso de ser necesario.

Paso 5 y 6. Armado de la solución final y entrega del trabajo final. En el paso 5 se utilizó un foro de debate sencillo a través del cual el estudiante a cargo del armado de la solución final presentó la solución para ser analizada por el equipo completo. Luego de recibir el visto bueno de todos, debían subir su propuesta la tarea definida en el *Paso 6*.

Paso 7. Devolución y evaluación del trabajo colaborativo. Finalmente se agrega un foro de solo lectura en el que el docente-tutor realiza una devolución. Se puso énfasis en la calidad y oportunidad de las participaciones de todos los estudiantes. Se realizaron observaciones sobre el diseño y construcción de la solución propuesta solo si fue necesario.

De los 65 inscriptos en el redictado de Elementos de Programación, 28 lograron regularizar la asignatura. Este resultado muestra una mejora significativa en el rendimiento de los estudiantes pues de un porcentaje histórico del 30% de regularización, se obtuvo un 43%. De los

28 estudiantes que regularizaron, un 80% de ellos participaron activamente en todo el trayecto del taller, mientras que el otro 20% participó en al menos alguna etapa de las actividades propuestas.

Un segundo taller se desarrolló en el marco del curso de ingreso de la cohorte 2017. Este taller, a diferencia del anterior cuyo objetivo fue mejorar los índices de rendimiento que impactan en la permanencia de los estudiantes, tuvo como principal objetivo motivar a los ingresantes en la resolución de problemas computacionales. Para esto se diseñó el taller utilizando un contexto lúdico.

El taller se llevó a cabo en dos encuentros de 3 horas cada uno. En el primer encuentro se realizó la presentación del software, incluyendo instrucciones y requisitos de instalación. Se introdujeron los conceptos de algoritmo y variables. Se presentaron dos estructuras de control: secuencia y alternativa. Se ejemplificaron situaciones problemáticas simples vinculadas a procesos matemáticos que les sean familiares usando las instrucciones de ingreso, asignación, salida y alternativa. Posteriormente, se propuso la formación de grupos de hasta 6 integrantes. Cada docente del taller (un total de 4 docentes) se hizo cargo de 5 grupos cada uno. Utilizando la herramienta Whatsapp se creó para cada equipo un grupo en la herramienta. Durante la semana entre el primer y segundo encuentro, cada equipo tuvo el desafío de resolver la mayor cantidad de problemas computacionales posibles. El desafío consistió en armar rompecabezas en donde cada pieza se corresponde con una instrucción del algoritmo que resuelve el problema computacional planteado por el docente. Los estudiantes debían ordenar de forma correcta las instrucciones de manera que éstas, al ejecutarse en el orden sugerido por el equipo, resuelvan la situación

problemática planteada. Como primer paso el docente envió un enunciado, con las respectivas piezas. Los integrantes del equipo trabajaron en la resolución de forma individual en sus celulares utilizando la aplicación Open Diagramar. Durante este proceso podían realizar consultas a través del grupo a sus pares; el docente fue un orientador del proceso en el que realizó sugerencias con el fin de orientarlos hacia una correcta solución sin proporcionar la solución en ningún momento. Una vez que acordaron una solución, el grupo la presentó al docente, quien les hizo los aportes necesarios y les entregó un nuevo desafío. La dificultad de los problemas fue incremental a medida que avanzaron en el juego.

Asistieron al taller 80 alumnos ingresantes de los cuales 55 participaron de los dos encuentros. Al finalizar el taller se implementó una encuesta de valoración relacionada al uso de la herramienta y la modalidad del taller. De este relevamiento se desprende que el 75% de los estudiantes no consideró difícil el uso de la herramienta ni la modalidad del taller y que tan solo 5 estudiantes no pudieron utilizar Open Diagramar por problemas técnicos con su dispositivo y 2 estudiantes por falta de celular. Un resultado muy interesante es que el 100% de los estudiantes manifestó que les gustó mucho el taller y que aprendieron nuevos conceptos. La razón más destacada en las respuestas es por haber sido un taller divertido al utilizar un juego, por incluir el uso de celulares y por incentivar el trabajo en equipo. Estos datos muestran que el juego y la tecnología motivan a los estudiantes a participar e interesarse por una temática académica porque aún cuando hubo estudiantes que no pudieron utilizar la herramienta, manifestaron que se sintieron motivados y que adquirieron de nuevos conocimientos.

Formación de Recursos Humanos

El proyecto de investigación está integrado por alumnos avanzados de la carrera de Licenciatura en Análisis de Sistemas lo cuales realizan programación con Android aplicando metodologías de desarrollo del software estudiadas en la carrera.

Referencias

- [1] MAC GAUL, M., et all. (2014). “Estrategia didáctica y recursos tecnológicos para la enseñanza de los sistemas de numeración”. Revista Iberoamericana TE&ET N°12, ISSN 1851-0086, RedUNCI – UNLP, pp. 81--91.
- [2] LITWIN, E. (2008). “Las configuraciones didácticas. Una nueva agenda para la enseñanza superior”. Paidós, Buenos Aires.
- [3] BUSANICHE, B, et all. (2005). “La contradicción fundamental de la Sociedad del Conocimiento” www.voltairenet.org/article125559.html
- [4] SARAVIA, D. (2005). “Ontología de la libertad del conocimiento y del software libre: luchas y debates.” <http://docs.hipatia.net/ontologia/>
- [5] AMATRIAIN, X. (2003). “Free software in education: a guide for its justification and implementation.” <http://www.iaa.upf.es/~xamat/FreeSoftware/FreeSoftwareEducation.pdf>.
- [6] COLLAZOS, C., et all. (2001) “Aprendizaje Colaborativo: un cambio en el rol del profesor”. Memorias del 3er Congreso de Educación Superior en Computación, Jornadas Chilenas de la Computación. Punta Arenas, Chile.

Educación con Tecnologías: la Gamificación Aplicada para el Aprendizaje de la Programación

Pedro A. WILLGING¹, Gustavo J. ASTUDILLO¹, Silvia BAST¹, Maricel OCCELLI^{2,3},
Leandro CASTRO¹ & Juan, DISTEL¹

¹Departamento de Matemática/FCEyN/UNLPam

²CONICET

³Departamento de Enseñanza de la Ciencia y la Tecnología /FCEyN/UNC

¹Av. Uruguay 151-Santa Rosa-La Pampa, 02954-425166

pedro@exactas.unlpam.edu.ar, astudillo@exactas.unlpam.edu.ar, silviabast@exactas.unlpam.edu.ar,
realiquensis@gmail.com, moccelli@efn.uncor.edu, leajcastro@gmail.com, disteljm@gmail.com

Resumen

Dentro de las líneas de investigación del proyecto “Educación con tecnologías: herramientas y tendencias”, se propone investigar el impacto sobre la enseñanza y el aprendizaje de la utilización de diversos recursos tecno-pedagógicos. Es así que se viene implementando un taller, para ingresantes, sobre introducción a la programación utilizando un enfoque lúdico basado en gamificación y juegos serios. A partir de la definición del estado del arte, que se llevó adelante durante 2014, se comenzó con el re-diseño de un curso existente sobre el entorno Moodle. Dicho taller ha mostrado un impacto positivo en el desempeño de los estudiantes que cursan Introducción a la Computación y continúa su implementación y mejora evolutiva año a año en función de las sucesivas evaluaciones.

Palabras clave: gamificación, programación, aprendizaje, ingreso

Contexto

El proyecto “Educación con tecnologías: herramientas y tendencias”, que inició en 2013, centra su investigación en el impacto sobre la enseñanza y el aprendizaje de la utilización de

diversos recursos tecno-pedagógicos y dentro de sus líneas de investigación hace foco en el impacto de un enfoque lúdico en el diseño y desarrollo de actividades, clases y materiales educativos. En este contexto se investiga el uso de la gamificación y los juegos serios.

El grupo de trabajo, cuenta con más de doce años de experiencia en la investigación sobre TIC aplicada a la educación.

Se trata de un proyecto acreditado (RCD N° 121/13) y financiado por la Facultad de Ciencias Exactas y Naturales de la UNLPam.

En el contexto del proyecto, el Grupo de Investigación y Desarrollo en Innovación Educativa (GrIDIE) y la cátedra Introducción a la Computación llevaron adelante el re-diseño e implementación del Taller de Introducción a la Programación (TIP). Los destinatarios del mismo son los ingresantes de las carreras Profesorado en Computación, Profesorado en Matemática y Licenciatura en Matemática (Facultad de Ciencias Exactas y Naturales, UNLPam).

Introducción

Las TIC permiten aprender tanto habilidades como conceptos, parte de los cuales, son adquiridos en el aprendizaje informal. Recuperar estos aprendizajes invisibles (Cobo *et al.*, 2013), como los obtenidos al utilizar

videojuegos, así como también hacer uso de metáforas que éstos usan para diseñar actividades/materiales e incorporarlas al aula es un desafío y un tópico que continúa siendo foco de las investigaciones a nivel mundial. Como muestran los estudios desarrollados por Felicia (2009) y Connolly *et al.* (2012) la actividad lúdica puede ser satisfactoria y segura, y puede incorporarse en el proceso de aprendizaje si se respetan hábitos de juego.

La inclusión de las TIC no produce, *per se*, mejoras en los procesos de aprendizaje, por tanto, su utilización debería estar enmarcada dentro de un proceso de investigación que permita la evaluación del impacto de las mismas en la enseñanza y/o en el aprendizaje.

En la actualidad los videojuegos o juegos digitales forman parte de la vida cotidiana de la mayoría de niños y adolescentes. Los avances tecnológicos propician este fenómeno, ya que posibilitan la utilización de este tipo de juegos, a través de distintos dispositivos y plataformas.

Se han planteado las consecuencias negativas del uso excesivo de los videojuegos (Felicia, 2009; Connolly *et al.*, 2012). Sin embargo, en la actualidad el énfasis está puesto en características tales como la motivación, el desafío o la obtención de competencias especiales. Aspectos, éstos, de relevancias en situaciones de aprendizaje.

El *Informe Horizonte* (Johnson *et al.* 2013) plantea un “horizonte de implantación” de dos a tres años para la *gamificación*. Por su parte, el TEC de Monterrey en su Reporte *Edu Trends* (Observatorio de Innovación Educativa, 2015) vaticina, para este enfoque, un tiempo de adopción un año o menos. Es posible corroborar el auge de la *gamificación*, sólo con una simple búsqueda en Google Académico (11700 resultados) o bien en algunos de los buscadores de repositorios como: IEEE Xplore:

495 publicaciones de las cuales 292 fueron desarrolladas entre 2015-2017, Elsevier: presenta 941 artículos, SciELO: 12, J-STOR presenta 129 resultados para “gamification” y 7

para el término “gamificación”, Redalyc presenta 92 artículos para “gamification”, 85 para “gamificación” y 11 para “ludificación”, PLOS expone 86 resultados para “gamification” y 3 para “gamificación.

La *gamificación* (también denominada *ludificación*, o en inglés *gamification*), hace referencia al “uso de elementos del diseño de juegos en contextos no lúdicos” (Deterding *et al.*, 2011, p.2). Esto es, aplicar metáforas que se corresponden con los juegos a otros contextos, como el educativo, en busca de mejorar la motivación y aumentar la participación. Como afirman Gallego *et al.* (2014) “gamificar es plantear un proceso de cualquier índole como si fuera un juego. Los participantes son jugadores y como tales son el centro del juego [...] y deben divertirse mientras consiguen los objetivos propios del proceso gamificado” (p. 2).

Se puede definir juego digital como “cualquier tipo de software de entretenimiento basado en computadora [...] usando cualquier plataforma electrónica como computadoras o consolas y que involucra a uno o varios jugadores en un ambiente físico o de red” (Frasca, 2001, p. 4).

Varios autores (Lifelong-Learning Programme, 2009; McGonigal, 2011; Kapp, 2012) coinciden en que los juegos digitales deben contar con las siguientes características: objetivos, reglas, desafíos o conflictos, competencia, colaboración y/o cooperación, retroalimentación y *re-playability* (re-jugabilidad). Las mismas deben ser tenidas en cuenta al llevar adelante un proceso de *gamificación*.

Los denominados juegos serios (como por ejemplo Kokori¹) son juegos digitales diseñados para educar, entrenar o informar (Michael & Chen, 2005).

¹ Juego digital en el cual a partir del control de nanobots que recorren la célula se resuelven problemas utilizando conceptos de biología celular (<http://www.kokori.cl/>).

El enfoque pedagógico

Un aspecto importante en toda propuesta educativa es lograr *buenos aprendizajes*. Es decir un aprendizaje caracterizado por “demostrar (a) un cambio duradero y (b) transferible a nuevas situaciones (c) como consecuencia directa de la práctica realizada.” (Pozo, 2008a, p. 162).

Sin embargo, deben darse las condiciones para que estos aprendizajes se produzcan: (i) que los estudiantes tengan los conocimientos suficientes para acceder al nuevo material, (ii) que la creación de actividades que (basadas en conocimientos previos) propicien los nuevos aprendizajes, (iii) que las actividades propuestas resulten un desafío (alcanzable) para los estudiantes. Además, debe plantearse (iv) cómo los conceptos serán presentados a los estudiantes (secuencia de aprendizaje) (Astudillo, Bast & Willging, 2016).

A lo anterior, es posible sumarle el razonamiento del experto y que tanto compañeros/as como juegos actúen como guía, es decir, como soporte permitiendo alcanzar nuevos conocimientos con mayor eficacia que si el estudiante emprendiera este desafío en soledad (Vygotsky, 1979).

El TIP

Definidos los aspectos pedagógicos de la propuesta, fue necesario llevar adelante un proceso de selección de juegos serios que serían incluidos en el TIP.

Con base en un conjunto de juegos serios enfocados en la enseñanza de la programación se indagó sobre fortalezas y debilidades de dichos juegos como recurso didáctico aplicables al contexto del TIP (conceptos de algoritmo, variables y estructuras de control), y se analizó si se adaptaban a los objetivos de aprendizaje propuestos y la posibilidad de definir con los mismos una secuencia didáctica coherente y significativa.

Como afirma Pozo (2008b) “La naturaleza dinámica de los procesos de aprendizaje tiene

[...] implicaciones para el diseño de situaciones más eficaces. Una de esas implicaciones tiene que ver con la importancia del orden temporal en las actividades de aprendizaje” (p. 165). Así se definió la siguiente secuencia en función de los conceptos que permite abordar cada juego y los conocimientos previos de los participantes del TIP:

1. lightbot. Lenguaje restringido e icónico. Con tutoría visual del sistema para el uso y la resolución de los problemas/ejercicios. Se aborda la noción de algoritmo. No se requieren conocimientos de programación.
2. code.org. Lenguaje restringido, icónico/textual. Se programa por encastre. Con tutoría visual del sistema para el uso y la resolución de los problemas/ejercicios. Se aborda la noción de algoritmo y estructuras de control.
3. Scratch. Lenguaje no restringido, icónico/textual. Se programa por encastre. Sin guía para la resolución de problemas. Se retoma la noción de algoritmo y estructuras de control, y se aborda el concepto de variables y operadores. Se trabaja sobre problemas abiertos.

Para *gamificar* el TIP (implementado sobre el entorno Moodle) se definieron cuatro niveles, donde por cada actividad completada se ganan puntos, información y se habilitan nuevos desafíos. Al finalizar cada nivel se gana una insignia y se habilita el siguiente nivel (Astudillo, Bast & Willging, 2016).

Líneas de Investigación, Desarrollo e Innovación

Las investigaciones se centran en la identificación y evaluación de metodologías y aplicaciones informáticas con un potencial

lúdico, y en el diseño, adaptación o re-utilización de estrategias que permitan incorporarlas a la clase (presencial/virtual), para posteriormente, analizar el impacto que éstas tienen en el diseño de materiales y/o en los procesos de enseñanza y de aprendizaje.

Los ejes principales en la línea de trabajo con gamificación y juegos serios son:

- Revisión de las técnicas y estrategias involucradas en la gamificación de un entorno educativo.
- Recopilación y evaluación de juegos serios de acceso libre.
- Desarrollo de estrategias y materiales compatibles con un entorno gamificado.
- Evaluación de módulos y recursos compatibles con el entorno Moodle para la implementación de la estrategia de gamificación.
- Evaluación del impacto del uso de gamificación y/o juegos serios en la motivación y en los procesos de enseñanza y de aprendizaje.

Resultados y Objetivos

En el contexto de la investigación sobre juegos serios y gamificación que lleva adelante el grupo GrIDIE se ha logrado localizar y evaluar un conjunto de juegos serios y con ellos definir una secuencia de aprendizaje basada en la estrategia de gamificación para el diseño de un taller de introducción a la programación.

Se han evaluado y/o instalado bloques y recursos compatibles con el Moodle tendientes a la implementar gamificación sobre dicho entorno.

Se han diseñado un conjunto de actividades educativas tendientes a favorecer el buen aprendizaje de las nociones básicas de programación de computadoras.

Se implementó el TIP durante las ediciones 2015, 2016 y 2017 del período de Ambientación Universitaria.

Para 2015 y 2016 se pudo observar un

mejoramiento en la participación y en la motivación de los participantes resolviendo las actividades propuestas, en comparación con ediciones anteriores. También, se pudieron corroborar buenos resultados en la evaluación conceptual al final del TIP (la mayoría de los participantes no cuentan con experiencia en programación). Asimismo, se aprecia un impacto positivo en el rendimiento de los estudiantes que completaron el TIP en la cursada de Introducción a la Computación. Cabe aclarar que el TIP es optativo. Al momento de escribir este trabajo se están recopilando los datos de 2017.

Finalmente, la implementación del TIP, da cuenta de la posibilidad de la utilización de una estrategia basada en gamificación y juegos serios para el aprendizaje de la programación.

Formación de Recursos Humanos

En este proyecto, trabajan actualmente dos investigadores formados y cuatro investigadores en formación, dos de los cuales alcanzaron el grado de Magíster durante 2016. Los investigadores pertenecen a dos universidades: UNC y UNLPam.

Referencias

- Astudillo, G. J., Bast, S. G., & Willging, P. A.** (2016). Enfoque basado en gamificación para el aprendizaje de un lenguaje de programación. *Virtualidad, Educación y Ciencia*, 7(12), 125–142.
- Cobo, C., & Moravec, J.** (2011). *Aprendizaje Invisible. Hacia una nueva ecología de la educación* (Laboratori de Mitjans Interactius / Publicacions i). Barcelona: Edicions de la Universitat de Barcelona.
- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M.** (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2),

661 – 686.

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). Gamification: Toward a definition. En *CHI 2011 Gamification Workshop Proceedings*. Vancouver, BC, Canada.

Felicia, P. (2009). *Videojuegos en el aula: manual para docentes*. Bruselas, Bélgica: European Schoolnet.

Frasca, G. (2001). Videogames of the oppressed: Videogames as a means for critical thinking and debate (Tesis). Georgia: Georgia Institute of Technology.

Gallego, F.; Molina, R. y Faraón, L. (2014). Gamificar una propuesta docente. Diseñando experiencias positivas de aprendizaje. Conferencia presentada en XX Jornadas sobre la enseñanza universitaria de la informática, Oviedo, España. En línea: <http://hdl.handle.net/10045/39195>

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013). *NMC Horizon Report: 2013 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Kapp, K. M. (2012). *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*. San Francisco: John Wiley & Sons.

Lifelong-Learning Programme (2009). *Production of Creative Game-Based Learning Scenarios - A Handbook for Teachers*.

Michael, D. & Chen, S. (2005). *Serious Games: Games That Educate, Train, and Inform*. Muska y Lipman/Premier-Trade.

McGonigal, J. (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Transform the World*. New York: The Penguin Press.

Observatorio de Innovación Educativa - Tecnológico de Monterrey (2015). *EduTrends. Radar de Innovación Educativa 2015*. Editorial Instituto Tecnológico y de Estudios Superiores de Monterrey.10-11

Pozo, J. I. (2008a). Capítulo 3. Las teorías del

aprendizaje: la integración entre diferentes niveles y sistemas de aprendizaje. En *Aprendices y maestros: la psicología cognitiva del aprendizaje* (pp. 121-148). Madrid, España: Alianza.

Pozo, J. I. (2008b). Capítulo 4. Los rasgos de un buen aprendizaje. En *Aprendices y maestros: la psicología cognitiva del aprendizaje* (pp. 159-175). Madrid, España: Alianza.

Vygotsky, L.S. (1979). *El desarrollo de los procesos psicológicos superiores*. Barcelona: Crítica.

Educación con Tecnologías: la Robótica Educativa Aplicada para el Aprendizaje de la Programación

Pedro A. WILLGING^{1,2}, Gustavo J. ASTUDILLO¹, Leandro CASTRO¹, Silvia BAST¹, Maricel OCCELLI^{2,3} & Juan, DISTEL¹

¹ Departamento de Matemática/FCEyN/UNLPam

² CONICET

³ Departamento de Enseñanza de la Ciencia y la Tecnología /FCEyN/UNC

¹ Av. Uruguay 151-Santa Rosa-La Pampa, 02954-425166

pedro@exactas.unlpam.edu.ar, astudillo@exactas.unlpam.edu.ar, leajcastro@gmail.com, silviabast@exactas.unlpam.edu.ar, moccelli@efn.uncor.edu, disteljm@gmail.com

Resumen

Desde el proyecto investigación “Educación con tecnologías: herramientas y tendencias”, se propone indagar sobre el impacto de la utilización de diversos recursos tecno-pedagógicos en la enseñanza y el aprendizaje. Es así que, desde 2014, se puso el foco en uso de robótica educativa como recurso didáctico para el aprendizaje de la programación. Principalmente en alumnos/as sin experiencia previa en la temática. La investigación está enfocada en tres aspectos: la motivación, el diseño de materiales de auto-aprendizaje y el uso de ambientes multi-plataforma de programación iconográfica.

El uso de la robótica educativa está ampliamente estudiado y, actualmente, existen una gran cantidad de proyectos (comerciales/libres) que permiten implementar actividades con robótica en el aula. Sin embargo, ¿Cuál es la posibilidad de hacerlo en las aulas de escuelas de nivel medio? ¿Qué recursos de software y hardware son los más apropiados para hacerlo? ¿Cómo se migra de la programación iconográfica a la textual? ¿Cuál sería la secuencia de didáctica que permitiría lograr aprendizajes transferibles (a otros lenguajes)? Estas preguntas guían esta línea de investigación y tienen sus respuestas preliminares en los resultados alcanzados hasta el momento.

Palabras clave: robótica educativa, programación, aprendizaje

Contexto

El proyecto “Educación con tecnologías: herramientas y tendencias”, que inició en 2013, centra su investigación en el impacto sobre la enseñanza y el aprendizaje de la utilización de diversos recursos tecno-pedagógicos. Es así que dentro de sus líneas de investigación, se enfoca en el impacto del uso de la robótica educativa en aprendizaje de los conceptos básicos de programación de computadoras.

El grupo de trabajo, cuenta con más de doce años de experiencia en la investigación sobre TIC aplicada a la educación.

El proyecto se encuentra acreditado (RCD Nº 121/13) y es financiado por la Facultad de Ciencias Exactas y Naturales de la UNLPam.

Introducción

Como se afirma en el *Horizon Report - Edición Educación Superior 2016* “La noción trabajar y vivir entre los robots es cada vez menos futurista y más práctica que nunca” (Johnson *et al.*, 2016, p. 46).

En este contexto una tendencia que cobra fuerza es el uso de robots en con fines educativos, la denominada robótica educativa. “La robótica se refiere al diseño y aplicación de robots, que son máquinas que realizan una serie de tareas automatizadas” (Johnson *et al.*, 2016, p. 46). Aplicado al contexto educativo “La robótica forma parte de un enfoque

pedagógico centrado en el alumno, que le permite construir objetos tangibles de su propio diseño y con sentido para él.” (Vaillant, 2013, p. 38). Para Monsalves (2011) y Ruiz-Velasco (2007) se trata de una disciplina que tiene como objetivo “generar entornos de aprendizaje heurístico” poniendo el foco en la participación activa de los estudiantes, donde los aprendizajes se construyen a partir de la experiencia del estudiante durante el proceso de construcción y programación de los robots. La robótica educativa entonces, convierte a la robótica en un medio para alcanzar ciertos aprendizajes.

Los robots son la conexión ideal entre una programación con una impronta lúdica y la representación de las instrucciones sobre un contexto real. Inicialmente, cuentan con el potencial de facilitar el aprendizaje de un lenguaje de programación, propiciar la experimentación y estimular las competencias asociadas a la resolución de problemas.

Desde esta línea de investigación, el grupo GrIDIE¹, busca desarrollar una secuencia de aprendizaje basada en el aprendizaje por descubrimiento de Bruner (1961) y en la teoría constructorista de Papert & Harel (1991) para el aprendizaje de nociones básicas de programación de computadoras.

Al enfocarse en el aprendizaje por descubrimiento se busca que el estudiante logre conocer a través de un razonamiento inductivo. Es decir, pasar de lo particular (detalles y ejemplos) a lo general (principios o conceptos). Así como también, requiere del involucramiento del estudiante en su propio proceso de aprendizaje. Para Bruner (1961) es necesaria la participación activa del estudiante en el proceso de aprendizaje, donde estudiante y docente cooperan en la resolución del problema.

El constructorismo (Papert & Harel, 1991) afirma que el conocimiento se construye. “El constructivismo es la idea de que el conocimiento es algo que construyes en tu cabeza. El constructorismo nos recuerda

que la mejor manera de hacerlo es construir algo tangible -algo fuera de tu cabeza- que es también personalmente significativo.”

(Papert, 1990, p. 14).

Con el fin de favorecer la participación activa de los estudiantes y una construcción de los conceptos de programación desde ejemplos concretos fue necesario, para el diseño de la propuesta, llevar adelante una búsqueda de aplicaciones (software) que permitiera programar los robots de forma sencilla.

Asimismo, dado que el fin último de la secuencia de aprendizaje es su utilización/aplicación en escuelas de educación secundaria (del medio local), se intenta reducir el costo de la implementación de la propuesta. Existen en el mercado robots comerciales pre-ensamblados (como Scribbler², Thymio³, Dash & Dot⁴ o Bee-Bot⁵, entre otros) o que requieren de ensamblado (como los kit de RobotGroup⁶, Robotis⁷ o Lego Mindstorms⁸, entre otros) diseñados para su uso didáctico. Sin embargo, pueden ser costosos o difíciles de adquirir para una Institución Educativa (Díaz *et al.*, 2012; González & Jiménez, 2009). Por tanto, en esta propuesta se hace uso software/hardware libre (lo que además permitir el abordaje de estos conceptos).

Otro aspecto que se tuvo en cuenta es el desarrollo de las actividades extra-clase. Dado el caso de que las Instituciones cuenten con los robots, los mismos, suelen estar disponibles de forma restringida. Además, el enfoque adoptado prioriza la experimentación por parte de los estudiantes. En este contexto, surge la posibilidad de incorporar los simuladores como parte de la secuencia. Se analizaron dos posibilidades: Laboratorios Virtuales (LV) y Laboratorios Remotos (LR).

Los primeros, están implementados

² Disponible en: <https://www.parallax.com/>

³ Disponible en: <https://www.thymio.org/es:thymio>

⁴ Disponible en: <http://www.dashdot.com/>

⁵ Disponible en: <https://www.bee-bot.us/>

⁶ Disponible en: <http://www.robotgroup.com.ar/es/>

⁷ Disponible en: <http://www.robotis.us>

⁸ Disponible en: <https://www.lego.com>

¹ Grupo de Investigación y Desarrollo en Innovación Educativa. Dpto. de Matemática. FCEyN. UNLPam.

íntegramente a través de herramientas informáticas, hacen uso de objetos/recursos que no tiene existencia física y no se cuenta con infraestructura de laboratorio, todo se soporta con herramientas informáticas (Revuelta, Massa & Bertone, 2016). “En términos generales, el [LV] permite manipular las variables del objeto, lo que favorece el aprendizaje por descubrimiento” (Revuelta, Massa & Bertone, 2016, p. 36).

Por su parte el LR permite el acceso al laboratorio real, en general, a través de la red, de manera no presencial (Revuelta, Massa & Bertone, 2016).

Arduino

Para llevar adelante la implementación de la secuencia se decidió utilizar placas Arduino.

Según el sitio oficial “Arduino es una plataforma electrónica de código abierto basada en hardware y software fáciles de usar.” (Arduino, 2017). Específicamente se trata de un Sistema Embebido (SE). Es decir “una combinación de hardware y software que trabaja junto con algún sistema mecánico o electrónico diseñado para cumplir una función específica” (Bordignon & Iglesias, 2015, p. 10).

Desde el punto de vista del hardware Arduino, como todo SE, cuenta con un microcontrolador, al que se puede conectar tanto sensores y como actuadores. “Un microcontrolador es un circuito electrónico programable, capaz de ejecutar las órdenes grabadas en su memoria” (Revuelta, Massa & Bertone, 2016, p. 45). “Los sensores le permiten [...] obtener información del mundo real; [...] y los actuadores realizan acciones con el mundo físico” (Bordignon & Iglesias, 2015, p. 11).

Las placas Arduino cuentan, además, con un entorno integrado de desarrollo o IDE (basado en *Processing*⁹) que permite la programación de la placa en lenguaje Arduino (basado en *Wiring*¹⁰).

Líneas de Investigación, Desarrollo e Innovación

Las investigación, en el contexto del proyecto, se focaliza en identificar, probar y evaluar nuevas (o *revisitadas*) estrategias de enseñanza y de aprendizaje, pero también aplicaciones informáticas que permitan potenciar un enfoque lúdico en el diseño y desarrollo de una clase (presencial/virtual). Esto permite, posteriormente, analizar el impacto que éstas (estrategias/aplicaciones) tienen en los docentes (su rol, el diseño de materiales, entre otros) y en los aprendizajes de los estudiantes.

Los ejes principales en la línea de trabajo con robótica educativa son:

- Revisión de experiencias que hagan uso de la robótica educativa para la enseñanza o el aprendizaje de conceptos y competencias.
- Relevamiento y evaluación de entornos de desarrollo, de acceso libre, que permitan la programación de placas Arduino. Focalizando en enfoques de tipo lúdico.
- Desarrollo de estrategias y materiales educativos que permitan el aprendizaje de conceptos de programación a través de la robótica educativa.
- Evaluación de sensores y actuadores, de bajo costo, compatibles con Arduino que permitan la generación de problemas en un contexto real.
- Evaluación del impacto del uso de la robótica educativa en la motivación y en los procesos de enseñanza y de aprendizaje.

Resultados y Objetivos

En el proyecto se han podido evaluar placas Arduino UNO, Arduino Mega, RaspBerry Pi, picoBoard¹¹, también un kit Kano¹² y robots N6/N8¹³. Éstos últimos se

⁹ Más información en: <https://processing.org/>

¹⁰ Más información en: <http://wiring.org.co/>

¹¹ Más información en: <http://www.picocricket.com>

¹² Más información en: <http://www.kano.me/>

¹³ Más información en: <http://www.robotgroup.com.ar>

comenzarán a utilizar a partir de la inclusión de actividades en la cátedra Informática (Licenciatura en Física) para propiciar el aprendizaje de las nociones de programación estructurada.

Para el diseño de la secuencia de aprendizaje se están utilizando placas Arduino UNO, junto con un conjunto de sensores y actuadores. De los primeros, hasta el momento, se han testeado: sensor ultrasónico (HC-SR04), de movimiento/PIR (HC-SR501), de humedad relativa y temperatura (DHT11), de presión¹⁴, de lluvia (FC-37), infrarrojo (AX1838HS) y módulo *bluetooth* HC-05. Respecto de los actuadores, se utilizaron: leds, servomotores (HXT900), parlantes de papel (reciclados de PCs), motores DC, display LCD 16x2 y control remoto.

Se han construido con los recursos descritos en el párrafo anterior: tres robots móviles (denominados *carrindanGO!*), uno que define sus movimientos en función del hallazgo de obstáculos (sensor ultrasónico, Figura 1), otro que lo hace a partir de las teclas del control remoto y, el último, a partir de órdenes recibidas a través de bluetooth desde una aplicación de celular. También, se implementaron una barrera de estacionamiento (servomotor) que se levanta a partir del sensor de presión (incluye sonido y luces), y una alarma a través de un sensor de presencia (incluye sonido y luces).

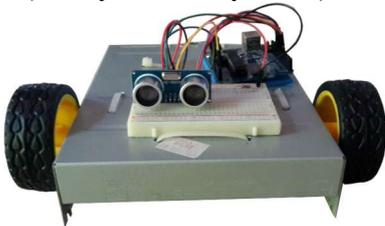


Figura 1. CarrindanGO! con sensor ultrasónico

Respecto a los LV se han analizado: Autodesk Circuits¹⁵, Virtual Breadboard (VBB4Arduino)¹⁶, UnoArduSim¹⁷ y Arduino

¹⁴ Construido con cartón y papel de aluminio.

¹⁵ Disponible en: <https://circuits.io/>

¹⁶ Disponible en: <http://www.virtualbreadboard.com>

¹⁷ Disponible en:

<https://www.sites.google.com/site/unoardusim/home>

Simulator¹⁸. Sólo se evaluó el LR Arduino Remote Lab¹⁹. Se adoptó para la secuencia *Autodesk Circuits* por sus características: *online*, de diseño visual y lo amigable del ambiente.

Desde el punto de vista de los IDEs, se analizaron: Scratch for Arduino (S4A)²⁰, ArduBlock²¹, mBlock²², Physical Etoys²³ y Visualino²⁴. Si bien, inicialmente se seleccionó mBlock, finalmente se cree que Visualino tiene la potencialidad de facilitar el paso de la programación icónica a la textual.

También se ha diseñado material de apoyo donde se muestra para cada uno de los sensores: los materiales, el esquema de conexión (diseñado con Fritzing²⁵, Figura 2) y un ejemplo básico en lenguaje Arduino para probar su funcionamiento.

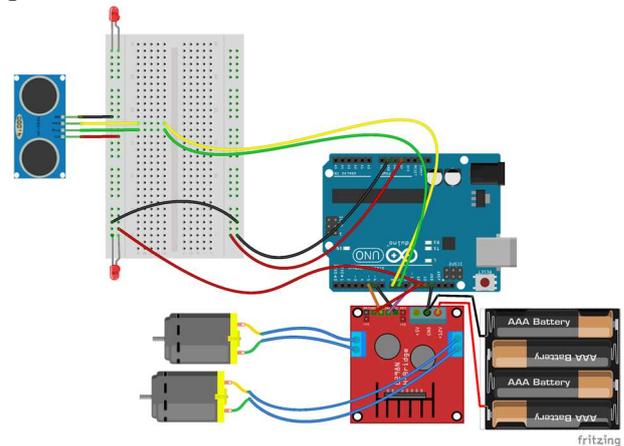


Figura 2. Esquema de conexión del CarrindanGO!

Asimismo, se diseñó la primera secuencia de aprendizaje, la cual será puesta a prueba, en primer lugar, con un conjunto de profesores de computación, con cursos en educación secundaria, con el fin de evaluar su pertinencia y hacer los ajustes necesarios. Para luego, llevar las propuestas a las aulas o a talleres extra-curriculares en las escuelas.

¹⁸ Disponible en: <https://arduinosim.sourceforge.io/>

¹⁹ Disponible en: <http://194.24.226.109/>

²⁰ Disponible en: <http://s4a.cat/>

²¹ Disponible en: <http://blog.ardublock.com/>

²² Disponible en: <http://www.mblock.cc/>

²³ Disponible en:

<http://tecnodacta.com.ar/gira/projects/physical-etoys/>

²⁴ Disponible en: <http://www.visualino.net>

²⁵ Disponible en: <http://fritzing.org>

Formación de Recursos Humanos

En este proyecto, trabajan actualmente dos investigadores formados y cuatro investigadores en formación, dos de los cuales alcanzaron el grado de Magíster durante 2016. Los investigadores pertenecen a dos universidades: UNLPam y UNC.

Referencias

- Arduino.** (2017). What is Arduino? [Página Web]. Recuperado a partir de <https://www.arduino.cc/en/Guide/Introduction>
- Bordignon, F. R. A., & Iglesias, A. A.** (2015). *Diseño y construcción de objetos interactivos digitales*. UNIPE: Editorial Universitaria. Recuperado a partir de <http://sedici.unlp.edu.ar/handle/10915/50448>
- Bruner, J. S.** (1961). The act of discovery. *Harvard Educational Review*, 4, 21-32.
- Díaz, F. J., Banchoff Tzancoff, C. M., Martín, E. S., & López, F.** (2012). Aprendiendo a programar con juegos y robots. En *VII Congreso de Tecnología en Educación y Educación en Tecnología* (p. 6). Buenos Aires, Argentina.
- González, J. J., & Jiménez, J. A.** (2009). La robótica como herramienta para la educación en ciencias e ingeniería. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, (10), 31-36.
- Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., y Hall, C.** (2016). NMC Informe Horizon 2016 Edición Superior de Educación. Austin, Texas: The New Media Consortium.
- Monsalves González, S.** (2011). Estudio sobre la utilidad de la robótica educativa desde la perspectiva del docente. *Revista de Pedagogía*, 32(90), 81-117.
- Papert, S., & Harel, I.** (1991). Situating constructionism. *Constructionism*, 36(2), 1-11.
- Papert, S.** (1990). A critique of technocentrism in thinking about the school of the future. Epistemology and Learning Group, MIT Media Laboratory. Memo No. 2.
- Revuelta, M. Á., Massa, S. M., & Bertone, R.** (2016). *Laboratorio Remoto en un Entorno Virtual de Enseñanza Aprendizaje* (Tesis de Maestría). Universidad Nacional de La Plata, La Plata, Argentina.
- Ruiz-Velasco, E.** (2007). *Educatrónica: Innovación en el aprendizaje de las ciencias y la tecnología*. Madrid: Díaz de Santos.
- Vaillant, D.** (2013). *Integración de TIC en los sistemas de formación docente inicial y continua para la Educación Básica en América Latina*. Argentina: UNICEF Argentina. Recuperado a partir de https://www.unicef.org/argentina/spanish/educacion_Integracion_TIC_sistemas_formacion_docente.pdf

Educación Mediada por Dispositivos Móviles

Mg. Roberto Bertone¹, Mg. José Luis Filippi², Lic. Guillermo Lafuente³, Mg. Carlos Ballesteros⁴,
Lic. Gustavo Lafuente⁵, I.S. Daniel Perez⁶, I.S. Sofía Aguirre⁷, A.S. Alejandra Mansilla⁸
LIAU⁹ - Facultad de Ingeniería – UNLPam.

pbertone@ada.info.unlp.edu.ar¹

{filippij², lafuente³, balleste⁴, gustavo⁵, perezd⁶, aguirres⁷, mansilla⁸}@ing.unlpam.edu.ar

⁹Laboratorio de Investigación de Ambientes Ubicuos

Resumen

El objetivo de este proyecto es mejorar los diferentes procesos que se efectúan en las instituciones educativas mediante la aplicación de las tecnologías de la información y la comunicación. En particular, se enfocará en la exploración y explotación de las tecnologías de la computación ubicua [1], direccionando la labor investigativa a mLearning.

Con el fin de explorar las posibilidades que brinda el cómputo móvil en el contexto educativo, se efectuará el trabajo experimental en la Facultad de Ingeniería - UNLPam, contando con un escenario de aplicación real. Como parte del trabajo de investigación se realizarán diferentes tareas:

- Indagar el potencial de la tecnología móvil en el ámbito educativo.
- Analizar diferentes herramientas tecnológicas móviles en ambientes educativos.
- Entrevistar a los diferentes actores en su ámbito laboral cotidiano.
- Identificar diferentes escenarios de trabajo que muestren aspectos móviles en las actividades de sus actores.
- Documentar el proceso de investigación.
- Presentar en diferentes congresos de carácter nacional e internacional los resultados alcanzados.
- Confeccionar aplicaciones móviles de interés educativo.

- Difundir y capacitar a instituciones interesadas en utilizar las aplicaciones desarrolladas.

El proyecto trata de determinar la utilidad real de los dispositivos móviles en el ámbito educativo con el objetivo de establecer una implementación a gran escala.

Palabras claves: Aprendizaje Móvil. Dispositivo Móvil. Aprendizaje Colaborativo.

Contexto

Tipo de Investigación: Aplicada

Campo de Aplicación Principal: 7 1802 Computación, 7 1803 Comunicaciones.

Campos de Aplicación posibles: 13 1040 Ciencia y Tecnología, 7 4399 Otras – Educación – Tecnología Aplicada a la Educación

Institución que Coordina el Proyecto: Facultad de Ingeniería, Universidad Nacional de La Pampa.

Introducción

En ambientes con necesidades de representación y manejo de conocimiento como el ámbito educativo existen múltiples necesidades: desarrollo de contenido bibliográfico digital, implementación de programas que permiten la comunicación a través de variados dispositivos tecnológicos móviles, almacenar y distribuir el material educativo en variados formatos y diferentes plataformas, atender permanentemente a los actores involucrados dando respuesta a las

necesidades de padres, alumnos, docentes, administrativos y entidades del medio.

Estas relaciones constituyen en sí conocimiento agregado de suma utilidad, rebrotan la inquietud de Vannevar Bush (considerado el precursor de la Hipermedia), escrita en “*As we may think*”, donde establecía las dificultades para almacenar y consultar eficientemente la gran cantidad de conocimiento acumulado. La clave para que el material pueda ser consultado eficientemente está en la posibilidad de contar con las relaciones apropiadas que conecten dicho material.

Con la aparición de las redes de computadoras y la conectividad que ofrecen los dispositivos móviles, se conforma un ambiente propicio para satisfacer las necesidades de interacción entre los diferentes actores. Mark Weiser [1] menciona que: “*vamos camino a ambientes ubicuos, ambientes poblados de numerosos sensores que gracias a la miniaturización de los dispositivos son invisibles al usuario y están en permanente rastreo de la actividad humana*”. Aquí radica el objetivo primario del proyecto: indagar diferentes herramientas tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el ámbito educativo.

Situación Actual del Problema

En la sociedad actual se pueden observar una gran cantidad de componentes electrónicos que rodean a las personas y que están integrados en la vida cotidiana, gracias a su reducido tamaño y a su capacidad de cómputo, por ejemplo, computadoras personales, portátiles, teléfonos móviles de última generación, iPod, etc. [2], que logran comunicarse entre sí gracias a la existencia de protocolos de comunicación inalámbrica; a través de redes de celulares, redes de área

local (LAN), redes de área extensa (WAN), y Bluetooth entre otras [3].

Los dispositivos móviles son cada vez más pequeños, más poderosos y a medida que aumenta la demanda del consumidor, más baratos, incorporando nuevas funcionalidades diseñadas para permitir a los usuarios trabajar y acceder a la información en cualquier situación.[4]

Ante la presencia de nuevos dispositivos de comunicación un nuevo escenario social y educativo se define acorde a los cambios en la nueva organización del conocimiento, que inciden directamente en los procesos mentales. Está cambiando la forma de conocer, de interactuar con los demás, el volumen de información, su accesibilidad, los nuevos medios de almacenamiento masivo de información, nuevas formas de comunicarse a través de diferentes redes entre personas que se encuentran distantes físicamente. Mark Weiser clasifica la era de la computación en tres momentos: en 1980 el uso de las grandes computadoras (mainframes), en 1990 las computadoras personales (PC), y a partir del 2000 la computación ubicua (UC).

En este sentido, la tercera generación o paradigma en la computación cuenta con múltiples aplicaciones y equipos de investigación que procuran su desarrollo [5]. El concepto de ubicuidad se refiere en general a la presencia de una entidad en todas partes; pero en la computación adquiere la característica de ser, además, invisible. Este paradigma pretende brindar sistemas de cómputo inteligentes que se adapten al usuario y cuyas interfaces permitan que éste realice un uso intuitivo del sistema. De allí que la meta de la computación ubicua, de integrar varias computadoras y dispositivos inteligentes al entorno físico busca habilitar los beneficios de éstas y de la información digitalizada en todo momento y en todas partes.

La influencia de la computación ubicua es en gran parte de carácter teórico, y dadas las condiciones de integración entre educación y sistemas de cómputo explicada antes, comparte créditos junto a otros conceptos como usabilidad, interactividad e hipertextualidad [6]. Una posible aplicación en educación a distancia sería que el docente presentara el conocimiento en varias formas que se diferencien en sus características técnicas y didácticas. Así el estudiante optará por una vía conveniente de aprehensión del conocimiento con el consecuente mejoramiento de la calidad en el aprendizaje. Didácticamente, esto implicaría el diseño de distintos materiales o herramientas teniendo en cuenta los resultados de los antecedentes educativos del alumno. De esta manera se podrán ofrecer cambios significativos en el aprendizaje al conocer los modos de enseñanza que prefiere el estudiante. Otro aspecto que tiene relación con el concepto de ubicuidad y dispositivos móviles se encuentra en la ciencia cognitiva. En este sentido las teorías de la percepción retomadas por Norman (1998), hablan de la existencia de una percepción periférica donde los objetos solamente son apreciables cuando el ser humano centra la atención en ellos. En los sitios educativos se utilizan lenguajes de programación que posibilitan la “invisibilidad” de los elementos no relevantes en las acciones que ejecuta el usuario en la interfaz. Algunos de estos lenguajes, como XML [7], permiten que el receptor de la información pueda acceder a ésta desde cualquier navegador. Todo esto sucede sin percatarse de los cambios en el diseño que son necesarios para presentar los datos. A modo de ejemplo y bajo el mismo concepto, los futuros buscadores de contenidos que sigan los lineamientos definidos por la Web Semántica [8] permitirán arribar a contenidos con mayor precisión para el usuario, que los

buscadores tradicionales actuales; actual desafío de la llamada Web 3.0.

En el diseño de materiales didácticos para la web serán también importantes otras características técnicas, como por ejemplo la diferencia de peso entre archivos, la facilidad de lectura de un formato, el ancho de banda de la red y sus posibilidades multimediales, etc. [9]. Este conjunto de opciones técnicas que permiten la ubicuidad de un sistema de cómputo y las posibilidades conceptuales de diseñar la información pueden considerarse diseño ubicuo. Trascendiendo teóricamente de los sistemas computacionales a la concepción de aplicaciones didácticas en sitios educativos, el diseño ubicuo brinda opciones para mejorar el desempeño de Internet y la Web como medios de comunicación y educación [10].

A partir de establecer una relación estrecha entre el diseño de interfaz y el diseño didáctico, bajo el enfoque de la computación ubicua, ambos ofrecerán al usuario o alumno el conocimiento que requiere sin realizar un gasto cognitivo importante en el medio que transmite [11]. Una integración recíproca entre la educación a distancia y el diseño ubicuo, además de potenciar el uso de Internet en procesos de enseñanza y de aprendizaje, contribuirá a la difusión y aplicación de las tendencias tecnológicas que facilitan su uso. En este contexto la escuela debe cambiar, añadiendo nuevas herramientas tecnológicas que se encuentran presentes en la vida cotidiana de las personas (por ejemplo los teléfonos inteligentes) e incorporar su potencial al servicio de los diferentes actores, generando así un espacio de acción profesional distinto al actual. El entorno educativo cambia [12,3], la educación se presenta como la formación de los educandos en competencias, destrezas, habilidades para desempeñarse en un nuevo espacio social, el digital [13]. Nuevos escenarios educativos,

los mismos actores con un nuevo rol, y la implantación de las tecnologías móviles es el desafío actual [14].

Línea de Investigación y Desarrollo

El plan de actividades corresponde al proceso de investigación aplicada, con objetivos de corto, mediano y largo plazo, y una duración prevista de cuatro años.

Primer Año:

- Entrevistar a los diferentes actores, directivos, docentes, alumnos, personal administrativo, en su ámbito laboral cotidiano.
- Censar la presencia de dispositivos móviles, el tipo de plataforma utilizada.
- Investigar los requerimientos básicos de TICs para incorporar al proyecto.
- Indagar trabajos de terceros sobre las aplicaciones móviles utilizadas en el ámbito educativo.
- Analizar diferentes herramientas tecnológicas orientadas al desarrollo de actividades educativas móviles (mLearning).
- Definir una metodología de desarrollo de trabajo a efectuar.

Segundo y Tercer Año:

- Estudiar la tecnología necesaria para el desarrollo de aplicaciones móviles.
- Definir prioridades según necesidades de inmediatez para el uso y desarrollo de aplicaciones móviles orientadas a la educación.
- Confeccionar nuevas aplicaciones que se ejecuten en dispositivos móviles y que sean de interés educativo.
- Probar las aplicaciones desarrolladas en el ámbito educativo adecuado y realizar las pruebas necesarias para la puesta a punto.

Cuarto Año:

- Implantar las aplicaciones en los dispositivos móviles que hayan alcanzado un funcionamiento aceptable.
- Registrar el trabajo desarrollado en sus diferentes etapas.
- Difundir los logros alcanzados a la comunidad, para su posterior distribución.
- Presentar en diferentes congresos de carácter nacional e internacional los resultados alcanzados.
- Difundir y capacitar a instituciones interesadas en aplicar la tecnología presentada.

Resultados Obtenidos/Esperados

El proyecto da inicio a su actividad de investigación y desarrollo durante el año 2015, por lo que sus integrantes se encuentran comenzando las tareas enmarcadas en el tercer año.

Entre los logros alcanzados durante los dos primeros años se destaca el desarrollo de una aplicación móvil de difusión denominada Facultad de Ingeniería Móvil, con el objetivo de crear una comunicación dinámica con los diferentes actores que transitan el ámbito universitario. La aplicación se encuentra conformada por varios módulos entre los que se destacan: noticias de los sucesos que acontecen en el ámbito universitario, agenda de contactos, galería de imágenes y videos, entre otras opciones.

Facultad de Ingeniería móvil está concebido para ser una ayuda al proceso comunicacional existente actualmente en nuestra facultad. El objetivo, hacer de la institución un ambiente educativo ágil, que cumpla con las características básica del aprendizaje móvil: ubicuo, flexible, portable, inmediato, motivante, accesible, activo, con conectividad permanente [15].

Con ello en mente, y a partir de diferentes prioridades que se evaluarán, se decide implementar aplicaciones móviles de interés a nuestro ámbito educativo, dando inicio a las tareas enmarcadas para el tercer año del proyecto.

Al concluir el proyecto se espera contribuir en la instalación, personalización y/o desarrollo de diversas aplicaciones móviles que se distribuirán en forma libre y gratuita entre los diferentes actores que conforman el ámbito educativo universitario.

Formación de Recursos Humanos

Director de Proyecto
Co-Director de Proyecto
6 Investigadores

Referencias

- [1] Weiser M. (1991), *The Computer for the Twenty-First Century*.
- [2] Vedar E, Evans W, Griswold, W, (2009) *Ubibot - Prototyping Infrastructure for Mobile Context-Aware Computing*. Ubicomp 2009.
- [3] Carmona M., González S., Castro Ruiz, *Innovación Tecnológica en Comunicaciones Móviles Desarrollada Con Software Libre: Campus Ubicuo* Novática: Revista de la Asociación de Técnicos de Informática, ISSN 0211-2124, N°. 190, 2007
- [4] de la Riva D. (2007) “Aplicaciones Web para celulares”, Tesis de Grado. Facultad de Informática. Universidad Nacional de La Plata.
- [5] Weiser, M. (1993), *Ubiquitous Computing*, IEEE Computer "Hot Topics", October 1993, <http://www.ubiq.com/hypertext/weiser/UbiCompHotTopics.html>
- [6] Nielsen J. (2000), “*Designing Web Usability: The Practice of Simplicity*”, New Riders Publishing, Indianapolis, ISBN 1-56205-810-X
- [7] W3C, Extensible Markup Language (XML), <http://www.w3.org/XML/>
- [8] W3C, W3C Semantic Web Activity, <http://www.w3.org/2001/sw/>
- [9] Girouard J, Horn H, Solovey S, Zibgelbaum, (2008), *Reality Based Interaction, a framework for post-WINP Interfaces*.
- [10] Krumm J., (2009), *Creating a New Book of Tutorials: Ubiquitous Computing Fundamentals*, Microsoft Research Redmond, Ubicomp 2009.
- [11] Bravo C., Redondo M., Ortega M., Bravo J., *Evolución de un Entorno Colaborativo de Enseñanza Basado en Escritorio hacia la Computación Ubicua*, (2002) COLINE'02 Investigación en Entornos de Interacción Colectiva, Workshop de Investigación sobre nuevos paradigmas de interacción en entornos colaborativos aplicados a la gestión y difusión del Patrimonio cultural. Granada, 11 y 12 de Noviembre del 2002.
- [12] Weiser, M. (1998) *The future of Ubiquitous Computing on Campus*. Communications of ACM, 41-1, January 1998, 41-42.
- [13] Malani R., Griswold W, Simon B, (2009) *Public Digital Note-Taking in Lectures*. Ubicomp 2009.
- [14] Richards M, Woodthorpe J, (2009), *Introducing TU100 "My Digital Life": Ubiquitous computing in a distance learning environment*. Ubicomp 2009.
- [15] M2learn, (2015). “M2learn, Framework for the development of mobile context-aware learning applications”. <https://code.google.com/p/m2learn/>

El Museo de Informática de la UNPA-UARG: Organización, Acciones y Difusión

Esteban GESTO, Osiris SOFIA, Verónica HAMMAR, Karim HALLAR, Fernanda ARGÜELLES, Fernanda BACHILIERI

Unidad Académica Río Gallegos (UARG)
Universidad Nacional de la Patagonia Austral (UNPA)
imuseo@uarg.unpa.edu.ar

RESUMEN

Los avances en informática son tan vertiginosos y constantes que en muy poco tiempo el equipamiento utilizado entra en rápida obsolescencia. Existen numerosas tecnologías que han estado en uso hasta hace pocos años, pero que las nuevas generaciones no conocen e incluso no han oído nombrar (cintas, tarjetas perforadas, etc.).

Por esto, la Universidad Nacional de la Patagonia Austral creó el Museo de Informática: un espacio para salvaguardar los objetos informáticos y utilizarlos como recursos educativos para los alumnos de la UARG -y también para agentes externos- con el fin de conocer la historia de la computación y la digitalización de la información para comprender muchos de los cambios sociales de hoy.

Palabras clave: *Museo, Informática, Historia*

CONTEXTO

Este trabajo ha sido realizado en el marco del Proyecto de Investigación PI 29/A368 “El Museo de Informática de la UNPA-UARG: organización, acciones y difusión” de la Universidad Nacional de la Patagonia Austral (2016-2018).

El Museo de Informática se constituyó en el año 2010 como un Proyecto de Extensión permanente, dentro del Ámbito de la Unidad Académica Río Gallegos de la Universidad Nacional de la Patagonia Austral. Durante los años 2014 a 2016, los miembros de este equipo de Investigación participaron en el Proyecto de Investigación PI 29/A323 “Pensando el Museo de Informática de la UNPA-UARG: Plan Museológico y Sistema Documental”.

1. INTRODUCCIÓN

Los museos, tal como los conocemos hoy, poseen una larga historia de cambios en su concepción. Aunque las definiciones sobre qué es un museo fueron modificándose a lo largo del tiempo, el interés principal por agrupar y conservar ciertos objetos ha continuado, manifestándose de distintas maneras en todo el mundo, de acuerdo a su contexto socio-cultural.

En sus inicios, los museos eran espacios atestados de objetos sin una clasificación definida, verdaderos depósitos de objetos raros. Con el crecimiento de las ciudades y el desarrollo de la actividad profesional, los museos fueron modificándose gradualmente. La práctica museológica fue abarcando nuevos campos de estudios y considerando aspectos antes desestimados.

Así comenzó un proceso continuo de revisión y búsqueda para definir qué debe hacerse en un museo y qué acciones deben implementarse para lograrlo. Mientras se creaban nuevos museos, y los ya existentes se actualizaban, crecía en los profesionales se preguntaban cómo debía organizarse un museo de acuerdo a sus funciones y tareas específicas, ya que el campo de estudio se ampliaba día a día con nuevos desafíos. Pero entonces, ¿a qué se denomina museo? ¿Cuáles son las funciones principales de un museo? Según la definición dada por el ICOM en 1974 y ratificada luego en 1989, un museo es una *“Institución permanente, sin fines lucrativos, al servicio de la sociedad que adquiere, conserva, comunica y presenta con fines de estudio, educación y deleite, testimonios materiales del hombre y su medio”*.

De esta manera, al definir las características de los museos, se observa que existen otras instituciones que sin ser museos responden también a dicha definición. Estas instituciones o centros son denominadas instituciones museológicas e incluyen a las bibliotecas, los archivos, los jardines botánicos, los parques naturales, incluso los zoológicos.

Los museos son espacios donde se aprende y donde se transmiten experiencias culturales de generación en generación. Son instituciones que aparecen en la sociedad para darle sentido a la vida. Por medio de la interpretación de los objetos se investiga el mundo, tratando de lograr un abordaje integral de los diferentes temas.

Pero, teniendo en cuenta la gran diversidad de museos existentes, y las distintas disciplinas que intervienen en el proceso, ¿cómo se logra organizar sus tareas? Existen enfoques

diferentes sobre este asunto, de acuerdo a la ideología de cada país. Pero hay algo en que lo diferentes autores acuerdan: la necesidad de definir ciertas funciones principales que guíen el quehacer museológico.

Según el ICOM, un museo debe cumplir, además de sus funciones de institución administrativa, con tres funciones básicas: la *preservación*, que incluye la documentación, la restauración, la conservación y el almacenaje; la *investigación*, que se refiere al estudio e interpretación de la información de la colección y por último la *comunicación*, que es la encargada de transmitir los conocimientos hacia los visitantes, por medio de las exhibiciones, los folletos informativos, la cartelería y los programas educativos, entre otros.

Por esto, la tarea de preservar el patrimonio desde la universidad significa un gran compromiso, ya que se debe tener en cuenta no sólo la forma de conservar y estudiar los objetos, sino también la manera en que éstos son comunicados o transmitidos y para ello se requiere de preparación profesional, planificación integral y de un arduo trabajo en conjunto e interdisciplinar.

Los museos surgidos desde una universidad son recursos indiscutidos de investigación y generación de contenidos educativos, importantes tanto para complementar la formación académica de los alumnos, como para constituirse como espacios de educación no formal en los cuales se refuerza y extiende el vínculo de la universidad con la comunidad.

2. LINEAS DE INVESTIGACION

El museo cuenta con una colección de aproximadamente 800 objetos (entre

maquinarias, accesorios y documentos) y espera en el futuro incrementar el acervo en función de los ofrecimientos de donación y las necesidades que se detecten al diseñar el guión de la exhibición.

Actualmente el museo se encuentra en la tarea de elaborar su guión museológico teniendo en cuenta la evolución de la informática y sus repercusiones sociales en Santa Cruz. Luego se espera continuar con la elaboración de un guión museográfico que organice la exposición respondiendo a las necesidades comunicacionales de los objetivos del proyecto. Se espera también la elaboración de diferentes recursos comunicacionales (gráficos y audiovisuales) a fin de darle forma y sentido al Museo de Informática como un proyecto permanente de investigación, conservación y extensión universitaria en la provincia de Santa Cruz.

3. OBJETIVOS PROPUESTOS

Objetivo Principal

Organizar las diferentes actividades del Museo de Informática de la UNPA-UARG para consolidar su funcionamiento como una institución donde se investiga, conserva y difunde la historia de la informática en la región.

Objetivos Específicos

- Proponer actividades educativas y acciones de difusión del Museo de Informática dirigidas a los niveles universitario, terciario y secundario, considerando las necesidades de los alumnos y los contenidos a desarrollar.

- Promover el correcto manejo de la colección del museo, atendiendo las recomendaciones internacionales para su conservación.
- Proponer lineamientos para el diseño de la exposición, considerando las características de los objetos, el público destinatario y los objetivos del museo.
- Establecer relaciones significativas entre los objetos del Museo de Informática y la sociedad actual para conocer el valor de la colección y su importancia a nivel local.
- Utilizar el lenguaje audiovisual para conservar y difundir los testimonios orales relacionados a la evolución de la informática y a sus usos en la sociedad.
- Continuar con el registro de los objetos en el inventario digital del museo y en la base de datos CONAR de la Secretaría de Cultura de la Nación.

4. MUESTRA ESPECIAL HOME COMPUTERS

Durante 2016 se dio inicio a una muestra especial dedicada a las *Home Computers*, cuyo auge se dio durante la década de los años '80 y principios de los '90, proponiendo una serie de actividades que tuvieron como objetivo la puesta en valor y la difusión de los objetos característicos de esa época con los que cuenta el Museo de Informática.

De esta manera, se realizaron durante el pasado año diversas actividades con alumnos de las carreras de Sistemas de la UNPA-UARG, a través de las cuales se acondicionaron, pusieron en marcha y utilizaron distintas *Home Computers*, culminando con una exhibición interactiva que inició durante las XIII Jornadas de Informática de la UNPA-UARG, y un campeonato de "Juegos Retro" organizado en

colaboración con los alumnos, en el cual los asistentes pudieron disfrutar de equipos informáticos y consolas de juegos icónicas en funcionamiento.

Esta actividad tuvo gran repercusión, tanto en la comunidad universitaria como en el público en general, logrando la participación de gran cantidad de asistentes.

Para este 2017 se prevén en este marco, otras actividades como son un taller/concurso de Retroprogramación y un taller de emulación de Sistemas Operativos que concluirá con la fabricación por parte de los alumnos de un *Arcade* que funcionará con emuladores de consolas de juegos antiguas.

5. FORMACIÓN DE RECURSOS HUMANOS

La línea de investigación asociada al proyecto tiene un carácter interdisciplinario, ya que es necesario el abordaje de la temática desde distintos puntos de vista: historia, comunicación, informática, medios audiovisuales. Es por ello que el grupo de trabajo está compuesto por personal del área de informática, comunicación social e historia. Durante 2016 se contó además con la participación de una alumna becaria del área de comunicación social. Finalmente el grupo está integrado por dos asistentes técnicos que colaboran en la elaboración de materiales gráficos y audiovisuales. Cabe destacar que una de las integrantes del proyecto posee el título de museóloga.

6. BIBLIOGRAFÍA

- AGUIRRE, Jorge y CARNOTA, Raúl Comp. (2009). Historia de la Informática en Latinoamérica y el Caribe: Investigaciones y testimonios. Río Cuarto. Ed. Universidad Nacional de Río Cuarto.
- AGUIRRE, Jorge, CARNOTA, Raúl, ROJO, Guillermo (2012). Línea de investigación en Historia de la Informática, el proyecto SAMCA-Salvando la Memoria de la Computación en Argentina. Actas del Workshop de Investigadores en Ciencias de la Computación, Córdoba, Argentina.
- AVENDAÑO, R y HAMMAR, V. (2015). El registro de la historia oral para el análisis de los procesos de cambio en el uso de tecnologías dentro de las áreas de la administración pública provincial y de la universidad en Río Gallegos, provincia de Santa Cruz. Río Gallegos. Ed. Universidad Nacional de la Patagonia Austral.
- AVENDAÑO, R. y HAMMAR, V. (2016). Puesta en valor de la colección del Museo de Informática de la UNPA-UARG por medio del análisis de la historia oral y su relación con la sociedad actual. Informes Científicos – Técnico UNPA. Vol. 8, núm. 2, 153-169.
- CASTILLOR., J. (2007). El futuro del patrimonio histórico: la patrimonialización del hombre. Revista electrónica de patrimonio histórico e-RPH, (1), 4-28.
- SOFIA, Osiris; HALLAR, Karim; HAMMAR, Victoria; FERNANDEZ, Andrea. La Comunicación Audiovisual, un recurso para la difusión del Museo de Informática de la UNPA-UARG. Caleta Olivia: UNPA (2016), 172-173. ISBN 978-987-3714-37-5
- SOFIA, Osiris; HALLAR, Karim; HAMMAR, Victoria. Un Museo de Informática en la Patagonia Austral.

- AGUIRRE, Jorge y CARNOTA, Raúl Comp. (2009). Historia de la Informática

Memorias del III Simposio de Historia de la Informática de América Latina y el Caribe (SHIALC 2014). 60-64.

El Uso de la Tecnología para la Inclusión de los Disminuidos Visuales en las Aulas

Barrios, Teresita Haydeé barriosth@gmail.com UTN FRRE GIESIN
 Marín, María Bianca mbiancamarin@yahoo.com.ar UTN FRRE GIESIN
 Torrente, Natalia nataliamtorrente@gmail.com UTN FRRE GIESIN

RESUMEN

En la actualidad la inclusión presenta nuevos retos, de cara a integrar a las instituciones educativas a personas con capacidades diferentes. Nuestra experiencia en proyectos anteriores con TICs nos dicta que las potencialidades de las herramientas virtuales nos brindarán una perspectiva interesante para facilitar la inclusión.

Es por ello que se ha encarado una línea de investigación para abordar esta problemática, enfocadas principalmente en personas con algún tipo de disminución visual. En el estudio se propone un diagnóstico de las necesidades especiales de estos alumnos, específicamente en la carrera de Ingeniería en Sistemas; como también de los docentes que los reciben. Además, se analizarán las herramientas disponibles que se adapten mejor a las estrategias a desarrollar para su inclusión. Se adoptarán dichas herramientas según los contenidos seleccionados para la muestra, se capacitará a los docentes involucrados y se hará seguimiento y evaluación de su aplicación. Si los resultados obtenidos son positivos, se realizará la transferencia a las otras carreras.

CONTEXTO

El artículo es una de las presentaciones que realiza el Grupo de Investigación sobre Ingeniería (GIESIN), de la Universidad Tecnológica Nacional, Facultad Regional Resistencia (UTN FRRE), para el proyecto denominado “El uso de las Tics para favorecer la inclusión, especialmente a los disminuidos

visuales en las carreras de Ingeniería”, que se inicia durante 2017.

1. INTRODUCCIÓN

Hace unas décadas el tema de la Integración Educativa solo estaba en boca de algunos profesionales de la educación y la psicología, que se empezaban a cuestionar acerca de los alcances y los límites de la educación especial y la común, en un intento de articular ambas ofertas generando, la más de las veces, resistencias importantes en ambas orillas.

El panorama actual, aún con sus debilidades y obstinaciones, dista totalmente de aquel. La idea de una escuela inclusiva ha logrado salir del restringido ámbito profesional al que se circunscribía para ubicarse en espacios que superan ese coto.

La nueva corriente supone reconocer el derecho de cada sujeto a satisfacer necesidades educativas, lo que impone normalizar las condiciones de vida de las personas con alguna minusvalía. No se trata de “normalizar” la persona con discapacidad sino de generar un entorno menos restrictivo y más potenciador de su desarrollo.

Hoy en día, la inclusión educativa de personas en situación de discapacidad es una obligación de las instituciones educativas, situación que ha sido liderada por organismos multilaterales como la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO, y cuyos principios se han expresado a través de diversas declaraciones, conferencias e informes (Parra, Pasuy & Flórez, 2012). El objetivo de éstos procesos de inclusión es la mejora de la calidad de vida de

las personas con discapacidad, favoreciendo aquellos factores personales y ambientales que impacten positivamente su nivel de vida, así como las relaciones laborales y familiares de éste grupo poblacional (Henaó & Gil, 2009). Lo anterior implica cambiar la visión asistencialista con que la sociedad usualmente ha tratado a los discapacitados, limitando sus posibilidades, por otra en la cual ellos puedan asumir su propio destino con autonomía (Piccolo & Mendes, 2013). Es así como los sistemas educativos en varios países de Latinoamérica buscan generar propuestas educativas incluyentes, que garanticen procesos efectivos de aprendizaje, acceso, permanencia, promoción y evaluación para las personas con Necesidades Educativas Especiales, NEE (Fabela & Robles, 2013).

En nuestro país, el sistema de educación superior, se encuentra regido bajo la ley N° 24.521, “LEY DE EDUCACION SUPERIOR”, que nuclea a las universidades e institutos universitarios, estatales o privados autorizados y los institutos de educación superior de jurisdicción nacional, provincial o de la Ciudad Autónoma de Buenos Aires, de gestión estatal o privada. Una de las obligaciones de estas instituciones, siendo la responsabilidad principal e indelegable del Estado nacional, las provincias y la Ciudad Autónoma de Buenos Aires, sobre la educación superior, implica: *“Establecer las medidas necesarias para equiparar las oportunidades y posibilidades de las personas con discapacidades permanentes o temporarias;”* (Art. 2 – inc d)

En el marco de esta tendencia y necesidad a nivel tanto nacional como internacional, es que el Grupo de Investigación Sobre Ingeniería (GIESIN) de la Universidad Tecnológica Nacional, Facultad Regional Resistencia (UTN FRRE), se encuentra iniciando un

proyecto de investigación sobre herramientas tecnológicas que puedan ayudar a personas con determinadas discapacidades a integrarse en las aulas, específicamente en la carrera de ingeniería en sistemas de información. En este proyecto de investigación se decidió abordar específicamente a los disminuidos visuales, quienes requieren un apoyo y soporte extra para poder acceder al material que ofrece la Universidad, adaptado a sus necesidades específicas.

Avanzando en este sentido, se reconoce la potencialidad de las herramientas virtuales, que brindan una perspectiva interesante para facilitar la inclusión de algunos estudiantes con capacidades diferentes, especialmente los disminuidos visuales. En este punto, es necesario hacer foco en la accesibilidad de las herramientas virtuales. La Accesibilidad es el grado en el que todas las personas pueden utilizar un objeto, visitar un lugar o acceder a un servicio, independientemente de sus capacidades técnicas, cognitivas o físicas. Es indispensable e imprescindible, ya que se trata de una condición necesaria para la participación de todas las personas independientemente de las posibles limitaciones funcionales que puedan tener. Impone el análisis y la debida consideración de toda barrera que impida o condicione la accesibilidad.

Para promover la accesibilidad se implementan facilitaciones y ayudas técnicas que posibilitan salvar los obstáculos o límites de accesibilidad al entorno, consiguiendo que la persona con discapacidad realice la misma acción que pudiera llevar a cabo una persona sin ningún tipo de discapacidad. Las más de las veces no son las barreras arquitectónicas o de comunicación las que condicionan la inclusión educativa de un alumno con discapacidad a la escuela común sino las barreras ideológicas y

pedagógicas con las que se maneja la comunidad.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La UTN FRRE es una facultad que tiene una amplia oferta académica, formada por carreras de grado, terciarios, cursos y posgrados. Es una necesidad creciente el hecho de adaptar la infraestructura técnica y educativa a las necesidades de la sociedad educativa. Es por ello que se iniciado un proceso de adaptación edilicia de la facultad que incluye, por ejemplo, incorporación de ascensores, aulas específicas acondicionados para hipoacúsicos, baños especiales, entre otros.

En este marco, no se puede descuidar la oferta académica y su preparación educativa (materiales y recursos, entre otros), que también debe adaptarse para hacer frente a esta necesidad. Es por ello que deben adecuarse los materiales de estudio, las herramientas tecnológicas, campus virtuales, y demás herramientas para que cualquier persona pueda acceder a ellos.

Para llevar adelante el proyecto de investigación “El uso de las Tics para favorecer la inclusión, especialmente a los disminuidos visuales en las carreras de Ingeniería.” se ha seleccionado la carrera de ingeniería en sistemas, con intención luego de utilizar la experiencia generadora para ampliarla a las demás ofertas y así lograr carreras que estén disponibles “para todos”.

Dentro de las discapacidades, se ha seleccionado la discapacidad visual, debido a que han presentado casos en la facultad, y se ha necesitado buscar estrategias sobre la marcha para poder responder a dicha necesidad.

Es por ello que se han definido como objetivos específicos del mismo:

- Describir las características de las distintas capacidades de los disminuidos visuales.
- Revisar las herramientas virtuales que podrían ser de utilidad para su uso con estos estudiantes.
- Implementar las soluciones encontradas en un grupo de prueba.
- Establecer conclusiones para futuras acciones de inclusión en la FRRe y eventualmente en otras instituciones.

Para alcanzar estos objetivos, el Grupo de Investigación aborda la problemática, iniciando un proceso de relevamiento y análisis del contexto social y educativo del escenario elegido; esto involucra indagar acerca de los casos de alumnos con discapacidades visuales y su paso por esta institución, haciendo hincapié en aspectos no solamente inherentes al alumno, sino también a los docentes, no docentes con los cuales se relacionó estos alumnos, de manera de recopilar todas las visiones y aspectos de cada uno de los casos.

Además, en el desarrollo de este proyecto se incluyen diferentes técnicas para analizar y evaluar las herramientas tecnológicas disponibles para dotar a la institución de recursos para proporcionar las mismas condiciones a estudiantes con discapacidades visuales. Sobre este punto el grupo ha avanzado sobre un primer borrador de algunas TICs que serán desarrolladas más adelante.

3. HERRAMIENTAS RELEVADAS

Sin dudas que la tecnología es una estrategia adecuada para hacer frente a esta estrategia de adaptación de herramientas educativas. Buscando la accesibilidad de las propuestas, es que se han investigado herramientas como: traductores braille, impresoras braille, ampliadores de texto, lectores de pantalla, entre otras.

De una investigación preliminar surge conveniente aclarar estos conceptos:

Impresoras braille: son dispositivos electrónicos que permiten imprimir textos e imágenes simples empleando puntos percutidos en papel y otros soportes parecidos. No difiere en aspecto y tecnología de interconexión que una impresora convencional; permiten imprimir braille convencional de 6 puntos, y algunas de ellas también permiten el de 8 puntos.

Al igual que en las impresoras de tinta se pueden realizar dibujos simples con los caracteres, en las impresoras braille podemos emplear los puntos para realizar dibujos en el papel de manera que la persona ciega pueda sentirlos al tocarlos.

La principal diferencia está en el mecanismo de impresión que en este caso se utilizan percutores. Los percutores realizan la misma función que un punzón para escribir braille. Estos son lanzados contra el papel con la fuerza suficiente para que lo marquen, sobresaliendo los puntos por la cara contraria a la percutida. La fuerza que se aplica al papel debe ser suficiente para que la persona ciega pueda sentir el punto, pero no puede ser excesiva para evitar perforar el papel o el desgaste prematuro de los puntos.

Si bien se encuentran disponibles en el mercado nacional, no así en el mercado local ni a precio de una impresora convencional.

Traductores de braille: traducen un documento electrónico para ser impreso por una impresora braille que imprime en relieve.

Existen en la web varias páginas que permiten realizar esta función de traductor, permitiendo hacerlo en varios idiomas y admitiendo diferentes formatos de entrada: HTML, XHTML, OOXML, ODF, todos los documentos de Microsoft Office, PDF, EPUB y archivos de texto plano (TXT).

Ampliadores de Texto: son programas informáticos que agrandan las imágenes en las pantallas de ordenador de modo que las personas con dificultades de visión puedan ver al tamaño deseado lo que se presenta, textos, imágenes, videos, etc.

Los hay integrados en los sistemas operativos, como el Ampliador Lupa de Microsoft, el Zoom de Apple en el IOS o el KMagnifier de Ubuntu en Linux.

En general, permiten: ampliar una porción de la pantalla, ampliar la pantalla entera, cambiar la combinación de colores, cambiar el contraste, seguir el foco, usarse mediante ratón y atajos de teclado

Los más populares:

- SuperNova Magnifier de Dolphin
- ZoomText de Aisquared

Los gratuitos:

- OneLoupe, que no requiere instalación, de Softwareok
- Zoom Lens de Eli Fulkerson

Lectores de pantalla: son un software que permite la utilización del sistema operativo y las distintas aplicaciones mediante el empleo de un sintetizador de voz que "lee y explica" lo que se visualiza en la pantalla.

Entre las herramientas disponibles tenemos las que son gratuitas:

- NonVisual Desktop Access
- Lector de pantalla Orca
- Sistema de Acceso ToGo

Y los lectores comerciales:

- Jaws
- SuperNova Screen Reader
- Windows Eyes

Sobre estas herramientas se elaborará una comparativa para determinar ventajas y desventajas de cada una de ellas, a fin de determinar un conjunto mínimo para implementar durante el proyecto

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación GIESIN es un Grupo UTN que posee varias líneas de investigación y proyectos vigentes. El proyecto “El uso de las Tics para favorecer la inclusión, especialmente a los disminuidos visuales en las carreras de Ingeniería”, es un proyecto dirigido por la Mgt. María del Carmen Maurel y co-dirigido por el Ing. Fernando Soria. Cuenta además con el asesoramiento científico externo de la directora del Grupo UTN Mgt. Nidia Dalfaro. Como integrantes del equipo además se cuenta con la Ing. Teresita Barrios y la Ing. Maria Bianca Marín, y como becarias las Ing. Sabrina Sotomayor y la AUS Natalia Torrente.

5. BIBLIOGRAFIA

AQUINO Zúñiga, Silvia Patricia; GARCÍA MARTÍNEZ, Verónica; IZQUIERDO, Jesús. (2012) “La inclusión educativa de ciegos y baja visión en el nivel superior. Un estudio de caso”; disponible en:
http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2012000200007
 BORSANI, María José (2016); “De la integración a la inclusión educativa”; Disponible en:
<http://elcisne.org/2016/07/27/de-la-integracion-a-la-inclusion-educativa/>
 FABELA, M. A. ROBLES, L. A. (2013). Educación inclusiva y preparación docente: percepciones y preocupaciones de docentes en el aula de educación regular. Simposio internacional: Aprender a ser docente en un mundo en cambio, Barcelona, España, 21- 22
 Ampliador de pantalla: implicaciones para la docencia. Disponible en:
<https://uayuda.ua.es/profesores-y-pas/adaptaciones-para-las-nee/ampliador-de-pantalla-implicaciones-para-la-docencia/>

de noviembre, 2013. Disponible en:
<http://som.esbrina.eu/aprender/>
 FERNANDEZ MORALES, F.; DUARTE, J.; GUTIÉRREZ, J.; (2015) “Estrategia pedagógica para la formación de ingenieros con discapacidad visual”. Disponible en:
http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1852-45082015000100003
 HENAO, C. P. & GIL, L. M. (2009). Calidad De Vida Y Situación De Discapacidad. Revista Hacia la Promoción de la Salud, Disponible en:
<http://148.215.2.10/articulo.oa?id=309126692005>.
 PARRA, A. M., PASUY, L. & FLÓREZ, J. A. (2012). Atención a estudiantes con necesidades educativas diversas: Clave para instituciones de educación superior inclusiva. Plumilla educativa
 PICCOLO, G. M. & MENDES, E. (2013). Contribuições a um pensar sociológico sobre a deficiência. Educação&Sociedade, Disponible en:
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-73302013000200008&lng=en&tlng=pt.10.1590/S0101-73302013000200008
 RIVERA, Catalina (2016); “La Accesibilidad: un derecho para todos y todas”; La izquierda diaria; Disponible en:
<http://www.laizquierdadiario.cl/La-Accesibilidad-un-derecho-ciudadano-para-tod-s>
 Ley N° 24.521 “LEY DE EDUCACION SUPERIOR”
 Ley N° 26.206 “LEY DE EDUCACION NACIONAL”
 Pagina Web:
<http://www.crmfalbacete.org/recursosbajocoste/facillectura> (consultada el 01/12/2016)

Elicitación de Requerimientos, Centrada en el Usuario, para el Desarrollo de un Serious Game

Adolfo Tomás Spinelli¹, Stella Maris Massa¹

¹Facultad de Ingeniería/ Universidad Nacional de Mar del Plata/Argentina
(7600) Av. Juan B. Justo 4302, +54-223-481660
spinelliadolfo@gmail.com, smassa4@gmail.com

Resumen

Los Serious Games (SG) contribuyen con la capacitación y entrenamiento en múltiples áreas tales como la defensa, la salud, la educación ó la política. Para que un SG's cumpla su objetivo, su construcción debe partir de una buena especificación.

Los requerimientos de SG involucran; objetivos educativos, objetivos del juego, requerimientos específicos del software y elementos de medición que nos permitan evaluar su eficacia.

El presente trabajo describe las consideraciones tomadas en cuenta para la definición de un Proceso de Elicitación para SG's, basado en la Modelo de Proceso MPOBA.

En dicho proceso se procura lograr una Especificación de Requerimientos precisa que exprese el equilibrio necesario entre las necesidades pedagógicas, lúdicas y del software. Para lograrlo el mismo pone énfasis en el aspecto multidisciplinario y el enfoque del proceso de Elicitación, basado en favorecer los aspectos positivos para el proceso de aprendizaje.

Actualmente, el Proceso de Elicitación se está testeando y validando a través de su aplicación en el desarrollo de un SG sobre el uso consciente y eficiente de la energía

Palabras clave: Serious game, Elicitación, Requerimientos.

Contexto

Este artículo presenta la investigación realizada en el marco de la tesis: "Elicitación de requerimientos centrada en el usuario para el desarrollo de Serious Games" de la Maestría en Ingeniería de Software de la Facultad de Informática de la Universidad de la Plata.

Se inserta además en el proyecto: "Tecnología e Innovación en Ambientes de Aprendizaje: Desarrollo y Gestión" (2016/2017) del Grupo de Investigación en Tecnologías Interactivas (GTI) de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Introducción

Autores como Abt [1], Sawyer [2] y Zyda [3], coinciden en definir un SG como un videojuego que además de entretener, pretende incorporar un nuevo conocimiento ó modificar alguna conducta en el jugador.

Rocha [4] afirma que el éxito de estos juegos depende de una Especificación de Requerimientos que incluya: los aspectos pedagógicos (contenido y tareas), la mecánica y trama del juego, los requerimientos del software, así como los mecanismos que nos permitan evaluar el aprendizaje, antes, durante y después del juego, así como su eficacia pedagógica.

Asi mismo Rocha [4] destaca que ningún SG será exitoso, ni cumplirá su objetivo educativo, sino consigue atrapar la atención del jugador y lograrlo implica le existencia de

un equilibrio entre los objetivos pedagógicos, y el estado de flow .

El "estado flow" según Csikszentmihalyi [5] es la sensación que experimenta el jugador en el acto de jugar, cuando es capaz de abstraerse del mundo real y considerar ciertos los estímulos que recibe. Para el jugador dichas sensaciones agradables o no, divertidas o no, son reales mientras dura el juego.

Algunos autores hacen referencia a los términos de inmersión y jugabilidad como sinónimos del "estado flow". Sin embargo para Murray [6] inmersión es la sensación de vivir (sensorialmente) la narrativa que propone el juego. En tanto que González Sánchez [7] denomina "jugabilidad" al conjunto de propiedades que describen la experiencia del jugador ante un producto cuyo objetivo principal es divertir y entretener. Basándonos en esta definición podemos asimilar a la jugabilidad como una extensión de la experiencia del usuario.

El "estado flow" implica un estado de inmersión y paralelamente un grado de jugabilidad que le de soporte. Por ello no podemos considerar a dicho estado como sinónimo de los anteriores, pues son abarcados por este.

Lograr este equilibrio es una premisa fundamental en el diseño de SG, si queremos productos exitosos y de calidad, deben existir procesos de desarrollo que contemplen en las etapas de Especificación de Requerimientos y Diseño tanto el proceso de aprendizaje como el "estado flow".

Dado el rol relevante del jugador en estos productos, los mismos deben ser diseñados siguiendo los principios del Diseño Centrado en el Usuario (DCU), este consiste en realizar las tareas de diseño teniendo presente las necesidades del usuario. Podemos identificar dichas necesidades cuando estamos en condiciones de establecer la experiencia del

usuario frente al producto, lo que implica en nuestro caso tomar en cuenta la jugabilidad y el "estado flow" [8].

En este proceso de diseño, es interesante incorporar el concepto de "usabilidad pedagógica" propuesto por Massa [9] pues su definición: "La facilidad de aprendizaje, eficiencia de uso pedagógico y la satisfacción con las que las personas son capaces de realizar sus tareas gracias al uso del producto con el que está interactuando", es una forma de integrar "estado flow", experiencia del usuario y objetivos pedagógicos.

No existe evidencia de la existencia de metodologías, directrices y mejores prácticas para desarrollar SG eficaces e integrados con las actividades de aprendizaje y la evaluación de su impacto en el proceso de aprendizaje [10].

Con relación a ello, en el marco del Proyecto de investigación, se propuso Modelo de proceso MPDSG [11] (combinación del modelo Lineas de Producción de Software (LPS) [12], el Diseño Centrado en el Usuario (DCU) [13] y el modelo de Proceso para el desarrollo de Objetos de aprendizaje (MPOBA)[9].

En este modelo los expertos vinculados al desarrollo de un SG (desarrolladores, jugadores, pedagogos, expertos en contenido) deben tomar decisiones en conjunto para lograr un SG atractivo e inmersivo. El proceso es iterativo con una fase Elicitación en cada ciclo. El proceso de Elicitación es una herramienta de la Ingeniería de Requerimientos, siendo el objetivo de esta la redacción y validación de la especificación de los productos de software [14]. Por ello es conducente considerar la extensión de la Ingeniería de Requerimientos a los SG para dar soporte a su proceso creativo [15], incluyendo la experiencia del jugador [16].

La metodología propuesta para el Proceso de Elicitación es una adaptación de la usada por el Modelo de Proceso MPOBA. Consiste en un mecanismo iterativo donde en cada ciclo se utilizan técnicas de Entrevista, Analisis de Contexto, Focus Group y Brain Storming. El uso de estas prácticas permiten recabar y clasificar la información obtenida a a partir de los conocimientos de los stakeholders (docentes, expertos, gamers, sonidistas, gráficos). Cada uno de ellos aporta una parte del todo, lograr un trabajo armónico exige la existencia de un lenguaje común. Para ello la información obtenida se expresa en forma escenarios de Leite [17], utilizando en el Lexico Extendido de Lenguaje (LEL) (Leite [17]), como lenguaje común y los escenarios como medio de intercambio de ideas y validación.

El proceso de Elicitación culmina en la elaboración y entrega de un documento que constituye la especificación del producto. Este consiste en la enumeración de los requerimientos que debe cumplir, redactados claramente y sin ambigüedad. En el caso de los videojuegos este documentos es el GDD (Game Design Documento), cuyo formato presento Scott [18].

Los escenarios, descriptos en lenguaje natural constituyen la narrativa del videojuego y la base a partir de la cual se ha de redactar el GDD, una vez que superen la etapa de validación en juicio de expertos.

Líneas de Investigación

- Elicitación de requerimientos para serious games.
- Diseño centrado en el usuario.
- Usabilidad pedagógica.
- Estado flow

Resultados y Objetivos

El objetivo general del proyecto de tesis que se presenta en este artículo es: proponer un Proceso de Elicitación de Requerimientos para SG's, que tome en cuenta los tres niveles de Especificación: pedagógico, del juego y del software. Para ello se plantean los siguientes objetivos específicos:

- Determinar las tareas y secuencias que han de conformar el proceso de Elicitación de requerimientos de los SG's.
- Construir los documentos y protocolos a cumplimentar el Proceso de Elicitación por cada una de las técnicas.
- Elaborar una estrategia para la definición de requerimientos de los SG's,
- Establecer los técnicas de Validación de requerimientos.
- Desarrollar un trabajo de campo en el que se analizará la viabilidad y posibilidades del Proceso de Elicitación de requerimientos para SG's.

Actualmente el grupo de investigación ha desarrollado una metodología para el desarrollo de SG's denominada: Modelo de proceso de desarrollo para Serious Games (MPDSG). La misma se está utilizando en la construcción de los SG: "Power down the Zombies" y "Fronteras" como estudios de caso. Esta metodología es descripta por Evans et. al [11],

La etapa de Elicitación que integra dicha metodología es descripta en Spinelli [19],[20] y [21], en ella participan los diferentes stakeholders (profesores, expertos del dominio, gamers, programadores, narradores, animadores y sonidistas). Con el objeto de lograr una Especificación de Requerimientos

que armonice los aspectos pedagógicos (contenido, monitoreo y evaluación) con la necesidad de un juego atractivo.

Formación de Recursos Humanos

La formación de recursos humanos se considera de vital importancia. Integrantes del proyecto se encuentran desarrollando y dirigiendo tres tesis de posgrado de la UNLP en el marco del proyecto de investigación, correspondientes a la Maestría en Ingeniería de Software y a la Maestría en Tecnología Informática Aplicada en Educación y una del Doctorado en Educación de la Universidad Nacional de Rosario.

Referencias

- [1] Abt, C. (1970). *Serious Games*. The Viking Press. New York, EEUU.
- [2] Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9), 25-32. IEEE.
- [3] Sawyer, B. y Smith, P. (2008). Serious games taxonomy. En *Game Developers Conference (GDC 2008)*, Baltimore, USA.
- [4] Rocha V.R., Isotani S. y Biitencourt I. (2015). Análise, Projeto, Desenvolvimento e Avaliação de Jogos Sérios e Afins: uma revisão de desafios e oportunidades, en nais do IV Simpósio Brasileiro de Informática na Educação e X Conferencia Latino Americana de Objetos e Tecnologias de Aprendizagem 26(1), pp. 692. Maceió, Alagoas, Brasil
- [5] Csikszentmihalyi M. (1975). Play and intrinsic rewards, *Journal of Humanistic Psychology*. (1975), 15, pp. 41–63.
- [6] Murray, J. H. (1997). *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Simon and Schuster.
- [7] González Sánchez, J. L. (2010). Jugabilidad. Caracterización de la experiencia del jugador en videojuegos. Tesis doctoral. Universidad de Granada, España.
- [8] Gonzalez Sanchez, J. L., Padilla Zea, N. y Gutierrez, F. L. (2009). From usability to playability – Introduction to player-centered videogame development process. En *1st Interantional Conference on Human Centred Design*, pp. 65-74. Springer-Verlag London. San Diego, California EEUU.
- [9] Massa, S. M. (2013). *Objetos de Aprendizaje: Metodología de Desarrollo y Evaluación de Calidad*, (Tesis Doctoral). Facultad de Informatica, UNLP, La Plata.
- [10] Catalano, C. E., Luccini, A. M. y Mortara, M. (2014). Best Practices for an Effective Design and Evaluation of Serious Games., *International Journal of Serious game* 1(1) 2014. Disponible en <http://journal.seriousgamesociety.org>.
- [11] Evans F., Spinelli A., Zaipirain E., Masa S., Soriano F. (2016), Proceso de desarrollo de Serious Games. Diseño centrado en el usuario, jugabilidad e inmersión, 3er Congreso Argentino de Ingeniería y 9no Congreso Argentino de Enseñanza de la Ingeniería (CAEDI). (2016), Resistencia, Chaco, Argentina.
- [12] Clements, P. et. Al (2001). *Software Product Lines: Practices and Patterns*. Addison-Wesley
- [13] Granollers, T. (2004). MPIu+a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares (Tesis Doctoral). Universitat de Lleida.
- [14] Loucopoulos, P. y Karakostas, V., (1995). *System Requirements Engineering*. Mac Graw-Hill.

- [15]Calelle, D., Neufeld, E. y Schneider, K. (2005). Requirements Engineering and the Creative Process in de Video Game Industry. En Actas de la 13a Conferencia Internacional de Ingenieria de Requerimientos, pp. 240-250. IEEE. Paris, Francia.
- [16]Calelle, D., Neufel, E. y Schneider, K. (2006). Emotional Requirements in Video Games. En Actas de la 14th IEEE International Requirements Engineering Conference RE06, pp. 299-302. IEEE. Minneapolis/St. Paul, Minnessota EEUU.
- [17]Leite, J. C. S. P., Hadad, G. D. S., Doorn, J. H. y Kaplan, G. N. (2000). "A Scenario Construction Process". *Requirement Engineering*, 5 (1), 38-61. Springer.
- [18]Scott, R. (2010). Level UP - The Guide to Great Game Design. John Wiley & Sons.
- [19] Spinelli A., Massa S., Evans F., (2016). El proceso de creación de un videojuego como herramienta para la toma de decisiones en el uso eficiente de energía. II Congreso Argentino de energías sustentables (CES). (2016) Bahía Blanca, Provincia de Buenos Aires, Argentina.
- [20] Spinelli A., Massa S., Zapirain E., kühn F., Rico C., (2016). Elicitación de requerimientos para un Serious Game. II Jornada Argentina de Tecnología, Innovación y Creatividad (JATIC). (2016), Mar del Plata, Buenos Aires Argentina.
- [21] Spinelli A., Massa S., Zapirain E.,2016. La Construcción Narrativa de un Serious Game. IV Congreso de Videojuegos y Educación (CIVE). (2016) Pontevedra, España.

Elicitación y Especificación de Requerimientos en Pervasive Serious Games

Felipe Evans, Stella M. Massa

Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Argentina.

(7600) Juan B. Justo 4302 +54-0223-4816600

{fevans, smassa}@fi.mdp.edu.ar,

Resumen

Este artículo describe una investigación sobre los Pervasive Serious Games (PSG) realizada en el marco de la tesis de Maestría de Ingeniería de Software.

Los PSG son videojuegos del tipo Serious Game (SG) adaptables al contexto que rodea al usuario permitiendo que el juego intervenga por razones aleatorias, por características definidas al nivel educativo y/o de inmersión, ofreciendo otras formas no convencionales de interactuar con el jugador.

Este tipo de videojuegos se vuelven muy interesantes, ya que invitan a repensar el juego, lo que obliga también a reformular algunas metodologías para llegar a los requerimientos del videojuego.

En esta investigación nos centraremos en los procesos de Elicitación y Especificación de los PSG.

Existen escasos procesos de Elicitación sistemáticos de construcción de SG profundizándose aún más en los PSG. Esta falta suele impactar fuertemente en la inmersión o su objetivo pedagógico, siendo su desarrollo más una actividad artesanal que profesional.

El contar con un modelo para el Proceso de Elicitación de PSG contribuirá con la difusión de buenas prácticas en un sector y mejorar la calidad.

Palabras clave: Videojuegos, Pervasive Serious Games, Elicitación.

Contexto

Este artículo presenta la investigación realizada en el marco de la tesis: “Elicitación y especificación de requerimientos para un Pervasive Serious Games” de la Maestría en Ingeniería de Software de la Facultad de Informática de la Universidad de la Plata.

Se inserta además en el proyecto: “Tecnología e Innovación en Ambientes de Aprendizaje: Desarrollo y Gestión” (2016/2017) del Grupo de Investigación en Tecnologías Interactivas (GTI) de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Introducción

El juego es una de las actividades inherentes al humano y de carácter universal. Trabajos como los de Connolly [1] y Urquidi [2] aportan fuerte evidencia tanto teórica como empírica, sobre el juego como medio para adquirir habilidades.

El videojuego se puede definir como juego electrónico que se visualiza en una pantalla (RAE) o como lo definen Salen y Zimmerman [3] y Crawford [4] un videojuego es un software donde los jugadores participan en un conflicto artificial definido por reglas que transcurre dentro de una realidad simulada donde el jugador pone sus emociones en juego. Más allá de su definición, el videojuego es una de las actividades humanas con mayor evolución y aceptación de los últimos años, esto se puede visualizar en el estudio de Theesa [5] donde: el 63% de los jefes de hogar juegan videojuegos,

hay un promedio de 1.7 jugadores de videojuegos en EEUU y el promedio de jugador tiene 35 años y se le dedican 3 a más horas por semana.

Esta evolución la explica McGonigal [6] como que el mundo real no tiene grandes desafíos, placeres cuidadosamente diseñados, y los vínculos sociales que brindan los entornos virtuales.

En el caso particular de Serious Game (SG) autores como Abt [7], Zyda [8], Sawyer y Smith [9], coinciden en definir como un videojuego donde el objetivo trasciende el mero entretenimiento y busca incorporar un nuevo conocimiento o modificar alguna conducta en el jugador. Para lo cual el desarrollo de un SG debe lograr combinar que los obstáculos sean aprendizaje de una o más habilidades necesarias para la vida real y que a su vez mantenga el interés del jugador realizando la actividad como voluntaria.

Para esto Rocha [10] afirma que el éxito de los SG depende de:

- Una especificación que cubra los aspectos: Educativo, Juego, Contexto, Software
- La existencia de un ambiente de desarrollo que favorezca la integración de conocimientos multidisciplinarios.
- La elección de metodologías de calidad que contemplen el rehúso de artefactos,
- Prever la evaluación de la rutina pedagógica embebida en el producto y la validación del mismo como herramienta pedagógica.

En esa línea, Padilla Zea [11] identifica los siguientes requerimientos pedagógicos para los videojuegos con intencionalidad educativa como los SGs : contenidos educativos, monitoreo del aprendizaje, servicios al profesor, evaluación del

aprendizaje y personalización del juego con relación a los contenidos educativos. Pero por otro lado es imprescindible lograr: “juegos que sean jugados voluntariamente”. En este sentido, Connolly et. al. [1] enumeran los aspectos a tener en cuenta: placer, competencia, diversión, desafío, interacción social, fantasía, excitación, percepción social y estado flow. Por su parte Murray [12] resume estos aspectos en el concepto de inmersión que define como: “la sensación de sumergirse completamente en otra realidad [...] que acapara toda nuestra atención y aparato sensorial.”

En resumen un SG exitoso debe balancear los objetivos pedagógicos y la inmersión. En esa línea y con el foco en difuminar los límites entre lo real y lo virtual surgen tecnologías con varias denominaciones : Ubiquitous Computer, Pervasive Computer, Calm Computing, Think that Thinks and Everywhere (MIT Media Labs). Este concepto es presentado en 1987 los miembros de Xerox Palo Alto Research Center (PARC) con el nombre de Ubiquitous computer o “Calm computer”. Como menciona Weiser [13], dispersar computadoras en todas partes, en forma invisible, en toda el ambiente. La IEEE, engloba esta noción dentro de “Pervasive computer”. Se visualiza en los objetivos y alcances del Journal “IEEE Pervasive Computing”¹ que propone la exploración del rol de la computadora en el mundo real a través de visiones como Internet Over think (IOT) y Ubiquitous computer.

Por último, existen un grupo de juegos (no videojuegos) denominados Pervasive games que extienden la experiencia del juego al mundo real[14].

Conjugando las nociones descritas en los párrafos anteriores, esta investigación se utilizará el término Pervasive Serious Games (PSG).

Los PSG no sólo se adaptan al contexto que rodea al usuario, sino pueden ir más allá, permitiendo que el juego intervenga por razones aleatorias, por características definidas a nivel educativo y/o de inmersión, ofreciendo otras formas no convencionales de interactuar con el jugador.

Es entonces que los PSG invitan a repensar el juego. Ya que a diferencia de un juego tradicional, donde las reglas son inmutables durante todo el transcurso del mismo, las mismas dependen de parámetros como el estado de uno o varios jugadores, su condición social, espacial y/o temporal. Según Montola [15] los PSG permiten visualizar las barreras entre lo lúdico y la vida real, llevando el placer del juego a la vida ordinaria.

Este replanteo del videojuego tradicional hace necesaria la búsqueda de metodologías para desarrollar y validar procesos de elicitación de un PSG siguiendo los principios del Diseño centrado en el usuario (DCU).

Ampatzoglou y Stamelos [16] destacan que existen pocos trabajos referidos a sistematizar los procesos propios de la Ingeniería de requerimientos para videojuegos. Esto se repite en la construcción de SG, y se profundiza más en los PSG, pues no se evidencia la existencia de metodologías, directrices y mejores prácticas para desarrollar productos eficaces e integrados en las rutinas de aprendizaje y procesos formativos, así como la evaluación de su impacto [17].

Manrubia Pereira [18] divide el proceso productivo de un videojuego en tres fases: pre-producción, producción y post-producción, la elicitación y especificación

del videojuego conforman la pre-producción. Se buscará que el producto de la elicitación tenga una forma parecida a la que toman los videojuegos comerciales que es el documento llamado GDD (Game Design Document).

Este documento contiene información detallada sobre los objetos, reglas, entornos, contexto, estructura, narrativa, condiciones de victoria/derrota y la estética del juego [19]. La construcción del GDD es fruto de un trabajo interdisciplinario donde participan todos los stakeholders interesados en el diseño creativo del videojuego, el documento resultante es un entregable evolutivo que se actualizará a medida que el desarrollo avance [20].

El GDD de un PSG será de apariencia y contenido similar al descrito, pues esencialmente es un videojuego, sin embargo existen varios aspectos que se deben tenerse en cuenta que el contenido del GDD debe validarse respecto de los objetivos pedagógicos, el dominio de inmersión, las habilidades a adquirir y los aspectos que hacen pervasivo un juego.

Finalmente, contar con un modelo para el Proceso de Elicitación de PSG contribuirá con la difusión de buenas prácticas en un sector en expansión y permitirá la apertura de líneas de investigación, en procura de técnicas y procesos de desarrollo en dominios donde el trabajo multidisciplinario es fundamental.

Líneas de Investigación

- Pervasive Serious Games
- Elicitación de requerimientos para PSG
- Diseño Centrado en el Usuario
- Validación de la especificación de los requerimientos de PSG

Resultados y Objetivos

El objetivo general del proyecto de tesis que se presenta en este artículo es: proponer un Modelo de Elicitación de Requerimientos para PSG's, que tome en cuenta los todos niveles de Especificación necesarios, prestando atención sobre todo a lo pedagógico, del juego y los atributos que hacen que el juego sea pervasivo. Para ello se plantean los siguientes objetivos específicos:

- Determinar las tareas y secuencias que han de conformar el proceso de Elicitación de requerimientos de los PSG's.
- Construir los documentos y protocolos a cumplimentar el Proceso de Elicitación
- Elaborar una estrategia para la definición y validación de requerimientos de los PSG's,
- Desarrollar un trabajo de campo en el que se analizará la viabilidad y posibilidades del modelo planteado.

Actualmente el grupo de investigación ha desarrollado una metodología para el desarrollo de SG's denominada: Modelo de proceso de desarrollo para Serious Games (MPDSG). La misma se está utilizando en la construcción de los SG: "Power down the Zombies" y "Fronteras" como estudios de caso. Esta metodología es descrita por Evans et. al [21], y La etapa de Elicitación que integra dicha metodología es descrita en Spinelli et. al. [22] y [23]. Por lo que se utilizará estas experiencias.

Formación de recursos humanos

Varios integrantes del Grupo GTI acreditan estudios de posgrado: una tesis doctoral: Doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP [24] y dos tesis de maestría en Tecnología Informática Aplicada en Educación [25], [26]. Integrantes del proyecto se encuentran desarrollando dos tesis de posgrado de la

UNLP correspondientes a la Maestría en Ingeniería de Software y una del Doctorado en Educación de la Universidad Nacional de Rosario.

Referencias Bibliográficas

- [1] Connolly, T. M., Boyle, E. A., Mac Arthur E., Hainey T. y Boyle J. M. (2012). A systematic literature review of empirical evidence on computer games an serious games. *Journal Computers & Education*, 52(2), 661-686. Elsevier. Filadelfia. Pensilvania, EEUU
- [2] Urquidi, M. & Tamarit Aznar, C. (2015). Juegos serios como instrumento facilitador del aprendizaje: evidencia empírica. *Revista Opción*, 31(3), 1201 - 1220. Universidad de Zulia, Venezuela.
- [3] Salen, K. y Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. The MIT Press.
- [4] Crawford, C. (2003). *Chris Crawford on game design* (pag. 31). New Riders, Indianapolis EEUU.
- [5] Theesa (2016), Essential facts about the computer and video game industry: 2016 sales, demographic and usage data <http://essentialfacts.theesa.com/Essential-Facts-2016.pdf>
- [6] Jane McGonigal (2011), ¿Por qué los videojuegos pueden mejorar tu vida y cambiar el mundo?. pp. 21
- [7] Abt, C. (1970). *Serious games*. The Viking Press. New York, EEUU.
- [8] Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9), 25-32. IEEE.
- [9] Sawyer, B. y Smith, P. (2008). *Serious games taxonomy*. Serious Game Summit 2008. San Francisco, USA.
- [10] Rocha V.R., Isotani S. y Biitencourt I. (2015). Análise, Projeto, Desenvolvimento e Avaliação de Jogos Sérios e Afins: uma revisão de desafios e oportunidades. IV Simpósio Brasileiro de Informática na Educação e X Conferencia Latino Americana de Objetos e Tecnologias de Aprendizagem , 26(1), pp. 692. Maceió, Alagoas, Brasil.
- [11] Padilla Zea N. (2011). Metodología para el diseño de videojuegos educativos sobre una arquitectura para el análisis del aprendizaje colaborativo. Tesis Doctoral. Departamento de Lenguajes y Sistemas Informáticos. Universidad de Granada, España.

- [12] Murray, J. H. (1997). Hamlet on the Holodeck: The Future of Narrative in Cyberspace. Simon and Schuster.
- [13] M. Weiser, R. Gold, J. S. Brown (1999) The origins of ubiquitous computing research at PARC in the late 1980s. IBM SYSTEMS JOURNAL, VOL 38, NO 4, 1999
- [14] Benford Steve, Carsten Magerkurth, AND Peter Ljungstrand (2005), March 2005/Vol. 48, No. 3 COMMUNICATIONS OF THE ACM, pag 54-59
- [15] Montola, M. and Stenros, J. and Waern, A., 2009. Pervasive Games: Theory and Design. Taylor & Francis
- [16] Ampatzoglou, A. y Stamelos I. (2010). Software engineering research for computer games: A systematic review. Information and Software Technology, 51(9), 888-901. Elsevier. Filadelfia Pensilvania, EEUU.
- [17] Catalano, C. E., Luccini, A. M. y Mortara, M. (2014). Best Practices for an Effective Design and Evaluation of serious games., International Journal of serious game 1(1). Disponible en <http://journal.seriousgamesociety.org>
- [18] Manrubia Pereira, A. M. (2014): El proceso productivo del videojuego: fases de producción. Historia y Comunicación Social ,19, 791-805. Universidad Complutense de Madrid.
- [19] Scott, R. (2010). Level UP - The Guide to Great Game Design. John Wiley & Sons.
- [20] Bethke, E. (2003). Game Development and Production. Wordware Publishing.
- [21] Evans F., Spinelli A., Zaipirain E., Massa S., Soriano F. (2016), Proceso de desarrollo de Serious Games. Diseño centrado en el usuario, jugabilidad e inmersión, 3er Congreso Argentino de Ingeniería y 9no Congreso Argentino de Enseñanza de la Ingeniería (CAEDI). (2016), Resistencia, Chaco, Argentina.
- [22] Spinelli A., Massa S., Evans F., (2016). El proceso de creación de un videojuego como herramienta para la toma de decisiones en el uso eficiente de energía. II Congreso Argentino de energías sustentables (CES). (2016) Bahía Blanca, Provincia de Buenos Aires, Argentina.
- [23] Spinelli A., Massa S., Zapirain E., Kühn F., Rico C., (2016). Elicitación de requerimientos para un Serious Game. II Jornada Argentina de Tecnología, Innovación y Creatividad (JATIC). (2016), Mar del Plata, Buenos Aires Argentina.
- [24] Massa, S. M. (2013). Objetos de aprendizaje: Metodología de desarrollo y Evaluación de la calidad. Tesis Doctoral. Facultad de Informática. UNLP. En Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/26207>
- [25] Bacino, G. (2015). Aula extendida en la educación superior en Ingeniería. Una propuesta de aplicación en el área tecnológica básica de electrotecnia. Tesis de Maestría. Facultad de Informática. UNLP. Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/45476>
- [26] Revuelta, M. (2016). Laboratorio Remoto en un Entorno Virtual de Enseñanza Aprendizaje. Tesis de Maestría. Facultad de Informática. UNLP. Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/55888>

Entorno Virtual Flexible como Apoyo al Aprendizaje en el Área de Ciencias Exactas

Myriam G. Llarena, Silvia L. Villodre

Departamento de Informática / Facultad Ciencia Exactas Físicas y Naturales Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O). Complejo Universitario "Islas Malvinas".

Rivadavia, San Juan

Teléfono:0264- 4234129

{myriam.llarena, svillodre} @gmail.com

RESUMEN

Este trabajo se encuentra en el marco de un proyecto de investigación que propone entre sus objetivos definir estrategias para integrar las herramientas web 2.0, que conforman los entornos personales de los alumnos, a los entornos virtuales institucionales utilizados como apoyo a la educación presencial.

La generación de un espacio flexible, n donde converjan aspectos de la educación formal e informal, propiciando en el alumno oportunidades de aprendizaje para desarrollar competencias y habilidades requeridas en su perfil profesional.

Esta investigación se diseña atendiendo las características del método cualitativo Investigación – Acción en tres etapas: diseño de la investigación, que incluye propuesta de estrategias a ser implementadas en las aulas de la UNSJ, correspondiente a carreras del área de las Ciencias Exactas, a fin de lograr el desarrollo de algunas competencias especificadas en el perfil profesional de

los alumnos de dichas carreras, puesta en práctica de las estrategias propuestas y evaluación de la calidad de las mismas, atendiendo la calidad del servicio y primordialmente la satisfacción del alumno.

Campo de Aplicación: Educación Universitaria del área de las Ciencias Exactas.

PALABRAS CLAVE

Educación Formal-Informal, LMS, PLE, comunidades virtuales de aprendizaje, convergencia tecnológica.

CONTEXTO

Este trabajo se enmarca en el proyecto de investigación “Diseño, Implementación y Evaluación de Entorno Virtual Flexible para el Aprendizaje” dentro de la línea de Investigación Tecnología y Educación.

Es un proyecto del Programa Permanente de Educación a Distancia y del Departamento de Informática de la Facultad de Ciencias Exactas de la UNSJ con aplicaciones en el Campus Virtual de la UNSJ. Integran el grupo de investigadores, docentes de la Facultad de Ciencias Exactas, Facultad de Ingeniería y de la Facultad de Filosofía, Humanidades y Artes. La investigación está financiada por la Universidad Nacional de San Juan.

INTRODUCCIÓN

Se ha producido en los últimos años una verdadera transformación tecnológica en virtud de la convergencia de tecnologías de microelectrónica, computación, telecomunicaciones, optoelectrónica. Para Castells, el actual proceso de transformación tecnológica se expande exponencialmente por su habilidad para crear una interfaz entre los campos tecnológicos a través de un lenguaje digital común en el que la información es generada, almacenada, recuperada, procesada y transmitida. Vivimos en un mundo que, siguiendo la expresión de Nicolás Negroponte, se ha vuelto digital. (Castells, 2015).

El impacto de las tecnologías se expande a todos los ámbitos y actividades del hombre, los dispositivos se han vuelto portables, las conexiones son inalámbricas, la información se encuentra digitalizada y accesible en diferentes formatos en la web. Los sujetos que se desarrollan en este contexto cumplen un rol activo en la producción de información y difusión de esa información a través de diferentes servicios y aplicaciones de la web 2.0 que se identifica como un espacio en donde el usuario deja de ser tan solo un consumidor pasivo y se transforma en un

productor de contenidos que comparte en diferentes redes, a través de aplicaciones y servicios disponibles en la web. Según Besada Estevez (2014),

En este contexto, los límites tradicionales que separan la Educación Formal de la Educación Informal se han vuelto permeables, en función del repertorio de contenidos digitalizados disponibles a todos los usuarios: el acceso al conocimiento es cada vez más sencillo, los procesos de enseñanza-aprendizaje se dan cada vez en más contextos y de una manera más dinámica en una fórmula próxima al aprendizaje bajo demanda.

Para esta investigación se retoma la conceptualización de Coombs en Sirvent (2006) que define la Educación Formal como “altamente institucionalizada, cronológicamente graduada y jerárquicamente estructurada”, mientras que la Educación Informal es aquella que se realiza durante toda la vida, “en la que se adquieren y acumulan conocimientos, habilidades, actitudes y modos de discernimiento mediante las experiencias diarias.

Las instituciones universitarias han incorporado a sus prácticas educativas tecnología web 2.0 tanto para soportar su oferta académica virtual como para el apoyo a la educación presencial. En la República Argentina, la mayor parte de las universidades han incorporado campus virtuales soportados en los sistemas conocidos como Learning Management System (LMS), plataformas de gestión de contenidos en las que se emula el aula tradicional. En ellas, el docente estructura su propuesta didáctica, el alumno accede a los materiales y se producen los intercambios entre docentes y alumnos a través de diferentes recursos de comunicación. Sin embargo no capitalizan todo el potencial que las tecnologías web 2.0 ofrecen para

fortalecer un aprendizaje centrado en el alumno, con grados de libertad en el tiempo, lugar y métodos para acceder al conocimiento.

Para Cabero (2010) el conocimiento en LMS es estático, declarativo (aseverativo) basado en la autoridad.

Por otro lado, han surgido nuevas corrientes pedagógicas que promueven la personalización del aprendizaje y desde esa perspectiva impulsan los PLE (Personal Learning Environment) que propician el autoaprendizaje por medio de la utilización de tecnología web a partir de las decisiones que toma el alumno y no el docente. Para Cabero es un sistema centrado en la figura del estudiante que le permite tomar el control de su propio proceso de aprendizaje de forma que pueda fijar sus propios objetivos, gestionar su actividad y comunicarse con otros. (Cabero,2010).

Tal como expresan en su libro Castañeda y Adell (2013) la idea de PLEs se remonta al año 2001 y durante los primeros años coexistieron dos corrientes de pensamiento y estudio en torno a ellos. Por un lado, la primera de ellas, heredera directa de las ideas esbozadas en el proyecto NIMLE que los entiende como un artefacto tecnológico y cuyo objetivo es crear y generalizar la mejor herramienta de PLE posible (Wilson, 2005, 2007; Van Harmelen, 2006, 2008;

Taraghi et al. 2009; Vavuola y Sharples, 2009; Casquero et al., 2008). Por otro, la que entiende que el PLE es una idea pedagógica sobre cómo aprenden las personas con tecnología, respaldada por diversos autores (Attwell, 2007, 2010; Waters, 2008; Downes, 2010, Adell y Castañeda, 2010, entre otros).

Indican Castañeda y Adell (2010) el tema de los PLE es a la vez un punto de inflexión y un nodo de confluencia en toda la discusión y prácticas referidas a aprender con tecnología. Decimos que se

trata de un punto de inflexión porque, una vez planteada la naturaleza del concepto y el contexto tecno-social en el que se desarrolla, propone claramente la urgencia de un cambio de dirección en casi todos los aspectos relacionados con la educación y el aprendizaje.

Si bien los PLE ofrecen la alternativa de un modelo de aprendizaje centrado en el alumno, libre, abierto que capitaliza y fortalece la convergencia de la Educación Formal e Informal, su creación exige de profesores y alumnos una fuerte capacitación conceptual y tecnológica (Cabero, 2011).

Por su parte Valero(2012) expresa.En los procesos de enseñanza que se redefinan, la importancia reside no sólo en qué quiere y debe enseñar el docente en términos de contenidos, cerrados e inamovibles, sino en qué necesita aprender el alumno y cómo el docente organiza en torno a él una experiencia/actividad/tarea que le permita aprenderlo.

La perspectiva teórica desde la que se aborda el análisis del proceso de Enseñanza Aprendizaje (E/A) en el marco del proyecto de investigación, es una teoría socio-constructivista. Ésta considera que el alumno es el centro del proceso de E/A, en el sentido de ser el responsable de su formación y en el que es fundamental su interacción con otros. En este escenario planteamos los interrogantes ¿Es posible la convergencia entre la Educación Formal, e Informal a través de servicios y aplicaciones de la web 2.0? ¿Es posible un entorno virtual de aprendizaje que combine el entorno virtual institucional con los entornos informales? ¿Es posible un enfoque pedagógico que permita equilibrar las componentes formal e informal que ofrece la tecnología para favorecer el aprendizaje? ¿Cómo definir estrategias

de uso de un entorno virtual flexible? ¿qué requisitos e indicadores son adecuados para validar la implementación y uso del entorno virtual flexible?

Esta investigación se propone la generación de un modelo de entorno virtual flexible que integre el entorno institucional -asociado a la Educación Formal- con el conjunto de aplicaciones y servicios de web 2.0 que los alumnos usan preferentemente- asociado al aprendizaje informal-. Se espera que el conocimiento resultante de la investigación favorezca la construcción de un espacio de trabajo más dinámico para fortalecer la participación del alumno en su proceso de aprendizaje y en la toma de decisiones que implica. Asimismo, que la estrategia de diseño e implementación sea una herramienta valiosa para aquellos docentes dispuestos a fortalecer la articulación de los espacios de trabajo institucionales con aquellos que el alumno recorre en su cotidianidad.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes abordados en esta investigación son, el conocimiento tecnológico, pedagógico y disciplinar, dado que integrar las TIC en el proceso de enseñanza/aprendizaje implica no solamente conocer las herramientas, sino también “reacomodar” las prácticas educativas, revisar y resignificar los conocimientos pedagógicos y disciplinares.

RESULTADOS Y OBJETIVOS

Se busca generar conocimiento acerca de la relación tecnología-educación, con el propósito de potenciar el aprendizaje de los alumnos del nivel universitario y el desarrollo de propuestas pedagógicas con TIC, en el área de Ciencias Exactas.

Para ello se han realizado las siguientes acciones:

- Diseño de un entorno personal que considera como una de sus componentes el aula virtual institucional (AV) frente a la opción de incorporar al AV algunas herramientas web 2.0.
- Definición de estrategias de implementación y uso de un Entorno Virtual flexible de Aprendizaje EVFA, en una cátedra de primer año de informática de la Facultad de Ciencias Exactas.
- Implementación de un EVFA (en 1er Año de la asignatura Procedural de FCEFYN- UNSJ)

Objetivos Futuros:

- Investigar e implementar nuevas estrategias
- Implementar un EVFA en distintos años y carreras del área de Ciencias Exactas de la UNSJ
- Validar los resultados obtenidos a partir de las experiencias de uso de los distintos EVFA.
- Conformar una comunidad virtual entre educadores de distintas unidades académicas de la U.N.S.J en relación a la problemática de entornos flexibles.

FORMACIÓN DE RECURSOS HUMANOS

El proyecto marco de este trabajo promueve la capacitación y actualización permanente de los miembros del equipo de investigación y a la consolidación del grupo de investigadores en la temática bajo estudio.

Se prevé continuar con la capacitación iniciada en 2016 en la temática a docentes Universitarios y de Nivel Medio, en el marco de la Diplomatura **Educación y Nuevas Tecnologías en tiempos de convergencia**, organizada por el Programa Permanente de Educación a distancia de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNSJ

Se encuentra en curso el desarrollo de una Tesis Doctoral de un integrante del proyecto.

BIBLIOGRAFIA

Besada Estevez Alba (2014) Las TIC y su impacto en la sociedad

<http://stellae.usc.es/red/blog/view/111106/las-tic-y-su-impacto-en-la-sociedad>

Cabero, J., Barroso, J.; Llorente, M. C. (2010). El diseño de Entornos Personales de Aprendizaje y la formación de profesores en TIC. Digital Education Review, 18, 26-37. Recuperado de <http://greav.ub.edu/der/index.php/der/article/view/169/298>

Cabero, J. ; Marín, V. (2011) Creación de un entorno personal para el aprendizaje: desarrollo de una experiencia. Edutec-e, Revista Electrónica de Tecnología Educativa, volumen (38) Recuperado de

http://edutec.rediris.es/Revelec2/Revelec38/creacion_entorno_personal_aprendizaje_desarrollo_experiencia.html

Cantilo, Valero, C; Roura, Redondo, M; Sánchez, Palacín, A. (2012). Tendencias actuales en el uso de dispositivos móviles en educación. [Versión electrónica]. La Educación Digital Magazine, núm. 147. Recuperado de http://educoas.org/portal/la_educacion_digital/147/pdf/ART_UNNED_EN.pdf

Castañeda, L. y Adell, J. (2013) Entornos personales de aprendizaje: claves para el ecosistema educativo en la red Editorial Marfil, S.A.I.S.B.N.: 978-84-268-1638-2. Recuperado de <http://www.um.es/ple/libro>

Castels, Manuel La era de la información

Recuperado de <http://comunicacion3unlz.com.ar/wp-content/uploads/2014/07/Castells-Manuel-La-Era-De-La-Informacion-Tomo-1.pdf>

Sirvent, M. T. y Otros (2006) “ Revisión del Concepto de educación No Formal, Cuadernillo de la Cátedra: Educación No Formal: Modelos y Teorías. Bs.As. Facultad de Filosofía y Letras de la UBA.

Espacios Dialógicos con Tecnologías de la Información y Comunicación (TIC) como Portadores de Conocimientos en Construcción en el Aprendizaje en Carreras Universitarias

González Liliana M., Ureta Laura, María José Marcovecchio, Viviana Margarit

{lilianamirna, laubureta, marajosm, viviana.margarit}@gmail.com

Programa Permanente de Investigación de Educación a Distancia
Dpto. de Informática. Dpto. de Biología
Facultad de Ciencias, Exactas, Físicas y Naturales
Universidad Nacional de San Juan

Resumen

En la actualidad para los estudiantes usar tecnologías es un hecho cotidiano; se comunican, relacionan, entretienen y estudian. Las tecnologías se han convertido en una extensión de la mente y del cuerpo. La participación de los estudiantes en redes sociales como Facebook, grupos de wassap, foros y otras formas de participación hace que se evidencie una actividad intensa con diferentes tecnologías. Área Moreira (2009) y Cabero Almenara (2011) coinciden plantear una problemática: Escuelas con estructuras organizativas del siglo XIX, profesores del siglo XX y alumnos del siglo XXI?. Es evidente el desfase entre las estructuras educativas, los actores participantes y el uso de las TIC. Cabe preguntarnos, ¿Cómo abordar esta realidad socio-educativa que enfrentan las generaciones actuales de alumnos y profesores en un contexto socio-educativo que no ofrece las estructuras pertinentes? ¿Cómo usan los docentes universitarios los artefactos tecnológicos de la web 2.0 y/o de sistemas de mensajería instantánea? ¿Los espacios dialógicos con TIC favorecen los procesos de construcción de conocimiento? En este proyecto se pretende analizar las interacciones y mediaciones que se generan en espacios dialógicos con TIC. Se optó por metodología cualitativa para la comprensión del uso de diferentes tecnologías en procesos de

construcción de conocimientos utilizando espacios dialógicos.

Palabras clave: Espacios Dialógicos, Tecnologías información y comunicación, Construcción Conocimiento

Contexto

El Proyecto “**Espacios dialógicos con tecnologías de la información y comunicación (TIC) como portadores de conocimientos en construcción en el aprendizaje en carreras universitarias**”, financiado por CICITCA, U.N.S.J, Código E1009, tiene como objetivo principal generar conocimientos que aporten a la problemática del uso de tecnologías en procesos de construcción de conocimiento en carreras universitarias.

Este Proyecto se nutrió del conocimiento científico resultante de los Proyectos: “Indicadores de Alfabetización Académica para alumnos de carreras de Informática” ejecutado en 2011-2013. Código 21E/ ¿Cómo se enseña y cómo se evalúa la lectura y escritura en carreras universitarias? Ejecutado en 2014-2015. Código 21E/946 Ambos Proyectos se encuentran enmarcados en el “Programa Permanente de Educación a Distancia creado por Resolución 08/05 de la Facultad de Ciencias Exactas, Físicas y Naturales. El Programa se constituye en un espacio académico y de investigación,

sustentado por la conformación de equipos de trabajo interdisciplinarios abocados a la tarea de generar experiencia, conocimientos, soporte tecnológico-comunicacional y organizativo a las propuestas académicas con soporte tecnológico de apoyo a la presencialidad”¹.

1. Introducción

El conocimiento es el resultado de un proceso permanente de construcción a través de la interacción y del diálogo con “otros”, situado en un contexto socio-histórico. Desde esta perspectiva los alumnos aprenden por la interacción con el objeto según sus estructuras cognitivas, conocimientos previos y su experiencia; “aprenden haciendo”, investigando, analizando ejemplos y análisis de casos concretos de la realidad.

Subyace el concepto de diálogo educativo, al que Burbules (1999) refiere como “la particular comunicación pedagógica: una interacción conversacional deliberadamente dirigida a la enseñanza y al aprendizaje”, que busca propiciar la comprensión de los temas.

En los procesos de aprendizaje un ingrediente importante es el rol de la tecnología en la interacción entre docentes/alumnos y entre alumnos/alumnos y en los últimos años con el advenimiento de las TIC se ha enfatizado su incidencia en estos procesos. La interacción a través del diálogo facilita los procesos de construcción del conocimiento. Esto nos permite definir el aprendizaje dialógico como: “El que resulta de las interacciones que produce el diálogo igualitario, es decir, un diálogo en el que diferentes personas aportamos argumentos en condiciones de igualdad, para llegar a consenso, partiendo de que queremos entendernos hablando desde pretensiones de validez” (Elboj, Puigdemívol, Soler & Valls en Castro Sandúa y otros, 2014).

El aprendizaje dialógico se enmarca en la sociedad de la información y pone énfasis en la dimensión intersubjetiva de los aprendizajes en contraposición a las concepciones de

aprendizaje propias de la sociedad industrial que centró su atención en la dimensión intrapsicológica del alumno. Aún coexisten diferentes concepciones y estructuras institucionales que siguen dando lugar a rupturas que problematizan la realidad que nos toca vivir.

Área Moreira (2009) y Cabero Almenara (2011) coinciden en reconocer y plantear una problemática que se evidencia al momento de implementar las nuevas tecnologías en las aulas. Este último expresa “encontramos una escuela articulada alrededor de diferentes siglos: escuelas con estructuras organizativas del siglo XIX, profesores del siglo XX y alumnos del siglo XXI”.

Área Moreira (2002) refiere a la escuela como una institución propia del siglo XIX sustentada en la imprenta, tecnología creada en el siglo XVI. Los docentes, formados y participes de fines del siglo XX y principios del siglo XXI, han incorporado las tecnologías pero no han logrado hacerlo de manera genuina; periodo histórico respaldado por tecnologías como los medios masivos de comunicación. Los alumnos del siglo XXI asisten a la revolución científico-tecnológica en la cual las tecnologías de la información y comunicación son el sello de la era digital o sociedad de la información. Hoy se da la naturalización de ciertas prácticas sociales y culturales que hasta hace poco tiempo no habían impactado en los vínculos sociales y en las formas de comunicación. Los jóvenes se ven beneficiados por el acceso a dispositivos personales móviles lo que permite que estén “conectados” de forma continua. La participación de los estudiantes en redes sociales como Facebook, grupos de WhatsApp, foros de temas de su interés y otras formas de participación hacen que se evidencie una actividad intensa con diferentes tecnologías. Es evidente el desfase entre las estructuras educativas, los actores participantes y el uso de las TIC.

Ante esta realidad, cabe preguntarnos, ¿Cómo abordar esta realidad socio-educativa que

¹Resolución 08/05 de la Facultad de Ciencias Exactas, Físicas y Naturales, de la U.N.S.J.

enfrentan las generaciones actuales de alumnos y profesores en un contexto socio-educativo que no ofrece las estructuras pertinentes? ¿Las Universidades promueven institucionalmente el aprovechamiento de diferentes espacios dialógicos con TIC? ¿Cómo usan los docentes universitarios los artefactos tecnológicos de la web 2.0 y/o de sistemas de mensajería instantánea? ¿Qué intencionalidad pedagógica plantean ante el uso de tecnologías que generan espacios dialógicos? ¿Los espacios dialógicos con TIC favorecen los procesos de construcción de conocimiento? ¿Los alumnos usan los espacios dialógicos de manera pertinente en sus procesos de aprendizaje?

Aquí es pertinente considerar la propuesta de Maggio (2012), quien expresa “lo primero que debemos hacer es reconocer lo que los alumnos hacen” (p.153), preguntarnos qué son capaces de hacer, cuáles son sus preferencias tecnológicas, en qué actividades las usan y cómo usan las tecnologías, qué valor les asignan en sus procesos de estudio y producción. Esto significa reconocer el valor de los “usos culturales y formas cognitivas del mismo modo como deberíamos tener en cuenta los estilos cognitivos propios de los sujetos culturales que son nuestros alumnos” (p.153)

A docentes e investigadores se nos plantea un desafío ineludible “reconocer la tendencia cultural, que en general impacta tempranamente en los más jóvenes” (p.157).

Esto contribuirá al análisis y diseño de espacios dialógicos presenciales/virtuales más flexibles y abiertos que permitan integrar saberes disciplinares, tecnológicos y socio-culturales.

Las tecnologías de la información y de la comunicación se incorporaron a la educación generando la necesidad de continuar investigando y profundizando desde la perspectiva de la Tecnología Educativa.

2. Líneas de Investigación, Desarrollo e Innovación

El Proyecto “Espacios dialógicos con tecnologías de la información y comunicación (TIC) como portadores de conocimientos en construcción en el aprendizaje en carreras universitarias”, se ubican en la línea de investigación en el Programa Permanente de

Educación a Distancia: Los medios, materiales o recursos de enseñanza.

En el mencionado Proyecto se busca generar conocimientos que aporten a la problemática del uso de tecnologías en procesos de construcción de conocimiento en carreras universitarias. Los ejes centrales de la investigación giran entorno de identificar modos de uso de espacios dialógicos con tecnologías en carreras universitarias y describir las potencialidades tecno-pedagógicas de los mismos.

Además, nuestro interés se centra en formar equipos interdisciplinarios que se aboquen al análisis y estudio de la temática desde la visión de las diferentes tramas disciplinares.

3. Resultados obtenidos/ Esperados y

- **Experiencia piloto.** Se creó un grupo con alumnos de una Tecnicatura en recursos Humanos, dictada en la ciudad de Las Flores, Jáchal en el primer semestre del año 2016. Se plantearon dos objetivos iniciales: 1) Diseñar algunos WOOC y poner en práctica algunas cápsulas de micro aprendizaje. 2) Detectar inconvenientes para llevar a cabo la experiencia.

- **Diseño e implementación de Encuesta en línea** a docentes referida al Uso sobre el uso de Tecnología en procesos de Enseñanza-Aprendizaje, de las carreras Licenciatura en Biología, Licenciatura en Astronomía y Licenciatura en Ciencias de la Computación y Licenciatura en Geofísica de la FCEFYN.

<https://docs.google.com/forms/d/1dPjbb74oBljyacLoAmzC65fdyfi6ZQxtcYcS50Bf-BU/edit>

- **Diseño de instrumentos de autoevaluación y evaluación: Rúbricas** para el análisis de producciones en espacios dialógicos en Cursos de Postgrado: Foros y Blogs.

- **Publicación:** en elaboración, “Una aproximación hacia las Concepciones –ideas- de los alumnos sobre la lectura y escritura con o sin tecnologías web”.

4. Formación de Recursos Humanos

El equipo está formado por investigadores con formación en Lengua extranjera, Lengua y Literatura, Sociología. Investigadores con Maestría en “Procesos Educativos mediados por Tecnologías”

Tesis de Posgrado: 1 en proceso de defensa 2017.

- Actividades de Postgrado propuestas/implementadas.

- Diseño e implementación de la Diplomatura “Educación y nuevas tecnologías en tiempos de convergencia tecnológica”. Aprobada por Ordenanza N° 5 CD-FCEF, Prof. Responsable: Mgter Liliana Mirna González. 1° versión mayo 2015-Diciembre 2015. 2° versión mayo 2016-Diciembre 2016. Conformada por los siguientes Cursos:
- ✓ Curso I “Tecnología Educativa en tiempos de convergencia”. Prof. Responsable: Mgter Liliana Mirna González, mayo-junio 2016, aprobado por Resolución n° 17/2016 - CD-FCEF

5. Bibliografía

- Angeli S., Cerdá E. y otros. (2006) Construcción de espacios dialógicos de formación docente para la integración de la tecnología en ámbitos educativos de la Universidad. 6to Seminario Internacional de la Educación a Distancia. Universidad de Río Cuarto. Córdoba.
- Barberá E, Mauri T y Onrubia J. (compiladores (2008). Cómo valorar la Calidad de la Enseñanza basada en las TIC. Pautas e instrumentos de análisis. Edit. GRAÓ. España.
- Cabero Almenara, J. (2007). (Coord.). Tecnología Educativa.: Nuevas Tecnologías Aplicadas a la Educación. Madrid, Mc Graw Hill
- Coll, C., Mauri, T. y Onrubia, J. (2008). Análisis de los usos reales de las TIC en contextos educativos formales: una aproximación sociocultural. Revista Electrónica de Investigación Educativa.
- Cols E. (2011). Estilos de Enseñanza. Edit. HomoSapiens. Argentina.
- De Benito B. y Salinas J. (2008). Los Entornos Tecnológicos en la Universidad. Universidad de Sevilla. N° 32 Pixel-Bit. Revista de medios y educación. julio 2008. España.
- De Luca A. (compiladora) (2012) Educación a Distancia. Problemáticas, Alternativas de Solución, Experiencias y algo más... Editorial Fundación de la Universidad Nacional de San Juan. Argentina.
- Frida Díaz Barriga en UBATIC+, 2012 <http://encuentroubatic.rec.uba.ar/index.php/vis-ta-previa-conferencias/item/414-frida-d%C3%ADaz-barriga>
- González L. (2012) Revista Digital: espacio para la producción escrita de textos académicos en la universidad. En De Luca A. (compiladora). Educación a Distancia. Editorial Fundación de la U.N.S. J. Argentina
- Litwin, E. (2005). Tecnologías educativas en tiempos de Internet. Amorrortu. Bs.As.
- Maggio M. (2012) Enriquecer la enseñanza. PAIDÓS. Bs. As.
- Pea R. (2001). Prácticas de inteligencia distribuida y diseños para la educación. En Salomon Gabriel (compilador). Cogniciones distribuidas. Amorrortu ediciones. Bs. As, 75-125.

Estilos y Paradigmas de Interacción Persona-Ordenador y sus Posibilidades para el Escenario Educativo

Sanz Cecilia^{1,3}, Gorga Gladys¹, Artola Verónica^{1,5}, Moralejo Lucrecia¹, Salazar Mesía Natali^{1,6}, Archuby Federico¹, Baldassarri Sandra², Manresa-Yee Cristina⁴, Pesado Patricia^{1,3}

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata

²GIGA AffectiveLab, Universidad of Zaragoza, España

³Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

⁴Departamento de Ciencias Matemáticas e Informática de la Computación. Universidad de las Islas Baleares, España

⁵Becaria doctoral CONICET

⁶Becaria TIPO A UNLP

{[csanz](mailto:csanz@lidi.info.unlp.edu.ar), [ppesado](mailto:ppesado@lidi.info.unlp.edu.ar), [ggorga](mailto:ggorga@lidi.info.unlp.edu.ar), [vartola](mailto:variola@lidi.info.unlp.edu.ar), [lmoralejo](mailto:lmoralejo@lidi.info.unlp.edu.ar), [nsalazar](mailto:nsalazar@lidi.info.unlp.edu.ar), [farchuby](mailto:farchuby@lidi.info.unlp.edu.ar)}@lidi.info.unlp.edu.ar, sandra@unizar.es, cristina.manresa@uib.es

RESUMEN

Las interfaces han ido evolucionando a lo largo de la historia de las Ciencias de la Computación. El diseño centrado en el usuario y el aprovechamiento del entorno físico, en el que las personas desarrollan sus actividades, se han vuelto eje de variedad de investigaciones (Weiser, 1991) (Ishii & Ullmer, 1993) (Milgram, Takemura, Utsumi & Kishino, 1994). Los paradigmas de interacción actuales proponen generar escenarios naturales de interacción, y tender puentes entre los objetos del entorno físico y los objetos virtuales, para aprovechar de esta manera los beneficios de ambos mundos. Estos paradigmas pueden ofrecer varios aportes para el escenario educativo, ya sea en su aplicación en actividades exploratorias, de asociación, de simulación. También permiten atender a diferentes fines didácticos, ya sea la motivación, el cambio del rol del alumno a uno más activo, los estímulos multimodales, entre muchos otros (McNeil & Jarvin, 2007).

En este trabajo se presenta una investigación vinculada a la integración de diferentes paradigmas y estilos de interacción persona – ordenador en el escenario educativo. Aquí se describen los principales resultados alcanzados y las líneas de trabajo que se persiguen.

Palabras claves: Realidad Aumentada, Interacción Tangible, Computación Afectiva, Dispositivos móviles.

CONTEXTO

Esta investigación forma parte del subproyecto “Escenarios educativos mediados por Tecnologías de la Información y la Comunicación (TIC)”, enmarcado en el proyecto “Tecnologías para Sistemas de Software Distribuidos. Calidad En Sistemas y Procesos. Escenarios Educativos Medrados Por TIC (período 2014-2017)”, perteneciente al Instituto de Investigación en Informática LIDI, de la Facultad de Informática de la UNLP, acreditado por el Ministerio de Educación de la Nación.

1. INTRODUCCION

El área de interacción persona – ordenador ha despertado el interés de los investigadores y convoca la mirada de diferentes disciplinas tales como la Psicología Cognitiva, la Educación, la Comunicación, el Diseño, y las Ciencias de la Computación. Se encuentra en continua evolución, teniendo como una de sus metas buscar las mejores alternativas para la interacción entre las personas y los entornos digitales.

Así, las interfaces de líneas de comando (*Command Line Interface-CLI*), dieron paso a las interfaces gráficas de usuario (*Graphical User Interface-GUI*) y hoy día ya se habla de un conjunto de paradigmas y estilos de interacción que han sido agrupados dentro de la categoría de Interfaces Naturales (*Natural User Interface-NUI*).

Las ideas de Weiser (1991) sobre la Computación Ubicua, las de Ishii & Ullmer (1993) sobre las Interfaces Tangibles y el surgimiento de las NUI han dado lugar a que el usuario pueda controlar una aplicación de cualquier tipo, a través de una interfaz invisible que resulte intuitiva, fácil de aprender y logre ser accedida a través de la voz, gestos, movimientos corporales, o movimientos de objetos del entorno que son interpretados como comandos de entrada para la aplicación (Sabino Moxo & Márquez Domínguez, 2014).

Las interfaces tangibles, la realidad aumentada, los mundos virtuales 3D, las capacidades de los dispositivos móviles y de los sensores, permiten hoy hablar de entornos aumentados, donde lo digital y lo físico se entrelaza para dar a la persona una experiencia mutisensorial y sumamente interactiva (Bujak, Radu, Catrambone, MacIntyre, Zheng, Golubski, 2013) (Chen & Tsai, 2012) (Dillenbourg & Evans, 2011).

También este tipo de interfaces ofrecen nuevas oportunidades para personas con necesidades educativas especiales (Bojanovic, Manresa-Yee, Abásolo, Larrea & Sanz, 2015) (Piper, O'Brien, Morris, & Winograd, 2006) (Battocchi et al., 2010). En (Bonillo, Cerezo, Marco & Baldassarri, 2016, p.1) se habla de “Interfaces de Usuario Distribuidas (IUD), que soportan trabajo colaborativo en habitaciones o espacios de trabajo digitalmente aumentados. Los usuarios en un ECI¹ pueden usar diferentes métodos de interacción basados en la manipulación física (paredes táctiles, *tabletops*, vestibles, objetos inteligentes...) para interactuar con una aplicación informática”.

Son varios los autores que han abordado el uso de este tipo de interfaces en escenarios educativos. Algunos motivados por el hecho de “...eliminar barreras tecnológicas del tipo física y mental, haciendo que las interfaces persona - ordenador sean más intuitivas y ampliando las maneras en que los usuarios pueden interactuar y experimentar con las Tecnologías de la Información y la

Comunicación, al trasladarlos de un modelo convencional utilizando dispositivos electrónicos artificiales como el teclado y el *mouse* [...], hacia un modelo de interacción más natural” (Duque & Vázquez, 2015, p.8).

En este trabajo se pone especial foco en la realidad aumentada, la interacción tangible, la computación ubicua en general, los mundos virtuales 3D, la computación afectiva y las posibilidades de los móviles para generar entornos aumentados que brinden oportunidades para el escenario educativo. Se diseñan y se desarrollan herramientas de software y de hardware, metodologías para el aprovechamiento de estas tecnologías, y se llevan adelante experiencias con docentes y alumnos, que son evaluadas para detectar problemas, realizar mejoras y avanzar en la investigación. A continuación se presentan las líneas de investigación y desarrollo que se abordan.

2. LINEAS DE INVESTIGACION / DESARROLLO

Las líneas de investigación, desarrollo e innovación en las que se trabaja son:

- Estilos y Paradigmas de Interacción Persona – Ordenador. Posibilidades para la educación, y con foco en la accesibilidad e inclusión.
- Realidad Virtual y Aumentada. Sus posibilidades para el ámbito educativo.
- Interacción tangible y multitáctil. Aplicaciones en educación y educación especial.
- Computación afectiva para el enriquecimiento de escenarios educativos.
- Dispositivos móviles y sensores para la generación de entornos aumentados.
- Herramientas colaborativas para la educación y educación especial en particular.

3. RESULTADOS ESPERADOS/OBTENIDOS

Durante el 2016 se profundizaron los avances en estas líneas de investigación. Se describen aquí diferentes acciones realizadas en este período.

Se trabajó en una revisión bibliográfica sobre las posibilidades de la realidad aumentada en escenarios educativos. Al mismo tiempo se profundizaron las experiencias realizadas utilizando

¹ ECI: Espacio Colaborativo Interactivo

EPRA, un material educativo con actividades basadas en realidad aumentada para la enseñanza de conceptos básicos de Programación (Salazar, Gorga & Sanz, 2015). En este sentido se realizaron experiencias con alumnos de las carreras de la Facultad de Informática de la UNLP y en la Universidad de Río Negro (estas últimas se dan en el marco de un proyecto de investigación en el que participan ambos grupos). Para evaluar estas experiencias se tomaron tres ejes de análisis, la motivación intrínseca de los alumnos, la satisfacción respecto de la experiencia y de las actividades de RA disponibles en EPRA, y el impacto en el aprendizaje sobre el tema abordado en el material. Se utilizaron diferentes técnicas para la recogida de datos: el cuestionario IMI (*Intrinsic Motivation Inventory*) para analizar la motivación intrínseca, cuestionarios de satisfacción, y se administraron un *pretest* y un *posttest*, antes y después de la experiencia con el material de RA para analizar el impacto en el aprendizaje. Este proceso ha sido motivo de varias publicaciones (Salazar, Sanz & Gorga, 2016 a, b y c). Al mismo tiempo, estos trabajos se vinculan con el tema de una tesis de maestría en Tecnología Informática Aplicada en Educación de uno de los miembros del subproyecto de investigación presentado aquí.

En relación a la investigación en temas de RA se ha continuado con el desarrollo de la herramienta de autor AuthorAR, para la creación de actividades educativas de RA. Esta herramienta cuenta con la posibilidad de crear un proyecto que incluya una serie de actividades. Las actividades pueden ser de dos tipos, por el momento: de exploración y de estructuración de frases (orientadas al escenario de educación especial). Aún resta la exportación de un proyecto para que las actividades puedan ser recorridas en forma independiente de la herramienta. Se espera avanzar en este sentido durante 2017 y también en la realización de las pruebas de usabilidad de AuthorAR con docentes (destinatarios de la herramienta).

Por otra parte, se ha profundizado la investigación en temas relacionados con interacción tangible. En primer lugar se avanzó en el desarrollo de TIES, un editor para la creación de actividades de interacción tangible (Artola, Sanz, Moralejo, Pesado & Baldassarri, 2015), que será evaluado durante 2017. Se continuó con el desarrollo de experiencias utilizando ITCol (un juego basado en interacción

tangible para la colaboración), que se integra como parte de un seminario de la Maestría en Tecnología Informática Aplicada en Educación de la Facultad de Informática de la UNLP. Para evaluar las sesiones con ITCol se diseñó una metodología que abarca una entrevista individual con los participantes, luego de recibir las pistas iniciales que son únicas para cada jugador, un registro filmico y de interacciones, a través de una matriz que completan los observadores del proceso, y una entrevista grupal final para realizar un análisis metacognitivo del proceso vivenciado (Zangara & Sanz, 2015). Los resultados de las sesiones realizadas en los últimos 3 años, serán publicados en un trabajo durante 2017. Sin embargo ya se han presentado resultados parciales en (Artola, Sanz, Pesado & Baldassarri, 2016) y en el marco de un libro² publicado en 2016 que refleja un trabajo conjunto con la Universidad de Guadalajara en el marco del Programa Pablo Neruda de la Red *IberoTICs*.

Durante 2017 se espera avanzar también en las evaluaciones del juego EPIT para la enseñanza de Programación utilizando interacción tangible. También se ha involucrado, a partir del corriente año, la participación de un pasante para trabajar en la arquitectura de la mesa interactiva utilizada en el marco de estas experiencias y realizar mejoras en este sentido.

En relación a la temática de mundos virtuales 3D, se han abordado diferentes tesis. Una de ellas se orienta al estudio de las posibilidades de estos mundos para personas con discapacidad auditiva (Fachal, Sanz & Abásolo, 2015). En este marco, se ha realizado un relevamiento bibliográfico de experiencias vinculadas; se ha estudiado también la situación del alumno sordo en Argentina y la utilización de TICs en instituciones de nuestro país para la enseñanza a estos alumnos. Vinculado a estas temáticas, también se desarrolló ComuniCA (como parte de un trabajo de grado). ComuniCA es un componente de software que integra los pictogramas de Comunicación Aumentativa y Alternativa (Boix & Basil, 2005) y videos de lenguaje de seña para la composición de mensajes y su publicación en espacios de escritura en la web

² Aplicación de la Tecnología en el Aprendizaje. Casos de Argentina y Mexico. Compiladores: Sara Catalina Hernandez Gallardo, Cecilia Verónica Sanz, Cornelio Cano Guzmán, y Ricardo Marcelo Naiouf. Publicado en 2016.

(Santillán, Sanz & Gorga, 2017). Este componente se ha integrado a CkEditor³ y se han realizado pruebas también con Facebook. Se cree que será de utilidad también en el marco de los mundos virtuales 3D y será evaluado en el marco de la tesis doctoral antes referida.

Al mismo tiempo, se finalizó una tesis de maestría dirigida por miembros del proyecto sobre “El impacto de los juegos de rol en *Second Life* en la práctica de las subcompetencias lingüística y discursiva en inglés a nivel oral”. Se ha realizado una publicación al respecto (Quintín, Sanz & Zangara, 2016).

Por otra parte, se ha continuado con la revisión bibliográfica relacionada con el área de Computación afectiva (Picard, 1997) y sus posibilidades para el escenario educativo (Huerta- Pacheco, Rebolledo-Mendez & Hernandez- Gonzalez, 2016) (Shen, Wang & Shen, 2009) (Qi Luo, 2006), a partir de dos propuestas de tesis de doctorado en estas temáticas dirigidas por miembros del proyecto en conjunto con la Universidad de Zaragoza.

Finalmente, se han desarrollado juegos educativos, uno de ellos sobre móviles orientado a la autoevaluación de los alumnos. Se trata de un juego de preguntas y respuestas del estilo Preguntados⁴ o *Geography*⁵, que se integra a un entorno virtual de enseñanza y aprendizaje (EVEA), en particular se ha realizado la integración a IDEAS⁶. El juego se llama Desafiate y permite que los docentes publiquen las autoevaluaciones generadas en el EVEA como una aventura dentro del juego, que presenta desafíos a los participantes. Se trata de una historia donde el personaje principal es un pirata que recorre islas en busca de tesoros, pero para obtenerlos deberá resolver un desafío en cada isla. Este trabajo forma parte de un trabajo de grado de uno de los miembros de este proyecto. También en esta línea se ha desarrollado un juego para la enseñanza de la Programación, como parte de una

³ CKEditor: se trata de un editor HTML para la escritura en la web. Disponible en: <http://ckeditor.com/>

⁴ Preguntados: es un videojuego de trivia para móviles iOS, Android y Windows Phone, lanzado a finales de 2013 por la empresa argentina Etermax.

⁵ *Geography*: Juego de Preguntas de Geografía 3D

⁶ IDEAS: <http://ideas.info.unlp.edu.ar> – Es una evolución del EVEA WebUNLP, desarrollado en la Facultad de Informática de la UNLP.

tesina de grado dirigida por investigadores del proyecto. El juego es en primera persona y plantea la construcción de algoritmos como parte de una historia interplanetaria. Se difundirán los resultados en 2017.

Es importante destacar que el equipo de trabajo para esta investigación se integra por investigadores del III LIDI y otros de España (Universidad de las Islas Baleares y Universidad de Zaragoza) que colaboran en el desarrollo de los temas y en la dirección de trabajos de tesis de maestría y doctorado.

En cuanto a los proyectos y acuerdos de cooperación vinculados con la temática, el III- LIDI participa en los siguientes:

- Acuerdo de colaboración en estos temas con la Universidad de Zaragoza. En particular con el grupo GIGA *AffectiveLab*.
- Acuerdo de colaboración en estos temas con la Universidad de Islas Baleares. En particular con la Unidad de Gráficos, Visión por Computador e IA.
- REDAUTI: Red Temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva.
- REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente. Proyecto que se desarrolla en el marco de los Proyectos De Innovación y Transferencia En Áreas Prioritarias de la Pcia. de Buenos Aires (PIT- AP-BA). Período: 2017-2018.
- Acuerdo de cooperación con la Universidad Nacional de Santiago del Estero. Se asesora en temas relacionados a la integración de dispositivos móviles en procesos educativos.

4. FORMACION DE RECURSOS HUMANOS

En el marco del subproyecto que aquí se presenta se participa en la formación de los siguientes recursos humanos:

- Un becario de UNLP con beca TIPO A, que está realizando la tesis de maestría en TIAE relacionada con la temática aquí expuesta, y trabaja en el marco del subproyecto.
- Un becario doctoral CONICET, que está realizando su tesis en temas vinculados con Interacción Tangible para el escenario educativo, también es integrante del subproyecto.

- Dos especialistas en TIAE, que está finalizando su tesis de maestría e integrante del subproyecto.
- Un becario de entrenamiento de la Comisión de Investigadores Científicas de la Provincia de Buenos Aires.

Los investigadores formados de este proyecto se encuentran dirigiendo tesis doctorales y de maestría vinculadas a los temas de esta investigación, así como también tesinas de grado.

En la siguiente sección se presenta bibliografía de referencia para la investigación en los temas presentados y los trabajos citados a lo largo del documento.

5. BIBLIOGRAFIA

- Artola V. & Sanz C. (2016). Informe técnico de las sesiones de evaluación realizadas con ITCOL.
- Artola V., Sanz C. V., Moralejo L., Pesado P. M. & Baldassarri S. (2015). Herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible. En XXI Congreso Argentino de Ciencias de la Computación. Junín, 2015.
- Battocchi A., Ben-Sasson A., Esposito G., Gal E. Pianesi F., Tomasini D., Venuti P., Weiss P. L. and Zancanaro M. (2010). Collaborative Puzzle Game: a Tabletop Interface for Fostering Collaborative Skills in Children with Autism Spectrum Disorders. *Journal of Assistive Technologies*. 4(1): pp. 4-14
- Boix J., Basil C. (2005). CAA en atención temprana. *Comunicación y pedagogía: NT y recursos didácticos*. ISSN: 1136-7733, N° 205, pp. 29-35.
- Bojanovic, F., Manresa-Yee, C., Abásolo, M.J., Larrea, M. & Sanz, C (2015). Get Closer to Activate It! An Educational Tool for People with Multiple Disabilities. XXI Congreso Argentino de Ciencias de la Computación. Junín, 2015. ISBN: 978-987-3806-05-6. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/50238>
- Bonillo C., Cerezo E., Marco J., Baldassarri S. (2016). Herramienta para el desarrollo de espacios colaborativos interactivos: arquitectura y prueba de concepto. *Actas del XVII Congreso Internacional de Interacción Persona-Ordenador – Interacción 2016*, pp. 193-200
- Bujak K.; Radu I.; Catrambone R.; MacIntyre B.; Zheng R.; Golubski G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*.
- Chen C. & Tsai Y. (2012). Interactive augmented reality system for enhancing library instruction in elementary schools. *Computers & Education*, 59 (2), (638652).
- Dillenbourg P. & Evans M. (2011). Interactive tabletops in education. *Computer-Supported Collaborative Learning*. 6, pp. 491–514.
- Duque E. & Vásquez A. (2015). NUI para la educación. Eliminando la discriminación tecnológica en la búsqueda de la Inclusión digital. Disponible en: <http://hdl.handle.net/123456789/3780>
- Fachal A., Sanz C., Abásolo M.J. (2015). Informe técnico sobre “Antecedentes en el desarrollo de software y herramientas educativas para la enseñanza de personas con hipoacusia”.
- Huerta-Pacheco N. S., Rebolledo-Mendez G., & Hernandez-Gonzalez S. (2016). Cognitive-affective Modelling Approach in Tutoring System. In *Proceedings of the 1st International Workshop on Emotion Awareness in Software Engineering* (pp. 18–21). New York, NY, USA: ACM. <https://doi.org/10.1145/2897000.2897005>
- IMI (*Intrinsic Motivation Inventory*). Disponible en: <http://www.selfdeterminationtheory.org/> 7
- Ishii, H. & Ullmer, B. (1993). Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'97)*, pp. 234-241
- McNeil N. & Jarvin L. (2007). When Theories Don't Add Up: Disentangling the Manipulatives Debate. *Theory Into Practice*. Vol. 46, no. 4, pp. 309 - 316
- Milgram P., Takemura H., Utsumi, A., & Kishino F. (1994). Augmented Reality: A class of displays on the reality-virtuality continuum.

- Telemanipulator and Telepresence Technologies, 2351, 11.
- Picard R. W. & Picard R. (1997). Affective computing. Vol. 252. Cambridge: MIT press.
 - Piper A. M., O'Brien E., Morris M. R. & Winograd T. (2006). SIDES: a cooperative tabletop computer game for social skills development. 20th Conference on Computer Supported Cooperative Work.
 - Qi Luo. Application of Affective Computing in e-Learning System. (2006). Open Education Research. Pp. 80-83.
 - Quintín E., Sanz C. & Zangara A. (2016). The Impact of Role-Playing Games through Second Life on the Oral Practice of Linguistic and Discursive Sub-competences in English. Proceedings of the 2016 International Conference on Collaboration Technologies and Systems. En cooperación con IEEE, ACM and IFIP. Orlando, Florida, USA – ISBN: 978-1-5090-2300-4/16 IEEE DOI 10.1109/CTS.2016.40. PP. 148 - 155
 - Sabino Moxo B., Márquez Domínguez J.A. (2014). Rehabilitación Virtual Mediante Interfaces Naturales De Usuario. Congreso Virtual sobre Tecnología, Educación y Sociedad CTES. ISBN: 2007-7475. Disponible en: <http://cenid.org.mx/memorias/ctes/index.php/ctes/article/view/102>
 - Salazar Mesía N., Gorga G., Sanz C. (2015). EPRA: herramienta para la enseñanza de conceptos básicos de programación utilizando realidad aumentada. X Congreso sobre Tecnología en Educación & Educación en Tecnología. Corrientes, 2015. ISBN: 978-950-656-154-3, 426-435. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/49071>
 - Salazar Mesía N., Sanz C., Gorga G. (2016 a) Experiencia de enseñanza de Programación con Realidad Aumentada. Jornadas de la Enseñanza Universitaria de la Informática JENUI 2016. Almería, España. Julio de 2016. Publicado en las Actas de las Jornadas ISBN: 978-84-16642-30-4 - Pp. 213 a 220.
 - Salazar Mesía N., Sanz C., Gorga G. (2016 b) Augmented Reality for Programming Teaching. Student Satisfaction Analysis. Proceedings of the 2016 International Conference on Collaboration Technologies and Systems. En cooperación con IEEE, ACM and IFIP. Noviembre 2016, Orlando, Florida, USA – ISBN: 978-1-5090-2300-4/16 IEEE DOI 10.1109/CTS.2016.43 – Pp. 165-171
 - Salazar Mesía N., Sanz C., Gorga G. (2016 c) Experiencia con Realidad Aumentada. Análisis de la incidencia en la comprensión de las estructuras de control. Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), ISBN: 978-987- 733-072-4. Pp. 1231-1241.
 - Santillán C., Sanz C, Gorga G. (2017). ComuniCA: un componente de software para la escritura de mensajes de comunicación aumentativa en espacios web. Tesina de grado de la Licenciatura en Sistemas. Facultad de Informática UNLP.
 - Shen L., Wang M., & Shen R. (2009). Affective e-Learning: Using “Emotional” Data to Improve Learning in Pervasive Learning Environment. Journal of Educational Technology & Society, 12(2), 176–189. Disponible en: <http://www.jstor.org/stable/jeductechsoci.12.2.176>
 - Weiser M. (1991). The computer for the 21st century. Scientific American, 265(3), (pp. 94-104)
 - Zangara A., Sanz C. (2015). Collaborative Learning Work. Evaluation Model Using The Ethnographic Methodology Of “Analysis Of Social Networks”. 7th International Conference on Education and New Learning Technologies .EDULEARN 2015. España-Barcelona. Pp. 5373-5380. ISBN: 978-84-606-8243-1 - ISSN: 2340-1117 – Editor: IATED. Disponible en: <http://library.iated.org/view/ZANGARA2015C OL>

Estudio de la Influencia del Uso de Simulación en la Enseñanza de Redes de Computadoras en el Nivel Universitario (resultados finales)

Daniel Arias Figueroa¹, Javier Díaz², Cecilia Gramajo¹, Gustavo Gil¹, Ernesto Sánchez¹,
Loraine Gimson¹, Álvaro Gamarra¹, Jorge Silvera¹, Diego Gil¹
Nelson Rodríguez³, María Murazzo³, Marcelo Moreno³, Miguel Guevara³

C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada)¹

Facultad de Ciencias Exactas de la Universidad Nacional de Salta

L.I.N.T.I. (Laboratorio en Nuevas Tecnologías Informáticas)²

Facultad de Informática de la Universidad Nacional de La Plata

Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.³

Fac. de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan

daaf@cidia.unsa.edu.ar, nelson@iinfo.unsj.edu.ar

RESUMEN

La implantación de las técnicas de simulación mediante computadora en el aula es una realidad en los actuales sistemas educativos. La necesidad de poder comprender los conceptos y fundamentos de redes de computadoras obliga a utilizar la computadora como instrumento para el aprendizaje. Esto, sumado a los costos elevados de equipos específicos necesarios para montar un laboratorio de red, y a los escasos recursos con los que cuenta el Departamento de Informática de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta (similares a los recursos con los que cuentan la mayoría de las universidades públicas en el interior del país), hacen considerar a las herramientas de simulación, como una posible solución para que las prácticas sean mejor aprovechadas por los estudiantes, posibilitando además la utilización de estas herramientas fuera de los horarios de clase.

Dentro del trabajo de investigación se evaluó la influencia de la utilización de software de simulación, en la enseñanza de contenidos de redes de computadoras en carreras de grado y en cursos de

postgrado. Para ello se planteó un estudio aplicado, longitudinal y experimental que, asumiendo la complejidad del fenómeno educativo, consideró aspectos cuantitativos y cualitativos.

Palabras clave: Protocolo TCP-IP, Simulación, Virtualización, Enseñanza, Redes.

CONTEXTO

La línea de investigación fue apoyada por el C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada) que depende de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta y el L.I.N.T.I. (Laboratorio en Nuevas Tecnologías Informáticas) que depende de la Facultad de Informática de la Universidad Nacional de La Plata, y, por lo tanto, se contó con toda la infraestructura disponible para esta investigación. El proyecto fue acreditado como Proyecto Tipo "A" N° 2230/0 en el CIUNSa – Consejo de Investigación de la Universidad Nacional de Salta y contó con su financiamiento.

Además se firmó convenio de colaboración con la Universidad Nacional de San Juan, con el objetivo de replicar

las experiencias en la Facultad de Ciencias Exactas Físicas y Naturales.

1. INTRODUCCIÓN

Uno de los objetivos claves en la enseñanza de las redes de computadoras en la Licenciatura en Análisis de Sistemas, fue la de transmitir fundamentos y conceptos básicos a los estudiantes de nuestra carrera de grado. Sin embargo, desde hace un tiempo hemos venido investigando cómo facilitar la relación entre la realidad y las teorías y modelos, es decir, entre lo concreto y lo abstracto.

Las computadoras personales, con la variedad de software que se ha desarrollado, tienen en sí mismas un gran potencial para mejorar el proceso de enseñanza-aprendizaje, a fin de que:

- El aprendizaje sea más interesante.
- El aprendizaje sea activo, no pasivo en las aulas.
- Los estudiantes estén más motivados.
- El aprendizaje sea al ritmo del estudiante en forma personalizada.
- La educación sea permanente.

El Departamento de Informática de la Facultad de Ciencias Exactas de la UNSa, ve a la simulación como una herramienta de apoyo pedagógico, que facilita la investigación y el desarrollo de una mejor calidad y pertinencia de la educación y de formación de saberes contextualizados entre teoría y práctica en el estudiante. Debido a esto, se está utilizando en diversas materias software de simulación.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación, evaluó la influencia de la utilización de software de simulación en la enseñanza de contenidos de redes de computadoras, en el ámbito de las asignaturas relacionadas con la

temática en la carrera de Licenciatura en Análisis de Sistemas de la UNSa y en el ámbito de cursos de postgrado organizados por el C.I.D.I.A. Para ello, se llevó a cabo un estudio aplicado, longitudinal y experimental que, asumiendo la complejidad del fenómeno educativo, considera aspectos cuantitativos y cualitativos.

Del resultado de la búsqueda de la mejora de la enseñanza realizada y la formación y experiencia adquiridas, surgieron las siguientes hipótesis causales de investigación, en relación a las correlaciones aprendizaje-métodos de simulación:

- Hipótesis I: A igualdad de recursos utilizados y experiencias, el alumno aprende mejor, más significativamente, cuando complementa esas experiencias con métodos de simulación.
- Hipótesis II: El uso de software de simulación como complemento a los prácticos tradicionales o de laboratorio, mejora el aprendizaje de conceptos y la comprensión.

Asimismo, la investigación persiguió en el marco del paradigma cualitativo los siguientes objetivos:

- Realizar entrevistas y estudios de casos con los alumnos durante los períodos previstos.
- Diseñar actividades prácticas y de investigación adecuadas para ser realizadas con los simuladores.
- Fomentar el aprendizaje cooperativo entre los estudiantes.
- Introducir a los estudiantes de nivel universitario en la cultura científica actual, que concibe la simulación por computadora como una herramienta fundamental para el estudio, la investigación y la experimentación.

3. RESULTADOS ALCANZADOS

La primera etapa consistió en la preparación del material, guías de trabajos prácticos para la utilización de los simuladores y realización de pruebas piloto para validación del mismo.

En la segunda etapa del proyecto se realizaron diversas experiencias con simulación en asignaturas de la Licenciatura en Análisis de Sistemas y cursos de postgrado de la Universidad Nacional de Salta y en la Universidad Católica de Salta a través del convenio firmado entre ambas instituciones.

En la tercera etapa del proyecto se analizaron los resultados obtenidos buscando responder a las preguntas planteadas.

Experiencias realizadas

A continuación se mencionan las experiencias realizadas en el período 2012-2016:

Año 2012

- Experiencia con el simulador Packet Tracer en la asignatura Conectividad y Teleinformática de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador KivaNS en la asignatura Conectividad y Teleinformática de la Licenciatura en Análisis de Sistemas de la UNSa.

Año 2013

- Experiencia con el simulador Packet Tracer en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador Packet Tracer en un curso de Postgrado organizado por el CIDIA.

Año 2014

- Experiencia con Packet Tracer en la asignatura Redes de Computadoras I

de la Licenciatura en Análisis de Sistemas de la UNSa.

- Experiencia con el simulador MACSim en la asignatura Redes de Computadoras II de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador Packet Tracer y el Protocolo ARP en la asignatura Redes de Computadoras II de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador MACSim en la asignatura Conectividad y Teleinformática de la Licenciatura en Análisis de Sistemas de la UNSa.

Año 2015

- Experiencia con el simulador Packet Tracer y el Protocolo DNS en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con GNS3 en un curso dictado en el marco de las Jornadas de Ingeniería de la UNSa.
- Experiencia con el simulador Packet Tracer y el Direccionamiento IP en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador Packet Tracer y el Ruteo IP en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador MACSim en la asignatura Redes I de la Ingeniería en Telecomunicaciones en la Universidad Católica de Salta.
- Experiencia con GNS3 en un curso de extensión organizado por el CIDIA.
- Experiencia con el simulador MACSim en la asignatura Redes de

Computadoras II de la Licenciatura en Análisis de Sistemas de la UNSa.

- Experiencia con el simulador Packet Tracer y el Protocolo ARP en la asignatura Redes de Computadoras II de la Licenciatura en Análisis de Sistemas de la UNSa.

Año 2016

- Experiencia con el simulador Packet Tracer y el Direccionamiento IP en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia con el simulador Packet Tracer y el Ruteo IP en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas de la UNSa.
- Experiencia sobre el Direccionamiento IP con Router Cisco 3600 en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas.
- Experiencia sobre el Ruteo IP con Router Cisco 3600 en la asignatura Redes de Computadoras I de la Licenciatura en Análisis de Sistemas.
- Experiencia con el Simulador MACSim en la asignatura Redes de Computadoras II de la Licenciatura en Análisis de Sistemas.
- Experiencia en un Curso para certificación Mikrotik organizado por el CIDIA.
- Experiencia con la herramienta GNS3 y Laboratorio real en la asignatura Redes I de la UCASAL.
- Encuesta a docentes de asignaturas de redes de computadoras de diversas universidades vinculadas al proyecto.

Respuestas obtenidas a las preguntas planteadas

¿Cómo influyen en la enseñanza y el aprendizaje, los TP realizados con un software de simulación?

Contribuyen poderosamente a afianzar los conocimientos que se imparten. El estudiante puede verificar en topologías virtuales, como responde una variable dependiente a las variaciones de otra, u otras.

Los métodos de simulación son una herramienta que, bien utilizada, permite expandir las posibilidades de la realidad física, potenciando el aprendizaje.

¿Qué efectos causan los programas de simulación, en lo que hace a motivar al estudiante a aprender a resolver los problemas analizados en las clases teóricas?

Los estudiantes se mostraron muy receptivos e interesados en las diferentes herramientas de software utilizados, manifestando decididamente en las encuestas realizadas, que el uso del mismo los ha motivado al aprendizaje.

¿Cuánto pueden contribuir los trabajos sobre el simulador a que se reafirmen los conceptos teóricos y la comprensión de los fundamentos y funcionamiento en general de las redes?

Según las evaluaciones, las experiencias realizadas contribuyeron a brindarles una comprensión acabada de los conceptos teóricos, en mayor medida que los trabajos prácticos tradicionales. La posibilidad de variar la configuración de los componentes y visualizar el cambio de comportamiento potencia el aprendizaje. También en las encuestas, los estudiantes manifestaron su satisfacción.

Dado que el alumno puede diseñar e inventar topologías y experimentar con ellas ¿hasta qué punto favorece la creatividad?

En la medida que el alumno se sienta motivado a utilizar las herramientas de simulación, tendrá flexibilidad y oportunidades, que no es posible en la práctica tradicional. No tiene prácticamente limitaciones para crear y probar diferentes topologías. Los estudiantes manifestaron su actitud favorable hacia esta posibilidad de diseñar e inventar, sin costo ni peligros.

¿Hasta qué punto contribuyen a reemplazar un aprendizaje memorístico por aprendizaje significativo?

En general puede decirse que el uso de métodos de simulación potencia el aprendizaje significativo de los conceptos que se aprenden en las clases, donde se enseñan los conceptos teóricos y se realizan ejemplos de aplicación.

¿En qué momento conviene aplicar estos métodos, antes y/o después de los trabajos prácticos tradicionales? ¿Qué condiciones deben darse para la aplicación de estos métodos?

Por el nivel alcanzado por los estudiantes en las evaluaciones realizadas de las actividades prácticas, se considera que es conveniente su utilización después de la presentación teórica y la práctica tradicional en aula, para permitir que el estudiante llegue al mismo con los conceptos más asimilados y más motivado. El trabajo práctico de laboratorio con equipo real le sirve, así, para terminar de aprender aspectos procedimentales, relacionados con la operación de dispositivos, como ya hemos mencionado.

¿Cómo estructurar una actividad práctica utilizando métodos de simulación?

Por lo expuesto, consideramos que las mismas deben realizarse en forma muy coordinada con las clases en las que se

explica la teoría y la práctica tradicional. Debe buscarse la forma de que el estudiante se confronte con los conceptos aprendidos. Consideramos que las experiencias realizadas avalan la postura inicial en el sentido que estas experiencias con simulación complementan las prácticas tradicionales, y ponen de manifiesto que son un recurso excelente para complementar la enseñanza.

Los resultados alcanzados fueron publicados en revistas internacionales, congresos nacionales e internacionales y también hemos publicado un libro denominado “Redes de Computadoras I con Packet Tracer”, que fue evaluado y aprobado por la Editorial de la Universidad Nacional de Salta – EDUNSa – ISBN 978-987-633-132-6-1.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación se conformó con un director y cinco miembros entre docentes, egresados y estudiantes avanzados. Además se sumaron docentes de otras dependencias de la universidad con trabajos concretos de tesis de especialidad en seguridad y redes. También se contó con el apoyo de un asesor.

- Especialidad en Seguridad y Redes – UNLP. “Virtualización para la enseñanza de redes IP”. Expte. 3300-3489/11. Miguel A. Aguirre. (En desarrollo).
- Especialidad en Seguridad y Redes – UNLP. “Herramientas de Software de Simulación para Redes de Comunicaciones”. Expte. 3300-5305/12. Juan Antonio Torres. (Aprobada en julio de 2015).
- Doctorado en Ciencias Informáticas – UNLP. “Estudio de la influencia de un entorno de simulación en la

enseñanza de redes en el nivel universitario”. Expte. 3300-5743/13. Daniel Arias Figueroa. (Tesis presentada para su evaluación).

- Magister en Redes de Datos – UNLP. “Un estudio comparativo en extensiones de seguridad para el sistema de nombre de dominio (DNS)”. Expte. 3300-002113/10-000. Ernesto Sánchez.
- Magister en Ingeniería de Software – UNLP. “Desarrollo basado en Conocimiento siguiendo prácticas ágiles”. Expte. 3300-006507/13-000. Loraine Gimson. (Aprobada en noviembre de 2015).
- Cursando la Maestría en Redes de Datos en la UNLP. Ing. Álvaro I. Gamarra.
- Cursando la Licenciatura en Análisis de Sistemas en la UNSa. Diego Gil.

5. BIBLIOGRAFIA

SAKAR, N. I. (2006) Teaching TCP/IP Networking Using Practical Laboratory Exercises, *International Journal of Information and Communication Technology Education*, Vol. 2, No. 4, pp. 39-50.

GOLDSTEIN, G., M LEISTEN, S, STARK, K., & TICKLE, A. (2005) Using a Network Simulation Tool to Engage Students in Active Learning Enhances Their Understanding of Complex Data Communications Concepts, *Proceedings of the 7th Australasian conference on Computing Education*, pp. 223-228.

JAVIDI, G. & SHEYBANI, E. (2008) Content-Based Computer Simulation of a Networking Course: An Assessment, *Journal of Computers*, Vol. 3, No. 3, pp. 64-72.

CAMERON, B. (2003): Effectiveness of simulation in a hybrid online networking course. *Quarterly Review of Distance Education*, 4(1), 51.

GATTO, D. (1993): The use of interactive computer simulations in training. *Australian Journal of Educational Technology*, 9(2), 144-156.

YAVERBAUM, G., & NADARAJAN, U. (1996): Learning basic concepts of telecommunications: an experiment in multimedia and learning. *Computers & Education*, 26(4), 215-224.

ZHU, S. Y. (2011). Teaching Computer Networks through Network Simulation Programs. Faculty of Business, Computing and Law – School of Computing. University of Derby. *Learning Teaching & Assessment Conference*.

KUROSE, J.F. & ROSS, K.W. (2015). Computer Networking: A Top-Down Approach. 6th Edition. *Pearson Education*. ISBN: 9780132856201.

AVILA BLAS, Orlando José (2003). Probabilidad y estadística inferencial: teoría y aplicaciones. ISBN: 978-987-9381-23-6. *Editorial: Univ. de Salta*.

GLIEM, J & GLIEM, R. (2003). Calculating, Interpreting, and Reporting Cronbach’s Alpha Reliability Coefficient for Likert-Type Scales. *Conference in Adult, Continuing, and Community Education. Midwest Research to Practice*.

CÁMARA, ALZUGARAY (2011). Trabajos Prácticos, Métodos de Simulación y aprendizaje significativo. *La Tecnología Educativa al servicio de la Educación Tecnológica*.

Formación Automática de Grupos Colaborativos Considerando Estilos de Aprendizaje y Rendimiento Académico

Alberto Palavecino¹, Germán Lescano^{1,2}, Rosanna Costaguta¹

¹Instituto de Investigación en Informática y Sistemas de Información (IISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)

²Consejo Nacional de Investigaciones Científicas y Técnica (CONICET)

palapaoli@gmail.com, {gelescano, rosanna}@unse.edu.ar

RESUMEN

El Aprendizaje Colaborativo (AC) se vincula con métodos de enseñanza y de aprendizaje donde los estudiantes trabajan en pequeños grupos para resolver una consigna común. El avance tecnológico que se ha producido en las últimas décadas permitió al AC adoptar herramientas computacionales que facilitan la colaboración, la coordinación y la comunicación transformándolo en Aprendizaje Colaborativo Soportado por Computadora (ACSC). Crear grupos e instar a sus miembros a resolver una consigna de manera colaborativa, no garantiza en forma alguna que el comportamiento y el rendimiento de esos grupos sean adecuados, ni que la experiencia de enseñanza y de aprendizaje sea exitosa. Una de las variables que influyen en los resultados es la constitución de los grupos. Es por esto que nuestra línea de investigación se enfoca en la formación automática de grupos colaborativos on-line. En particular, se propone una herramienta software que crea grupos de estudiantes teniendo en cuenta sus estilos de aprendizaje, combinándolos de tal forma que se incrementen o maximicen sus rendimientos académicos. En este artículo se describen los objetivos de la investigación, el enfoque

propuesto para concretar la formación automática de grupos colaborativos, y algunos antecedentes relevantes.

PALABRAS CLAVE

Aprendizaje Colaborativo Soportado por Computadora, estilos de aprendizaje, grupos de estudiantes colaborativos, algoritmo genético, formación automática de grupos.

CONTEXTO

En este artículo se presenta una de las líneas de investigación del proyecto “Mejorando escenarios de aprendizaje colaborativo soportado por computadora”, acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT - UNSE) para el período 2017-2020. El proyecto es una continuación de la línea de investigación *Sistemas Adaptativos Inteligentes*, iniciada en 2005-2009 por el proyecto 23/C062 “Herramientas conceptuales, metodológicas y técnicas de la Informática Teórica y Aplicada”, continuada en 2009-2010 por el proyecto P09/C002 "Personalización en Sistemas de Enseñanza Virtual", en 2010-2011 por el proyecto 23/C089 "Fundamentos Conceptuales

y Soportes Tecnológicos de la Informática Educativa”, y en 2012-2016 por el proyecto 23/C097 “Sistemas de información web basados en agentes para promover el Aprendizaje Colaborativo Soportado por computadoras (ACSC)”. Todos los proyectos citados fueron acreditados y financiados por SECyT – UNSE.

1. INTRODUCCIÓN

Un grupo de aprendizaje se define como una estructura formada por personas que interactúan para lograr objetivos específicos de aprendizaje a través de su participación [1]. Uno de los factores que afectan el resultado de una experiencia de aprendizaje colaborativo es la composición del grupo [2]. Este tópico ha ganado gran interés en la comunidad de E-Learning en los últimos años [3].

Las tres formas más comunes aplicadas para la formación de grupos son: la selección al azar de los miembros, dejar que ellos mismos se elijan, o permitir que el profesor lo haga siguiendo algún criterio preestablecido. La asignación al azar puede generar grupos muy desequilibrados que probablemente no resulten eficaces, mientras que la autoselección puede causar discriminación entre los estudiantes con dificultades para mantener vínculos sociales. Sin embargo, la tercera alternativa pareciera ser la más prometedora hasta la fecha. En este sentido, un criterio a considerar podría ser tener en cuenta los estilos de aprendizaje de los alumnos. Un estilo de aprendizaje se define por la forma en que los estudiantes perciben, interactúan y responden en las situaciones de aprendizaje. El modelo de Dunn & Dunn [4] considera 21 variables para describir el estilo de aprendizaje del alumno, abarcándolas en los siguientes dominios: a) caracterización del

ambiente inmediato (sonido, iluminación, temperatura, diseño de mobiliarios); b) emocionalidad del alumno (motivación, persistencia, responsabilidad, la necesidad de contar con una estructura impuesta externamente o la oportunidad para hacer cosas a su propio modo); c) preferencias sociológicas (aprender sólo, trabajando en diadas, en grupo pequeños, con un alumno colegiado o autoritativo, y la necesidad de variedad como oposición a patrones y rutinas); d) características fisiológicas (fortaleza perceptual, niveles de energía durante el día, la necesidad de comer o moverse mientras aprende); y, las inclinaciones de procesamiento (global/analítico, zurdo/diestro, e impulsivo/reflexivo). Otro modelo es el propuesto por Kolb [5] quien identifica dos dimensiones principales del aprendizaje: la percepción y el procesamiento. Considera que el aprendizaje es el resultado de la forma como las personas perciben y luego procesan lo que han percibido. Sin embargo, el modelo más utilizado es el propuesto por Felder y Silverman [6], esto se debe posiblemente a que fue evaluado con éxito en distintos dominios [3]. Según estos autores un estilo de aprendizaje resulta de la combinación de cuatro dimensiones: comprensión, percepción, entrada y procesamiento. Cada una de estas dimensiones puede tomar uno de dos valores posibles. Para la dimensión comprensión el valor resultante puede ser global o secuencial, para la dimensión percepción puede ser intuitivo o sensitivo, para la dimensión entrada verbal o visual, y finalmente, el valor para la dimensión procesamiento puede ser activo o reflexivo. Así, combinando estos valores posibles para las cuatro dimensiones resultan 16 estilos de aprendizaje diferentes.

La creación manual de grupos tomando como criterio de agrupamiento los estilos de aprendizaje es prácticamente inviable cuando el número de estudiantes es alto, puesto que encontrar la condición óptima para agruparlos es un problema NP-Hard [7] que requiere tiempos exponenciales para resolverlo. Por esta razón en estos casos suelen utilizarse algoritmos de aproximación que, si bien no necesariamente llegan a la solución óptima se acercan considerablemente, constituyéndose así en una solución apropiada. Uno de los algoritmos más usados para resolver problemas de este tipo son los algoritmos genéticos [8].

En ACSC se han propuesto algunos enfoques para formar grupos automáticamente con la aplicación de algoritmos genéticos. En [9] se describe un algoritmo genético para la formación de grupos heterogéneos que busca maximizar las habilidades en los grupos formados. En [10] se presenta el uso de un algoritmo genético y dos funciones de evaluación: una intragrupal y otra intergrupala. Los autores buscan heterogeneidad intragrupal y homogeneidad intergrupala para que la agrupación sea óptima. En [11] se introduce un sistema de agrupación heterogéneo que hace uso de factores psicológicos. Este sistema utiliza un algoritmo genético para formar grupos heterogéneos con el mismo tamaño y mismo nivel de diversidad de integrantes. En [12] se propone el uso de un algoritmo genético para la creación de grupos bien balanceados con respecto a nueve roles de equipo. En [13] se describe un algoritmo genético, donde los miembros de los grupos se definen considerando sus habilidades de programación en Java.

La exploración bibliográfica realizada nos permite afirmar que no existen investigaciones vinculadas con la formación automática de grupos mediante un algoritmo genético que considere estilos de aprendizaje y rendimiento académico como en la presente propuesta.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta línea de trabajo se abocará a conformar grupos de manera automática mediante un algoritmo genético que considera los estilos de aprendizaje individuales de los estudiantes, intentando maximizar la cantidad de grupos con buen rendimiento. En particular, la línea tiene por objetivos generales: promover el uso de la computación evolutiva como método para generar grupos de estudiantes en ACSC, e innovar en el uso de tecnologías que asistan al docente en la formación automática de grupos en entornos de aprendizaje. El objetivo específico se definió como: crear un algoritmo genético que considerando los estilos de aprendizaje de los estudiantes permita conformar grupos que obtengan un buen rendimiento académico en las sesiones colaborativas en las que esos grupos participen.

La herramienta a desarrollar aumentará las funcionalidades de un entorno colaborativo desarrollado con anterioridad por integrantes del proyecto 23/C097. Este entorno permite el trabajo colaborativo de estudiantes en grupos síncronos y asíncronos, y fue desarrollado por [14] encontrándose disponible desde el dominio de la FCEyT (<http://fce.unse.edu.ar/seacunse/ProjectChatDB/>). Actualmente la creación de los grupos en este entorno debe ser realizada por el administrador en base a las indicaciones e información suministrada por los profesores. Así, el desarrollo de nuestra herramienta para

conformar grupos de manera automática acrecentará sus funcionalidades. En Palavecino *et al.* [15] se describen con mayor detalle las funcionalidades propuestas para la herramienta.

Otro punto importante en el presente trabajo es la experimentación con estudiantes reales. La validación experimental se realizará mediante el uso de la herramienta en sesiones colaborativas reales, en las que participarán estudiantes y docentes de la carrera de Licenciatura en Sistemas de Información (LSI) de la Universidad Nacional de Santiago del Estero (UNSE).

3. RESULTADOS OBTENIDOS Y ESPERADOS

Este trabajo pretende, desde la perspectiva del proyecto de investigación en el que se encuentra inmerso, favorecer el desarrollo de conocimiento científico-tecnológico de relevancia para el desarrollo de sistemas de información web personalizados en el área del ACSC. Desde la perspectiva puntual de este trabajo se esperan obtener los siguientes resultados:

- Estado del arte referido a la formación automática de grupos en contextos de ACSC, y en particular, sobre la aplicación de algoritmos de computación evolutiva para la realización de esta tarea.
- Una herramienta que ayude al profesor a generar de manera automática el agrupamiento de los estudiantes de acuerdo con sus estilos de aprendizaje y tendiendo a obtener mejores resultados.
- El desarrollo de un algoritmo genético para la formación de grupos considerando buenas combinaciones de

los estilos de aprendizaje individuales de los estudiantes.

- Datos experimentales obtenidos mediante la aplicación del algoritmo propuesto en situaciones reales de formación de grupos de estudiantes.

A la fecha se desarrolló el algoritmo genético que realiza la formación automática de grupos en base a los estilos de aprendizaje de sus miembros para maximizar el rendimiento académico [16].

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está compuesto por un docente investigador formado, un investigador en formación y un integrante estudiante. El investigador en formación es becario CONICET y está elaborando su propuesta de tesis para obtener el título de Doctor en Ciencias de la Computación. El integrante estudiante está desarrollando su trabajo final para obtener el título de Licenciado en Sistemas de Información. Ambos realizan sus trabajos en el marco de esta investigación.

5. BIBLIOGRAFÍA

- [1] M. Souto, *Didáctica de lo grupal*. Ministerio de Educación y Justicia, INPAD, 1990.
- [2] R. Costaguta, "Algorithms and Machine Learning Techniques in Collaborative Group Formation," *Adv. Artif. Intell. Its Appl.*, vol. 9414, pp. 249–258, 2015.
- [3] A. Monteserin, S. Schiaffino, P. Garcia, and A. Amandi, "Análisis de la formación de grupos en Aprendizaje Colaborativo Soportado por Computadoras," in *Simpósio Brasileiro*

de Informática na Educação, 2010.

- [4] R. Dunn, "Understanding the Dunn and Dunn Learning Styles Model and the Need for Individual Diagnosis and Prescription," *J. Reading, Writing, Learn. Disabil. Int. Overcoming Learn. Difficulties*, vol. 6, no. 3, pp. 223–247, 2006.
- [5] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*. New York, United States: Prentice Hall, 1984.
- [6] R. Felder and L. Silverman, "Learning and Teaching Styles in Engineering Education Application," *J. Eng. Educ.*, vol. 78, no. 7, 1988.
- [7] Y.-T. Lina, Y.-M. Huang, and S.-C. Cheng, "An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization," *Comput. Educ.*, vol. 55, no. 4, pp. 1483–1493, 2010.
- [8] S. Sivanandam and S. Deepa, *Introduction to Genetic Algorithms. Chapter 1 Evolutionary Computation*. Berlin, Heidelberg: Springer-Verlag, 2008.
- [9] A. Sukstrienwong, "Genetic Algorithm for Forming Student Groups Based on Heterogeneous Grouping," in *3rd European Conference of Computer Science (ECCS '12)*, 2012, pp. 92–97.
- [10] M. M. B. Jozan, F. Taghiyareh, and H. Faili, "An Inversion-based Genetic Algorithm for Grouping of students," *Proc. 7th Int. Conf. Virtual Learn.*, vol. 1, no. 1, pp. 152–161, 2012.
- [11] D. Wang, S. S. J. Lin, and C. T. Sun, "DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups," *Comput. Human Behav.*, vol. 23, no. 2007, pp. 1997–2010, 2010.
- [12] V. Yannibelli and A. Amandi, "A deterministic crowding evolutionary algorithm to form learning teams in a collaborative learning context," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 8584–8592, 2012.
- [13] Z. C. Ani, A. Yasin, M. Z. Husin, and Z. A. Hamid, "A Method for Group Formation Using Genetic Algorithm," *Int. J. Comput. Sci. Eng.*, vol. 2, no. 9, pp. 3060–3064, 2010.
- [14] D. Yanacón Atía, "Desarrollo de un chat con interfaz semiestructurada para Moodle," in *XVI Workshop de Investigadores en Ciencias de la Computación (WICC)*, 2014.
- [15] A. Palavecino, G. Lescano, R. Costaguta, and D. Missio, "Formación automática de grupos colaborativos: Una propuesta que considera estilos de aprendizaje y rendimiento académico," in *Quinto Congreso Argentino de la Interacción Persona Computador@, Telecomunicaciones, Informática e Información Científica (IPCTIIC 2016)*, 2016.
- [16] G. Lescano, R. Costaguta, and A. Amandi, "Genetic Algorithm for Automatic Group Formation Considering Student's Learning Styles," in *2016 8th Euro American Conference on Telematics and Information Systems (EATIS)*, 2016.

Gamificación en el Ámbito Universitario – Análisis e Implementación de Elementos de Juegos

Pablo Martín Vera, Edgardo Javier Moreno, Rocío Andrea Rodríguez,
Myriam Carina Vázquez, Federico Ezequiel Valles, Javier Gustavo Cescon

GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software)
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1903, San Justo, Provincia de Buenos Aires
{pvera, emoreno, rocio.rodriguez, mcvazquez, fvalles, jcescon} @unlam.edu.ar

RESUMEN

La gamificación consiste en agregar elementos de juegos a espacios no lúdicos. En este caso se aplica al ámbito universitario. Se toma como punto de partida un proyecto anterior del equipo de trabajo donde se desarrolló e implementó una plataforma de gamificación para la enseñanza de programación en la cual se proponen una serie de desafíos al alumno, que consisten en resolver distintos problemas codificando una función y asignado de forma automática puntaje según los errores cometidos ya sean sintácticos o lógicos. Esa asignación de puntos es un ejemplo de lo que se conoce como componentes en la gamificación. Con dichos puntos se generó un ranking general y un ranking por cursos lo que hace que el alumno se vea motivado por la competencia (mecánica de juego) y la autosatisfacción de lograr un puesto mejor (dinámica de juego). Este proyecto se basa en analizar y determinar nuevos elementos de juegos que pueden ser incorporados más allá de los puntos y rankings ya incorporados previamente. El agregado de nuevos componentes permitirá generar nuevas mecánicas y dinámicas que harán más atractiva a la plataforma fomentando la utilización de la misma por parte de los estudiantes.

Palabras clave: Gamificación, Programación, Educación, Elementos de Juego, Motivación

CONTEXTO

El presente proyecto de Investigación y Desarrollo corresponde al grupo de investigación GIDFIS (Grupo de Investigación, Desarrollo y Formación en Innovación de Software) perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza. El grupo es interdisciplinar al estar formado por docentes de distintos departamentos y en él se encuentran formándose en actividades de I+D alumnos de la universidad de las carreras de Ingeniería.

1. INTRODUCCIÓN

Con un claro objetivo de mejora en el proceso de enseñanza-aprendizaje, actualmente se está considerando implementar nuevas maneras de lograr una mayor atención e interés por parte de los estudiantes.

“Existen prácticas docentes que estimulan una mayor participación de los estudiantes, dando lugar a un trabajo motivador que estimula el pensamiento creativo e innovador, que potencia su autonomía y facilite el aprendizaje de competencias transversales y profesionales” [1]. De hecho en las materias actualmente se asegura un balance entre la teoría y la práctica, apoyándose en recursos provistos por las TIC (Tecnologías de la Información y Comunicación) [2] junto con diversas estrategias que permiten tomar en cuenta las características de los actuales estudiantes [3], [4]. Actualmente los jóvenes

en su mayoría “nativos digitales” tienen facilidad de uso de la tecnología y la han integrado como parte de su vida cotidiana. Pero además cabe destacar el interés de los jóvenes, y en algunos casos no tan jóvenes, por los videojuegos, la facilidad con que interactúan, dedicando sin darse cuenta muchas horas frente a estos desafíos, que poseen niveles y premios a medida que el usuario adquiere mejor manejo del producto. Es importante lograr transferir este interés espontáneo por los juegos y los desafíos que estos plantean al ámbito académico.

En cuanto al pensamiento lúdico, podemos decir que a través de los juegos se pueden recorrer todas las fases del proyecto, analizando desde las etapas de planteamiento de objetivos y análisis de la situación actual, pasando por la divergencia o generación de ideas, hasta la convergencia en soluciones prácticas. Permiten además, romper las barreras y mejorar la comunicación en el grupo y generar nuevas ideas, visiones y estrategias. Es por tanto, una potente herramienta para trabajar el pensamiento innovador y desarrollar la creatividad” [1].

Este proyecto se basa en la Gamificación que consiste en aplicar elementos de juegos a espacios no lúdicos. Aplicar elementos de juegos a prácticas cotidianas como por ejemplo asignando puntos por la resolución de ejercicios, permite motivar a los alumnos para mejorar su posicionamiento en los rankings y en muchos casos esto pasa a ser un desafío personal que los impulsa a la resolución de problemas. Pero la asignación de puntos es solo uno de los elementos posibles. Es importante poder enriquecer las herramientas con diversos elementos basados en juegos que permitan generar otros espacios de interacción.

La Gamificación puede definirse como:

- “el proceso de usar el pensamiento y las mecánicas de juegos en contextos distintos a estos con el fin de resolver problemas y comprometer a los usuarios” [5].

- “intentar aprovechar el poder motivacional de los juegos y aplicarlo al mundo real” [6].

Kevin Werbach [7], analista en tecnología, desarrolló la “Pirámide de Elementos de Gamificación” que se muestra en la Figura 1.

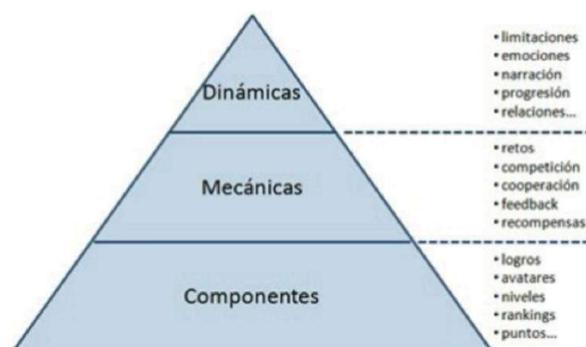


Figura 1. Pirámide de Elementos de Gamificación

A través de la gamificación se puede trabajar los procesos cognitivos de manera más creativa, atractiva y favoreciendo el aprendizaje. “...la gamificación no es convertir las asignaturas en un juego, ni simplemente poner una puntuación a cada tarea. El proceso de gamificar, al igual que el proceso de virtualizar una asignatura, supone un esfuerzo de planificación y seguimiento, diseño y realimentación constantes y adaptados a cada asignatura, grupo y curso”[8].

Actualmente hay una clara necesidad por parte de la industria de contar con ingenieros. Puede observarse que el mayor índice de deserción se produce en los primeros años de la carrera. La universidad ha implementado distintos mecanismos para disminuir esta deserción entre ellos las tutorías. Por otra parte, se contempla la incorporación a través de las materias, de herramientas apoyadas en la tecnología que permitan motivar a los alumnos en el estudio dentro de la cursada. El concepto de gamificación trae consigo la principal ventaja de la motivación, recompensas de logros personales y sentimiento de progreso. “La gamificación es la aplicación de mecánicas y dinámicas de juego en ámbitos que normalmente no son

lúdicos” [9]. “Se puede lograr que las personas se involucren, motiven, concentren y se esfuercen en participar en actividades que antes se podrían clasificar de aburridas y que con la gamificación pueden convertirse en creativas e innovadoras” [7].

“La gamificación puede ser una estrategia de gran alcance que promueva la educación entre las personas y un cambio de comportamiento” [10].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Se investigará principalmente sobre los siguientes temas:

- Análisis de los elementos de juego a incorporar a la plataforma incluyendo componentes, mecánicas y dinámicas.
- Nivel de aceptación de los nuevos elementos y su impacto en la motivación de uso.

En cuanto al desarrollo se prevé

- Mejorar la plataforma para el análisis de desafíos.
- Diseño y desarrollo de nuevos componentes en la plataforma web.

3. RESULTADOS TENIDOS/ESPERADOS

OB

Debido a que este proyecto se basa en uno anterior, se cuenta con resultados obtenidos en una prueba piloto de la plataforma, realizada a tres cursos durante el segundo cuatrimestre del 2016. Al finalizar el cuatrimestre se realizó una encuesta a los alumnos mediante la cual pudo establecerse que de todos los alumnos de dichos cursos, el 40% utilizó la plataforma. De ese porcentaje el grado de uso declarado por los alumnos, fue medio-alto como puede verse en la figura 2.

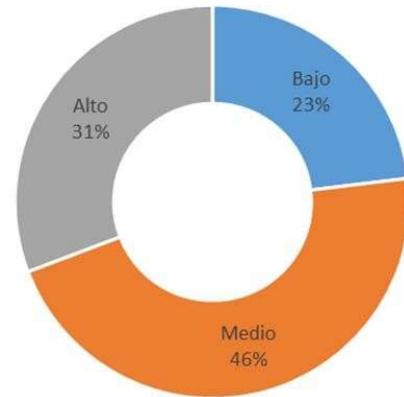


Figura 2. Grado de uso de la plataforma

Además, analizando la información disponible como trazabilidad en la plataforma pudo establecerse que el 62% de los alumnos registrados utilizaron la aplicación. La cantidad máxima de desafíos eran 37, cabe destacar que un 17% de los alumnos realizaron todos los desafíos y la cuarta parte de ellos han realizado más de 20 desafíos. En la tabla 1 se muestra el porcentaje de alumnos según la cantidad de desafíos que realizaron.

Tabla 1. Porcentaje de Alumnos según la cantidad de Desafíos Completados

Cantidad de Desafíos	% de Alumnos
37	16,67
> 20	25,00
<=20 y >16	8,33
<=15 y >11	4,17
<=10 y >5	8,33
<=5 y >1	12,50
1	41,67

Mediante este proyecto se busca aumentar la motivación para lograr un mayor uso de la plataforma esperando que no solo mayor cantidad de alumnos usen la plataforma, sino que también aumente el porcentaje de desafíos realizados por cada uno de ellos.

El objetivo principal es incorporar nuevos componentes de juego para generar diversas mecánicas y dinámicas que enriquezcan la plataforma actual de gamificación aplicada a la resolución de ejercicios de programación.

Esto deriva en los siguientes objetivos específicos:

- Analizar componentes de juegos.

- Generar estrategias para incluir nuevos componentes de juego a la plataforma.
- Desarrollar e incorporar los nuevos componentes a la plataforma.
- Incluir mejoras para análisis de código fuente.
- Establecer el impacto de los componentes incorporados en las mecánicas y dinámicas de juego.

Si bien los componentes a incorporar serán determinados luego del análisis detallado de los mismos, determinando la factibilidad de implementación, impacto en la motivación y análisis de las mecánicas y dinámicas empleadas algunas de las posibilidades son:

- Agregar distintos rankings más detallados y posiblemente un medallero general para generar mayor competitividad.
- Incorporar insignias por ejemplo al superar todos los ejercicios de un nivel, por constancia, etc.
- Utilizar herramientas sociales, permitiendo la creación de grupos de amigos para competir o vincular con redes sociales externas para publicar los resultados.
- Dar premios al completar cada ejercicio dando beneficios para la resolución de ejercicios posteriores, como tiempo adicional comodines para no contabilizar errores.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo está formado por 12 personas:

- 5 docentes pertenecientes a diferentes departamentos conformando un grupo interdisciplinario.
- 1 asesora, externa en el área de gamificación perteneciente a la Universidad de La Laguna, España.
- 2 graduados de Ingeniería.
- 4 alumnos de ingeniería quienes se están formando en actividades de Investigación y Desarrollo.

Vinculado con esta temática se encuentra en realización una tesis de maestría.

5. BIBLIOGRAFIA

[1] González, C. (2014). Estrategias para trabajar la creatividad en la Educación Superior: pensamiento de diseño, aprendizaje basado en juegos y en proyectos. RED, Revista de Educación a Distancia. Número 40. <http://www.um.es/ead/red/40/>

[2] Giulianelli, D. A., Rodríguez, R. A., Vera, P. M., Trigueros, A., Cruzado, G. S., Moreno, E. J., & Marko, I. (2014). Incorporación de recursos audiovisuales como apoyo al aprendizaje en carreras de ingeniería. In XX Congreso Argentino de Ciencias de la Computación (Buenos Aires, 2014). <http://hdl.handle.net/10915/42373>

[3] Moreno, E. J., Vera, P. M., Rodríguez, R. A., Giulianelli, D. A., Dogliotti, M. G., Cruzado, G. (2013) El Trabajo Colaborativo como Estrategia para Mejorar el Proceso de Enseñanza-Aprendizaje—Aplicado a la Enseñanza Inicial de Programación en el Ambiente Universitario. <http://conajisi.frc.utn.edu.ar/PDFsParaPublicar/1/schedConfs/4/204-481-1-DR.pdf>

[4] Rodríguez, R. A., Vera, P. M., Giulianelli, D., Valles, F. E., Dogliotti, M., Valles, G., Cruzado, G. (2014). Encouraging students participation in the classroom by taking advance of mobile devices and Ad Hoc networks. In Interactive Mobile Communication Technologies and Learning (IMCL), 2014 International Conference on (pp. 185-188). IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=7011128

[5] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: defining gamification,” Proc. 15th ..., pp. 9–15, 2011.

[6] Erenli K., “The impact of gamification: A recommendation of scenarios for education,” in 2012 15th International Conference on Interactive Collaborative Learning (ICL), 2012, pp. 1–8.

[7] Werbach, K (2013), Gamificación. Fundació Factor Humà. Unidad de Conocimiento.

[8] Barragán Piña, A. J., Ceada Garrido, Y., Andújar Márquez, J. M., Irigoyen Gordo, E., Gómez Garay, V., & Artaza Fano, F. (2015). Una propuesta para la motivación del alumnado de ingeniería mediante técnicas de gamificación.

[9] Arenas Paris R. (2014). Modelo para la Motivación del Aprendizaje de la Programación utilizando Gamification. Pontificia Universidad Javeriana. Bogotá D. C.

[10] LEE, Joey J., et al. Greenify: A Real-World Action Game for Climate Change Education. Simulation & Gaming, 2013
<http://tcgameslab.org/wp-content/uploads/2013/02/Lee-et-al.-Greenify-Simulation-and-Gaming-2013.pdf>

Herramienta de Evaluación de Hardware Pedagógico para la Educación Superior

Luis Mariano Mongelo, Bernardo Pedro Alcaraz, Marcelo Jorge Levi, Marcela Fabiana Dávila, Claudio Fabian Garabato, Clara Eva Sandoval y Gabriel Edgardo Panik.

Universidad Nacional de la Matanza (UNLaM)

luis.mongelo@gmail.com

RESUMEN

Al comenzar este proyecto nos propusimos analizar, proyectar y desarrollar una herramienta de evaluación de hardware pedagógico, destinada a colaborar con la clasificación, calificación y selección de herramientas de hardware; para su adquisición y posterior utilización por establecimientos de nivel superior, terciario y/o universitario. La misma permitirá a los profesores de las materias que incluyen recursos informáticos en su dictado, cargar un perfil de su materia y obtener una evaluación de la misma, recomendando o rechazando aquellos elementos de hardware que se deberían utilizar para la óptima impartición de clases. La gestación de esta herramienta contemplaba originalmente a una población de alumnos regulares del ciclo superior, pero se amplió en nuestro proyecto a contemplar también las necesidades de alumnos diversos funcionales (que presentan limitaciones visuales, auditivas o motrices).

CONTEXTO

Este proyecto se encuentra estructuralmente diseñado para su uso dentro del ámbito universitario de nuestra región en general, pero fuertemente orientada para responder a las necesidades geográficas y socio-culturales de los alumnos de la currícula de las carreras de grado en la UNLaM. Dicha universidad nacional ejecuta, coordina, evalúa y financia el proyecto en un 100%.

Palabras clave

Herramienta Pedagógica / Medición / Ponderación / Usabilidad Educacional.

1. INTRODUCCIÓN

Con la presente investigación, nuestro grupo espera sentar las bases para el desarrollo de una herramienta o aplicación, destinada a la clasificación y calificación del moderno hardware disponible para soporte en actividades educativas.

A fin de brindar una educación superior de calidad, es necesario no solo una clasificación de lo nuevo que existe en el área de dispositivos digitales para el uso en pedagogía, sino también la implementación de algún método de calificación de estos dispositivos, por categorías, alcances, costos y beneficios esperados.

Por ello, se investigaron diversas metodologías de evaluación para el desarrollo de planillas, tablas y diagramas, destinados a resolver esta problemática.

A partir de estas investigaciones, se tomó la decisión de desarrollar una aplicación de software que aunara todas las metodologías y mecanismos analizados, y nos permitiera desarrollar un primer prototipo que, cargado con la batería de datos necesaria, consiguiera determinar cuáles son los conjuntos de maquinarias o hardware pedagógicos necesarios que debemos adquirir para determinadas materias o cursos, cuyos datos hemos ingresado en esa batería y ponderado por mecanismos

automatizados, hasta brindarnos un informe presentable ante las autoridades de adquisición correspondientes.

Intentaremos reseñar cual fue el proceso de génesis de dicha herramienta, la cual mediante un sistema de metodologías paso a paso, bases de datos automatizadas y algoritmos de cálculo, permitirá clasificar en grupos específicos y evaluar rendimientos y alcances de los diversos dispositivos de hardware, que en la última década han llegado a los ámbitos de enseñanza superior y universitaria; y que son factibles de utilizar para la educación por vía informática.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto se inscribe en una línea de investigación inaugurada por la Universidad Nacional de La Matanza en el año 2014, referida a las Tecnologías de la Información y la Comunicación Aplicadas a la impartición de clases a alumnos diverso funcionales, mediante los trabajos previos “Tecnologías de la Información y la Comunicación Aplicada a las Personas con Discapacidad Motora Severa en la Universidad” (PIDC-C2-ECO-019) y “Técnicas de Evaluación y Selección de Hardware Pedagógicos para la Educación Superior” (C2-ING-017).

Se efectuó un relevamiento, evaluación y análisis de las diferentes propuestas tecnológicas en el área de dispositivos digitales compatibles con plataformas PC (la maquinaria más utilizada para la enseñanza superior a nivel nacional). A partir de ellas, se investigó como clasificar, ponderar, calificar y posteriormente eliminar aquellos dispositivos de hardware que no son aptos, para arribar al mejor modelo de hardware pedagógico a adquirir en la especialidad puntal para la cual ha sido evaluado.

Los dispositivos analizados podrían dividirse en dos grandes grupos:

- Hardware básico para el dictado de clases.
- Hardware pedagógico especializado.

En el primer grupo se encontraban:

- Impresoras
- Monitores
- Apuntadores (Mouse, Joystick, Trackball, Owl, Lápiz Óptico)
- Teclados
- Almacenamiento (Disco Rígido, Memorias, Grabadoras de CD, DVD, BluRay)
- Scanners

En el segundo se encontraban:

- Lápiz óptico
- Mimio o Pizarra Interactiva
- Tabletas
- Impresoras 3D
- Cascos de Realidad Virtual
- Dispositivos de Realidad Aumentada
- Elementos Cibernéticos o de Robótica
- Tarjetas Gráficas Digitalizadoras
- Capturadores y Digitalizadores Industriales
- Sintetizadores de Sonido o Mezcladoras de Audio
- Islas de Edición Digital

La aplicación de software de evaluación, por lo tanto, está inspirada en trabajos ya probados y ejecutados, como las investigaciones sobre el buen uso y adaptación de dispositivos periféricos digitales de Arnau Llombart (2012) para la Universidad de Cataluña, el estado actual de la ponderación informática, según los análisis de Wesley Colbert (2011) para el American Council for Education, y la valoración inteligente de dispositivos, propuesta desde un análisis HCI (Human - Computer Interaction) por Tur Costa del Departamento de Ciencias de la Computación y Matemáticas de la

Universitat Rovira I Virgili de Tarragona, que aúna de manera multidisciplinal las ciencias computacionales, la psicología cognitiva, la sociología y el diseño industrial. Estos aportes hacen a nuestro proyecto viable especialmente, para el desarrollo sostenible de recursos de aprendizaje focalizados en la región, con plena inclusión social dada la racionalización de sus gastos de implementación.

Siguiendo los criterios de clasificación de Eric Zabre (2011), combinados con la estandarización de Tur Costa (2009), llegamos a la conclusión de que estábamos hablando de cinco ítems a cumplimentar, bien definidos:

- Selección de parámetros (indicadores). En este paso se definieron una serie de parámetros que facilitan y mejoran el desarrollo de los sistemas pedagógicos para que la comparación y evaluación de los equipos y herramientas sean objetivos y útiles.
- Asignación de valores a cada parámetro.
- Asignación de pesos a los parámetros.
- Selección de herramientas a evaluar. En esta etapa se realizó un filtrado y selección de equipos, basados en plantillas pre-elaboradas, con valores de ponderación extraídos de las reglas de ponderación de Wesley Colbert (2011) y Tur Costa (2009) combinadas.
- Análisis y evaluación de cada equipo, con sus herramientas de software que los acompañan, sub-ponderando estos programas, como formando parte de un todo constitutivo del elemento de hardware.

Para la evaluación de cada herramienta de software y hardware se aplicaron algoritmos y fórmulas como las siguientes:

$$T = \sum T_p$$

$$T_p = C * P$$

Donde T es la calificación total obtenida por el equipo o herramienta de software que se está evaluando, T_p es la calificación obtenida por el equipo o herramienta de software para cada parámetro establecido, C es la calificación otorgada por las características del parámetro, y P es el peso asignado al parámetro que se está evaluando.

Con nuestro proyecto, hemos querido tomar estos conocimientos adquiridos, y llevarlos a la ponderación de conjuntos, aplicable en este caso a grupos de hardware, según la agrupación que consignamos anteriormente. Así, para el grupo de hardware “impresoras”, podríamos tener una valoración numérica del tipo (por ejemplo):

- Para todo el conjunto de impresoras a seleccionar, le asignaremos un valor del 1 al 100 dividido en 4 sub- conjuntos de modelos posibles.
- Del 1 al 25 se encontrarían las impresoras de carro o matriz de punto, o bien impresoras de chorro de tinta semi-obsoletas o de baja performance.
- Del 26 al 50 se encontrarían las impresoras de chorro de tinta de gama baja y media.
- Del 51 al 75 se encontrarían las impresoras de chorro de tinta de alta gama, impresoras de cera y por técnicas laser o de sublimación.
- Del 76 al 100 se encontrarían las impresoras industriales o plotters.

El usuario que utilizará la herramienta para la selección de hardware, irá punteando características del mismo, dentro de un conjunto de características que se presentarán en pantalla, y esto le irá asignando un valor o número de ponderación al conjunto de hardware correspondiente, para que un usuario superior (por ejemplo del Departamento de

Compras), pueda determinar a cual impresora del subgrupo ponderado corresponde este puntaje y aprobar o no la adquisición del mismo. Como se le requerirán datos extra, además de características del hardware (como por ejemplo, si se entregan apuntes impresos a los alumnos, si la impresión es predominantemente a color o blanco negro, si se utilizan transparencias, si se imprimen diagramas o planos, etc.), contribuye a determinar un perfil de la materia o curso, orientado a la selección fina del componente de hardware a adquirir.

3. NUESTRO PROTOTIPO

Hemos desarrollado un prototipo de herramienta de evaluación, en donde la pantalla de punteo correspondiente a nuestro ejemplo, es similar a la de la **Fig. 1**.

Fig. 1. Pantalla de punteo del cuestionario del conjunto Impresoras.

Allí se puntúan las respuestas que, sumadas en su ponderación, nos dará la clasificación del grupo correspondiente; y nos permitirá encuadrar, por la comparación contra una base de datos preexistente, cual es el hardware pedagógico más adecuado para adquirir. Existen pantallas de punteo diferentes para cada dispositivo de la clasificación original. Los datos de los dispositivos se cargan en otras pantallas similares.

Se realizará luego el procesamiento de estos datos, para formar un informe ponderado, que

mediante los algoritmos y cálculos correspondientes, determinará los elementos de hardware más adecuados para cubrir el perfil de la materia o curso a equipar.

Finalmente, esta evaluación podrá elevarse a las autoridades correspondientes, a las que llamaremos superusuarios. Los superusuarios son los que toman la decisión final, sobre el informe generado por la aplicación, de aprobar o no el presupuesto de compra de equipamientos, pudiendo realizar ajustes sobre lotes de posibilidades cercanas a los grupos de clasificación resultantes (por ejemplo, impresoras de prestaciones similares en resolución, sistema de impresión, costos de insumos, rapidez y confiabilidad a la que el sistema eligió como óptima, con un precio cercano por arriba y por debajo del de aquella).

4. RESULTADOS OBTENIDOS/ESPERADOS

Se espera, mediante el presente proyecto:

- Diseñar una herramienta de evaluación, que clasificará en grupos específicos y evaluará rendimientos y alcances de los diversos dispositivos de hardware necesarios para dar soporte a la enseñanza de la informática en la UNLaM u otros organismos de educación superior.
- Relevar, evaluar y analizar las diferentes propuestas tecnológicas en el área de dispositivos digitales compatibles con plataformas de computadoras por medios automatizados.
- Evaluar con la herramienta ProHardQual (generada como prototipo en investigaciones anteriores y que se pretende rediseñar en la presente investigación), el hardware pedagógico necesario para impartir la enseñanza en las materias que conforman la currícula de las carreras de grado en UNLaM.
- Evaluar y clasificar con la herramienta ProHardQual el hardware pedagógico

necesario y disponible en nuestro país para alumnos diversos funcionales.

- -Informar los modelos, especificaciones técnicas incluidas, pre-evaluadas y calificadas a fin de determinar los costes generales, promediados con los valores del mercado informático de nuestra región y actualizados mediante bases de datos, de la compra o adquisición de dichos hardwares.

5. FORMACION DE RECURSOS HUMANOS

Nuestro equipo es eminentemente multidisciplinario, y se encuentra constituido por docentes investigadores de los Departamentos de Ingeniería e Investigaciones Tecnológicas, Ciencias Económicas y la Secretaría de Pedagogía Universitaria. Se trata de ingenieros y licenciados en sistemas, programadores y técnicos, contando en nuestro equipo con personal altamente capacitado en el

área de estudio, como un Magister en Gestión Ambiental, un Licenciado en Sistemas de Computación con una especialización en Tecnología Educativa, y una Licenciada en Gestión Educativa con una especialización en Docencia de la Educación Superior; y cuya Tesis de Grado, aprobada dentro de los últimos dos años, versa sobre “La Discapacidad Motora Severa y la Alta Tecnología”. La misma especialista se encuentra realizando una Maestría en Gestión Educativa, con tesis a entregar este año titulada “Inclusión y Accesibilidad en la Cátedra de Computación Transversal de la UNLaM” y analiza las condiciones de accesibilidad relacionadas con la atención de la diversidad funcional en la UNLaM para garantizar la inclusión educativa en la cátedra de Computación Transversal común a todas las carreras de nuestra casa de altos estudios.

BIBLIOGRAFIA

- BLACKMAN, R. (2009). *Nuevos Desarrollos para el Nuevo Mundo Digital*. México: Ediciones Orbe.
- DEPIRENNE, A. (2008). *Administración de la Educación Virtual*. México: Publicaciones Planeta Inteligente.
- ANDERSON, R.E. (2011). *Social impacts of computing: Codes of professional ethics*. Social Science Computing Review. Vol. 10, No. 2, pp.453-469. Estados Unidos.
- ARNAU LLOMBART. V. (2012). *Periféricos y Dispositivos Digitales*. España: Universidad de Cataluña.
- SANCHEZ SERANTES. V. (2012). *La Biblia del Hardware*. Argentina: Ediciones Micropunto.
- COTTINO, DAMIAN. (2012). *Hardware desde Cero*. Argentina: Ediciones Micropunto.
- WESLEY COLBERT. S. (2011). *Educational Hardware Today*. American Council for Education. Estados Unidos.
- TUR COSTA, A. (2009). *Valoración de Dispositivos de Hardware*. España: Ediciones Planet-Rv.
- ERIC ZABRE, B. E ISLAS, P. (2011). *Evaluación de herramientas de hardware y software para el desarrollo de aplicaciones*. España: Wiley-Interscience.
- MUÑIZ, L. (2012). *Diseñar Cuadros de Mando con Excel utilizando Tablas Dinámicas*. España: Profit Editorial.
- SZNAJDLEDER, P. (2013). *Algoritmos a fondo*. México: Alfaomega Grupo Editor.

Implementación de un EVEA Institucional para Enriquecer la Enseñanza de Pregrado, Grado y Posgrado de la UNTDF

Depetris Beatriz, Zangara Alejandra, Feierherd Guillermo, Pendenti Horacio, Rodríguez Juan Manuel, Romano Lucas, Blanco Claudio, Rojas Sergio

Instituto de Desarrollo Económico e Innovación
 Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
 Hipolito Irigoyen 880 - Ushuaia - Tierra del Fuego
 {bdepetris, gfeierherd, hpendenti, jrodriguez, lromano, cblanco}@untdf.edu.ar
 alejandra.zangara@gmail.com, sergiorc.ush@gmail.com

RESUMEN

Los Sistemas de Administración del Aprendizaje (LMS, por sus siglas en inglés) o EVEAs (Entornos Virtuales de Enseñanza y de Aprendizaje) utilizados en sus inicios para resolver los problemas que planteaba la Educación a Distancia, fueron rápidamente adoptados para complementar la Educación Presencial, dando origen a diferentes modalidades que extienden las posibilidades del aula presencial, incorporando opciones de enseñanza mediadas por las herramientas ofrecidas en estos entornos.

En esta evolución han ido surgiendo un número significativo de EVEAs, que han logrado distintos niveles de trascendencia.

Dadas las características particulares de la UNTDF, el uso de estas tecnologías resulta imprescindible. La selección de un EVEA institucional constituye el primer paso para avanzar ordenadamente en ese sentido.

El proyecto propone analizar un subconjunto de los EVEAs disponibles, en particular los de código abierto, considerando aspectos pedagógicos y tecnológicos para poder recomendar el que mejor se adecue al modelo

pedagógico planteado por la UNTDF [1]. Además, a partir de un relevamiento de los conocimientos del cuerpo docente sobre el tema, se contempla desarrollar un plan de capacitación y realizar un conjunto de experiencias piloto en distintas cátedras con asistencia del equipo de proyecto.

Palabras clave: EVEAs, evaluación de EVEAs, criterios de selección de EVEAs, mediación de la enseñanza

CONTEXTO

El presente proyecto forma parte del “Área de Investigación 8 - Desarrollo Informático”, del Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego (UNTDF).

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2016 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral en trámite).

El financiamiento requerido es suministrado por la UNTDF.

El proyecto se desarrolla formalmente desde el 01/03/2017 hasta el

28/02/2019.

1. INTRODUCCIÓN

Un Entorno Virtual de Enseñanza y de Aprendizaje (EVEA) es una aplicación informática diseñada para facilitar la comunicación pedagógica entre los actores del proceso educativo [2]. Combina una variedad de herramientas con la finalidad de dar soporte a profesores y estudiantes y, al mismo tiempo, optimizar las distintas fases del proceso de enseñanza y de aprendizaje. Fueron diseñados con el propósito de facilitar la comunicación pedagógica entre los participantes en un proceso educativo, sea éste completamente a distancia, presencial, o de naturaleza mixta, es decir, un proceso que combina ambas modalidades en diversas proporciones (blended). Funcionan utilizando redes telemáticas como soporte, principalmente Internet. [3]

Si bien en un primer momento los EVEAs se utilizaron casi con exclusividad en la educación a distancia, hace mucho tiempo que se han incorporado muy fuertemente a la educación presencial, generando así las modalidades conocidas como aula extendida, aula invertida y, con un mayor grado de mediación, propuestas de “blended learning”. El propósito de un EVEA es mediar una propuesta educativa, organizar los contenidos en diferentes materiales de estudio, facilitar la comunicación entre los participantes de un proceso educativo (fundamentalmente entre docentes y alumnos, alumnos entre sí y docentes entre sí), realizar un constante seguimiento de la realización de actividades intermedias y, si se decide, de la evaluación final.

Con el tiempo, la primera generación de entornos, cuyo objetivo

principal (y casi único) era permitir la distribución de materiales y la evaluación mediante pruebas automáticas, ha dado paso a una segunda generación, inspirada en los nuevos conocimientos sobre cómo se aprende usando recursos en línea.

Por medio de estos entornos se ha logrado aprovechar las características de accesibilidad y cooperación entre los usuarios de la red, orientando fundamentalmente el proceso educativo hacia nuevas maneras de aprender, apoyadas en el trabajo colaborativo, acompañadas de una serie de recursos, más allá de la mera presencia del docente. [4]

En la actualidad existe un amplio abanico de entornos virtuales orientados a la educación superior. Los hay con características de software propietario, de software libre y en la nube (aunque en este último caso no siempre son considerados como plataformas LMS o EVEAs propiamente dichas, pues su mayor utilidad es la de permitir el apoyo a la clase presencial así como el desarrollo de MOOC (Cursos online abiertos y masivos). [5]

Las universidades y los docentes pueden elegir sus EVEAs para impartir un modelo de enseñanza E-Learning. Desde la aparición de la llamada WEB 2.0 todos estos EVEAs han evolucionado incluyendo nuevas herramientas colaborativas, como blogs, foros, wikis, etc. [6]

Cada uno de ellos cuenta con sus propias herramientas y funcionalidades. No es posible encontrarlas a todas en una única plataforma y, por esta razón, es importante que evaluemos cuáles resultan prioritarias teniendo en cuenta el modelo educativo que queremos implementar.

Si se consideran los desarrollos propios de distintas universidades, el

número de EVEAs asciende a varias centenas. En algunos casos se habla de más de ciento treinta. Si nos limitamos a los que han logrado cierta preponderancia en el ámbito internacional encontramos al menos 40 de ellos. [7][8]

Todos ofrecen un conjunto común de prestaciones y otras que los diferencian. También evolucionan de distinta manera.

Teniendo en cuenta los distintos módulos, componentes, recursos o herramientas incorporadas, los entornos privilegian distintas funcionalidades. Actualmente es posible observar tres tendencias bien definidas en sus prestaciones: [9] citado en [3]

- Entornos centrados en la creación, gestión y distribución de contenidos que además incorporan algunas herramientas de comunicación.
- Entornos centrados en la comunicación y las actividades de enseñanza / aprendizaje que incluyen además herramientas para gestionar materiales.
- Entornos de trabajo en grupo para comunidades académicas que agregan algunas funcionalidades utilizables en la enseñanza.

Es evidente que detrás de estas características subyacen distintas concepciones del proceso pedagógico. Por lo tanto, entre las actividades a desarrollar se ha planteado indagar sobre las prácticas de enseñanza que realizan los docentes de la UNTDF, en especial las referidas a la inclusión de TICs. Esta indagación conllevará a un modelo de enseñanza en el que se basará la capacitación docente y la implementación del entorno seleccionado.

Por lo tanto un buen entorno debe ser lo suficientemente flexible como para no imponer un modelo pedagógico

particular, sino posibilitar la aplicación de la mayoría de ellos, a fin de que los docentes y los estudiantes lo experimenten como un aumento de posibilidades y no como una limitación, un elemento de distorsión o un freno a su creatividad y a su capacidad de innovación. [9] citado en [3]

Resulta claro entonces que no debe restringirse la selección solamente a los aspectos técnicos. Al respecto señala Clarenc: *“Es necesario contemplar a los LMS desde la perspectiva del aprendizaje y no desde la tecnológica, ya que los aspectos técnicos deben ser considerados recursos a ser utilizados en el proceso de implementación.”* [10]

Teniendo en cuenta esta diversidad la decisión de elegir un EVEA para uso institucional no es sencilla.

Por otra parte, debe tenerse en cuenta que, si bien existen muchos EVEAs de uso gratuito, la inversión que deben realizar las instituciones que los adopten es significativa. Uno de los rubros a los que deben destinarse muchos recursos es la preparación del cuerpo académico para un uso adecuado y eficiente del entorno. Por su parte, los integrantes del cuerpo académico también deben invertir una parte de su tiempo en la reformulación de sus cursos (a fin de aprovechar las ventajas de los EVEAs y las TICs en general), preparación de los materiales adecuados, nuevas actividades para los alumnos, etc. etc.

Eventualmente podría requerirse el agregado de alguna funcionalidad particular, lo que implica realizar todas las tareas que demanda un desarrollo de software.

La conclusión es que, más allá de la “supuesta” gratuidad inicial, la inversión es importante y, en consecuencia, debe estar precedida por un proceso de

evaluación destinado a valorar la calidad del EVEA y a determinar objetivamente las potencialidades de cada uno. Mediante esta evaluación podremos determinar su filosofía pedagógica, su flexibilidad didáctica, sus funcionalidades y su usabilidad. [3] Los principales indicadores que pueden tenerse en cuenta a la hora analizar un entorno virtual están relacionados con el tipo de actividad educativa a desarrollar (cursos, carreras, comunidades virtuales), la modalidad, las características de los destinatarios y las posibilidades y limitaciones técnicas disponibles (ancho de banda, equipamiento informático y disponibilidad de tiempo, entre otros).

A su vez, a fin de formular un plan de implementación que permita el éxito del proyecto de introducción del EVEA elegido, debe tenerse en cuenta el estado actual de conocimientos del cuerpo académico. Esto permitirá desarrollar un plan de actualización de esos conocimientos así como los recursos necesarios para acompañar a las experiencias piloto.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Entornos virtuales de enseñanza y aprendizaje. Proyectos que los incorporan, metodologías, desarrollos, evaluación de su calidad y experiencias.
- Formación de recursos humanos en el uso de TICs.
- Trabajo colaborativo mediado por TICs. Conceptualización, análisis y

desarrollo de software y metodologías.

- Dispositivos móviles como soporte para el trabajo colaborativo y para prácticas de aula extendida y blended-learning.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto se ha iniciado el 1º de marzo de 2017 por lo que a la fecha no pueden enumerarse resultados obtenidos. El grupo de investigación ha comenzado a trabajar en la 1ra. actividad del proyecto que consiste en el relevamiento de las funcionalidades de los entornos a partir de la bibliografía existente y la experimentación con aquellos que estén disponibles y que sean considerados de interés para el proyecto.

El objetivo general del proyecto es seleccionar un entorno virtual de enseñanza y aprendizaje (EVEA) a fin de que sea utilizado como entorno institucional de la UNTDF y contribuir al proceso de implementación.

Algunos de los objetivos específicos son:

- Analizar las distintas funcionalidades provistas por los EVEA.
- Establecer el conjunto de criterios para la evaluación de los EVEAs.
- Aplicar los criterios a los entornos analizados.
- Recomendar un entorno, acompañado de una fundamentación relacionada con el modelo de enseñanza y sus prestaciones como herramienta de mediación educativa.
- Indagar sobre las prácticas de enseñanza del cuerpo académico de la UNTDF, y en particular sobre sus conocimientos de TICs y el modo en

- que las aplican.
- Sugerir un plan de actualización de conocimientos para su uso.
- Implementar instancias de capacitación.
- Brindar soporte a un conjunto de experiencias de aplicación.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está formado por docentes - investigadores, de los cuales uno (1) es Magister en Política y Gestión de la Ciencia y la Tecnología, dos (2) son Especialistas en Docencia Universitaria, tres (3) son Licenciados en Informática, uno (1) es Analista Programador Universitario y uno (1) es alumno avanzado de la carrera Licenciatura en Sistemas.

La Mg. Zangara está desarrollando su tesis para acceder al Doctorado en Ciencias Informáticas en la Facultad de Informática de la UNLP.

El licenciado Romano se encuentra cursando la maestría en Tecnología Informática Aplicada en Educación en la Facultad de Informática de la UNLP.

El licenciado Pendentí está cursando la maestría en Ingeniería de Software en la Facultad de Informática de la UNLP.

El alumno Rojas está desarrollando su tesis de grado, bajo la dirección de las docentes Depetris y Zangara sobre temas afines al proyecto.

Uno de los beneficios del proyecto es la consolidación de parte del grupo de I/D, que podrá asesorar y capacitar al personal docente en la implementación del EVEA seleccionado.

REFERENCIAS

[1] UNTDF. (2012, Accedido el 3 de

marzo de 2017). Estatuto de la UNTDF [Online]. Disponible en www.untdf.edu.ar/reglamentacion.

[2] Patricia Avila M., Martha Diana Bosco H.. (2001, Accedido el 27 de febrero de 2017). Virtual Environment for Learning, A New Experience [Online]. Disponible en

http://investigacion.ilce.edu.mx/panel_control/doc/c37ambientes.pdf.

[3] Ariel Ferreira Szpiniak, Cecilia V. Sanz. (2009, Marzo 2017) Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje. La importancia de la usabilidad. Revista TEyET [Online] N° 4 pp. 10 - 14. Disponible en: <http://teyet-revista.info.unlp.edu.ar/numero-4-octubre-de-2009/>

[4] M. A. López Carrasco, Aprendizaje, competencias y TIC (eBook). Pearson, 2012.

[6] Alvarez Diego. Plataformas de enseñanza virtual libres y sus características de extensión: Desarrollo de un bloque para la gestión de tutorías en Moodle. Tesis de Ingeniería. I.I. Universidad de Alcalá. Alcalá de Henares, España . Disponible en <http://www3.uah.es/libretics/files/Tutorias.pdf>. Accedido el 14 de octubre de 2016

[7] Baumgartner P, Häfele H., et al. E-Learning Praxishandbuch: Auswahl von erntplattformen. Marktübersicht - Funktionen Fachbegriffe. Innsbruck-Wien, StudienVerlag. 2002.

[8] Hernández Schäfer (2002, marzo 2017) L.E. Estado actual y futuro de las plataformas e-learning: estándares y especificaciones IMS. Universidad Católica del Norte, Chile. [Online] Disponible en <http://repositorial.cuaed.unam.mx:8080/jspui/bitstream/123456789/2974/1/1-cl-Luis%20Hern%C3%A1ndez%20Sc%C3%A4fe>

[r-Gim%20Ivy.pdf](#)

[9] Adell J. (2004, marzo 2017) Selección de un entorno virtual de enseñanza/aprendizaje de código fuente abierto para la Universitat Jaume I. Centre d'Educació i Noves Tecnologies de la UJI con la colaboración del Servei d'Informàtica y del Gabinet Tècnic del Rectorat. [Online].

Disponible en:
http://cent.uji.es/doc/eveauji_es.pdf

[10] Clarenc, C. A. Instrumento de evaluación y selección de sistemas de gestión de aprendizaje y otros materiales digitales: Medición y ponderación de LMS y CLMS, recursos educativos digitales y herramientas o sitios de la WEB 3.0. Presentado en Congreso Virtual Mundial de e-Learning: Grupo GEIPITE. Octubre 2013

Disponible en scribd:

<http://es.scribd.com/doc/175057118/Instrumentoevaluacion->. Accedido el 17 de marzo de 2017.

Iniciativa 3D en la Escuela

Villagra A., Pandolfi D., Varas V., García M., Serón N., Mercado V., Orozco S., Valdéz J.

Laboratorio de Tecnologías Emergentes (LabTEM)

Instituto de Tecnología Aplicada (ITA) - Unidad Académica Caleta Olivia

Universidad Nacional de la Patagonia Austral

{avillagra, dpandolfi, vvaras}@uaco.unpa.edu.ar, mariadel_garcia@yahoo.com.ar, {nseron, vmercado, sorozco, jcvaldez}@uaco.unpa.edu.ar

Resumen

Los últimos avances tecnológicos están transformando muchos ámbitos de la sociedad y su impronta tiene un especial eco dentro del marco del conocimiento y la educación, convirtiéndose en uno de los requerimientos básicos para el desarrollo de los procesos de enseñanza y aprendizaje. La incorporación de la impresión 3D al aula pone al alcance de los alumnos las mismas tecnologías de vanguardia que encontrarán a lo largo de su vida. Es un puente a los desafíos del mañana. Las impresoras 3D en las aulas mejoran la capacidad de resolver problemas y estimulan la creatividad de los estudiantes. En esta línea de trabajo se pretende aplicar y estudiar los resultados de la utilización de esta tecnología en el ámbito de la educación como una herramienta de apoyo al proceso de enseñanza-aprendizaje.

Palabras clave: aprendizaje, enseñanza, tecnología 3D, arduinos, hardware y software libre.

Contexto

La línea de trabajo se lleva a cabo en el Laboratorio de Tecnologías Emergentes (LabTEM), Instituto de Tecnología Aplicada (ITA) de la Unidad Académica Ca-

leta Olivia Universidad Nacional de la Patagonia Austral, en el marco del proyecto "Iniciativa 3D en la escuela".

Introducción

"Todos los niños/as y jóvenes del mundo, con sus fortalezas y debilidades individuales, con sus esperanzas y expectativas, tienen el derecho a la educación. No son los sistemas educativos los que tienen derecho a cierto tipo de niños/as. Es por ello, que es el sistema educativo de un país el que debe ajustarse para satisfacer las necesidades de todos los niños/as y jóvenes". (B. Lindqvist, UN-Rapporteur, 1994).

La UNESCO define la educación inclusiva en su documento conceptual así: "La inclusión se ve como el proceso de identificar y responder a la diversidad de las necesidades de todos los estudiantes a través de la mayor participación en el aprendizaje, las culturas y las comunidades, y reduciendo la exclusión en la educación. Involucra cambios y modificaciones en contenidos, aproximaciones, estructuras y estrategias, con una visión común que incluye a todos los niño/as del rango de edad apropiado y la convicción de que es la responsabilidad del sistema regular, educar a todos los niño/as. La educación para la inclusión se basa en el principio de que cada niño/a tiene características, intereses, capacidades y necesidades de aprendizaje

distintos y deben ser los sistemas educativos los que están diseñados, y los programas educativos puestos en marcha, teniendo en cuenta la amplia diversidad de dichas características y necesidades. Se ocupa de aportar respuestas pertinentes a toda la gama de necesidades educativas en contextos pedagógicos escolares y extraescolares.

Las escuelas especiales ayudan en el aprendizaje de alumnos con capacidades diferentes y la inclusión de diferentes tecnologías mejora el proceso de aprendizaje de los alumnos. La tecnología avanza a pasos agigantados y está transformando muchos ámbitos de la sociedad y en particular en la educación, convirtiéndose en uno de los requerimientos básicos para el desarrollo de los procesos de enseñanza y aprendizaje. De la irrupción de las tecnologías en la educación surge el reto de conocer, entender e implementar el modo en que éstos medios tecnológicos (notebooks, netbooks, tablets, plataformas en línea, webs, blogs, entre otros) nos llevan a dar soporte a las actividades de enseñanza-aprendizaje que buscan el desarrollo de capacidades y habilidades en los alumnos, de manera que los medios disponibles sean utilizados como una vía para adquirir conocimiento e interpretar la información adquirida.

Los estudios realizados en los distintos países (Reino Unido, Portugal, Suecia, Alemania y España), dentro del marco del proyecto europeo de JuxtaLearn, indican que las actividades que emplean el vídeo como herramienta educativa consiguen que el aprendizaje sea progresivo y reflexivo, a la vez que fomentan la creatividad y curiosidad de los alumnos. En todas las experiencias educativas realizadas con estudiantes de Primaria, Secundaria, Bachillerato y Universidad se resalta el poder innovador de este modelo de aprendizaje. En este caso y basándose en estas ideas [5]

se incorpora la tecnología 3D en la escuela.

Líneas de investigación y desarrollo

En esta sección se describe la línea de investigación que se lleva a cabo en el proyecto.

Los rápidos avances en las tecnologías electrónicas han dado lugar al desarrollo de una variedad de nuevas y económicas placas electrónicas con capacidades de automatización y control que ofrecen oportunidades para su implementación en múltiples áreas de aplicaciones.

La placa arduino se desarrolló originalmente en el año 2005 en el Interaction Design Institute, de Ivrea Italia, como una plataforma de hardware de código abierto. Los fundadores del proyecto fueron Massimo Banzi y David Cuartielles. Las placas arduino consisten de un micro-controlador programable, con capacidad de expansión a través de una placa denominada shield (donde se pueden conectar uno o varios sensores) y un entorno de desarrollo de programación para la creación de software personalizado. Tanto la placa de circuito electrónico, así como el software de programación, son de código abierto y de libre acceso.

Originalmente, las placas arduino fueron utilizadas para el desarrollo de una red de sensores en un campo de cultivos para controlar el estado de humedad de suelo, la programación del riego, el estudio del uso de agua en los cultivos, para medir los niveles diarios de agua por evaporación, y para controlar los parámetros ambientales en condiciones boscosas. Estos estudios ofrecen una guía para otros investigadores en el desarrollo de sensores de bajo costo,

sistemas de monitoreo y automatización en diversos campos de aplicación [7].

La tecnología avanza a pasos agigantados y está transformando muchos ámbitos de la sociedad y en particular en la educación. La impresión tridimensional (impresión 3D) para usuarios comunes fue un concepto introducido por el software libre de las impresoras 3D con el proyecto RepRap que se inició en 2004 [8] y el Fab@home [6].

El bajo costo de las componentes (como son las placas arduinos) para su fabricación ha impulsado aún más la migración de esta tecnología para el usuario doméstico. Permitiéndole hacer las cosas ellos mismos teniendo acceso a la impresión 3D. Además, el bajo costo de estas impresoras ha abierto la puerta para el uso extendido entre las pequeñas empresas y en las instituciones educativas.

La irrupción de las tecnologías en la educación se ha convertido en uno de los basamentos básicos para el desarrollo de los procesos de enseñanza y aprendizaje innovadores. Surge el reto de conocer, entender e implementar el modo en que éstos medios tecnológicos (notebooks, netbooks, tablets, plataformas en línea, webs, blogs, entre otros) nos llevan a dar soporte a las actividades de enseñanza aprendizaje que buscan el desarrollo de capacidades y habilidades en los alumnos, de manera que los medios disponibles sean utilizados como una vía para adquirir conocimiento e interpretar la información adquirida. En el aula los modelos diseñados en la computadora pueden ser impresos y prototipados en tres dimensiones (3D). La tecnología de impresión 3D en la educación es importante, por varias razones: (a) Ayuda a los profesores, al poder visualizar en 3D conceptos que son difíciles de ilustrar de otra forma, y les permite generar más interés mostrándole a los alumnos objetos reales;

(b) Posibilita producir diseños que son imposibles de construir. Los alumnos ya no están más limitados a la hora de diseñar y pensar nuevos objetos; (c) Alienta y mejora el aprender haciendo y abre una ruta al diseño iterativo: los estudiantes pueden diseñar partes en 3D, imprimirlas, testearlas, evaluarlas y, si no funcionan, trabajar con ellas de nuevo. Esto incrementa la innovación en los diseños producidos; (d) Permite clases interactivas y fomenta el trabajo en equipo y el diseño colaborativo, permitiendo realizar trabajos en conjunto con otras materias o especialidades; (e) Está fuertemente vinculado con el aprendizaje personal y las habilidades de pensamiento. Se forman estudiantes creativos y reflexivos.

En [1, 2, 3, 4, 9, 10] se puede profundizar sobre diversas experiencias de uso de impresoras 3D en diferentes niveles educativos.

Este proyecto combina dos estrategias innovadoras: a) implementar el proceso a partir de las consignas de crear, estudiar y compartir en un modelo de aprendizaje en el aula y b) usar tecnologías (impresoras 3D) altamente motivadoras en el proceso de aplicación del conocimiento.

Resultados obtenidos/esperados

Hacer que los estudiantes entiendan las aplicaciones y el potencial de esta nueva tecnología es, sin dudas, clave para su desarrollo. Además, una de las mayores fortalezas de las impresoras 3D es su habilidad para solucionar problemas reales en el mundo físico, lo cual posibilita lograr un cambio de mentalidad en los alumnos, permitiéndoles cambiar su entorno resolviendo problemáticas reales.

Las impresoras 3D en las aulas pueden generar espacios comunes de creación donde los estudiantes se reúnen para crear objetos nuevos. Esto puede ser un nuevo incentivo

en el proceso creativo de los jóvenes, extendiendo la cultura de aprender haciendo. Algo muy importante para mencionar es que las impresoras 3D posibilitan el trabajo en conjunto entre diferentes materias. Pudiendo lograr cruzar proyectos de arte con clases de matemática, o de física con materiales, siendo el producto impreso la conjunción de los conocimientos adquiridos en ambos estudios.

Otra ventaja que surge de la utilización de las impresoras 3D en la educación tiene que ver con los chicos con problemas de atención: al ver resultados rápidos y tangibles en sus clases, los alumnos mantienen el interés en la lección. Además, la posibilidad de hacer diseños más complejos hace que los chicos se involucren mucho más y estén incentivados a seguir creando.

En esta línea de trabajo se pretende: a) Transferir el conocimiento sobre el modelo de aprendizaje basado en crear, estudiar y compartir a docentes y alumnos de la escuela especial para que ellos puedan aplicar el uso de nuevas tecnologías en el aula. b) Transferir el conocimiento sobre la utilización de esta tecnología a docentes y alumnos de la escuela especial para que ellos puedan aplicar el uso de nuevas tecnologías en el aula.

En particular se intenta: i) Generar en el ámbito de la Escuela Especial Nro. 8 "Ventana a la vida" de la ciudad de Caleta Olivia, provincia de Santa Cruz, una capacidad propia y el saber necesario para el desarrollo de dispositivos, para el aprovechamiento de los mismos en distintos proyectos y asignaturas. ii) Transferir conocimiento científico tecnológico que permita

a los integrantes del proyecto replicar procesos de aprendizaje, creatividad y cooperación.

iii) Fomentar el trabajo colaborativo y participativo de todos los integrantes del proyecto.

iv) Transferir para lograr el auto sustento de miembros de la comunidad educativa para su utilización en el proceso de aprendizaje de los alumnos.

v) Elaborar material didáctico como forma de divulgar el conocimiento socialmente construido, para su distribución y apropiación en esa comunidad y en otras.

Formación de recursos humanos

Un integrante de este proyecto está desarrollando su Tesis de Maestría en temáticas afines.

Se cuenta con tres becarios alumnos de las carreras de Sistemas.

Referencias

- [1] Colegrove, P. T. (2012). Beyond the Trend: 3D Printers Transforming Learning and Knowledge Creation. *Internet Librarian, October*.
- [2] Czapka, J. T., Moeinzadeh, M. H., & Leake, J. M. (2002). Application of rapid prototyping technology to improve spatial visualization. *age, 7, 1*.
- [3] Gonzalez-Gomez, J., Valero-Gomez, A., Prieto-Moreno, A., & Abderrahim, M. (2012). A new open source 3d-printable mobile robotic platform for education. In *Advances in autonomous mini robots* (pp. 49-62). Springer Berlin Heidelberg.
- [4] Johnson, W. M., Coates, C. W., Hager, P., & Stevens, N. (2009). Employing rapid

prototyping in a first-year engineering graphics course. In *Proceedings of 2009 ASEE southeast section conference, Marietta, USA*.

[5] Llinás, P., Haya, P., Gutierrez, M. A., Martín, E., Castellanos, J., Hernán, I., & Urquiza, J. (2014, September). ClipIt: Supporting social reflective learning through student-made educational videos. In *European Conference on Technology Enhanced Learning* (pp. 502-505). Springer International Publishing.

[6] Malone, E., & Lipson, H. (2007). Fab@Home: the personal desktop fabricator kit. *Rapid Prototyping Journal*, 13(4), 245-255.

[7] Melgar, E. R., & Diez, C. C. (2012). *Arduino and kinect projects: Design, build, blow their minds*. Apress.

[8] Pearce, J. M., Blair, C. M., Laciak, K. J., Andrews, R., Nosrat, A., & Zelenika-Zovko, I. (2010). 3-D printing of open source appropriate technologies for self-directed sustainable development. *Journal of Sustainable Development*, 3(4), 17.

[9] Scalfani, V. F., & Sahib, J. (2013). A model for managing 3D printing services in academic libraries. *Issues in Science and Technology Librarianship*, 72(Spring), 1-13.

[10] Valero-Gomez, A., Gonzalez-Gomez, J., Gonzalez-Pacheco, V., & Salichs, M. A. (2012, April). Printable creativity in plastic valley uc3m. In *Global Engineering Education Conference (EDUCON), 2012 IEEE* (pp. 1-9). IEEE.

Instalación de UMremoto3D

Sattolo Iris¹, Lipera Liliana¹, Romero Juan Carlos¹, Cos Susana¹, Altube Alejandro¹,
Minutella Darío¹

Facultad de Informática Ciencias de la Comunicación y Técnicas Especiales (FICCTE) ¹
Universidad de Morón

Cabildo 134, (B1708JPD) Morón, Buenos Aires, Argentina. TE 56272000 Int 189
iris.sattolo@gmail.com, lipera@unimoron.edu.ar, juancarlosromer@gmail.com, sualcos@gmail.com,
alealtube@hotmail.com, minutelladario@gmail.com

Resumen

Las nuevas tecnologías basadas en la multimedia e Internet ofrecen modos novedosos de aprender y enseñar. Una tendencia que avanza en las Universidades del mundo, es la construcción de espacios virtuales tridimensionales. Estamos transitando un nuevo escenario mundial en donde el valor agregado de la información y del conocimiento resultan determinantes. En consonancia con los retos que implican la introducción de nuevas tecnologías, la Universidad de Morón, está haciendo frente a esos desafíos. Disponer de un metaverso propio que permita montar ambientes virtuales inmersivos, es la meta fijada a corto plazo. Se presentan en este artículo algunos desarrollos de universidades españolas, y los logros alcanzados en la Universidad de Morón en cuanto a Mundos Virtuales Inmersivos. Se exponen las metas fijadas para el presente año.

Palabras clave:

Mundos virtuales inmersivos, metaversos, interacción.

Contexto

El presente trabajo se enmarca dentro del proyecto de investigación y desarrollo PID 01-

003-16, llevado a cabo por la FICCTE (Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales), el cual fuera aprobado por la Secretaria de Ciencia y Tecnología (SECYT) de la Universidad de Morón. Este proyecto, “Ambientes virtuales inmersivos, elementos interactivos” se encuentra en ejecución desde comienzos del año 2016, durante el cual los resultados obtenidos fueron plasmados en los congresos TE&ET 2016 (Tecnologías en Educación y Educación en Tecnologías año 2016) y CACIC 2016 (Congreso Argentino de Ciencias de la computación)

Introducción

Un mundo virtual es una simulación o representación tridimensional de un espacio en un ordenador. Estos espacios, reales o imaginarios, pueden simbolizar escenografías geográficas, ciudades, hogares, escuelas, etc. Dentro de un mundo virtual la persona se desplaza e interactúa con el resto de los usuarios a través de un avatar.

Actualmente, los mundos virtuales inmersivos, ambientes virtuales inmersivos o metaversos, de ahora en más MVI, están siendo utilizados en reconocidas instituciones universitarias de nivel mundial.

Probablemente la plataforma virtual más utilizada en el mundo sea Second life pero tiene algunas limitaciones cuando es implementada en

educación. El primer limitante es su condición de entorno cerrado, debiendo pagar un alquiler para pertenecer al mismo. Por otro lado, se pueden encontrar espacios inapropiados en su contenido para estudiantes.

Otra plataforma que se utiliza en estos desarrollos es OpenSim. Este proyecto surgió en el año 2007 con la propuesta de crear un servidor de aplicaciones 3D, analizando la estructura del cliente de Second Life (ingeniería inversa), cuyas características resultaron interesantes a nivel educación. Ser software libre (Licencia BSD), tener una estructura modular, soportar múltiples visores o clientes, y estar escrito en C#, entre otros rasgos, representó para las universidades poder construir sus propios espacios (islas) sin tener que pagar por los terrenos y tampoco por las texturas y objetos que ofrecía Second Life. [3].

Como antecedentes de MVI montados sobre OpenSim se puede citar a la Universidad Autónoma de Madrid, que a través de su página VirtUAM [1], ofrece los avances y temas desarrollados sobre mundos virtuales en los cuales están trabajando.

La Universidad Politécnica de Madrid posee una plataforma virtual de nombre GridLabUPM, la cual consta de trece laboratorios virtuales y dos regiones institucionales, una de ellas como punto de encuentro virtual y otra como tutorial para adaptarse al entorno. Sergio López et al [2] cuenta la experiencia en su trabajo “Laboratorio remoto eLab3D: un entorno innovador para el aprendizaje de competencias prácticas en electrónica”, laboratorio remoto en el cual se realizan prácticas de electrónica y diseño de circuitos.

A nivel nacional, la Universidad Nacional del Noroeste de Bs. As., está realizando investigaciones las que se reflejan en “Entornos virtuales 3D, una propuesta educativa innovadora” [12]

En el ámbito de la Universidad de Morón se comenzó a trabajar con MVI desde el año 2013, en el cual se constituyó un grupo de autoformación que investigó las plataformas existentes. Éste presentó los avances en el

Congreso Argentino de Ciencias de la Computación (CACIC 2013) [4]. Desde el año 2014 y por espacio de dos años se trabajó en el proyecto de investigación y desarrollo PID 01-002-14 “Ambientes virtuales inmersivos de aprendizaje, orientados a la capacitación y al desarrollo de simulaciones en situaciones de riesgo”. Los logros fueron expuestos en los congresos de Tecnología en Educación y Educación en Tecnología, TE&ET 2014 [5], TE&ET 2015 [6] y JATIC 2015 [7].

Desde el año 2016 se está trabajando en el proyecto “Elementos interactivos en los mundos virtuales 3D”. Se presentaron trabajos en TE&ET 2016 [8] y CACIC 2016 [9].

En cuanto a la forma de instalación se realizaron pruebas en las máquinas de cada integrante del proyecto, utilizando el modo Standalone, el cual permite tener varias regiones ejecutándose al mismo tiempo.

Se probó también el modo grid (o servidor) que separa en diferentes procesos varios aspectos de la simulación.

Otra implementación realizada por el grupo de trabajo fue el acceso a las regiones a través de Internet con una conexión hogareña con 3 Mbps de bajada y 512 Kbps. Se eligió para la instalación del metaverso la distribución OpenSim Diva Standalone. La misma tiene como ventaja una interfaz Web que permite la generación de usuarios (avatares) de forma autónoma, gestión de los mismos, gestión de regiones, obtener información del inventario. La versión instalada fue Diva-r22458. [10] [11] En esta región se logró la interacción de 6 avatares. (Figura 1). Ésta versión se utilizó durante los dos años del PID 2014-2016, la cual facilitó la comunicación con acceso remoto durante todo el proyecto de los integrantes del mismo.



Fig. 1 (captura de pantalla de interacción con versión Diva)

En esta etapa se logró la conexión al servidor, montado en Bs. As., desde la ciudad de Mar del Plata, a través de la conexión satelital de un celular con 3G, capturada por nuestra notebook en la cual corría el cliente “Imprudence”. Allí se comprobó que las óptimas conexiones a Internet son sumamente importantes para la visualización de los avatares. (Demostración en CACIC 2013).

Otros desarrollos, llevados a cabo por nuestro equipo, incursionaron en la animación de avatares y las posibilidades que brindan los bots como agentes autónomos.

Para la segunda etapa de este PID, se pretende evaluar ventajas y desventajas en el uso de un VPS (servidor virtual privado) o un servidor dedicado, para instalar la plataforma utilizando OpenSim permitiendo disponer una conexión segura y que permita la interacción de más de 20 avatares en simultáneo.

Para la difusión general del proyecto se construirá un portal que informe de manera fácil e integrada los avances y desarrollos al usuario. Otra tecnología que se está utilizando en estos momentos para el aprendizaje, es Arduino, definida como “plataforma de creación de prototipos de código abierto basada en hardware y software fáciles de usar” [13]. Investigar las posibilidades que ofrece la misma para la interconexión de los mundos virtuales con el mundo real, abre otro camino en la investigación.

Líneas de investigación, desarrollo e innovación

Desde el proyecto de investigación (PID 01-003-16), planteado para los años 2016-2018, “Elementos interactivos en los mundos virtuales 3D”, se desprendieron diferentes líneas de las cuales aún se encuentran en desarrollo:

- 1) Uso de técnicas de Diseño Gráfico para incrementar la estética en los mundos virtuales.
- 2) Montar un servidor dedicado para tener acceso online con una carga de avatares que permita trabajar desde un curso.
- 3) Plantear casos de estudio articulando con cátedras de la Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales, utilizando el MVI.
- 4) Evaluar y desarrollar un portal para UMremoto3D.
- 5) Formar grupos de estudio en el ámbito de los MVI y la construcción de prototipos con Arduino.

Resultados y objetivos

Con este proyecto se pretende:

- Ofrecer un espacio virtual a la Universidad de Morón, distinto de la plataforma de educación a distancia, para desarrollar las características primordiales de los mundos virtuales:
 - a) Comunicación verbal entre residentes del mundo virtual.
 - b) Transmisión simultánea de imágenes, diapositivas y videos en el mundo virtual.
 - c) Control de prototipos Arduino.

- Mejorar el ambiente creado durante la investigación anterior para la cátedra de Higiene y Seguridad.
- Dejar zonas de capacitación para que los estudiantes participen en el Mundo.
- Dejar espacios que favorezcan el desarrollo de las cátedras que se interesen en los metaversos utilizándolos como laboratorios virtuales.

Como resultado de las diferentes etapas se elaborará material para participar en distintos congresos tales como WICC (Workshop de Investigadores en Ciencias de la Computación), CACIC (Congreso Argentino de Ciencias de la Computación), TE&ET (Tecnología en educación y Educación en Tecnología) como así también en otros congresos, en los cuales el tema sea de interés.

Promover la construcción y dictados de cursos de formación o asesoramiento.

Formación de Recursos Humanos

Este proyecto se encuentra financiado por la Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales de la Universidad de Morón. A su vez propicia la formación de recursos, con la participación de estudiantes de grado para la continuación de las líneas de investigación relacionadas. Este PID está compuesto por cuatro investigadores y docentes de dicha facultad y cuatro tesis de las carreras de Informática. A su vez están trabajando estudiantes de los últimos años de las carreras de Informática.

Referencias

[1] <http://aida.ii.uam.es/virtuam/>

[2] Serio López, Antonio Carpeño, Jesús Arriaga, Guillermo de Arcas, Ariel Contreras, José

Salazar, Miguel Barras,

Laboratorio remoto eLab3D: un entorno innovador para el aprendizaje de competencias prácticas en electrónica CINAIC-2013 2do Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. Pág. 41-46.

<http://www.cinaic.com>

[3] <http://opensimulator.org/wiki/Wifi>

[4] Sattolo Iris, y otros, *Primeros pasos en el desarrollo de ambientes virtuales inmersivos de aprendizaje utilizando software libre*. ISBN 978-987-23963-1-2 P 566-573.

[5] Sattolo Iris y otros, *Modelo de análisis de un ambiente virtual de aprendizaje inmersivo para el desarrollo de planes de evacuación*.

Disponible en:

<http://www.teyet2014.undec.edu.ar/Libro-de-ActasTEYET2014.pdf>

[6] Altube y otros. *Desarrollo de planes de evacuación utilizando un ambiente virtual inmersivo interactivo*. Disponible en:

<http://teyet2015.unne.edu.ar/files/TEyET2015.pdf>

[7] Lipera y otros, *Ambiente Virtual inmersivo interactivo* Disponible en: Libro de actas Jatic 2015 ISBN 978-987-23963-2-9

http://jatic2015.ucaecemdp.edu.ar/libro_actas.php

[8] Alejandro Altube y otros. *Bot guía para un ambiente virtual inmersivo*. ISBN 978-987-733-072-4 Libro de actas CACIC 2016.

[9] Darío Minutella y otros, *Animación de Avatares en un ambiente virtual inmersivo*.

ISBN 978-987-3977-30-5 Libro de actas TE&ET 2016

[10] <http://www.marlonj.com/blog/2012/04/instalando-diva-distroweb-0-7-3-en-ubuntu-server-11-10/#sthash.r79wI8CD>

[11] <http://metaverseink.com/Downloads.html>

[12] Claudia Ruso, Mónica Sarobe, Leonardo Esnaola, Nicolás Alonso. 2015 “*Entornos virtuales 3 D una propuesta educativa innovadora*” Campus Virtuales, Vol. IV, num. 1, pp. 32-42. www.revistacampusvirtuales.es

[13] <http://www.arduino.org/>

Juegos Interactivos en ARDUINO y Java, para Motivar y Despertar el Interés en Informática

Lic. Angela Belcastro 1, Roger David Alanes 2, Macarena Quiroga 3, Juan Giménez 2, Santiago Santana 2, Pablo Dibez 3, Mg. Rodolfo Bertone 4

1 Ciencia y Técnica. Facultad de Ingeniería (FI). UNPSJB. Comodoro Rivadavia.

2 Auxiliar de segunda. Facultad de Ingeniería. UNPSJB. Comodoro Rivadavia.

3 Estudiante. FI. UNPSJB. Comodoro Rivadavia.

4 III LIDI – Facultad de Informática – UNLP. (1900) La Plata, Buenos Aires, Argentina

1 (angelab) @ing.unp.edu.ar

2 rogerdavidalanes@yahoo.com.ar, jm.gimenez.inf@gmail.com, santana.santiago@gmail.com

3 msinf.quiroga@gmail.com, pdibez@gmail.com

4 pbertone@lidi.info.unlp.edu.ar

RESUMEN

Esta línea de investigación, se inició en 2003, en la UNPSJB. Este trabajo destaca elementos de Tecnologías de la Información y la Comunicación (TICs), empleados dentro de un contexto educativo, que pueden apoyar a los estudiantes, y favorecer el autoaprendizaje, la autonomía y la creatividad. Introduce recursos de la Web 2.0, características del aprendizaje significativo (AS), herramientas de trabajo colaborativo (TC), que se emplearon dentro del proyecto y apoyan el carácter adaptativo del trabajo del desarrollador de software (DS). Identifica condiciones para poder generar conocimiento, y AS. Presenta el concepto de motivación y distintas percepciones. Destaca las características de los juegos, y la organización y resultados de un evento de transferencia, con el desarrollo de juegos interactivos, en ARDUINO y java, empleados dentro de una estrategia de motivación, tendiente a fortalecer el avance en carreras de Informática (CI), la incorporación de alumnos en proyectos de investigación, y el desarrollo de juegos. Uno de los objetivos del evento, fue el de realizar prueba de uso del juego denominado: juegoDSSB “Juego de símbolos. Sistema binario”, desarrollado dentro del proyecto, implementado por integrantes que realizaron prácticas profesionales en “Programación orientada a objetos”, en 2014.

Palabras claves:

Aprendizaje significativo (AS), Motivación, Juegos, Recursos de la Web 2.0 (RW2.0)

CONTEXTO

El trabajo forma parte del proyecto de investigación acreditado: “Casos de estudio de sistemas, TICs y aprendizaje”, UNPSJB-UNLP, evaluado y financiado por Ciencia y Técnica de la UNPSJB. Pertenece al Área Tecnología Informática Aplicada en Educación. Es continuación de proyectos de investigación finalizados, que apoyaron la formación de recursos humanos, en carreras de grado, y generaron, avances en carreras de postgrado, mejoras en procesos de enseñanza aprendizaje (PE-A), y publicaciones a nivel nacional e internacional. Continúan estudios alternativos, para generar motivación en nuevos estudiantes de Informática, AS, y mejorar PE-A, en materias de Informática, empleando RW2.0.

1. INTRODUCCION

Con la aparición de internet, en especial la Web 2.0, se "promueven la interacción para involucrarse en procesos de compartición y construcción conjunta del conocimiento, que hacen posible la interacción grupal". [1] En un contexto educativo sólido, las TICs, pueden ayudar a los estudiantes a adquirir las capacidades necesarias. María Pinto sostiene que las exigencias de la Sociedad de la Información y del Conocimiento son, entre otras: adquirir nuevas capacidades, habilidades y competencias transversales y específicas a lo largo de los estudios de grado y postgrado; adoptar un modelo proactivo de AS que favorezca el autoaprendizaje, la autonomía y la creatividad de los estudiantes; rediseñar los métodos de docencia

universitaria; y concientizar a los docentes y estudiantes universitarios acerca de los nuevos retos que supone adaptar estos procesos a los nuevos ambientes. [2]

Desde el punto de vista pedagógico, los recursos tecnológicos comunicativos que ofrecen las TIC representan ventajas para el proceso de aprendizaje colaborativo, ya que permiten: estimular la comunicación interpersonal; el acceso a información y contenidos de aprendizaje; el seguimiento del progreso del participante, en lo individual y grupal; la gestión y administración de los alumnos; la creación de escenarios para la coevaluación y autoevaluación, y principalmente la construcción de significados comunes en un grupo social determinado. [1] Bereiter considera la colaboración como un discurso progresivo (DP). El foco del DP persigue la comprensión y ve el conocimiento como algo problemático, que debe ser explicado. Los errores y las malas interpretaciones se ven como oportunidades para ir más allá. Los participantes negocian sus ideas con los otros, como miembros de una comunidad de aprendizaje, producen conocimiento. [3]

El AS es aquel que ocurre cuando la información nueva por aprender se relaciona con la previa, ya existente en la estructura cognitiva del alumno, de forma no arbitraria ni al pie de la letra; para alcanzar esto debe existir disposición favorable del aprendiz, así como significación lógica en los contenidos o materiales de aprendizaje. [3]

Ausubel indicó tres condiciones básicas para que se produzca el AS, que los materiales de enseñanza estén estructurados lógicamente con una jerarquía conceptual, situándose en la parte superior los más generales, inclusivos y poco diferenciados, que se organice la enseñanza respetando la estructura psicológica del alumno, es decir, sus conocimientos previos y su estilo de aprendizaje, y que los alumnos estén motivados para aprender. [4]

Schlechy (2002) caracteriza al nuevo paradigma como aquel en el cual el estudiante es el trabajador, y en el cual el docente es el diseñador del trabajo del estudiante. Se busca

que el estudiante logre ser autodirigido y motivado. [5]

La motivación incluye procesos que dan energía, dirigen y mantienen la conducta. Distintas perspectivas psicológicas explican la motivación de diversas maneras. Mientras que la perspectiva conductual considera que la motivación del estudiante (ME) es una consecuencia de incentivos externos, la perspectiva cognitiva (PsC) argumenta que se debe dar menor importancia a las presiones externas, y recomienda que los alumnos reciban mayores oportunidades y responsabilidades para controlar su propio rendimiento. La (PsC) de la ME, coincide con las ideas de R. W. White (1959), quien propuso el concepto de motivación de competencia, la idea de que la gente está motivada a enfrentar de manera efectiva su ambiente, a dominar su mundo y a procesar la información de manera eficiente. La perspectiva humanista destaca la capacidad del estudiante para el crecimiento personal, la libertad para elegir su destino y cualidades positivas. La social enfatiza en la necesidad de establecer, mantener y restablecer relaciones personales cálidas y cercanas. [6]

Armar mapas conceptuales, puede ayudar a promover el AS. El mapa es conceptual más por ser el resultado de una concepción, que por contener conceptos. [7][8]

La Web 2.0 involucra a todos los sitios Web que permiten crear contenidos, diseñar y compartir información de manera colaborativa e interactuar de modo no presencial, en la Word Wide Web, creando comunidades virtuales. [9]

Los RW2.0 pueden favorecer el AS, entre ellos encontramos, los Entornos Virtuales de Aprendizaje, blogs, Google Drive, las herramientas para la gestión de TC, como Git y Trello, y otros recursos, algunos de los cuales mencionaremos a continuación.

CmapTools, es una herramienta colaborativa para elaborar esquemas conceptuales. [10]

Git apoya al DS, es un software de control de versiones que soporta el almacenamiento de los documentos. [11]

Haciendo referencia a las estrategias educativas para el desarrollo del pensamiento

computacional encontramos: ser incremental e iterativo (diseñar un proyecto no es un proceso limpio y secuencial en el que primero se identifica un concepto para el proyecto, luego se desarrolla un plan de diseño y posteriormente, se implementa el diseño del código. Es un proceso adaptativo, uno en el que el plan puede cambiar para responder a un enfoque de solución por pequeños casos), ensayar y depurar, reusar. [12]

Las herramientas de gestión colaborativa apoyan al DS, en el proceso adaptativo.

Las CI del país, en general, han disminuido el número de ingresantes. Diversos esfuerzos, están orientados a acercar a los jóvenes, a las CI. [13][14]

El juego es una actividad universal. Se podría decir que no hay ningún ser humano que no haya practicado esta actividad en alguna circunstancia. Mediante el juego se pueden crear situaciones de máximo valor educativo y cognitivo que permitan experimentar, investigar, resolver problemas, descubrir y reflexionar. Las implicaciones de tipo emocional, el carácter lúdico, el desbloqueo emocional, la desinhibición, son fuentes de motivación que proporcionan una forma distinta a la tradicional de acercarse al aprendizaje. [15] [16]

2. LINEA DE INVESTIGACION Y DESARROLLO

Este proyecto es continuación de dos proyectos de investigación finalizados, sus objetivos primordiales, son:

- formar personas competentes en materias seleccionadas de Informática, de la UNPSJB,
- confeccionar y analizar, sistemas interactivos de apoyo al aprendizaje,
- fomentar el AS con propuestas de actividades que requieren de RW2.0 seleccionados, y fortalecer la vinculación entre niveles universitarios y secundarios, mejorando los conocimientos previos de los futuros alumnos de Informática.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los antecedentes, encontramos los trabajos presentados en WICC 2015, TE&ET 2015, WICC 2016 y CONAIISI 2016 [17][18][19][20]. Se realizaron acciones tendientes a mejorar PE-A, con propuestas de uso de RW2.0 para promover el AS en materias de una CI. Los temas analizados, son: clasificación de juegos, y juegos educativos, características y componentes del software educativo, actividades cognitivas que despliegan en los usuarios, temas seleccionados para entrenamiento, en los juegos interactivos, patrones de diseño y factores de calidad del software, herramientas de TC, y ARDUINO.

Las placas microcontroladoras ARDUINO son dispositivos electrónicos que previa programación desde una computadora, con un software determinado, permiten interactuar con sensores y actuadores de naturaleza tanto digital (por ejemplo, leds, interruptores), como analógica (por ejemplo, sensores de temperatura o presión), para la automatización y control de procesos de diversa índole. En la figura 1, vemos los elementos del ciclo de trabajo con ARDUINO. [21]

Su diseño es open hardware. Respeta la libertad de los usuarios y la comunidad; los usuarios pueden ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software. [22]



Figura 1. Elementos de trabajo con ARDUINO Tal y como postula Stallman (2010), la figura de un centro educativo como transmisor no solo de conceptos y procedimientos, sino también de valores como la libertad, y el espíritu de colaboración y ayuda al vecino, tiene mucho que ver con los principios del software libre. [22]

El equipo de proyecto organizó y desarrolló, en el laboratorio de la UNPSJB, el evento de transferencia (JIVISB): “juego interactivo en java. Introducción al sistema binario”. Confeccionó dos formularios de google, de manera colaborativa, para inscripciones, y para la evaluación de uso del juego de símbolos. Sistema binario. (juegoDSSB). Armó material

para difusión, realizó la inscripción, para alumnos ingresantes de Informática y Electrónica, y para estudiantes sin concepto en “Elementos de Informática”, en 2016.

Objetivos principales del evento:

- Realizar una prueba de uso del juegoDSSB, y recabar información de retroalimentación, en base a lo programado, descrito en [19].
 - Motivar a los alumnos, a la participación de proyectos de investigación al alcanzar en la carrera; y al desarrollo de juegos interactivos
- Las etapas del evento fueron:

- Bienvenida, con presentación de todos los participantes, e identificación de objetivos.
- Interactuando con el juego y con los alumnos que lo implementaron, que participan en el proyecto. En esta etapa se acercó al alumno uno de los juegos diseñados en el proyecto.
- Llenado del formulario.
- Vislumbrando java. En esta etapa, los alumnos interactuaron con un juego simple, desarrollado en java, preparado dentro del proyecto, que permite al usuario desplazar una nave, en diversas direcciones, y cambiar su velocidad. Recibieron una explicación, de elementos claves involucrados, en el uso, y en su programación y desarrollo. Uno de los objetivos fue el de anticiparles la existencia del patrón modelo vista controlador, y sus componentes claves, motivándolos.
- Aplicación ARDUINO. Se brindó una explicación breve de ARDUINO e interactuaron con un sistema desarrollado por integrantes del proyecto, cuyo objetivo era construir un teclado binario capaz de interpretar pulsos electrónicos, y mostrar en pantalla un carácter codificado en ASCII.
- Despedida.

La figura 2, muestra una imagen de uso del sistema ARDUINO. Se inscribieron diecisiete alumnos.



Figura 2. Uso del sistema ARDUINO

Algunos resultados: todos los alumnos que participaron en el evento, llenaron el formulario, todos eran ingresantes de Informática. El 50%, egresaron del perito Moreno. El 40% del Instituto Austral, y un 10% del Colegio Cervantes. En su mayoría,

consideraron que los enlaces a sitios educativos de sistemas de numeración, del juegoDSSB, ayudaban a mejorar la comprensión, en una “muy buena” medida. Un 50% de los alumnos emplea juegos para esparcimiento y para aprendizaje, un 40% sólo para esparcimiento, y un 10%, no usa. Los tipos de juegos más utilizados por los participantes, fueron: “de lógica y matemática”, “de acción”, “simuladores y constructores”, en igual medida. Emplean juegos de mesa y juegos deportivos. Y en menor medida, emplean otros juegos y juegos de preguntas. Al analizar el empleo de recursos de la Web 2.0, mayormente emplearon blog; menos del 50%, publicó en el blog, el 25% de los participantes, armaron un blog. Un 70% de los encuestados, trabajó con archivos compartidos. Algunos alumnos destacaron haber empleado otras aplicaciones, como software de cálculo y dibujo, skype, geogebra, redes sociales, photoshop, WhatsApp. Un 50%, consideró haber trabajado en forma “excelente” con el sistema, con armonía, y satisfacción, y un 40 %, en forma “muy buena”. Manifestaron gran interés en desarrollar talleres de lenguaje java, y en menor medida, de desarrollo web y aplicaciones móviles, y ARDUINO. Al 100% de los encuestados le resultó entretenido el juegoDSSB. La última pregunta era: ¿Porque te resulta o no, entretenido?, los participantes destacaron que les resultó entretenido, algunas respuestas, son:

- Me resulta entretenido por la interacción que puede tener el usuario a través de la interfaz.
- Me resulta entretenido porque ayuda a comprender y resulta más fácil aprender.
- Porque te enganchas mucho calculando, hasta le agarrás la mano de memoria.
- Porque aprendí y me divertí.
- Porque aprendí un contenido que nunca en mi vida supe.
- Porque me llama a seguir aprendiendo.

4. FORMACIÓN DE RECURSOS HUMANOS

El director externo, es docente de grado y postgrado en la UNLP, en la especialidad de la investigación. En el equipo de proyecto

intervienen docentes de carreras de grado de Informática, dos de ellos, han aprobado todas las materias del Magister, en Tecnología Informática aplicada en Educación, de la UNLP. Los alumnos del equipo de proyecto, están avanzando en la Licenciatura en Informática de la UNPSJB.

5. BIBLIOGRAFIA

- [1] Edith Inés Ruiz Aguirre, y otros. Aprendizaje colaborativo en ambientes virtuales y sus bases socio constructivistas como vía para el AS. Revista de innovación educativa. 2012.
- [2] Mg. González De Doña, y otros. Desarrollo de competencias para la gestión de información y construcción de conocimientos: TICs y Nuevos Ambientes Educativos. 2009.
- [3] Fernando Santamaría González. Herramientas Colaborativas Para La Enseñanza Usando Tecnologías Web: Weblogs, Redes Sociales, Wikis, Web 2.0. 2005.
- [4] Richard E. Mayer. Psicología de la Educación. Enseñar para un AS. Volumen II. Pearson. Prentice Hall. 2004.
- [5] Sanz Cecilia y otros. Tendencias en procesos educativos mediados por TIC. LIDI. WICC 2015. ISBN: 978-987-633-134-0. Salta, Argentina. 2015.
- [6] John W. Santrock. Psicología de la educación. Mc Graw-Hill. 2004.
- [7] Frida Díaz. Barriga Arceo. Gerardo Hernández Rojas. Estrategias docentes para un AS una interpretación constructivista. Mc Graw-Hill. 2002.
- [8] Ángel Díaz-Barriga. Tic en el trabajo del aula. Impacto en la planeación didáctica. RIES. 2013.
- [9] Miriam Kap. Conmovidos por las Tecnologías. Pensar las prácticas desde la subjetividad docente. Prometeo. 2014.
- [10] Hugo M. Castellano. Enseñando con las TIC. Integración de la tecnología educativa en el aula. Cengage Learning. 2011.
- [11] Andrés Cely Jiménez. Desarrollo de prototipo de portal para el apoyo a la gestión de grupos de investigación enfocado en el trabajo colaborativo. UD. Bogotá, Colombia. 2015.
- [12] Elisenda Eva Espino. Estudio sobre diferencias de género en las competencias y las estrategias educativas para el desarrollo del pensamiento computacional. RED. Número 46 15-Sep-2015
- [13] Anuario de Estadísticas Universitarias Argentina 2011. ISSN 1850-7514.
- [14] Dapozo, Gladys y otros. Universidad Nacional del Nordeste Promoción de Vocaciones en TIC. Identificación de factores que influyen en el interés de los alumnos. WICC 2015
- [15] Deninse Farias y otros. Estrategias lúdicas para la enseñanza de la matemática en estudiantes que inician estudios superiores. Venezuela. 2010.
- [16] Laura Muñiz-Rodríguez, Pedro Alonso, Luis J. Rodríguez-Muñiz. El uso de los juegos como recurso didáctico para la enseñanza y el aprendizaje de las Matemáticas: estudio de una experiencia innovadora. Revista Iberoamerica de Educación Matemática. 2014.
- [17] Belcastro, Angela, y otros. Recursos de la Web 2.0, y juegos interactivos, alternativas viables para propiciar el aprendizaje en Informática. WICC 2015.
- [18] Belcastro, Angela y otros. Método para mejorar un software educativo en desarrollo. TE&ET 2015.
- [19] Belcastro, Angela y otros. Construcción y mantenimiento de un juego, para mejorar conocimientos previos de futuros ingresantes, y despertar el interés en Informática. WICC 2016.
- [20] Belcastro, Angela y otros. Modelos, Estrategias, y Recursos para la Enseñanza Mediada, en una Asignatura de una Carrera de Grado. CONAIISI 2016.
- [21] Daniel Ruiz Corres. Estudio sobre la implantación de la herramienta ARDUINO en centro de formación profesional. 2016.
- [22] Javier Díaz Sánchez. Open hardware y software, herramientas para el desarrollo de competencias educativas. benemérita Revista iberoamericana de producción académica y educativa. 2015.

Laboratorio de Escritorio para Enseñanza de Electricidad y Electrónica

Marcos A. Aranda^{1,3}, Paola Beltramini^{1,2}, Sergio Gallina², Eduardo Cano², Maria V. Poliche³, Marcelo D'amore², Matias Ferraro², Gabriel Díaz² & Luis Schneider²

(1) *Laboratorio de Sistemas Embebidos, Facultad de Tecnología y Ciencias Aplicadas, UNCa, Argentina.*

markosdarioaranda@gmail.com, palobel2012@gmail.com

(2) *Departamento Electrónica, Facultad de Tecnología y Ciencias Aplicadas, UNCa, Argentina.*

jesuseduardocano@gmail.com, sgallina@tecno.unca.edu.ar

(3) *Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas, UNCa, Argentina.*

vpoliche@tecno.unca.edu.ar

AREA TEMATICA: Tecnología Informática Aplicada en Educación

RESUMEN

Los estudiantes universitarios de hoy aprenden de manera muy diferente de lo que solían hacerlo hasta hace pocos años. La educación superior ya no se centra en la consulta a libros de texto y en asistencia a presentaciones expositivas y prolongadas en el tiempo. Actualmente el estudiante tiene acceso en línea a mucha información y múltiples plataformas de aprendizaje, por lo que ha perdido el interés en disponerse y centrarse en estos métodos de enseñanza tradicionales.

Para satisfacer estas exigencias, se propone el desarrollo tecnológico y pedagógico de un Laboratorio Portátil de Escritorio que permita al estudiante “ver” lo que estudia mediante la construcción de los ensayos y el uso de una interfaz gráfica como apoyo al proceso de visualizar señales y parámetros no visibles al ojo. Este desarrollo implica la construcción del hardware necesario y los aspectos didácticos de su utilización y al nuevo rol del alumno y del profesor. Finalizado el proyecto, se deberá confirmar en la práctica que esta nueva herramienta permitirá cambio real y favorable en el proceso de enseñanza-aprendizaje. Todo esto implicará además analizar: Metodologías y enfoques crítico-aplicativos para el autoaprendizaje - Actualización de programas - Construcción personalizada de aprendizajes significativos - Trabajo colaborativo.

Palabras Claves: TIC, Laboratorio de escritorio, Enseñanza, Electrónica

CONTEXTO

Esta investigación, a través de la cual se propone desarrollar y construir un laboratorio de escritorio portátil que resuma la instrumentación básica y necesaria para el desarrollo de prácticas experimentales guiadas en la enseñanza de la Electrónica básica, fue aprobada en dos instancias. En primer lugar en la convocatoria 2016 de Proyectos de Investigación y Desarrollo de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca (UNCA) y también en el marco de Proyectos Catamarqueños de Instrumentación Científica y Tecnológica – ProCaICyT 2016 de la Subsecretaría de Ciencia y Tecnología de la Provincia de Catamarca.

El organismo beneficiario es la Universidad Nacional de Catamarca, siendo las carreras de grado y pregrado de la Facultad de Tecnología y Ciencias Aplicadas los espacios donde se implementarán los prototipos. Concluidos estos ensayos, el Laboratorio de Escritorio será puesto a disposición de los colegios técnicos de la provincia que se encuentren interesados en esta nueva modalidad de enseñanza.

El ámbito de desarrollo es la Facultad de Tecnología y Ciencias Aplicadas, específicamente el Laboratorio de Sistemas Embebidos (LaSE) y el Laboratorio de

Electrónica, contribuyen con los materiales, equipos e instrumental necesarios para su realización.

1. INTRODUCCIÓN

“Padres y educadores afrontamos un difícil reto ante la nueva generación que quiere aprender por vías no tradicionales y siempre empleando nuevas tecnologías. La computadora e internet propician la actividad independiente, la observación, exploración y búsqueda; la comparación, ordenamiento y clasificación; la toma de decisiones y el procesamiento de la información. ¿Cómo satisfacerlo todo?”
Generación NET 2014 – Conclusiones - Ramón F. Ferreiro (2006)

En las carreras de ingeniería, esta situación adquiere especial relevancia, ya que, en general, las clases se desarrollan de manera expositiva, incluyendo prácticas de laboratorio y proyectos prácticos cuyo ámbito en la mayoría de los casos no va más allá de los contenidos del curso. Muchos docentes, motivados por el florecimiento del proceso enseñanza-aprendizaje y la necesidad de enseñar por vías no tradicionales para afrontar el reto que representan las nuevas generaciones de estudiantes, realizan innovaciones de enseñanza constructivista en sus clases y adicionalmente algunos alumnos se incorporan a proyectos de investigación donde complementan su formación.

La implementación de estas nuevas técnicas de enseñanza se basa en la disponibilidad de elementos tecnológicos que la faciliten o de costosos laboratorios de difícil mantenimiento en el tiempo, pocas veces acorde a la cantidad de alumnos.

Mediante este *“Laboratorio de Escritorio”* se pretende generar metodologías de enseñanza, que impacten positivamente en los estudiantes que se introducen en la electrónica práctica (ya sea que provengan de carrera de tecnicaturas,

ingenierías o de colegios o cursos de formación técnica), dando solución a diversas problemáticas relacionadas con el aprendizaje basado en métodos de memorización.

Para la concreción del proyecto será necesario encarar el desarrollo de tres aspectos, según el siguiente detalle:

1.1 Construcción física del laboratorio (hardware – software)

Centrada en la Computadora Industrial Abierta Argentina (CIAA), consta de cinco bloques diferentes que le brinda potencialidad al laboratorio (Figura 1):

- Fuente de alimentación:

Provee las tensiones necesarias para la CIAA y una variedad de valores que podrán ser utilizadas por el alumno en sus prácticas.

- Display gráfico y táctil:

Permitirá ver el valor medido en los instrumentos o señales en un osciloscopio y configurar las señales de salida del generador.

- Generador de señales:

Entregará señales de onda cuadrada, triangular y senoidal, todas de frecuencia y amplitud variable entre rango preestablecidos.

- Instrumentos

Utilizando las cuatros entradas analógicas que posee la CIAA y el desarrollo de hardware adecuado para cada caso, se implementan tres instrumentos: Voltímetro digital (con resolución desde 100 mV a 99.9 V, autorango e indicador de polaridad en CC) – Amperímetro digital (con rango desde 0,1 mA a 100 mA) – Osciloscopio de dos canales.

- Conectividad

Para la carga de los prácticos y el control del profesor, adicionalmente consta de un conector donde se tienen disponibles las tensiones de alimentación del circuito bajo prueba.

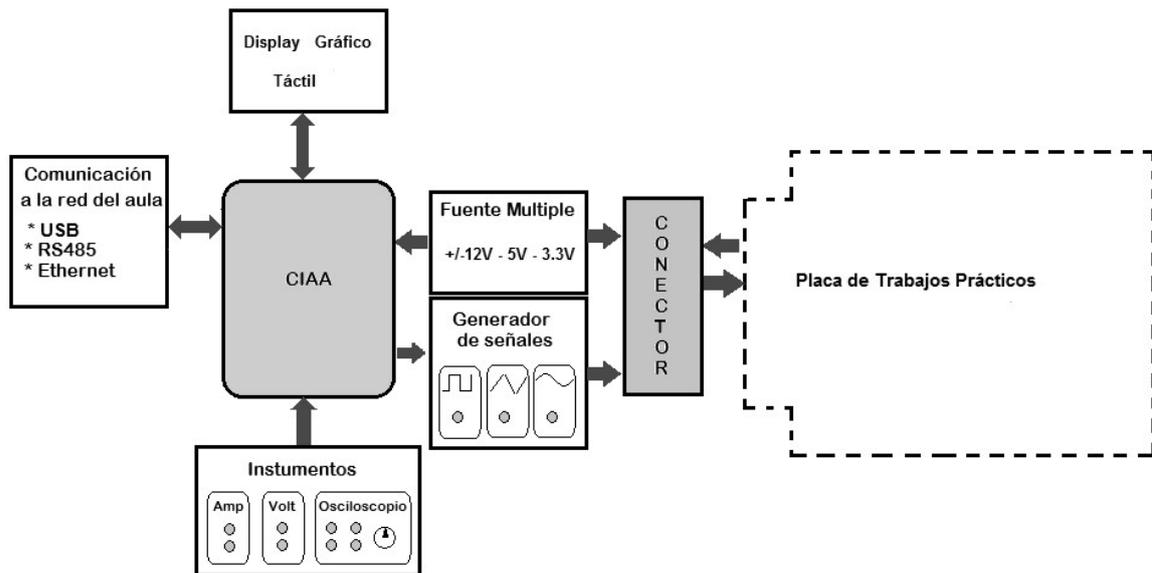


Figura 1: Diagrama en bloques del laboratorio de escritorio

1.2 Construcción física de los prácticos destinados a clases (hardware-software).

- Placa intercambiable

Aquí el alumno conectará su desarrollo para la realización de diversas actividades que permitirán construir el conocimiento y posteriormente su evaluación.

En la fig. 2 se observa un esquema donde se puede visualizar un bosquejo físico de distribución de estos sistemas enunciados y la forma final que tendría el gabinete o “valija” que contendrá todos los componentes del Laboratorio de Escritorio.

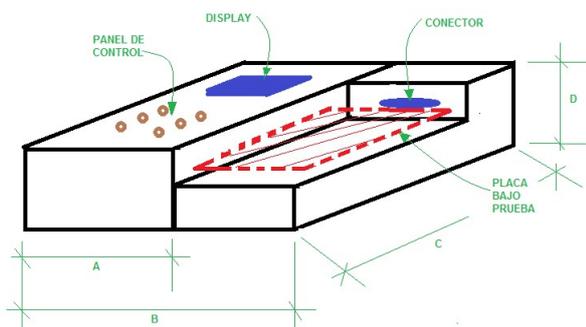


Figura 2: Esquema del gabinete del Laboratorio de Escritorio

1.3 Propuesta de clase

Finalizado el desarrollo se dispondrá de un elemento económico y portátil que permitirá trabajar sobre el 3er aspecto que consiste en la implementación de la *Enseñanza mediante la*

exploración. En base al conocimiento de las necesidades y exigencias de las nuevas generaciones de estudiantes, es posible realizar transformaciones educativas que no sean solo enunciación de teorías sino una modificación real de la actividad en el aula.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Este proyecto tiene como principal objetivo el desarrollo y construcción de un laboratorio de escritorio portátil que resuma la instrumentación básica y necesaria para el desarrollo de prácticas experimentales guiadas. La principal línea de investigación de este proyecto es el uso de las nuevas tecnologías de la información y comunicación, y en particular, mediante un doble desafío: en primer lugar adoptar un nuevo paquete tecnológico, que está promovido a nivel nacional, como lo es la CIAA y en segundo lugar aplicarlo fuera su ámbito natural que es la industria, para introducirlo en el ámbito educativo.

El desarrollo del mismo requiere de tres ingenierías muy relacionadas, la electrónica, la informática y las comunicaciones, de esta manera el proyecto se divide en tres fases de desarrollo:

- Desarrollo electrónico, que provee los equipos o dispositivos con los cuales se

ingresa, extrae, presenta o procesan datos, información o servicios.

- Desarrollo de software, uno para la interfaz que permita leer datos directamente de la placa electrónica y el software para la computadora o PC que permitirá interactuar docentes con alumnos en forma amigable, simple y eficaz.
- Desarrollo de las comunicaciones, se deberán manejar los sistemas de comunicación entre la interfaz y la computadora del alumno y además todo el grupo de interfaces o laboratorios con el profesor.

Una segunda línea de investigación es el correcto uso de estas nueva TIC en la educación. La misma se está abordando desde un inicio por especialistas docentes de las cátedras a beneficiarse. Consiste en ir investigando y diseñando las clases que se propondrán, una vez finalizado el desarrollo del prototipo, su ensamble y correspondientes pruebas de funcionamiento. Considerando necesario acotar este punto por cuanto la variedad de clases es muy amplia, en una primera etapa se abordará el desarrollo de tres clases o tópicos: - polarización de diodos zener, polarización de transistores y respuesta en frecuencia de amplificador con transistores. Para cada clase o tema se deberán desarrollar: aprendizaje teórico, prácticas a realizar por el alumno y mecanismos de evaluación.

3. RESULTADOS OBTENIDOS / ESPERADOS

La disponibilidad del laboratorio de escritorio implicará algunos cambios importantes:

- Metodologías y enfoques crítico-aplicativos para el autoaprendizaje. Ahora el problema de los estudiantes ya no es el acceso a la información (que está casi omnipresente) sino la aplicación de esta información para la resolución de problemas reales, los estudiantes pueden trabajar de manera independiente, y construir su conocimiento a partir de la realización de diversas actividades, apoyándose en las orientaciones y asesoramientos del profesorado.
- Actualización de los programas. El profesor ya no podrá desarrollar un programa

obsoleto, se puede consultar en Internet lo que se hace en otras universidades e incorporarlo a su práctica docente.

- Construcción personalizada de aprendizajes significativos. Los estudiantes pueden, de acuerdo con los planteamientos constructivistas y del aprendizaje significativo, realizar sus aprendizajes a partir de sus conocimientos y experiencias anteriores porque tienen a su alcance muchos materiales formativos e informativos alternativos entre los que escoger y la posibilidad de solicitar y recibir en cualquier momento el asesoramiento de profesores.
- Trabajo colaborativo. Los estudiantes se pueden ayudar más entre ellos y elaborar trabajos conjuntos con más facilidad. La ingeniería es una profesión eminentemente práctica que desarrolla diferentes tecnologías, energías, materiales e información.

El objetivo de una educación técnica es preparar a los alumnos para el estudio de una ingeniería, particularmente para tratar con las fuerzas y materiales presentes en la naturaleza. Por tanto, desde los albores de la educación para ingenieros, los laboratorios de enseñanza han sido una parte esencial de los programas de licenciatura y durante los años de carrera universitaria.

De hecho se puede decir que la mayoría de la enseñanza de las ingenierías se llevaba a cabo en el laboratorio. El énfasis sobre el trabajo en laboratorio ha variado a lo largo de los años. A pesar de prestar mucha atención a los currículos y metodologías de enseñanza, existe muy poca literatura sobre la enseñanza en los laboratorios. Por ejemplo, en las encuestas publicadas en la *Journal of Engineering Education* desde 1993 a 1997, se reveló que solo el 6.5% de los papers utilizaban la palabra "laboratorio" como una palabra clave.

De esta manera la propuesta de desarrollo representará un aporte novedoso y primordial ya que va a ir más allá del mero diseño tecnológico y abarcará también el fundamental aspecto pedagógico de implementación para el desarrollo de las clases asegurando de que se

cumplan los objetivos de mejoras del proceso de enseñanza aprendizaje.

4. FORMACIÓN DE RECURSOS HUMANOS

Todos los integrantes del proyecto son docentes de la Facultad de Tecnología y Ciencias Aplicadas de las carreras Ing. Electrónica e Ing. en Informática. Cinco de los ellos están categorizados en el Programa de Incentivos de la Secretaría de Políticas Universitarias (SPU) perteneciente al Ministerio de Educación de la Nación.

La diversidad del equipo de investigación es muy importante. Por un lado los formados en el área de educación, uno de los cuales posee Título de Especialista en Docencia Universitaria de Disciplinas Tecnológicas, con el Trabajo Final titulado “Pertinencia de las estrategias de enseñanza en las asignaturas de formación específica de la carrera de Ingeniería Electrónica”.

Por otro lado, aquellos que poseen una vasta experiencia de divulgación y de formación en temas de desarrollo tecnológico con microprocesadores de última generación. Incluso, algunos integrantes han participado en el desarrollo de la Computadora Industrial Abierta Argentina (CIAA) lo que brinda el conocimiento fundamental y experiencia (know how) para el desarrollo del dispositivo. Además, los docentes de Electrónica se desempeñan en asignaturas de electricidad y electrónica, por lo que se cuenta con un conocimiento detallado de la problemática de la educación técnica en el área.

En función de ello, para el desarrollo del proyecto, el equipo de trabajo se organiza en dos grupos: a) del hardware y b) del software. Ambos grupos coinciden en las tareas de montaje y testeo del equipo.

Al proyecto se ha incorporado un estudiante avanzado de Ingeniería Electrónica, estando abierta la convocatoria para incorporar un alumno avanzado de Ing. en Informática.

5. BIBLIOGRAFIA

- Aranguren G. *Nuevos métodos de enseñanza: una experiencia en diseño electrónico*. IEEE-RITA Vol. 3, Núm. 1 Páginas 39 a 46, Mayo 2008.
- Sánchez Ruvirosa R., *Enseñar y aprender con nuevos métodos. La revolución cultural de la era electrónica*, UAM-X – MÉXICO, pp. 321-326. 2002.
- Monge Nájera, J.; Méndez Estrada, V. H. *Ventajas y desventajas de usar laboratorios virtuales en educación a distancia: la opinión del estudiantado* Revista Educación, ISSN: 0379-7082, vol. 31, núm. 1, pp. 91-108 Universidad de Costa Rica San Pedro, Montes de Oca, Costa Rica. 2007.
- Delacôte, Goéry. (1997) *Enseñar Y Aprender Con Nuevos Métodos*. Barcelona: Editorial Gedisa, S.A., 251 p.
- Martínez Delgado A., *Constructivismo Radical, Marco Teórico de Investigación y Enseñanza de las Ciencias*, UAB, Enseñanza de las ciencias, 1999, 17 (3), 493-502. 1999.
- Ferreiro R., *Nuevos ambientes de aprendizaje*. INPUT. No. 21, Barcelona. 2006.
- Tapscott, D. (1998) *Growing up digital: The rise of the Net generation*. McGraw-Hill. New York.
- Montero J. M., Ferreiros J., Guarasa J. M., de Córdoba R. y Romeral J. D. *Enseñanza en laboratorios de electrónica: una filosofía basada en diseños no guiados del mundo real* Departamento de Ingeniería Electrónica. ETSI Telecomunicación. UPM
- El Aprendizaje Basado en Problemas como técnica Didáctica, Dirección de Investigación y Desarrollo Educativo, Vicerrectoría Académica, Instituto Tecnológico y de Estudios Superiores de Monterrey. <http://sitios.itesm.mx/va/dide/documentos/inf-doc/abp.pdf>, Mayo, 2016.
- La metodología del Aprendizaje Basado en Problemas, http://www.ub.edu/dikasteia/LIBRO_MURCIA.pdf, mayo, 2016
- Proyecto CIAA, <http://proyecto-ciaa.com.ar/>, marzo 2017.

Las Tecnologías de la Información y la Comunicación en los Procesos de Enseñar y Aprender

Sanz Cecilia^{1,5}, Madoz Cristina¹, Gorga Gladys¹, Gonzalez Alejandro¹, Zangara Alejandra¹, Depetris Beatriz², Ibáñez Eduardo¹, Artola Verónica^{1,4}, Violini Lucía^{1,3}, Salazar Mesía Natalí^{1,3}, Sanchez Mariano¹

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata

²Universidad Nacional de Tierra del Fuego

³Becaria TIPO A UNLP

⁴Becaria Doctoral CONICET

⁵Investigador Asociado de CIC

{csanz, cmadoz, ggorga, agonzalez, eibanez, vartola, lviolini, nsalazar, msanchez}@lidi.info.unlp.edu.ar,
alejandra.zangara@gmail.com, depetrisb@gmail.com

RESUMEN

El presente trabajo describe las principales líneas de investigación y resultados obtenidos del subproyecto “Escenarios educativos mediados por Tecnologías de la Información y la Comunicación (TIC)”. Se enfoca principalmente en compartir los avances realizados durante 2016 e inicios de 2017. El subproyecto tiene como meta realizar aportes al área disciplinar que integran las posibilidades de las Ciencias de la Computación en la Educación. Se investiga y se desarrollan aplicaciones, herramientas y metodologías que permiten enriquecer los procesos de enseñar y aprender, a partir de su mediación y diálogo con tecnologías digitales. Se cuenta con un equipo de trabajo interdisciplinario, en el que participan investigadores formados, en formación, y becarios. Es importante hacer notar que varios de los participantes son docentes de las carreras de postgrado de la Facultad de Informática que se vinculan con estas temáticas. Esto facilita la sinergia entre el proyecto y estas carreras, y converge en el desarrollo de tesis y en la formación de recursos humanos para el área disciplinar en cuestión.

Palabras clave: Trabajo y aprendizaje colaborativo, Entornos digitales, Materiales educativos digitales, Objetos de aprendizaje.

CONTEXTO

Este subproyecto forma parte del proyecto “Tecnologías para Sistemas de Software Distribuidos. Calidad En Sistemas y Procesos. Escenarios Educativos Mediados Por TIC (período 2014-2017)”, perteneciente al Instituto de Investigación en Informática LIDI, de la Facultad

de Informática de la Universidad Nacional de La Plata y acreditado por el Ministerio de Educación de la Nación.

1. INTRODUCCION

Las actuales investigaciones en tecnologías digitales y educación marcan la necesidad de reconocer las características de nuestra sociedad, “...donde la información crece a un ritmo exponencial, las herramientas cambian constantemente, se crean nuevas apps de una manera prácticamente viral, y el software se encuentra en un modo de mejora permanente” (Kuhn, 2017, p.11). Esto permite vislumbrar que se hace creciente la necesidad de fortalecer la investigación sobre cómo las personas se relacionan con la tecnología. “...Las personas necesitan poner al día sus destrezas digitales y conocimientos. Esto está ocurriendo en marcos tanto formales como informales... Los estudiantes tendrán éxito si pueden reflexionar sobre cómo aprenden, planear su trayecto de aprendizaje y seleccionar las herramientas y los recursos que prefieran. Resulta esencial que sean capaces de filtrar el (excesivo) flujo de información y el sinfín de recursos que están disponibles online” (Kuhn, 2017, p.11 y 12).

Así, el aprendizaje autónomo es un tema fundamental en las diferentes etapas educativas y especialmente, en la vida universitaria, ya que el valor del saber cobra un factor determinante en el futuro laboral, que le exige a una persona a aprender a lo largo de la vida, convirtiéndolo en un patrón de habilidades de uso cotidiano en el entorno académico y laboral. El aprendizaje autónomo es un proceso en el que el estudiante autorregula su aprendizaje y toma conciencia de sus propios procesos cognitivos, meta cognitivos y socio afectivos. Es decir que la formación de los sujetos está centrada en resolver aspectos concretos de su

propio aprendizaje, y no sólo en resolver una tarea determinada. En otras palabras, busca orientar al estudiante a que se cuestione, revise, planifique, controle y evalúe su propia acción de aprendizaje (Martínez, 2004).

Son varios los autores y los reportes que indican que la colaboración, la generación de habilidades de autorregulación de los aprendizajes, y las destrezas para la utilización de las TIC, son y serán factores claves en la educación actual y futura (Redecker, Leendertse, Punie, Gijbers, Kirschner, & Hoogveld, 2011) citado en (Khun, 2017), (Griffin et al., 2012).

En este sentido, el subproyecto que aquí se presenta, se focaliza a través de sus líneas de investigación en esos aspectos, ya que indaga sobre los procesos, herramientas y metodologías para la colaboración en entornos mediados por TIC, los materiales educativos digitales como un puente entre docentes y alumnos para compartir conocimiento, las posibilidades de los entornos digitales que ofrecen oportunidades para generar entornos personales de aprendizaje, fomentar destrezas digitales y de autonomía, y los diferentes paradigmas de interacción persona-ordenador que buscan facilitar las interacciones entre el mundo digital y el entorno físico de las personas integrando diferentes canales de entrada para el aprendizaje de las personas.

2. LINEAS DE INVESTIGACION / DESARROLLO

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- Entornos digitales: campus virtuales, entornos virtuales de enseñanza y aprendizaje, entornos virtuales 3D, entornos simulados y laboratorios, y redes sociales, entre otros. Proyectos que los incorporan, metodologías, desarrollos, evaluación de su calidad y experiencias.
- Materiales educativos digitales. Metodologías para su diseño y producción. Objetos de aprendizaje. Multimedia e hipermedia en escenarios educativos. Nuevos entramados de medios, soportes y lenguajes.
- Formación de recursos humanos en el uso de TICs.
- Trabajo colaborativo mediado por TICs. Autorregulación y capacidades metacognitivas como factores claves para su desarrollo. Conceptualización, análisis y desarrollo de software y metodologías. Actividades colaborativas aprovechando dispositivos móviles.

3. RESULTADOS ESPERADOS/OBTENIDOS

Se presentan aquí los principales resultados que se han obtenido en el subproyecto, que son el foco de este trabajo y corresponden al período 2016 - inicios de 2017.

Entornos digitales: campus virtuales, entornos virtuales de enseñanza y aprendizaje (EVEAs), entornos virtuales 3D, entornos simulados y laboratorios, redes sociales.

Durante el 2016 se han realizado varios avances en relación a esta línea de I+D+I. Se investiga aquí sobre las posibilidades que ofrecen diferentes tipos de entornos, principalmente, centrados en la web, para los procesos de enseñar y aprender (Jaggars & Xu, 2016). Se analizan las tendencias tanto en investigación como en las experiencias que se llevan a cabo sobre estas temáticas. Se construyen marcos teóricos a partir de una lectura crítica de los fenómenos que se estudian y la literatura de referencia. Por ejemplo, se ha trabajado desde hace ya varios años, en una diferenciación conceptual entre campus virtual y entornos virtuales de enseñanza y aprendizaje. Se elaboran estrategias para llevar adelante diseños de prácticas educativas en estos entornos digitales.

En el marco del Seminario de Educación a Distancia de la Maestría en Tecnología Informática Aplicada en Educación, algunos de los investigadores del proyecto, llevan adelante un taller de EVEAS, en el que se realizan estudios comparativos sobre estos sistemas y se visualizan tendencias.

Se ha presentado la propuesta de tesis doctoral sobre temas vinculados a las habilidades de autorregulación de los alumnos en procesos educativos mediados por EVEAS. Se ha concluido un trabajo final de especialización, dirigido por un miembro del proyecto, y relacionado con la temática de tutorías virtuales (Zampar & Gonzalez, 2016).

Durante 2016, se abordó la reingeniería del entorno virtual de enseñanza y aprendizaje WebUNLP, que dio lugar al nacimiento de una evolución del mismo, denominada IDEAS. Esta reingeniería tuvo dos ejes principales, por un lado, atender a varias cuestiones que los docentes y alumnos fueron aportando a lo largo de los últimos años para enriquecer y mejorar el EVEA, y por otro lado, proveer un entorno de experimentación en el que se comiencen a integrar las investigaciones que se están llevando a cabo en relación a este tipo de sistemas. Para ello se trabajó en la generación de una arquitectura que permita la escalabilidad del entorno virtual a largo plazo, y que pueda nutrirse

del aporte de diferentes investigadores. En enero de 2017 se realizó la migración de los cursos de WebUNLP a IDEAS, y actualmente se encuentra en la etapa de estabilización del sistema.

A inicios del 2016, se finalizó una tesina de grado sobre “Desarrollo de una plantilla para Moodle que oriente a los docentes en la incorporación de herramientas de la Web 2.0” (Romanut, Gonzalez, Madoz, 2016).

También, se continúa investigando sobre la integración de estos entornos con repositorios de Objetos de Aprendizaje para la implementación de cursos abiertos masivos on-line (MOOC). Además, se analizan las posibilidades de los MOOC y se está iniciando un trabajo de Especialización sobre esta temática.

Respecto de la investigación sobre mundos virtuales 3D y entornos inmersivos, se ha finalizado la tesis de maestría de (Quintín, Sanz & Zangara, 2016). Esta tesis da continuidad al trabajo previo de Escobar, Sanz & Zangara (2015). La tesis investiga sobre el uso de *Second Life* para el desarrollo de sub-competencias lingüísticas y discursivas en inglés

Diseño, producción y evaluación de Materiales educativos digitales. Objetos de aprendizaje y sus Repositorios.

En 2016 se consolidaron los avances de 2015 en relación a la temática de Objetos de Aprendizaje (OA). Se aplicó la metodología de diseño de OA (CROA), elaborada durante 2014/2015, y se realizaron publicaciones al respecto (Sanz, Barranquero & Moralejo, 2016).

Se aprobó una tesis de maestría sobre sistemas recomendadores y ensambladores de OA (Astudillo, Sanz & Santacruz Valencia, 2016). La dirección de esta tesis ha estado a cargo de un miembro del proyecto y ha permitido fortalecer el vínculo con otras universidades, acorde a la procedencia de los investigadores participantes. Se han realizado publicaciones conjuntas como fruto de esta tesis (Astudillo, Sanz & Santacruz Valencia, 2016b).

También se ha avanzado en el desarrollo de una tesis doctoral que propone la investigación y desarrollo de *frameworks* para la creación de OA, y se ha realizado una comparativa de herramientas que posibilitan la creación de este tipo de materiales (Violini, Sanz & Pesado, 2016).

Por otra parte, se ha avanzado en el desarrollo de una tesis de maestría sobre materiales hipermediales en la enseñanza de la Matemática. La tesis analiza los marcos teóricos que fundamentan la integración de este tipo de materiales en la enseñanza de la Matemática. Lleva adelante una experiencia en donde se analizan las actitudes de los alumnos y su opinión sobre la utilización de un material

hipermedial, en el marco de una cátedra de Matemática en una carrera de Ingeniería (Del Río, Bucari & Sanz, 2016). La tesis ha sido entregada en febrero de 2017, y se espera su evaluación. Relacionados con esta misma temática y también con materiales multimediales, se han desarrollado y concluido tres trabajos de especialización dirigidos por miembros del proyecto (Balanta & Zangara, 2016) (Amiconi & Zangara, 2016) (Martorelli, Pesado & Sanz, 2016).

Se continuó con la investigación sobre un simulador que permite introducir los conceptos iniciales de la programación concurrente y *Cloud Computing* (De Giusti, Chichizola, Eguren, Sanchez, Paniego & De Giusti, 2016).

Se ha profundizado la investigación sobre la utilización de materiales educativos y actividades de Realidad Aumentada en procesos educativos. Se llevaron adelante experiencias con alumnos y docentes y se trabajó en su evaluación. Se avanza en una tesis de maestría con una becaria del equipo de investigación (Salazar, Sanz & Gorga, 2016), y se participa en diferentes proyectos con distintas universidades del país y del exterior en este tema

Trabajo colaborativo mediado por TICs.

En esta línea se abordan metodologías específicas para llevar adelante procesos colaborativos y educativos mediados por TICs. Se implementan herramientas para mediar estos procesos. Se ha desarrollado una metodología para el seguimiento de la colaboración, a partir de la definición de indicadores, que surgen de un estudio bibliográfico en profundidad. Se está realizando una tesis doctoral vinculada a estos temas de un miembro del proyecto y dirigida por otro miembro del mismo (Zangara & Sanz, 2016).

Se finalizó una tesis de doctorado, dirigida por un miembro del proyecto, que se focaliza en el uso de dispositivos móviles para el aprendizaje, y se han implementado una serie de juegos cooperativos y colaborativos, utilizando el marco de diseño *MADe-mLearn*. Se han concretado experiencias *ad-hoc* para el testeo y análisis de posibilidades de dichos juegos (Herrera, Sanz & Fennema, 2016). Uno de los miembros del proyecto que dirigió esta tesis doctoral, es a su vez asesor en estas temáticas en un proyecto de la Universidad Nacional de Santiago del Estero.

Se realizó durante 2016 un estudio sobre indicadores de actividades educativas en foros virtuales y se analizaron indicadores de seguimiento. Como fruto de estos avances se publicó un trabajo que da cuenta de los resultados (Sanz, Zangara & Dieser, 2016).

En cuanto a los proyectos vinculados con la temática y los acuerdos de cooperación, el III- LIDI participa en los siguientes:

- Se ha firmado un acuerdo de colaboración en estos temas con la Universidad de Zaragoza y se trabaja en forma conjunta. Durante el 2016 uno de los miembros del proyecto participó en un proyecto de innovación docente de la Universidad de Zaragoza.
- Se participa en la Red constituida por universidades de Iberoamérica en el marco del programa "Pablo Neruda" dentro del Espacio Iberoamericano del Conocimiento (EIC) y de la Organización de los Estados Iberoamericanos (OEI) orientada a la movilidad de estudiantes y docentes de doctorado. En este ámbito se ha realizado un libro con 3 capítulos relacionados con temas de investigación del proyecto y 3 capítulos vinculados a las investigaciones abordadas por un grupo de investigadores de México. El libro se imprimió en 2016.
- RedAUTI: Red temática en Aplicaciones y Usabilidad de la Televisión Digital Interactiva. En este ámbito se analizan materiales educativos para la TVDI.
- REFORTICCA: Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente. Proyecto que se desarrolla en el marco de los Proyectos De Innovación y Transferencia En Áreas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA). Período: 2017-2018.

4. FORMACION DE RECURSOS HUMANOS

En este proyecto se participa en la formación de recursos humanos a través de la dirección de tesis de doctorado, maestría y trabajos finales de especialización y tesinas de grado. Al mismo tiempo participan del proyecto, becarios que están realizando sus tesis de postgrado.

En 2016, se han aprobado 2 trabajos de Maestría y 4 de Especialización en el área de Tecnología y Educación, dirigidos por miembros de este proyecto. Hay 1 tesis doctoral vinculada al proyecto, finalizada en 2016, y también 2 tesinas de grado finalizadas.

Hay 1 trabajo de Maestría y otro de especialización terminados y esperando su evaluación, todos dirigidos por miembros del proyecto. Uno de los miembros del proyecto está terminando su tesis de doctorado.

5. BIBLIOGRAFIA

- Amiconi, D. & Zangara, A. (2016). Softwares educativos orientados a la comprensión lectora

en alumnos de escuela primaria. Trabajo final Integrador de la carrera Especialización en Tecnología Informática Aplicada en Educación. Disponible en:

<http://sedici.unlp.edu.ar/handle/10915/56485>

- Astudillo, G., Sanz, C. & Santacruz Valencia, L. (2016). Estrategias de diseño y ensamblaje de Objetos de Aprendizaje. Tesis de Maestría en Tecnología Informática Aplicada en Educación. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/53442>
- Astudillo, G., Sanz, C. & Santacruz Valencia, L. (2016b). Criterios para evaluar metodología de ensamblaje de objetos de aprendizaje. XVIII Simposio Internacional de Informática Educativa (SIIE). Proceedings con ISBN 978-84-9012-630-1, pp. 339 – 344. Desarrollado en Salamanca, España. Septiembre de 2016.
- Balanta, A., Zangara A. “Diseño, desarrollo y evaluación de material multimedia interactivo para la enseñanza de español como lengua extranjera a niños”. Trabajo de Especialización en TIAE. 2016. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/55470>
- De Giusti, L., Chichizola, F., Eguren, S., Sanchez, M., Paniago, J.M. & De Giusti, A. (2016). Introduciendo conceptos de Cloud Computing utilizando el entorno CMRE. Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016), ISBN: 978-987-733-072-4, págs. 1357-1365, octubre de 2016. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/56303>
- Del Río, L., Bucari, N. & Sanz, C. (2016). Del Río, L.; Bucari, N.; Sanz, C. Uso de Materiales Hipermediales para la enseñanza y el aprendizaje de la Matemática. 2do Congreso Internacional de Enseñanza de las Ciencias y la Matemática. 6 al 9 de Septiembre de 2016, UNICEN, Tandil, Bs. As. Argentina. ISBN 978-950-658-396-5 pp. 153-158
- Escobar Gutierrez, M., Sanz, C., Zangara, A (2015). Posibilidades educativas del entorno 3D Second Life para docentes. Estudio de caso con docentes de un postgrado de la Universidad Nacional de la Plata. Tesis de Maestría en TIAE. Aprobada en 2015. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/49862>
- Griffin, P., McGaw, B. & Care, E. (Eds.) (2012). *Assessment and teaching of 21st century skills*, Springer, Dordrecht.
- Herrera, S., Sanz, C. & Fennema, M. (2016). *MADe-mlern: marco para el análisis, diseño y evaluación de experiencias de m-learning*. Tesis doctoral en Ciencias Informática. Facultad de Informática. UNLP. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/55492>

- Jaggars, S. & Xu, D. (2016). How do online course design features influence student performance? *Computers & Education*. Volume 95, 270-284.
- Kuhn, Caroline. *Are Students Ready to (re)-Design their Personal Learning Environment? The Case of the E-Dynamic.Space*. *Journal of New Approaches in Educational Research*, [S.l.], v. 6, n. 1, p. 11-19, jan. 2017. ISSN 2254-7339. Available at: <<https://naerjournal.ua.es/article/view/185>>. Accedido el: 13 mar. 2017. doi:<http://dx.doi.org/10.7821/naer.2017.1.185>
- Martínez, J. (2004). *Concepción del aprendizaje, metacognición y cambio conceptual en estudiantes universitarios de Psicología*. Tesis doctoral, Universidad de Barcelona
- Martorelli, S., Pesado, P., Sanz, C. (2016). Microscopios virtuales: estudio y comparación. Trabajo de Especialización en TIAE. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/52695>
- Quintín, E., Sanz, C. & Zangara, A. (2016). El impacto de los juegos de rol en Second Life en el desarrollo de las competencias lingüística y discursiva en inglés a nivel oral. Tesis de Maestría en Tecnología Informática Aplicada en Educación. Junio de 2016. Facultad de Informática. UNLP. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/53622>
- Romanut, L., Gonzalez, A., Madoz, C. (2016). Desarrollo de asistencia en el selector de actividades de Moodle para facilitar la incorporación de herramientas de trabajo colaborativo. Tesina de grado aprobada en 2016. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/59092>.
- Salazar Mesía, N., Sanz, C. & Gorga, G. (2016). *Augmented Reality for Programming Teaching. Student Satisfaction Analysis. Proceedings of the 2016 International Conference on Collaboration Technologies and Systems*. En cooperación con IEEE, ACM and IFIP. 31 Octubre - 4 de Noviembre 2016, Orlando, Florida, USA – ISBN: 978-1-5090- 2300-4/16 \$31.00 © 2016 IEEE. DOI 10.1109/CTS.2016.43 – PP. 165-171
- Sanz, C., Barranquero, F. & Moralejo, L. (2016). CROA: a learning object design and creation methodology to bridge the gap between educators and reusable educational material creation. *EDULEARN 2016 - 8th annual International Conference on Education and New Learning Technologies*. Barcelona (Spain). 4th - 6th of July, 2016. Pages: 4583- 4592. ISBN:978-84-608-8860-4
ISSN:2340-1117. DOI: [10.21125/edulearn.2016.2101](https://doi.org/10.21125/edulearn.2016.2101)
- Sanz, C., Zangara, A. & Dieser, P. (2016). Modelo de seguimiento de una actividad de foro en la enseñanza universitaria. *Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*, ISBN: 978-987-733-072-4, págs. 431-440
- Violini, L. & Sanz, C. (2016). Herramientas de Autor para la creación de Objetos de Aprendizaje. *Actas del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*, ISBN: 978-987-733-072-4, págs. 353-362. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/55813>
- Zampar, M. & Gonzalez, A. (2016). Análisis de estrategias de tutorías virtuales. Revisión de la utilización de herramientas asincrónicas. Trabajo final integrador de la Especialización en TIAE. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/56270>
- Zangara, A. & Sanz, C. (2016) Building traceability indicators in collaborative processes using digital technology in educational settings. Presentation and comparison of educational and technological perspectives. 10th International Technology, Education and Development Conference. *INTED2016 Proceedings*. isbn 978-84-608- 5617-7, issn = 2340-1079. DOI 10.21125/inted.2016.0753. Valencia, Spain.

Mejorando Escenarios de Aprendizaje Colaborativo Soportado por Computadora

Rosanna Costaguta¹, María de los Ángeles Menini¹, Daniela Missio¹, Pablo Santana-Mansilla^{1,2}, Germán Lescano^{1,2}, Cecilia Acosta¹, Diego Yanacón Atía¹, Pablo Varas, Carlos¹ Pérez Crespo¹, María Martha Pérez Crespo¹, Alberto Palavecino¹ y Edgar Concha Medina¹

¹Instituto de Investigación en Informática y Sistemas de Información (IISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE)

²Consejo Nacional de Investigaciones Científicas y Técnica (CONICET)
{rosanna, marameni, dmissio, psantana, glescano}@unse.edu.ar; {valeriacecilia.acosta, pablo.varas10, palapaoli}@gmail.com; {diegopunk27, charlyfpc, edgard_707}@hotmail.com; mariamarthapc@yahoo.com.ar

RESUMEN

El término Aprendizaje Colaborativo Soportado por Computadoras (ACSC) se asocia con situaciones de enseñanza y de aprendizaje organizadas a través de actividades colaborativas, mediadas por computadora, en las que interactúan grupos de estudiantes y también docentes.

En ambientes de ACSC los estudiantes muestran un determinado modo de actuar cuando encaran las diferentes actividades grupales. Este comportamiento describe una forma de colaborar a través de sus habilidades de colaboración, pero además muestra otras características personales como podría ser cierta inclinación hacia un estilo de aprendizaje específico, o la natural predisposición a desempeñar un rol de equipo, o talvez, un estado emocional. Es evidente que se requieren mecanismos que permitan tener en cuenta estas características para monitorear e incentivar a los estudiantes a desempeñarse adecuadamente en sus actividades de aprendizaje.

En los ambientes de ACSC la intervención oportuna del docente o e-tutor resulta imprescindible. Corresponde al e-tutor estructurar y guiar los procesos colaborativos de los estudiantes, con el fin de generar las interacciones que propicien la construcción colaborativa del conocimiento. Los e-tutores requieren capacitación y formación para desenvolverse adecuadamente, por tanto,

resulta indispensable plantear mecanismos que les permitan adquirir las habilidades necesarias.

Este artículo describe un proyecto que propone el desarrollo de módulos para sistemas de ACSC que, mediante el análisis de interacciones, tanto de estudiantes como de docentes, puedan promover conductas adecuadas que beneficien los procesos de enseñanza y de aprendizaje, y tiendan a propiciar así, el éxito de las experiencias de colaboración. Los módulos a desarrollar utilizarán técnicas de Inteligencia Artificial y de Aprendizaje de Máquina, tanto para efectuar el análisis de las interacciones, como para promover las conductas que se consideran adecuadas. Los desarrollos se validarán mediante sesiones experimentales especialmente diseñadas en las que participarán grupos de estudiantes y docentes universitarios reales. Los datos recabados, producto de dichas experiencias, serán procesados utilizando tanto técnicas estadísticas como métricas específicas del área, lo que posibilitará comprobar o refutar la pregunta de investigación que se plantea.

Palabras clave: *Aprendizaje Colaborativo Soportado por Computadora, técnicas de Inteligencia Artificial, técnicas de Aprendizaje de Máquina, estudiantes colaborativos, e-tutores colaborativos.*

CONTEXTO

En este trabajo se presenta el proyecto “Mejorando escenarios de aprendizaje colaborativo soportado por computadora”, acreditado y financiado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SECyT-UNSE) para el período 2017-2020. El proyecto es una continuación de la línea de investigación *Sistemas Adaptativos Inteligentes*, llevada a cabo entre 2005-2009 por el proyecto 23/C062, “Herramientas conceptuales, metodológicas y técnicas de la Informática Teórica y Aplicada”, continuada entre 2009-2010 por el proyecto P09/C002, “Personalización en Sistemas de Enseñanza Virtual”, entre 2010-2011, por el proyecto 23/C089 “Fundamentos Conceptuales y Soportes Tecnológicos de la Informática Educativa”, y entre 2012-2016 por el proyecto 23/C097, “Sistemas de información web basados en agentes para promover el Aprendizaje Colaborativo Soportado por computadoras (ACSC)”. Todos los proyectos citados fueron acreditados y financiados por SECyT-UNSE.

1. INTRODUCCIÓN

El término Aprendizaje Colaborativo Soportado por Computadora (ACSC) se usa para designar situaciones en las cuales el aprendizaje y la enseñanza se organizan a través de actividades colaborativas mediadas por computadora, que involucran a grupos de estudiantes, orientados y asistidos por el docente o e-tutor. La principal característica de esta modalidad es que los participantes pueden trabajar colaborativamente ubicados en puntos geográficos distantes, e incluso, contribuyendo en momentos diferentes en el tiempo, acorde con sus respectivas disponibilidades.

En ACSC el concepto de grupo es fundamental. Un grupo colaborativo se define como un conjunto dinámico de estudiantes que trabajan juntos para alcanzar eventualmente alguna meta prefijada [Del2001], siendo además asistidos por un docente o e-tutor. Cada estudiante es responsable por sus

acciones, pero todos trabajan juntos sobre el mismo problema respetando las habilidades y contribuciones de todos, y donde los talentos de cada individuo sirven como recursos para los otros miembros del equipo, por esto, el éxito de uno, ayuda al éxito de los restantes integrantes. Es una premisa fundamental que los estudiantes que conforman grupos en ACSC son responsables de su propio aprendizaje, pero también del de sus compañeros. Cuando un estudiante no se comporta adecuadamente, su participación es deficiente y perjudica al desenvolvimiento del grupo como equipo. La aparición de comportamientos individuales disfuncionales impacta de manera negativa en el rendimiento grupal e impide alcanzar un ACSC adecuado [Mai1998], [Rom1994], [Pie2005].

Investigaciones recientes en ciencias de la computación, neurociencia, educación y psicología demostraron que la cognición y la emoción, están intrínsecamente relacionadas [Fra2010, Cial2003, Ekm2004], por ende, pueden influir sobre el rendimiento de una persona [Dam1994] y juegan un rol esencial en la atención y en la memoria, en el juicio, en la toma de decisiones y en la resolución de problemas creativos [Ise2000]. Las emociones son significativas en la interacción entre las personas, por lo tanto, es esencial poder reconocer cómo éstas se manifiestan en los estudiantes colaborativos y analizar su impacto sobre el aprendizaje del grupo.

En particular, en el contexto del ACSC, se requiere que el tutor actúe como un moderador o facilitador que promueva y coordine las discusiones grupales de modo tal de alcanzar los resultados de aprendizaje esperados [San2016]. El e-tutor tiene la responsabilidad de crear una atmósfera adecuada promoviendo la participación y la colaboración entre los estudiantes [San2016]. Dado que las interacciones que promueven el aprendizaje no ocurren espontáneamente, es imprescindible la participación de los e-tutores para fomentar las interacciones y resolver los conflictos grupales que pudieran producirse [Onr2012]. Aun, cuando las funciones de un e-tutor en ambientes virtuales no son una extensión y/o transferencia de las funciones requeridas en las

clases presenciales, entre los docentes persiste la tendencia de utilizar en ACSC prácticas educativas tradicionales, lo cual se traduce en desempeños inadecuados en el rol específico de e-tutor y este comportamiento impacta de manera desfavorable en el aprendizaje de los grupos que tienen a su cargo.

En resumen, el éxito o fracaso de una experiencia colaborativa está influenciado, entre otras variables, por la tendencia a comportarse, contribuir e interrelacionarse de una determinada manera de cada estudiante con respecto al resto de miembros del grupo y, por las habilidades que los e-tutores manifiesten en el seguimiento y moderación del trabajo grupal de sus estudiantes.

El desarrollo de sistemas de ACSC que utilicen técnicas de Inteligencia Artificial y de Aprendizaje de Máquina, permitirá ofrecer sistemas con la capacidad de monitorear las interacciones y adaptarse al comportamiento de sus usuarios (estudiantes y e-tutores). Estas capacidades de monitoreo y adaptación contribuirán de manera positiva en los procesos de enseñanza y de aprendizaje, ya que, permitirán detectar situaciones conflictivas y sugerir estrategias correctivas ajustadas a cada situación detectada y a cada integrante involucrado (estudiante o e-tutor). Por lo expuesto, en este proyecto se propone investigar teórica, metodológica y experimentalmente las contribuciones que puedan realizarse en el desarrollo de sistemas de ACSC que incluyan técnicas de Inteligencia Artificial y de aprendizaje de máquina para promover un ACSC exitoso.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO/OBJETIVOS

El éxito de las experiencias en ACSC depende de la conjunción de una serie de variables que caracterizan el comportamiento tanto de los e-tutores como de los estudiantes que participan en los procesos de enseñanza y de aprendizaje colaborativos, siendo esta conjunción fundamental para alcanzar los objetivos de enseñanza y de aprendizaje que se prevean. Dichas variables y sus influencias vienen

siendo analizadas por este grupo de investigación desde el año 2005 (como se expresó en la sección Contexto). En particular, la concreción de este nuevo proyecto permitirá realizar contribuciones puntualmente vinculadas con el desarrollo de sistemas de ACSC que incluyan técnicas de Inteligencia Artificial y de aprendizaje de máquina para promover un ACSC exitoso. Por lo expuesto, trabajos previos del grupo de investigación servirán de base para alcanzar los resultados esperados. Entre otros, se citan: [Cos2008], [Cos2009], [Cos2011], [San2011], [Cos2013], [Cos2014], [Cos2015] y [Les2016].

Para la presente investigación se establecieron los siguientes objetivos generales:

- Favorecer la generación de conocimiento científico-tecnológico de relevancia, para el desarrollo de sistemas en el área del ACSC.
- Realizar propuestas metodológicas y desarrollos de módulos para sistemas de ACSC, que usen técnicas de Inteligencia Artificial y de Aprendizaje de Máquina a fin de promover los procesos de enseñanza y de aprendizaje.
- Transferir conocimientos y ofrecer servicios al medio a través del asesoramiento y la capacitación de estudiantes de grado y posgrado, y también por medio del dictado de cursos y seminarios.
- Difundir los resultados obtenidos mediante la realización de publicaciones en revistas especializadas y de presentaciones en congresos, simposios y jornadas nacionales e internacionales.
- Formar recursos humanos mediante la realización de tesis de grado y de posgrado, y la dirección de becarios e investigadores en formación.

La pregunta central que guía el desarrollo de este proyecto es la siguiente: *¿Es posible mejorar el rendimiento académico de los estudiantes universitarios y el desempeño de los docentes, en los sistemas de ACSC, mediante el uso de técnicas de Inteligencia Artificial y de Aprendizaje de Máquina?*

Acorde con este interrogante, se fijaron los objetivos específicos que se enuncian a continuación:

- Determinar el estado actual de conocimiento y desarrollo de sistemas que apliquen técnicas provenientes de la Inteligencia Artificial y del Aprendizaje de Máquina en el ámbito del ACSC.
- Diseñar y construir módulos que apliquen técnicas de Inteligencia Artificial y de Aprendizaje de Máquina considerando características de los estudiantes tales como: habilidades de colaboración, roles de grupo, emociones, y estilos de aprendizaje, para favorecer los procesos de enseñanza y de aprendizaje en el ámbito del ACSC universitario.
- Diseñar y construir módulos que apliquen técnicas de Inteligencia Artificial y de Aprendizaje de Máquina para mejorar el desempeño de los docentes, considerando, por ejemplo, las habilidades que éstos manifiestan, a fin de favorecer los procesos de enseñanza y de aprendizaje en el ámbito del ACSC universitario.
- Evaluar el impacto producido por la inclusión de técnicas de Inteligencia Artificial y de Aprendizaje de Máquina, en los procesos de enseñanza y de aprendizaje, en sistemas de ACSC en contextos universitarios.

3. RESULTADOS ESPERADOS

Con la concreción de este proyecto se esperan obtener los siguientes resultados principales:

- Estado del arte vinculado con el uso de técnicas de Inteligencia Artificial y de Aprendizaje de Máquina en sistemas de ACSC.
- Especificación de las características consideradas para los estudiantes en sistemas de ACSC.
- Especificación de las características consideradas para los docentes en sistemas de ACSC.
- Diseño de módulos que incorporen técnicas de Inteligencia Artificial y Aprendizaje de Máquina para mejorar el rendimiento

académico de los estudiantes en sistemas de ACSC.

- Diseño de módulos que incorporen técnicas de Inteligencia Artificial y Aprendizaje de Máquina para mejorar el desempeño de los docentes en sistemas de ACSC.
- Implementación de los módulos diseñados.

Por otra parte, el desarrollo de este proyecto permitirá consolidar las actividades de investigación que llevan a cabo los integrantes de este grupo de investigación. Por otra parte, también se verán consolidados los esfuerzos realizados y resultados obtenidos con anterioridad en esta temática, derivados de los proyectos de investigación anteriores citados en la sección Contexto.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de este proyecto está constituido por un docente investigador formado y dos en formación. También integran el equipo dos becarios doctorales CONICET y un maestrando, los tres desarrollando sus tesis en el marco de esta investigación. Además, se contribuirá a la formación de estudiantes, ya que seis integrantes desarrollarán sus Trabajos Finales de Graduación en este proyecto.

5. BIBLIOGRAFÍA

- [Bul2001] Bull S., Greer J., McCalla G., & Kettel, J. 2001. User Modeling in I-Help: What, Why, When and how. En *Proc 8th. International Conference on User Modeling (UM'01)*, pp. 117-126.
- [Cha2001] Chandler, H. 2001. The Complexity of Online Groups: A Case Study of Asynchronous Distributed Collaboration. *ACM Journal of Computer Documentation*, Vol. 25 (1), pp. 17-24.
- [Cial2003] Cialdini, R. 2003. *Influence: Science and Practice*. Allyn Bacon.
- [Cos2008] Costaguta, R. & Amandi, A. 2008. Training collaboration skills to improve group dynamics, *Proceedings ACM of the 2008 Euro*

- American Conference on Telematics and Information Systems*, Aracajú, Brasil.
- [Cos2009] Costaguta R., Durán E., y Gola M. 2009. "El modelo B-learning y el aprendizaje colaborativo en una experiencia innovadora de enseñanza-aprendizaje". *VI Encuentro Nacional "La Universidad como Objeto de Investigación"*. Córdoba, Argentina.
- [Cos2011] Costaguta, R., García, P. y Amandi, A. 2011. Entrenando las habilidades de colaboración de los estudiantes mediante agentes. *IEEE Latin America Transactions*. Vol. 9(7), pp. 1118-1124.
- [Cos2013] Costaguta, R., Menini, M., Missio, D., Santana Mansilla, P. y Yanacón Atía, D. 2013. Modelos Multiagentes Recomendadores Adaptativos en Grupos. *V Congreso Internacional de Ambientes Virtuales de Aprendizaje Adaptativo y Accesible*, 2013.
- [Cos2014] Costaguta, R. & Menini M. 2014. An Assistant Agent for Group Formation in CSCL based on Student Learning Styles. *7th Euro American Conference on Telematics and Information Systems (EATIS 2014)*, Valparaiso, Chile.
- [Cos2015] Costaguta, R., Menini, M., Missio, D., Santana Mansilla, P., Lescano, G. y Yanacón Atía, D. 2015. Potenciando el Aprendizaje Colaborativo Soportado por Computadora: Algunas ideas traducidas en acciones. *XVI Virtual Educa*, Guadalajara, México.
- [Les2016] Lescano, G., Costaguta, R. & Menini, M. 2016. Applying Data Mining to Discover Successful Collaborative Groups Styles. *8th Euro American Conference on Telematics and Information Systems (EATIS 2016)*, Cartagena de Indias, Colombia.
- [Dam1994] Damasio, A. R. 1994. *Descartes' error: emotion, reason, and the human brain*. Grosset/Putnam Publications.
- [Del2001], Delgado, A.; Olgún, C. & Ricarte, I. 2001. Monitoring Learners Activities in a Collaborative Environment. *Proceedings of the 7th IEEE International Workshop on Groupware*. Germany.
- [Dil1999] Dillenbourg, P. 1999. Introduction: What do you mean by collaborative learning? In Dillenbourg, P. (ed.), *Collaborative learning: Cognitive and Computational Approaches*, Pergamon, Elsevier, pp. 1-19.
- [Ekm2004] Ekman, P. 2004. *Emotions Revealed: Recognizing Faces & Feelings to Improve Communication and Emotional Life*. Henry Holt & Co Publications.
- [Ell1991] Ellis, C. & Gibbs, S. 1991. Groupware: Some Issues Experiences. *Communic. of ACM*, Vol. 34 (1), pp. 38-58.
- [Fra2010] Frasson, C. & Chalfoun, P. 2010. Managing Learner's Affective States in Intelligent Tutoring Systems. *Advances in Intelligent Tutoring Systems*, vol. 308, Springer, pp. 339-358.
- [Ise2000] Isen, A. 2000. Positive Affect and Decision Making(ed). *Handbook of Emotions*, 2da Edición, Guilford Publications, pp. 417-435.
- [Onr2012] Onrubia, J. & Engel, A. 2012. The role of teacher assistance on the effects of a macro-script in collaborative writing tasks. *International Journal of Computer-Supported Collaborative Learning*, vol. 7(1), pp. 161-186.
- [Pie2005] Piezon, S. & Donaldson, R. 2005. Online groups and social loafing: Understanding student-group interactions. *Online Journal of Distance Learning Administration*, Vol. 8 (4).
- [Rom1994] Romero, R. 1994. *Grupo: Objeto y Teoría (Vol. II)*. Lugar Editorial, Buenos Aires, Argentina.
- [San2016] Santana Mansilla, P.; Costaguta, R. & Schiafino, S. 2016. Towards E-Tutors Training in On-Line Collaborative Learning. *8th Euro American Conference on Telematics and Information Systems (EATIS 2016)*, Cartagena de Indias, Colombia.

Minería de Datos y una Aplicación en la Educación Superior

Mag. Graciela Elida Beguerí, Mag. María Alejandra Malberti

Instituto de Informática / Departamento de Informática / Facultad de Ciencias Exactas
Físicas y Naturales / Universidad Nacional de San Juan

Av. Ignacio de la Roza 590 (O), Complejo Universitario "Islas Malvinas", Rivadavia, San Juan, Teléfonos:
4260353, 4260355 Fax 0264-4234980, Sitio Web: <http://www.exactas.unsj.edu.ar>
grabeda ,amalberti @gmail.com

RESUMEN

Se propone abordar la problemática que afecta a la educación en general y en particular a la universitaria, en cuanto a deserción, rezago y abandono. La idea, si bien ambiciosa, es contemplar dos líneas de acción. Por un lado las características de los alumnos, extrayendo conocimiento a partir del análisis de instrumentos propios por medio de algoritmos de aprendizaje de máquina, para contribuir con las acciones de apoyo de los tutores asignados. Por otro lado, el análisis y revisión curricular de las carreras de grado del Departamento de Informática-Facultad de Ciencias Exactas, Físicas y Naturales- Universidad Nacional de San Juan DI- FCEFNU-UNSJ con el fin de vincular las áreas disciplinares e incorporar, integrar y explotar simultáneamente distintas herramientas libres, provenientes de los aportes de las Tecnologías de la Información y la Comunicación –TIC-.

Palabras clave: Desempeño Académico, Aprendizaje Automático, TIC, Software Libre.

CONTEXTO

La línea de investigación se enmarca en el proyecto bianual 2016-2017 “La Ciencia de Datos en grandes colecciones de datos” código E1014 que se desarrolla en el ámbito de la FCEFNU-UNSJ, aprobado por el Consejo de Investigaciones Científicas Técnicas y de Creación Artística (CICITCA), financiado por la propia Universidad y ajustado a evaluación externa.

En el ámbito de la misma se pretende dar continuidad a los proyectos “Minería de Datos en la Determinación de Patrones de Uso y Perfiles de Usuario” y “Búsqueda de Conocimientos en Datos Masivos” llevados adelante a partir de 2012.

Es idea del grupo para la presente propuesta, profundizar el conocimiento y utilización de algoritmos específicos de aprendizaje de máquina que permitan la determinación de perfiles de alumnos universitarios, utilizando técnicas de caracterización multidimensional. Respecto al abordaje de problemas actitudinales se integran a la presente propuesta, participantes del proyecto “El comportamiento académico de los alumnos de primer año de las carreras de informática de la FCEFNU de la UNSJ Estrategias para mejorar su rendimiento” desarrollado en el periodo 2011-2013, Cod. 21/E875 res037/11CS, quienes han realizado experiencias teóricas y de campo en la difícil tarea de la caracterización de comportamientos de alumnos, detectando una compleja confluencia de factores que tienen que ver con la personalidad, aspectos vocacionales y su situación social, económica y familiar.

El estudio realizado por Clara Patricia Buentello y un equipo de colaboradores en 2013, titulado “Deserción Escolar, Factores que Determinan el Abandono de la Carrera Profesional” Estrategias y Condiciones para el Desarrollo del Estudiante, demostró que el principal factor es la distancia que existe entre las expectativas personales y las metas de la institución, así como los planes de estudio, el

tipo de docentes que son considerados como tradicionalistas y rígidos.

1. INTRODUCCIÓN

Algunos modelos psicológicos relativos a la deserción señalan que existen diferencias entre los alumnos que terminan sus estudios universitarios respecto de aquellos que no logran finalizarlos. La diferencia se halla en los rasgos de la personalidad. La teoría de la Acción Razonada, desarrollada por Ajzen y Fishbein (1975), explica que el debilitamiento de las intenciones iniciales de los estudiantes en pos de concluir sus estudios es -en rasgos generales- lo que provoca la deserción. Esta teoría de la Acción Razonada es considerada un predictor (Breva Asensio y Carpi Ballester, 2002), un juicio probabilístico en el que puede estimarse la intención de la conducta directamente mediante una escala de probabilidad (Reyes Rodríguez, 2007).

Gabriela Edith Huesca Ramírez y Belén Castaño Corvo, en “Causas de Deserción de Alumnos de Primeros Semestres de una Universidad Privada” citan la declaración de Vincent Tinto: *“Es importante reconocer que ciertas personas deciden a veces abandonar las instituciones de educación superior, sin alcanzar la graduación, sencillamente porque no tienen la meta o la intención de permanecer en ella hasta lograr esa graduación”*. También señalan que Romo y Fresán (Romo y Fresan, 2001) identifican como elementos explicativos de las dificultades para integrarse al medio académico y social de la institución: la falta de personalidad y madurez intelectual del estudiante, así como la falta de conocimientos y habilidades previas necesarias para realizar estudios superiores.

En la actualidad habilidades de aprendizaje de máquina son fuertemente requeridas en el mercado laboral, a raíz de que las empresas tienden cada vez más a construir sistemas de decisión automatizados. Esto conlleva a promover la incorporación en las currículas de las carreras dependientes del DI, de los saberes requeridos para la formación de

habilidades inherentes a un científico de datos. Datos concernientes a la identificación de las causas potenciales y los principales rasgos de personalidad que puedan interpretarse como riesgos para el buen desempeño académico serán relevados, en particular, entre los alumnos ingresantes y del primer año de las carreras del DI-FCEFNU-UNSJ, Para el tratamiento y análisis de los datos, se prevé utilizar, entre otras herramientas libres, Weka, R y RapidMiner.

- Weka: Waikato Environment Knowledge Analysis es un desarrollo en Java de la Universidad de Waikato, Hamilton, Nueva Zelanda desde 1993. Se descarga de <http://www.cs.waikato.ac.nz/~ml/weka/>
- R: (librería Rattle): desarrollado en la Universidad de Auckland, Nueva Zelanda, en 1993 y la librería Rattle, creada por el Dr. Graham Williams. www.r-project.org
- RapidMiner: (originalmente YALE) versión: 5.3.15, desarrollado, en Java (plataforma Eclipse) en la Universidad de Dortmund, Alemania, en lenguaje Java, desde 2001 <http://sourceforge.net/projects/rapidminer/>

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En el marco del proyecto se pretende vincular las áreas disciplinares e incorporar, integrar y explotar simultáneamente distintas herramientas libres, provenientes de los aportes de las Tecnologías de la Información y la Comunicación –TIC

También se procura:

- Analizar y describir el conjunto de estudios y prácticas requeridos en Ciencia de Datos.
- Construir y validar instrumentos tendientes a recabar datos inherentes a los estudiantes.
- Estudiar y analizar diferentes conjuntos de datos masivos a procesar. Evaluar herramientas de software libre para arquitecturas secuenciales, paralelas y distribuidas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Con el propósito de analizar las posibles relaciones existentes entre desempeño académico y tipos de personalidad, se aplica Minería de Datos a datos de 112 alumnos de una asignatura de primer año de las carreras dependientes del Departamento de Informática de la FCEF. Estos datos contienen, entre otros, la situación de cada alumno en la asignatura (Aprobado, Reprobado, Abandono, Promocionó) y los resultados de encuestas que caracterizan su personalidad en 9 tipos (Vargas, 2015)

La deserción y el rezago son problemáticas que preocupan cada vez más a autoridades y a docentes de las instituciones de educación superior. Con el propósito de generar acciones tendientes a prevenir situaciones que alteren el “normal” desempeño académico, se emplean algoritmos de clasificación para descubrir posibles relaciones entre las tendencias actitudinales según el tipo de personalidad y el desempeño académico en las carreras tratadas. Dada una colección de registros, en el que cada registro contiene un conjunto de atributos, siendo uno de ellos la clase, los algoritmos de clasificación hallan un modelo para el atributo de clase como una función de los valores de otros atributos. En este caso el atributo de clase es la situación académica del alumno después de cursar la asignatura, mientras que los restantes atributos tienen que ver con la caracterización del tipo de personalidad (desde 1 hasta 9).

Al momento de la elaboración de este documento, se ha obtenido, usando algoritmo J48 de Weka, lo expuesto en la figura 1. En el modelo se muestra la posible relación existente entre el tipo de personalidad y situación académica, plausible de ser validada análogamente en otros ambientes.

```

=== Run information ===
Scheme:   weka.classifiers.trees.J48 -U -M 2
Relation: DataTable-weka.filters.unsupervised.attribute.Remove-R1-10,12-21,23-296
Instances: 112
Attributes: 2
          Estado
          Dimension tipo
Test mode: 10-fold cross-validation
=== Classifier model (full training set) ===
J48 unpruned tree
-----
Dimension tipo = 1: REPROBADO (12.0/6.0)
Dimension tipo = 2: REPROBADO (10.0/4.0)
Dimension tipo = 3: REPROBADO (13.0/6.0)
Dimension tipo = 4: APROBO-BOLETA (2.0/1.0)
Dimension tipo = 5: PROMOCIONA (19.0/12.0)
Dimension tipo = 6: REPROBADO (14.0/5.0)
Dimension tipo = 7: REPROBADO (24.0/9.0)
Dimension tipo = 8: REPROBADO (8.0/2.0)
Dimension tipo = 9: REPROBADO (10.0/5.0)

```

Figura 1- Modelo de Clasificación generado por el algoritmo J48.

Algunos de los resultados esperados son:

Sugerir un conjunto de saberes convenientes para la formación de un científico de datos, para que los mismos sean incorporados, según corresponda, en las currículas de las carreras pertenecientes al DI-FCEF-UNSJ.

Realizar aportes con los datos tratados en las diferentes aéreas del saber.

Encontrar perfiles de alumnos universitarios utilizando técnicas de caracterización multidimensional que oriente en la posibilidad de analizar la existencia o no, de relación entre tipos de personalidad, conducta y desempeño académico y desde allí contribuir con acciones de apoyo.

Obtener una segmentación con el listado de los alumnos, según el riesgo en la regularidad académica.

Caracterizar a los estudiantes de acuerdo al tipo de personalidad o en términos de posibles dimensiones.

Contribuir, ante una situación de contingencia, con el trabajo del tutor en el área de tutorías de carreras.

4. FORMACIÓN DE RECURSOS HUMANOS

Los miembros del equipo de investigación, son docentes pertenecientes a las áreas curriculares de Ciencias Básicas, Algoritmos y Lenguajes, Teoría de la Computación e Ingeniería de Software-Base de Datos- Sistemas de Información, de las carreras del

DI, e integrantes de sucesivos proyectos de investigación en el área temática de minería de datos, desarrollados en el ámbito de la FCEF. Se destaca también que para el proyecto en ejecución se cuenta con tres adscriptos: una Lic. en Psicología y dos alumnos avanzados.

5. BIBLIOGRAFÍA

- Buentello Martínez, Clara Patricia "Deserción escolar, factores que determinan el abandono de la carrera profesional" Estrategias y condiciones para el desarrollo del estudiante. XVI Congreso Internacional sobre Innovaciones en Docencia e Investigación en Ciencia.
- Kargupta, H., Han, J., Philip, S. Y., Motwani, R., & Kumar, V. (Eds.). Next generation of data mining. CRC Press. (2008).
- Merlino, A., Ayllón, S., & Escanés, G. (2011). Variables que influyen en la deserción de estudiantes universitarios de primer año. Construcción de índices de riesgo de abandono/Variables that influence first year university students' dropout rates. Construction of dropout risk indexes. *Actualidades Investigativas en Educación*, 11(2).
- Peña, Daniel (2002), *Análisis de datos multivariantes*. Ed. Mc. Graw Hill, España.
- R <http://www.r-project.org/>
- Rapid-I <http://rapid-i.com/api/rapidminer-5.1/com/rapidminer/tools>. 2011.
- Ramírez, G., & Corvo, M. (2007). Causas de deserción de alumnos de primeros semestres de una universidad privada. *Revista Mexicana de Orientación Educativa*, 5(12), 34-39.
- Vargas, A. (2015). Eneagrama (Nueva edición):¿ Quién soy?. Alamah.

M-learning con Realidad Aumentada

Susana I. Herrera, María I. Morales, Rosa A. Palavecino, Marilena Maldonado, Ivana Irurzun, Alvaro J. Carranza

*Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
{sherrera, rosypgg, marilena}@unse.edu.ar, mines_morales@yahoo.com.ar,
ivanairurzun@gmail.com, carranza1903@hotmail.com*

Resumen

La investigación que se propone constituye una continuación del proyecto sobre Computación Móvil iniciada en el año 2012 en la Universidad Nacional de Santiago del Estero, en el que se lograron importantes resultados en m-learning.

Actualmente, en Argentina, el contexto móvil está cambiando con la implementación de dispositivos móviles 4G por parte de los usuarios y el ingreso de smartphones de alta gama. El monopolio del SO Android está desapareciendo y surge la necesidad de desarrollar aplicaciones nativas para diversos SO.

En esta propuesta se pretende investigar sobre plataformas o entornos de desarrollo que generen aplicaciones móviles nativas que puedan ser optimizadas mediante la modificación del código, y desarrollar aplicaciones móviles para m-learning, para la rehabilitación de personas con discapacidad y RA. Se buscará determinar, además, un método de desarrollo ágil adecuado para su elaboración usando herramientas de generación automática de código.

Palabras clave: m-learning, realidad aumentada, aplicaciones móviles.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado “Computación Móvil: desarrollo de aplicaciones y análisis forense”, financiado por el Consejo de Ciencia y

Técnica de la Universidad Nacional de Santiago del Estero, durante el período enero 2017-diciembre 2018.

El equipo de investigación del proyecto marco está conformado por el grupo de docentes que ya viene trabajando en el proyecto anterior y que tienen antecedentes en la temática, y que pertenecen al Instituto de Investigaciones en Informática (IISI) de la Facultad de Ciencias Exactas de la Universidad Nacional de Santiago del Estero (UNSE). También continuarán asesorando los siguientes investigadores: Dra. Cecilia Sanz (Laboratorio de Informática LIDI, Facultad de Informática de la Universidad Nacional de La Plata) y Mg. Sergio Rocabado (de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta). A través de los asesores del proyecto, el equipo de investigación interactuará con los mencionados centros de investigación. Asimismo, se trabajará en forma coordinada con el proyecto de investigación en Computación Móvil de la Facultad de Matemática Aplicada de la Universidad Católica de Santiago del Estero (se firmará un convenio específico de colaboración entre facultades).

Además, participan alumnos de Licenciatura en Sistemas de Información de la UNSE que ya se han venido formando en desarrollo e investigación en aplicaciones móviles.

2 Introducción

El m-learning es considerado principalmente como una nueva

modalidad de aprendizaje que surge de la mediación de las tecnologías móviles en el proceso de aprendizaje. Está relacionado con otras modalidades surgidas en el continuum “educación presencial–educación a distancia”, por ejemplo, con el e-learning y el u-learning [25].

En cuanto a las características del m-learning, las principales son: el ecosistema, los modos de interacción, los enfoques para la implementación de prácticas (donde se destaca la colaboración), la relación con la vida cotidiana y cuestiones generales de aprendizaje que son importantes para el diseño de prácticas de m-learning.

En cuanto a las experiencias desarrolladas, tanto en el plano internacional [6, 17, 18] como nacional [3, 24], cada vez son más diversas y se aplican a todos los niveles educativos desde el primario en adelante, acentuándose las experiencias en el nivel superior. En realidad, en Argentina están surgiendo con mucha fuerza, recién a partir del año 2015 [13, 4, 7].

En la UNSE, en el marco de la investigación vigente sobre Computación Móvil, principal antecedente de esta propuesta, se han obtenido importantes resultados. Se desarrollaron dos aplicaciones móviles para aprendizaje y se diseñaron e implementaron diversas prácticas de m-learning en escuelas primarias rurales, en escuelas secundarias, en carreras de grado, en carreras de posgrado, abarcando el aprendizaje en

Tecnologías,

Programación y Matemática [9, 20].

Dada la diversidad de experiencias, en los últimos años, algunos autores han propuesto frameworks que constituyen guías para el análisis de aplicaciones de m-learning; son ejemplos las propuestas de [22, 16, 19]. Estos marcos permiten clasificar las aplicaciones de m-learning, desde un punto de vista pedagógico y/o tecnológico. Sin lugar a dudas, uno de los resultados más importantes del proyecto

vigente, consiste en haber elaborado un marco sistémico y ecológico para el análisis, diseño y evaluación de experiencias de m-learning, denominado MADE-mlearn [10].

En cuanto al eje de la RA, esta es una tecnología que complementa la percepción e interacción con el mundo real y permite al usuario estar en un entorno real aumentado con información adicional generada por la computadora o algún dispositivo móvil [5]. Esta tecnología está introduciéndose en nuevas áreas de aplicación que van desde el esparcimiento hasta la investigación.

Los objetos 3D permiten a los estudiantes la interacción con objetos virtuales que complementan los objetos reales del entorno. Esto es importante dado que, en ciertas ocasiones, los objetos reales no pueden ser manipulados de manera directa, debido a distintas razones tales como: tamaño, coste, peligrosidad o distancia. La representación de objetos en 3D, permite una exploración espacial que no incluyen otros medios de representación en 2D. El nivel de abstracción que se exige en este último caso, dificulta que algunos estudiantes puedan entender realmente el espacio, tal como explica [15]. Los objetos 3D simulan objetos con volumen en el monitor y permiten al usuario una mejor abstracción, interacción y exploración de las características de un objeto [21].

En el ámbito educativo, donde esta propuesta pretende aplicar esta tecnología, según estudios recientes, tanto los profesores como los alumnos muestran mayor motivación cuando trabajan en actividades basadas en RA [8]. Una de las características claves de la RA en el ámbito educativo consiste en que involucra mayor interactividad entre el usuario y la aplicación informática, obliga la participación de la persona. Entonces, la RA es atractiva, dado que se alinea con el aprendizaje activo [15, 12] recalcan la importancia de los objetos 3D en el aprendizaje, ya que proveen:

sensación de tocar, sentido, orientación y posición en el espacio. Permiten ver y experimentar información que es dinámica e interactiva.

Si bien es notoria la importancia del uso de objetos 3D y RA en el aprendizaje, no es fácil la gestión y reutilización de estos objetos, al menos en el ámbito educativo. Según [1], si bien existen algunos repositorios de objetos 3D, no son de calidad o no presentan los recursos que se necesitan. Al mismo tiempo, es deseable contar con repositorios de objetos 3D educativos, que estén descriptos a partir de metadatos estandarizados para facilitar su almacenamiento, búsqueda y recuperación.

Algunas aplicaciones más recientes, realizadas con esta tecnología son: proyectos turísticos orientados al patrimonio, y que tienen como principal objetivo la reconstrucción, animación o visualización virtual (mediante el uso de modelos 3D) de monumentos y edificios emblemáticos que se encuentran actualmente en ruinas [26], sistemas de información geográfica a nivel catastral [23], en medicina para autopsia de patólogos forenses [11], aplicación móvil de compras de supermercado AR-asistida [2] entre otras.

Es precisamente ahora, cuando casi todo el mundo dispone de dispositivos móviles, que la RA llegó a prácticamente todos los usuarios de estas nuevas tecnologías de bolsillo. En otra línea del mismo proyecto de investigación se estudia sobre la optimización del desarrollo de aplicaciones móviles con RA, considerando diferentes alternativas [14].

3 Líneas de investigación y desarrollo

Esta investigación se organiza en torno a la siguiente línea:

- M-learning: Diseño de experiencias con RA para diferentes niveles educativos, diseño de repositorio de experiencias de m-learning y de repositorio de objetos 3D para RA.

4 Objetivos y resultados

Los objetivos de esta investigación son:

- Diseñar e implementar actividades de m-learning que involucren el uso de aplicaciones móviles multiplataforma basadas en RA, usando MADE-mlearn [10], tanto en nivel secundario como universitario.
- Diseñar un repositorio para las experiencias de m-learning basadas en MADE-mlearn.
- Analizar y clasificar herramientas (bibliotecas de funciones, herramientas de autor) para implementar RA en móviles en diversos formatos (texto, imágenes, objetos 3D).
- Caracterizar los objetos 3D que pueden ser usados en actividades educativas con RA en móviles.
- Diseñar un repositorio para la gestión de objetos 3D educativos para RA.

A partir de esta investigación los resultados que se esperan son:

- Experiencias de m-learning con RA para distintos niveles educativos.
- Repositorio de objetos 3D para RA.
- Repositorio de experiencias de m-learning.

5 Formación de recursos humanos

Esta investigación contribuye al desarrollo de dos tesis de Especialización en Enseñanza en Tecnologías (UNSE), una tesis de Maestría en Informática Educativa (UNSE) y tres trabajos finales de grado de la Licenciatura en Sistemas de Información (UNSE).

6 Referencias

1. Adair, M., Carina, F., Claudia, A., Susana, P., & Re. (2016). Computación Aplicada: Búsqueda y Desarrollo de Nuevas Estrategias de. XVIII Workshop de Investigadores en Ciencias de la Computacion. Concordia.
2. Ahn j. (2015). Supporting Healthy Grocery Shopping via Mobile Augmented Reality, journal ACM,

- volumen 12 Issues 1s, octubre 2015, article N° 16.
3. Arce, R. A. (2013). Mobile learning: aprendizaje móvil como complemento de una estrategia de trabajo colaborativo con herramientas Web 2 y entorno virtual de aprendizaje WebUNLP en modalidad de blended learning. En: Primeras Jornadas Nacionales de TIC e Innovación en el Aula. Recuperado el 19 de diciembre de 2015, de: <http://hdl.handle.net/10915/26538>
 4. Ascheri, M., Testa, O., Pizarro, R., Camiletti, P., Díaz, L., Di Martino, S. (2015). Inclusión de dispositivos móviles con sistema operativo Android en la enseñanza aprendizaje de temas de Matemática en el nivel medio. XVII Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-987-633-134-0. Salta. Disponible en: <http://hdl.handle.net/10915/46342>. Fecha de acceso: 19/12/2015.
 5. Azuma, R. (2001). *Augmented Reality: Approaches and Technical Challenges. Fundamentals of Wearable Computers and Augmented Reality*. Mahwah, New Jersey: W. Barfield, Th. Caudell.
 6. DeWitt, D., Siraj, S., & Alias, N. (2014). Collaborative mLearning: A Module for Learning Secondary School Science. *Educational Technology & Society*, 17 (1), 89–101.
 7. Díaz, F., Banchoff Tzancoff, C., Martín, S., Lanfranco, E. (2015). Software libre para dispositivos móviles. XVII Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-987-633-134-0. Salta. Disponible en: <http://hdl.handle.net/10915/46430>. Fecha de acceso: 19/12/2015.
 8. García Cabezas, S. (2013). *ugmented Learning Enviroments to enrich the classroom*. Londres.
 9. HERRERA, S., SANZ, C. (2014-a). Collaborative m-learning practice using Educ-Mobile. Estados Unidos de América. Minneapolis. Libro. Artículo Completo. Conferencia. International Conference on Collaboration Technologies and Systems (CTS) 2014. University of Minnessota.
 10. HERRERA, S., SANZ, C., FENNEMA, C. (2013-a). MADE-mlearn: un marco para el análisis, diseño y evaluación de experiencias de m-learning en el nivel de postgrado. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*. La Plata: Universidad Nacional de La Plata. vol. n°10. p7 - 15. issn 1850-9959.
 11. Kilgus T and. All. (2015). Mobile markerless augmented reality and its application in forensic medicine, May 2015, Volume 10, Issue 5, pp 573–586, *International Journal of Computer Assisted Radiology and Surgery*
 12. Krange, I., Fjuk, A., Larsen, A., Ludvigsen, S. (2002). Describing construction of. In *Proceedings of the Conference on Computer* (págs. 82-91). CSCL Community.
 13. Massé Palermo, M., Reyes, C., Ramirez, J., Trenti, J., Vargas, C., Espinoza, C., Figueroa, W., Tapia, C., Soria, M., Barrientos, O. (2015). Dispositivos móviles como soporte para el aprendizaje colaborativo de Programación en el nivel universitario inicial. XVII Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-987-633-134-0. Salta. Disponible en: <http://hdl.handle.net/10915/46426>. Fecha de acceso: 17/12/2015.
 14. Mocholí A. (2014). Claves y herramientas para desarrollar aplicaciones móviles de Realidad Aumentada, en: <https://www.yeeply.com/blog/desarro>

- llar-aplicaciones-movilesde-realidad-aumentada/
15. Moralejo, L. (2014). Análisis comparativo de herramientas de autor para la creación de actividades de realidad aumentada. La Plata.
 16. Navarro, C., Molina, A., Redondo, M., Juárez-Ramírez, M. (2015). Framework para Evaluar Sistemas M-learning: Un Enfoque Tecnológico y Pedagógico. VAEP-RITA Vol. 3, Núm. 1. ISSN 1932-8540. Pp. 38-45.
 17. Nouri, J. (2012). A theoretical grounding of learning mathematics in authentic real-world contexts supported by mobile technology. IADIS Mobile Learning, pp. 35-42. ISBN 978-972-8939-66-3. 2012.
 18. Nouri, J., Cerratto-Pargman, T., Eliasson, J., Ramberg, R. (2011). Exploring the Challenges of Supporting Collaborative Mobile Learning. International Journal of Mobile and Blended Learning, Volume 3 Issue 4, pp. 54-69. IGI Publishing Hershey, USA. ISSN: 1941-8647.
 19. Pachler, N., Bachmair, B., Cook, J. (2010). Mobile learning: structures, agency, practices. New York: Springer.
 20. PALAVECINO, R., HERRERA, S., SANZ, C., IRURZUN, I., CARRANZA, A. (2016). M-learning: aprendizaje de estructuras de datos con Ima-Colab. XI Congreso de Tecnología en Educación y Educación en Tecnología. RED UNCI Argentina. Morón. 2016.
 21. Paredes, R., Sánchez, J., Rojas, L., Strazzulla Martínez. (2009). Interacting with 3D learning objects. In Web Congress, (págs. 168-168). Latin American.
 22. Park, Y. (2011). A Pedagogical Framework for Mobile Learning: Categorizing Educational Applications of Mobile Technologies into Four Types. International Review of Research in Open and Distance Learning. ISSN 1492-3831. Vol. 12.2.
 23. Ramirez Navarro S. (2015). Estado del Arte-Desarrollo de una aplicación móvil basada en sistemas de realidad aumentada para la validación de sistemas de información geográfica a nivel catastral, Redes de ingeniería, Vol 6, Universidad distrital Francisco José de Caldas.
 24. Sanz, C., Cukierman, U., Zangara, A., Gonzalez, A., Santángelo, H., Rozenhauz, J. Iglesias, L., Ibañez, E. (2007). Integración de la tecnología móvil a los entornos virtuales de enseñanza y de aprendizaje. II Congreso TE&ET, Argentina.
 25. Zangara, A. (2014). Apostillas sobre los conceptos básicos de educación a distancia o...una brújula en el mundo de la virtualidad. Maestría en Educación a Distancia. Facultad de Informática de la Universidad Nacional de La Plata. La Plata.
 26. Zugazaga Echebarria, S. (2015). Realidad Aumentada aplicada a la tecnología móvil en el sector turístico, Memoria de Proyecto Final de Máster, Universidad Oberta de Catalunya, 2015.

Modelo para el Diseño de Pruebas Personalizadas

Huapaya Constanza, Guccione Leonel, Lazurri Guillermo

Grupo de Investigación en Inteligencia Artificial Aplicada a Ingeniería / Departamento de Matemática/ Facultad de Ingeniería/ Universidad Nacional de Mar del Plata
 Juan B. Justo 4302, Mar del Plata, Pcia. de Bs. As. , Argentina
 {constanza.huapaya,leonel.guccione,guillesky}@gmail.com

Resumen

En este artículo se presenta el diseño y el avance de la implementación de un sistema de Gestor de Evaluaciones utilizando el Modelo de Perturbación (GEMP). El núcleo del GEMP es un modelo de perturbación cuya finalidad es tomar decisiones sobre la conformación de pruebas personalizadas. Estas pruebas son utilizadas en una plataforma de aprendizaje. En este artículo se expone el mecanismo principal para diseñar las pruebas creadas a partir del árbol que representa el dominio a enseñar. Estas pruebas tendrán un nivel de personalización creciente. Inicialmente, el evaluador/usuario podrá seleccionar los ítems que forman una prueba, “podando” el árbol del dominio. Con estos ítems se construyen pruebas para plataformas de aprendizaje, como Moodle. Los resultados de las pruebas activan el modelo de perturbación. Se comienza el ciclo de personalización con sucesivas pruebas hasta que el estudiante alcance el nivel de conocimiento apropiado.

Palabras clave: pruebas personalizadas, modelo del estudiante, sistemas de gestión de pruebas, aprendizaje adaptativo, lógica difusa.

Contexto

Nuestra línea de I/D/I se encuentra inmersa en el desarrollo de Ambientes Virtuales de Aprendizaje (AVAs) orientados a la enseñanza/aprendizaje de la ingeniería. En particular, el proyecto actual en el cual está incluida la investigación se denomina “Adaptación en un ambiente virtual de aprendizaje: pruebas y materiales personalizados“. El proyecto fue aprobado por la Universidad Nacional de Mar del Plata.

Introducción

La adaptabilidad es una propiedad que define la habilidad de cambiar ajustándose a diferentes condiciones. Algo es adaptativo si es capaz de cambiar para ajustarse, por sí mismo o no, a las circunstancias que lo rodean.

Cada estudiante es un individuo único con diversos intereses, experiencias, y logros. Los materiales de estudio tradicionales, en general, no consideran estas diferencias. Actualmente se procura que los materiales asistan, enriquezcan y extiendan el currículo del estudiante a fin de mejorar sus habilidades y conocimiento bajo la consideración de sus características individuales.

La **adaptación** en sistemas web AVA puede verse, primero, como la adaptación de los contenidos (la cual trata de seleccionar los contenidos más relevantes a las necesidades del estudiante), y en

segundo lugar, la adaptación de la presentación de los contenidos (esto es, decidir cómo presentarlos de la forma más efectiva para el aprendizaje del alumno).

La **personalización** es entendida como la adecuación de los contenidos o visualización del sistema a la individualidad del usuario. En particular, la personalización web trata con la forma de proveer los contenidos, esto es, que estilo y formato usará para *cada estudiante*.

El proceso de adaptación está basado en información almacenada en el modelo del usuario y organizada en modelos específicos. Estos modelos dan la posibilidad de distinguir entre los usuarios y proveer al sistema de la habilidad de adecuar su reacción a la actividad del usuario (Brusilovsky y Maybury 2002). En nuestro caso, para alcanzar un buen nivel de adaptación del sistema en un AVA estamos desarrollando un modelo, GEMP, a fin de facilitar la construcción de pruebas personalizadas.

Representación del dominio como base del modelo de perturbación

La representación del dominio que usamos está basado en el modelo FR-CN (FR-CN, Fuzzy Related-Concepts Network, Chrysafiadi y Virvou 2015), pero, por ahora, solo hemos utilizado la relación es-parte-de. El modelo toma una estructura de árbol. En las figuras 2 y 3 se aprecia dos secciones del dominio de programación orientada a objetos propuestos por dos docentes/evaluadores distintos para el mismo tema. En la figura 4 se muestra la edición de un árbol del dominio en el sistema que se está desarrollando.

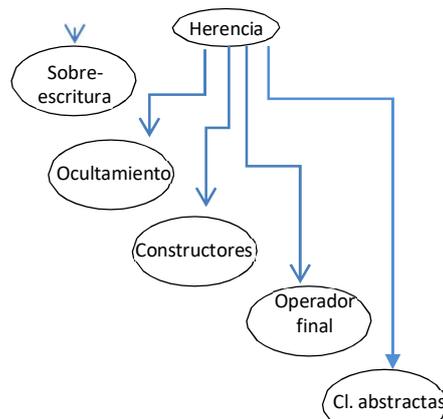


Fig 2: una sección del dominio de POO

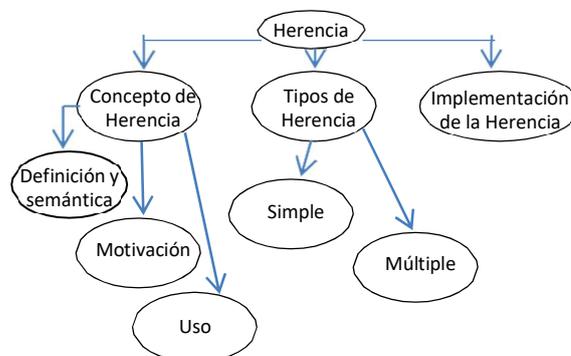


Fig. 3: otra forma de una rama del dominio de POO

Modelo de perturbación

El modelo de perturbación representa el grado de conocimiento que tiene el estudiante sobre cada tópico del dominio mediante el uso una medida cualitativa. Se han utilizado los siguientes cuatro conjuntos difusos para describir el conocimiento del estudiante en cada nodo del dominio: *desconocido*, *insatisfactoriamente conocido*, *conocido* y *aprendido*.

A cada nodo se le asocia una 4-upla formada por los valores de cada una de las funciones de pertenencia ($\mu_{desc}(x)$,

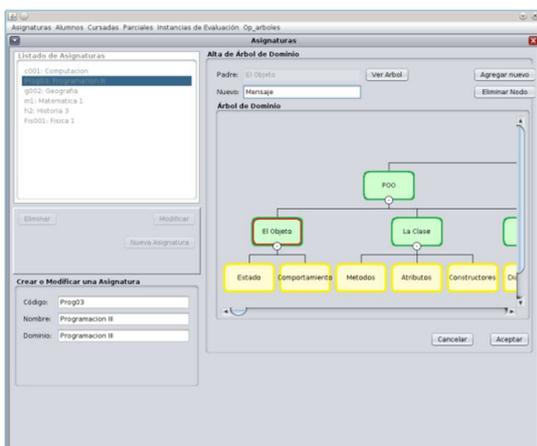


Fig 4: creación del árbol del dominio

$\mu_{insast}(x), \mu_{conoc}(x), \mu_{aprend}(x)$ a fin de expresar el conocimiento del estudiante sobre el concepto en evaluación, esto es, para valor de x , se evalúan las cuatro funciones de pertenencia.

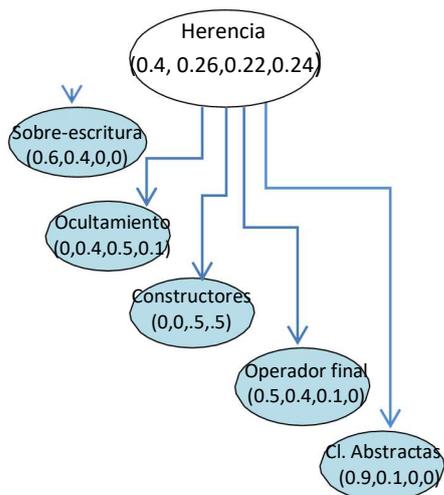


Figura 5: Modelo de perturbación

Para ilustrar su uso en un tópico como “herencia” presentado en la figura 5 se expresa que el estudiante desconoce el tema en un 40 %, posee un conocimiento insatisfactorio en un 26%; conoce el tema en un 22 % y lo aprendió en un 24 %. Estos valores de los conjuntos difusos se

recibieron de las hojas del árbol (en celeste), las cuales conformaron una o varias pruebas. Luego, con el programa cuya interface se aprecia en la figura 7, se calculó los valores del nodo “herencia”. La activación se ha calculado, en este caso, como promedios de los hijos de cada nodo.

Nuestro modelo mantiene una copia del árbol del dominio para cada estudiante con el progreso que muestra a lo largo de un periodo académico. Esto es, procesa un modelo de perturbación dinámico.

Construcción de una prueba

Para sustentar el modelo de perturbación se necesita construir pruebas (tanto para ejercicios o exámenes). Estas pruebas se conforman a partir de algunos nodos del árbol del dominio. El evaluador puede decidir eliminar tantos nodos como desee (ver figura 6). Esta “poda” deja un subconjunto de nodos del Árbol de Dominio, esto es, el sistema permite borrar del Árbol de Dominio los nodos (hojas o intermedios) que no forman parte de la prueba correspondiente.

Inicialmente, se construyen las mismas pruebas para todos los estudiantes del curso. Posteriormente, seleccionando los nodos con nivel de conocimiento más bajo, se arman nuevas pruebas, logrando de este modo pruebas personalizadas para cada estudiante. Las pruebas están conformadas por modelos de (Scalise y Guiford, 2006). Estos autores introducen una taxonomía o categorización de 28 tipos de pruebas de “restricción intermedia” muy útiles para la evaluación basada en computadoras.

Estas pruebas son utilizadas en un AVA (como Moodle). Los resultados

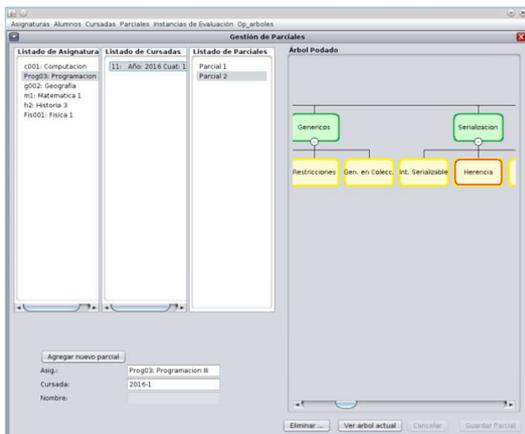


Fig. 6: Recorte de temas del dominio para armar pruebas

proveen información al modelo de perturbación (ver figura 7).

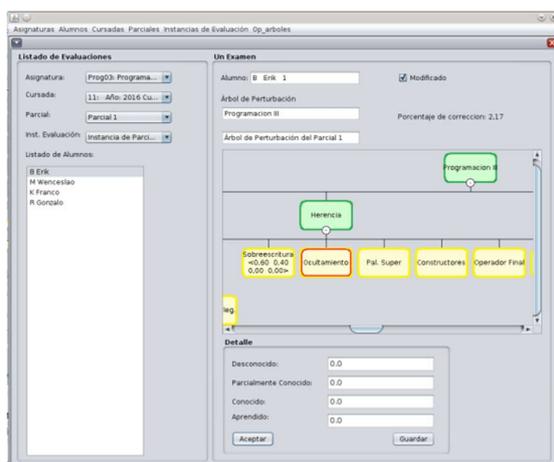


Fig. 7: Actualización de los conjuntos difusos en el modelo de perturbación

Sistema GEMP

El sistema (GEMP, Gestor de Evaluaciones utilizando el Modelo de Perturbación) tiene como objetivo la administración de evaluaciones de estudiantes universitarios. Posee editores para manipular los arboles del dominio y de perturbación, administra las pruebas, cursos y asignaturas. Próximamente se un sistema de reglas para decidir automáticamente la “poda” de nodos del

dominio según los resultado del modelo de perturbación.

Líneas de Investigación, Desarrollo e Innovación

- ✓ **Modelos eficientes de representación del dominio y pruebas personalizadas:** La adaptación se puede alcanzar con información individualizada del estudiante. Una forma de mejorar la adaptación tratar la incertidumbre con diversas técnicas. Aquí hemos utilizado la Lógica Difusa. Otro enfoque, actualmente seriamente explorado es analizar y volcar en los sistemas adaptativos la personalidad del estudiante (Okpo et al., 2016)
- ✓ **Adaptación en E-learning:** un sitio web adaptativo para la enseñanza/aprendizaje se caracteriza por su nivel de adaptación en su organización y presentación de contenidos al estudiante. Actualmente los entornos de aprendizaje de e-learning pueden vincularse tanto a un modelo de trabajo centrado en el aula y a la organización de los contenidos como al acceso flexible bajo demanda (sin tiempo ni restricciones de dispositivos) a los recursos educativos desde cualquier lugar (Sharples and Roschelle, 2010).

Resultados y Objetivos

Resultados

Se está desarrollando un sistema Gestor de Evaluaciones utilizando el Modelo de Perturbación (GEMP). En GEMP se aplica el paradigma de programación orientada a objetos, utilizando lenguaje JAVA. De esta manera, el sistema es multiplataforma. La interfaz se basa en la

librería Swing. Para lograr la representación gráfica de los árboles de perturbación se desarrolló una librería que permite una visualización específica de acuerdo a los atributos de cada nodo de perturbación. Para la persistencia se utiliza el motor de bases de datos MySQL. Se diseñó en base a una arquitectura Model View Controller. Se utilizaron las siguientes herramientas: JDeveloper versión 12 y MySQL Workbench versión 6.

Objetivo general: aumentar el nivel de personalización en pruebas para ambientes e-learning.

Objetivos específicos:

- ✓ Explorar nuevas técnicas para aumentar la personalización del aprendizaje.
- ✓ Analizar el modelo de personalidad FFM (Five-Factor Model: Openness, Conscientiousness, Extraversión, Agreeableness and Neuroticism) compuesto por los factores Apertura a la experiencia, Responsabilidad, Extraversión, Cordialidad y Estabilidad emocional (McCrae et al. 1996).
- ✓ Desarrollar un sistema de gestión de pruebas personalizadas.

Formación de Recursos Humanos

En el equipo de investigación hay especialistas en educación, informática, ingeniería y psicólogos cognitivos.

Afines de 2016 un integrante del grupo terminó el Magíster en Tecnología Informática Aplicada en Educación de la Facultad de Informática de la Universidad Nacional de La Plata. Actualmente hay tres postgrados cursándose en la Facultad de Informática de la UNLP: un Magíster en Tecnología Informática Aplicada en Educación a punto de finalizar y dos

Magísteres en Ingeniería de Software en curso.

Referencias

Brusilovsky P. y Maybury M.T. (2002) From adaptive hypermedia to the adaptive web. *Communications of the ACM*, vol. 45, no. 5 p.p. 30–33.

Chrysafiadi K.; Virvou M., (2015). *Advances in Personalized Web-Based Education*. Springer Cham Heidelberg.

Okpo, Juliet, Dennis, Matt, Masthoff, Judith, Smith, Kirsten A. y Beacham, Nigel (2016) Exploring requirements for an adaptive exercise selection system At PPALE 2016: 6th International Workshop on Personalization Approaches in Learning Environments, Canada. 16 Jul 2016. 6 pp.

McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the Five-Factor Model. In J. S. Wiggins (Ed.), *The Five-Factor Model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford.

Sharples, M., Roschelle, J. (2010). Guest editorial: special issue on mobile and ubiquitous technologies for learning. *IEEE Transactions on Learning Technologies* 3 (1), 4–5.

Scalise, K. & Gifford, B. (2006). Computer-Based Assessment in E-Learning: A Framework for Constructing “Intermediate Constraint” Questions and Tasks for Technology Platforms. *Journal of Technology, Learning, and Assessment*, 4(6).

Modelo para la Evaluación del Repositorio de Objetos de Aprendizaje

Cuenca Pletsch, Liliana R.; Maurel, María del Carmen; Cernadas, María Alejandra;
Sandobal Verón, Valeria C.

GIESIN (Grupo de Investigación Educativa Sobre Ingeniería) / Facultad Regional Resistencia /
Universidad Tecnológica Nacional
French N° 414, 54-0362) 4432683 / 4432928
cplr@frre.utn.edu.ar; mmaurel_38@yahoo.com.ar; ma_cernadas@hotmail.com, vsandobal@frre.utn.edu.ar

Resumen

El Proyecto “Evaluación del Repositorio de Objetos de Aprendizaje (ROA) de la Facultad Regional Resistencia” es una continuación de un proyecto anterior desarrollado por el mismo grupo de investigación GIESIN denominado “Modelización y desarrollo de un ROA para la gestión del conocimiento en la UTN-FRRE”. Dicho proyecto finalizó en diciembre de 2015 con un prototipo de Repositorio institucional y una herramienta para evaluar los Objetos de Aprendizaje según criterios y taxonomías seleccionados. La idea central del actual proyecto es establecer criterios de calidad contextualizados a la realidad local e institucional para evaluar el repositorio modelizado. La definición de criterios de calidad permitirá a las autoridades institucionales, conocer la eficiencia y calidad del servicio de ROA, además de detectar posibles fallos en cualquiera de los aspectos que lo forman. La calidad en ROA debe evaluarse en tres dimensiones: Organizacional, Formativa y Técnica. La primera está vinculada con la alineación del Repositorio a los objetivos institucionales. La segunda, está referida a la calidad de la formación on line ofrecida. La última está vinculada con la calidad de las plataformas, herramientas y objetos digitales puestos a disposición de la comunidad científica y académica.

Palabras clave: Objetos de Aprendizaje, Repositorios Institucionales, Repositorio de Objetos

de Aprendizaje, criterios de calidad, evaluación de calidad

Contexto

La línea de I/D presentadas en este trabajo forman parte de las actividades definidas en el marco del proyecto “Evaluación del Repositorio de Objetos de Aprendizaje de la Facultad Regional Resistencia”, homologado por la Secretaría de Ciencia y Tecnología de la Universidad Tecnológica Nacional (UTN).

El proyecto tiene como objetivo general proponer una metodología de evaluación de Repositorios de Objetos de Aprendizaje (ROA) ajustado al modelo de OA. Donde los objetivos específicos son: relevar el estado del arte en lo vinculado con la evaluación de calidad de los Repositorios Institucionales, de los recursos educativos y de la gestión educativa; definir los indicadores a incluir en la propuesta de evaluación de la calidad del ROA, teniendo en cuenta tres aspectos: organizacional, formativo y técnico; desarrollar nuevas, o adaptar herramientas existentes, para automatizar la generación de los indicadores que conforman las métricas de evaluación de la calidad del ROA; evaluar el ROA institucional y elaborar informes tendientes a su mejoramiento; realizar transferencias tecnológicas y/o de formación de recursos humanos en temas relacionados con los ROA y la evaluación de calidad de los mismos.

Se trata de una línea de investigación derivada de un proyecto previo, donde se

modelizó el repositorio de OA para la FRRe y se propuso una herramienta de evaluación para OA.

Introducción

Cada vez más instituciones incorporan las TIC a sus procesos educativos, ya sea porque tienen como propósito final ofrecer cursos no presenciales o semipresenciales o como complemento a la formación presencial. Tal como se afirma en [1], la incorporación de TIC en la educación se presenta en diferentes formas: simuladores, cursos de formación de cualquier área del conocimiento desarrollados en un entorno virtual, aplicativos multimedia, tutoriales que pretenden simular al maestro y fomentar el autoaprendizaje, documentos interactivos.

Los estándares para el desarrollo del elearning marcan la pauta para crear sistemas que integren aplicaciones para los procesos de enseñanza y aprendizaje en línea, en las que los contenidos puedan ser reutilizados y compartidos, entre personas y entre sistemas. A estos contenidos se denominan Objetos de Aprendizaje (OA) y con ellos se están formando bibliotecas de aplicaciones para el ámbito educativo, llamados Repositorios de Objetos de Aprendizaje (ROA) [2].

La concepción original de los OA proponía una pieza de software interactivo con una estructura de tres componentes: un objetivo educativo, materiales instruccionales para cumplir con el objetivo, y una evaluación para identificar el nivel de progreso de los aprendices que usaron el objeto [3]. El Comité de Estandarización de Tecnología Educativa de la IEEE establece que los objetos de aprendizaje son "una entidad, digital o no digital, que puede ser utilizada, reutilizada y referenciada durante el aprendizaje apoyado con tecnología".

Según Lowerison, Gallart y Boyd [4] los beneficios que los objetos de aprendizaje pueden tener en un contexto educativo son: flexibilidad, ya que el mismo recurso puede

utilizarse en distintos contextos; administración del contenido, que se facilita porque los recursos están descritos con metadatos que permiten su control; adaptabilidad, que facilita al diseñador poder seleccionar y componer recursos según la aplicación; y código abierto que elimina los problemas de incompatibilidad entre plataformas.

Evaluación de Repositorios

La principal motivación para la evaluación de un repositorio institucional se basa en su propósito final: dar servicios a usuarios de una comunidad designada (o comunidad destinataria) y, en este sentido, es a todas luces necesario evaluar la calidad del trabajo realizado. La construcción de un repositorio implica un largo proceso hasta alcanzar el nivel de satisfacción esperado, por lo que resulta necesario monitorear el avance o el progreso. Una vez alcanzado el nivel de satisfacción esperado por la comunidad designada se debería avanzar en la certificación de la calidad del Repositorio.

Durante el relevamiento bibliográfico se ha detectado que la investigación sobre evaluación de calidad de los Repositorios Institucionales se encuentra en pleno desarrollo. Existen guías (Stellenbosch University, 2013 [5]; Barrueco Cruz et al . 2010 2011 [6]; Harvard Open Access Project, 2014 [7]), directrices OpenAIRE Plus 2013, [8], y rankings web para evaluar la calidad de los mismos. Asimismo, existen recomendaciones para la certificación de confiabilidad de repositorios digitales (CCSDS 652.0 M1,2011)[9].

Se considera relevante mencionar la Guía de Evaluación de Repositorios Institucionales de investigación, propuesta por Barrueco Cruz, que toma como referencia a OpenAIRE (Open Access Infrastructure for Research in Europe) basado en las directrices DRIVER donde se toman como referencia los siguientes aspectos: implementación del protocolo OAI/PMH, obligatoriedad de la utilización de sets (agrupaciones) que definen las colecciones accesibles y lo relacionado a metadatos donde se establece como estándar el uso de Dublin

Core. Una opción analizada es la de Evaluación para Repositorios Institucionales lo presenta Cassella (2010) [10] que toma como base el Balance Score (BSC) propuesto Kaplan y Norton, el cual plantea evaluar la gestión del negocio a partir de cuatro perspectivas: la del cliente, financieros, los procesos internos y el crecimiento y aprendizaje. Otra posibilidad de evaluación es tomar los indicadores de naturaleza cibernética, donde se deben tener en cuenta las siguientes consideraciones [11]: "El repositorio es una ventaja competitiva para la universidad, pero ésta no se provee con su mera implantación: se necesitan estrategias. Un repositorio no debería ser simplemente un depósito, sino una herramienta para mejorar la descripción, organización y recuperación de la herencia de la institución: se proporcionan servicios. Debe ser tan importante la producción de conocimiento como su difusión, impacto y uso, que son aspectos cuantificables y evaluables".

Estas guías y recomendaciones aplican específicamente a los Repositorios que almacenan los resultados de Investigación (artículos, Tesis, Libros). Si se desea aplicar a los ROA, deberían realizarse adaptaciones según el modelo y taxonomía adoptados en la Institución.

Atento a que los OA implican objetivos de enseñanza y aprendizaje, su evaluación debe encararse como una tarea constante y, por lo tanto, debe realizarse durante las fases de Análisis y Obtención, Diseño y Desarrollo, con el fin de corregirlo y perfeccionarlo, así como durante su Gestión y Uso real por parte de los usuarios, para juzgar su eficacia y los resultados que con él se obtienen [1].

Antecedentes para la Evaluación del ROA de la FRRe

Para la evaluación de OA se tomará como referencia la propuesta realizada en el marco del Proyecto "Modelización y desarrollo de un Repositorio de Objetos de Aprendizaje para la gestión del conocimiento en la UTN-FRRe" que finalizó en diciembre de 2015, plasmado en [12], el cual se basa en lo planteado por Álvarez

Rodríguez y Muñoz Arteaga [13], que facilita la integración de OA y se ajusta al estilo de creación de estos recursos en la Facultad Regional Resistencia (FRRe) de la Universidad Tecnológica Nacional (UTN).[14].

La metodología propuesta en [15] considera los siguientes aspectos: Objetivos del aprendizaje, Contenidos, actividades, y Tecnológico. Cada uno de los ítems antes enumerados es evaluado según sea Necesario, Deseado o No se aplica, para cada objeto de la taxonomía.

Atento a que las metodologías de evaluación de repositorios institucionales consultadas son genéricas y se ajustan a objetos digitales (OD), la calidad en Repositorios de OA debe evaluarse en tres dimensiones: Organizacional, Formativa y Técnica. La primera está vinculada con la alineación del Repositorio a los objetivos institucionales. La segunda, está referida a la calidad de la formación on line ofrecida. La última está vinculada con la calidad de las plataformas, herramientas y objetos digitales puestos a disposición de la comunidad científica y académica.

Asimismo, el Grupo de Investigación en Ingeniería y Calidad del Software (GICS) de la FRRe, ha propuesto un modelo para evaluar la calidad de aplicaciones Web y ha desarrollado herramientas que se integran en un Framework que permiten la evaluación de la calidad del Software. Teniendo en cuenta que un repositorio es una aplicación web y el framework desarrollado por el grupo GICS permite la evaluación de este tipo de aplicaciones, resulta adecuado analizar la utilización del mismo para la evaluación de la dimensión técnica. De esta forma quedarán dos dimensiones pendientes de evaluación: la organizativa y la formativa, directamente relacionada con los aspectos educativos de un ROA. En este sentido, se pretende realizar aportes de herramientas de evaluación adicionales, que permitan evaluar al ROA en su totalidad.

Líneas de Investigación, Desarrollo e Innovación

Las líneas de investigación que aborda el proyecto están vinculadas con:

- El desarrollo de una metodología de evaluación de Repositorios de Objetos de Aprendizaje ajustado al modelo OA definido para la FRRe.
- El desarrollo o adaptación de una herramienta para automatizar la generación de los indicadores que conforman las métricas de evaluación de la calidad del ROA.

Resultados y Objetivos

Los resultados del proyecto aportarán al desarrollo tecnológico en el ámbito de la Tecnología aplicada a la Educación. Aportará a la eficiencia en el diseño y construcción de material educativo para la formación de profesionales y la oferta de formación continua, como así también para la disponibilidad de un Repositorio Institucional que cumpla con los criterios de calidad vigentes a nivel nacional e internacional y adaptado a las características de los OA incluidos en el mismo.

La concreción de los objetivos del proyecto permitirá continuar con el desarrollo de actividades de investigación en un área de vacancia en la región y completar el proyecto previo, en que se propuso un modelo de OA y se implementó un prototipo de ROA, mediante la evaluación de la calidad del mismo desde el punto de vista Organizacional, Formativo y Técnico.

Los resultados directos esperados del Proyecto son:

- Definición de una lista de indicadores de evaluación para el ROA de la FRRe y aplicación al ROA de la FRRe.
- Aporte de indicadores que se adapten al contexto local, pero que puedan ser utilizados en otros contextos.
- Transferencia de los resultados de la investigación a la Universidad y al sector educativo de la región, a través del trabajo

que se realiza en el marco del Programa de articulación con Escuelas Técnicas del Chaco y del Programa de formación Docente que incluye a los Institutos de Educación Superior de la Provincia.

Con respecto a los resultados indirectos:

- Integrar, reutilizar y compartir OA entre diferentes cátedras de la Facultad, enmarcado dentro de los estándares y criterios establecidos para el repositorio institucional bajo los lineamientos de contribuir a la mejora constante del proceso de enseñanza y aprendizaje.
- En el caso de la FRRe (también se verifica la escasez de) se aportará al incremento de proyectos de investigación relacionados con la carrera Ingeniería en Sistemas de Información, específicamente en la línea Ingeniería y Calidad del Software, por lo que la realización de este proyecto impactará positivamente en los estándares de calidad establecidos para la carrera.

Como primeros resultados puede mencionarse lo plasmado en [15], donde se presentan las características a tener en cuenta para la evaluación de ROA, como una herramienta en su primera versión manual de evaluación. Las características propuestas son resultados de una comparación y selección de los siguientes autores: Guía Recolecta [6], los Indicadores propuestos por Orduña-Malea [11] y el realizado por Casella [10], en donde se evalúan las dimensiones organizacional y técnica.

Formación de Recursos Humanos

La estructura del equipo de trabajo está conformada por:

- Director: Cuenca Pletsch, Liliana Raquel
- Investigadores de Apoyo: Maurel, María del Carmen y Cernadas, María Alejandra
- Investigador tesista: Sandobal Verón, Valeria Celeste
- Becario BINID: Branca, Fernando
- Becario Alumno UTN- SCYT: Romero, Clarisa

- Becario Alumno UTN- SAE: Seniquiel, Matías

La formación de recursos humanos es otro aspecto al cual puede aportar el proyecto, dado que constituye un nuevo ámbito para el desarrollo de los Proyectos Finales de Ingeniería en Sistemas de Información, Tesinas de la Licenciatura en Tecnología Educativa, Trabajos Finales de Especialización y Tesis de Maestría para obtener la titulación de posgrado en la Especialización en Ingeniería en Sistemas de Información, proyecto acreditado por la CONEAU y en la Maestría en Educación en Entornos Virtuales dictada por la Universidad Nacional de la Patagonia Austral (UNPA) mediante convenio con la UTN-FRRe. Asimismo, una de las integrantes de este proyecto, Valeria Sandobal Verón, desarrolla su tesis de Doctorado en el marco del mismo.

Esta iniciativa permitirá también incorporar becarios alumnos o recientes egresados en las becas de investigación que para tal fin propone la Secretaría de Ciencia y Tecnología de la UTN, como así también a los alumnos avanzados beneficiarios de las Becas de Investigación y Servicios (BIS) ofrecida por la Secretaría de Asuntos Estudiantiles de la Universidad.

Referencias

- [1] Bianchi, G. S.; Saenz Lopez, M. S.; Rosanigo, Z. B. (2008) "Objetos de Aprendizaje: Análisis de la Calidad". En XIV CACIC 2008. Universidad Nacional de La Rioja.
- [2] López Guzmán, Clara. (2005) "Los repositorios de OA como soporte para los entornos de elearning". Tesis doctoral. Universidad de Salamanca. España.
- [3] Barrit C. y Alderman F.L.. (2004) "Creating a Reusable Learning Objects Strategy". John Wiley & Sons, San Francisco, CA
- [4] Lowerison G.; Gallant, G. & Boyd, G. (2003). "Learning Objects in Distance Education: Addressing issues of Quality, Learner Control and Accessibility". CADEACED Conference.
- [5] Stellenbosch University, 2013. "Guía para la puesta en marcha de un repositorio institucional con Software Dspace".
- [6] Barrueco Cruz, José Manuel, coord. (2010). "Guía para la evaluación de repositorios institucionales de investigación". FECYT, RECOLECTA, CRUE y Rebiun
- [7] Harvard Open Access Project(2014). "Good practices for university openaccess policies".
- [8] OpenAIRE Plus(2013) "Directrices OpenAIRE 1.1: Directrices para proveedores de contenido del espacio de información".
- [9] Audit and Certification of Trustworthy digital repositories. Recommended Practice. CCSDS 652.0 M1,2011. The Consultative Committee for Space Data Systems.
- [10] Casella, Maria(2010). "Institutional Repositories: an Internal and External Perspective on the Value of IRs for Researchers" Communities. University of Turin.
- [11] Ordueña Malea, Enrique (2011). "Visibilidad de los repositorios institucionales argentinos en la Web: Indicadores y buenas prácticas". Universidad Politécnica de Valencia (UPV). Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación.
- [12] Cernadas, María Alejandra; Cuenca Pletsch, Liliana R.; Sandobal Verón, Valeria C. (2014) "Una propuesta de evaluación para Objetos de Aprendizaje". III Jornadas de Investigación del NEA y países limítrofes. Resistencia, Chaco, Argentina.
- [13] Álvarez Rodríguez, F.J.; Muñoz Arteaga, J. "Fundamentos del enfoque de objetos de aprendizaje". Cap.I, pp 20 a 29. Disp: <http://ingsw.ccbas.uaa.mx/sitio/images/libro/>
- [14] ASTD y SmartForce. (2003, sep. 30). "A field guide to learning object". Hacia la construcción de la sociedad de aprendizaje. Apertura
- [15] Sandobal Verón, Valeria C.; Cernadas, María A.; Cuenca Pletsch, Liliana R.; Maurel, María del C.(2016). ¿Pueden aplicarse las herramientas de evaluación web a la evaluación de Repositorios de Objetos de Aprendizaje?. III Congreso Argentino de Ingeniería. IX Congreso Argentino de Enseñanza de la Ingeniería. ISBN: 978-950-42-0173-1. Resistencia, Chaco. Argentina.

MOOC: Propuesta y Medición de la Calidad Percibida

Nombre del autor: Tumino, Marisa Cecilia;
Bournissen, Juan Manuel y Carrión, Federico
Universidad Adventista del Plata - Argentina

marisa.tumino@uap.edu.ar
juanbournissen@doc.uap.edu.ar
federicocarrion@al.uap.edu.ar

Abstract: El presente trabajo tiene como objetivo analizar comparativamente las principales plataformas MOOC disponibles bajo licencias abiertas, contribuyendo a la toma de decisiones de quienes pretenden desarrollar sus propios cursos masivos online, acerca de la plataforma a usar. El trabajo de investigación incluye (a) un análisis comparativo de las principales plataformas MOOC, (b) la selección de las dos plataformas que ofrecen más prestaciones, (c) la construcción e implementación de un módulo de curso bajo las dos plataformas seleccionadas, (d) la construcción y aplicación de un instrumento de medición de la percepción de calidad de los usuarios y (e) la medición de la motivación mediante el instrumento estandarizado a partir del Instructional Materials Motivation Survey (IMMS) y adaptado por Castaño, Maiz y Garay (2015). A partir del estudio se arribará a las conclusiones que permitan brindar información pertinente a quienes manifiesten interés en la construcción de MOOC.

Palabras clave: MOOC, Tecnologías educativas; medición de la calidad

CONTEXTO

El estudio se inserta dentro de la línea de investigación institucional: TIC aplicadas a la educación, dentro de la disciplina general: Contenidos Digitales Interactivos.

1. INTRODUCCIÓN

Las siglas MOOC aluden a cuatro características de la modalidad educativa: (a) es masivo puesto que es capaz de asimilar gran número de alumnos registrados y al mismo tiempo los contenidos tienen alcance global, (b) es abier-

to dadas las características de gratuidad y la inexistencia de requisitos para acceder a ellos, (c) es online puesto que los cursos que se desarrollan en modalidad habitual utilizan Internet y (d) es un curso que debe contar con una serie de elementos estructurados a fin de orientar los contenidos hacia el aprendizaje. El desarrollo de cursos abiertos masivos online (MOOC) es un fenómeno en pleno auge a nivel mundial. Estos cursos se asientan en plataformas tecnológicas cuyas características funcionales conviene tener en cuenta, más aún, tratándose de desarrollos de código abierto (Pereira, 2014). En este sentido, son importantes los detalles de carácter educativos tales como: (a) principales funcionalidades, (b) formas de analizar las prestaciones formativas para el alumnado en función del uso de la plataforma o (c) procedimientos para medir la percepción de la calidad de las herramientas utilizadas en los MOOC.

Las iniciativas de generación de contenido por parte del usuario, el aumento de las prácticas educativas abiertas (OEP), los cursos abiertos en línea y masivos (MOOC), y la creación de nuevos proveedores de soluciones de autoaprendizaje como OER University, Peer2Peer University o University of the People, están transformando escenarios conocidos en otros dominios de una naturaleza mucho más incierta. Esta tendencia plantea un reto para las instituciones conservadoras, especialmente las universidades. El aprendizaje es cada vez más informal a medida que las personas desarrollan redes complejas para ayudarse entre sí, de modo que intensifican la percepción de que la enseñanza y la formación ya no están proporcionadas exclusivamente por las instituciones, sino que ambas pueden también resultar de una colaboración

entre individuos y sus redes especializadas (Sangrá, 2013, p. 110)

Tal como lo señalan Koutropoulos y Hogue (2012), los MOOC proporcionan una nueva metodología y modalidad para la enseñanza y el aprendizaje. Esta novedad no solo plantea algunos problemas para los estudiantes, sino que también proporciona nuevas e interesantes posibilidades. Algunas de las habilidades de aprendizaje adquiridas por los estudiantes a través de cursos en línea y programas tradicionales son transferibles al aprendizaje MOOC. Sin embargo, la naturaleza distribuida del MOOC y la enorme cantidad de participantes no permite el método tradicional de acudir al maestro para obtener una explicación. Los MOOC requieren de los estudiantes que sean más proactivos en su educación y en la construcción de sus planes. Todos pueden tener éxito en un MOOC si se toman ciertas medidas y diseñan estrategias antes, durante, y después de un MOOC. Asimismo, de acuerdo con estos autores, las oportunidades que pueden ofrecer los MOOC irán en paulatino y veloz aumento.

1.1. Descripción de los principales tipos de MOOC

Cabero Almenara, Llorente Cejudo y Vázquez Martínez (2014), luego de un estudio pormenorizado de las tipologías de los MOOC, y sus implicaciones, señala que, independientemente de las amplias propuestas de MOOC, la tendencia se dirige a considerar dos tipos básicos de MOOC, que suelen denominarse xMOOC y cMOOC. Asimismo un último tipo de MOOC utilizado es el centrado en las tareas, tal como el tMOOC.

Los xMOOC comúnmente se corresponden con los cursos universitarios tradicionales de e-learning, mientras que los cMOOC se apoyan en la filosofía del aprendizaje conectivista de George Siemens y Stephen Downes.

Los cMOOC no se centran tanto en los contenidos, sino más bien en comunidades discursivas que crean el conocimiento de forma conjunta (Lugton, 2012). Los MOOC diseñados bajo esta perspectiva están basados en el aprendizaje distribuido en red y se fundamentan en la teoría conectivista y en su modelo de

aprendizaje (Siemens, 2007; Ravenscroft, 2011).

El modelo tMOOC responde a un híbrido que adopta los planteamientos de las propuestas xMOOC y cMOOC, y hace especial hincapié en la resolución progresiva de tareas y actividades del estudiante. Las actividades de los tMOOC incluyen variadas propuestas tales como la resolución de casos, lectura y análisis de documentos, construcción de recursos, análisis de sitios web, elaboración de blog y wikis, entre otros, todas ellas actividades que pueden ser desarrolladas tanto individual como colaborativamente. (Cabero & Román, 2005).

Los xMOOC vienen marcados desde la posición objetivista, mientras que los cMOOC desde la conectivista. Por su parte, los MOOC tienen como directriz la visión constructivista.

1.2. Implicaciones de la tipología de MOOC en los pilares de la educación de Delors
Cabero Almenara, Llorente Cejudo y Vázquez Martínez (2014) destacan la potencia de los xMOOC y los cMOOC en la adquisición de diferentes competencias de los estudiantes. Desde esta idea, Moya (2013) ha analizado ambos tipos de diseños, a la luz de los cuatro pilares de la educación rescatados del informe de Delors (1996), tal como son aprender a ser, aprender a aprender, aprender a hacer y aprender a convivir.

1.3. La medición de la motivación
La teoría de la educación a distancia de Holmberg (1989, citado en Cheng, 2014) se basa en la empatía que podría promover la motivación de los participantes para aprender y mejorar los resultados del aprendizaje. Holmberg (2003, citado en Cheng, 2014) argumentó que tales sentimientos de empatía se verían reforzados por un enfoque de diseño más conversacional, lo que permite una amistosa interacción mediada entre alumnos, tutores, consejeros y otros miembros del personal en la organización de apoyo. En MOOC, estos representantes podrían desempeñar un rol diferente, dado que un gran número de participantes podría requerir atención; sin embargo la comunicación entre los participantes podría

tener una influencia más poderosa en la creación de la empatía en MOOC.

La motivación ya ha sido identificada previamente por Milligan, Littlejohn y Margaryan (2013, citado en Castaño, Maiz y Garay, 2015), como una variable que contribuye a aumentar la participación y el éxito de los estudiantes.

1.4. Definición del problema

El desarrollo de cursos abiertos masivos online (MOOC) es un fenómeno en pleno auge a nivel mundial. Sin embargo, en varias instituciones educativas, la iniciativa se ha visto demorada por limitaciones que obedecen, entre otros aspectos, a la falta de recursos humanos capacitados y dedicados a este propósito.

La exploración de plataformas MOOC contribuye con planteos de nuevos desafíos como la elaboración de propuestas en pro de una transposición didáctica más efectiva.

El trabajo pretende brindar un instrumento de evaluación de las diferentes propuestas de MOOC adaptado y validado a los fines de optimizar estas iniciativas.

1.5. Justificación científica, académica-institucional y social

En respuesta a las demandas sociales de oportunidades de capacitación o formación profesional adaptadas a sus necesidades, el estudio propone ofrecer información relevante que conduzca a la capacitación de los docentes en el desarrollo de MOOC. Las limitaciones de los cursos online masivos, demanda una herramienta de evaluación que contemple las principales dimensiones a tener en cuenta a la hora de construir un MOOC.

1.6. Objetivos

1. Impulsar el desarrollo y la implementación de MOOC en los docentes de la UAP.
2. Adaptar y validar un instrumento de evaluación de la calidad de cursos virtuales adaptado a los cursos MOOC construidos.
3. Ofrecer capacitación a los docentes en el desarrollo de MOOC.
4. Brindar lineamientos básicos de los aspectos que inciden en la motivación de los cursos MOOC.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se ha trabajado en las siguientes líneas de investigación y desarrollo: a) Análisis de plataformas MOOC, b) Selección de herramientas, c) Adecuación de instrumento de medición de calidad percibida de MOOC, d) Adecuación del instrumento de motivación en usuarios de MOOC, e) Diseño de directrices comunes a la construcción de MOOC y f) Implementación del MOOC y ejecución de las pruebas estadísticas a partir de la administración del instrumento de evaluación. Esta última etapa se encuentra actualmente en desarrollo.

3. RESULTADOS

A partir del estudio se arribará a las conclusiones que permiten brindar información pertinente a quienes manifiesten interés en la construcción de MOOC.

El estudio se lleva a cabo en una universidad de la Mesopotamia Argentina, con una muestra conformada por estudiantes de un curso de la institución.

Como se observa se trata de un estudio de campo con un abordaje cuantitativo dado que se analizan las puntuaciones obtenidas de la evaluación y de la medición de la motivación de los participantes, una vez implementado el MOOC a partir de los procedimientos de selección, análisis y evaluación previos.

3.1. Análisis comparativo de plataformas

Para realizar la comparación de herramientas de MOOC se han elegido aquellas que son libres y de código abierto, lo que permite ser instaladas y configuradas sin costos para una institución educativa. Asimismo, estas características evitan la dependencia de terceras instituciones para publicar los cursos, si bien algunas de ellas brindan las dos opciones, como lo es Edx que se encuentra disponible para subir cursos en el MIT y que también puede descargarse e instalar la plataforma en servidores propios.

De las plataformas disponibles en el mercado, las que cumplen con los requisitos mencionados son las siguientes: (a) Lernanta: P2PU (Peer to Peer University de California), (b)

OpenMOOC: UNED (Universidad Nacional de Educación a Distancia), (c) edX: MIT (Massachusetts Institute of Technology) y (d) CourseBuilder: Google.

Si bien existen otras plataformas abiertas, estas cuatro, según Pereira, Sanz-Santamaría y Gutiérrez (2014) son las de mayor impacto en el mercado internacional de los MOOC. Para este trabajo se descarta CourseBuilder debido a que Google discontinuó el proyecto de publicar nuevas versiones y aunó sus esfuerzos con el MIT, apoyando la plataforma edX. Por su parte, se incluye Moodle, como plataforma para desarrollo, dado que es la plataforma de educación virtual LMS (Learning Management Systems) que ya posee instalada la universidad que financia la investigación. Aunque no consiste en una plataforma específica, fue la plataforma que se utilizó para el primer MOOC dado que puede adaptarse para soportarlo, como lo menciona el trabajo del Gobierno Vasco (2014).

Se procedió a describir comparativamente las características de las herramientas basadas en distintos trabajos de investigación y en las experiencias del equipo de investigación para edX y Moodle.

3.2. Selección de las plataformas

A partir del estudio comparativo, la selección de las plataformas para construir MOOC se fundamentó en el cumplimiento de las condiciones deseadas para el trabajo. Las plataformas que reúnen las mencionadas condiciones son edX y Moodle, plataformas disponibles y utilizadas para diseñar un mismo MOOC a ser evaluado desde la calidad percibida por los usuarios.

El MOOC diseñado en ambas plataformas, apela a las directrices fundamentales que deben ser atendidas en el diseño y construcción de un MOOC. Se destaca que el principio fundamental que se contempló en los MOOC fue el de la co-construcción de los saberes entre docentes y estudiantes. Para ello resulta indispensable la incorporación de un curador de contenido en el equipo administrador de un MOOC.

Por su parte, se selecciona la tipología del cMOOC, dado que refleja claramente el

aprendizaje del ser, del aprender, del hacer y el de convivir, ya que implica la conexión con el resto de los participantes, sus interacciones, el crecimiento y el desarrollo de las personas, manteniendo la esencia del aprendizaje continuo y la relación con los demás. Construcción de un módulo de curso bajo las plataformas seleccionadas.

En este apartado se sintetizan las consideraciones que se tienen en cuenta a la hora de construir los MOOC propuestos y los componentes constituyentes del módulo de MOOC implementado “Construyendo un MOOC”.

3.3. Adecuación de un instrumento de medición de la calidad percibida de MOOC

El instrumento aplicado para evaluar los MOOC resulta de una adaptación del instrumento validado por Mengual-Andrés, Lloret Catalá y Roig Vila (2015), al que se le incorporaron dimensiones propuestas por Baldomero, Salmerón y López (2015).

El instrumento así adaptado, se sometió al juicio de cuatro expertos con el propósito de obtener evidencia de validez de contenido de las dimensiones adoptadas para el estudio. Finalmente se incorporaron las recomendaciones de los expertos, quedando conformado por:

1. Dimensión 1: Calidad en la comunicación y elementos multimedia con 19 ítems.
2. Dimensión 2: Coherencia curricular y adaptación al usuario con 9 ítems.
3. Dimensión 3: Planificación didáctica con 18 ítems.
4. Dimensión 4: Desempeño de los docentes con 9 ítems.

3.4. Adecuación del instrumento de motivación en usuarios de MOOC

Los datos relacionados con la motivación se obtienen mediante la escala IMMS (Instructional Materials Motivation Survey) que los participantes del MOOC completan al finalizar el curso. Se trata de un cuestionario tipo Likert formado por 36 ítems divididos en cuatro categorías (atención, confianza, satisfacción y relevancia) basado en la propuesta de Di Serio, Ibáñez y Delgado (2013, citado en Castaño, Maiz y Garay, 2015), con un coeficiente de fiabilidad documentado de 0,96, y

adaptada ligeramente al campo de los MOOC por Castaño, Maiz y Garay (2015).

Cabe señalar que las primeras tres dimensiones responden a los aspectos a ser evaluados por los diseñadores de MOOC, mientras que la cuarta dimensión responde a las expectativas del estudiante respecto al desempeño docente, por lo que se reserva esta dimensión para ser evaluada por los estudiantes, una vez finalizado un curso de MOOC.

Una vez recolectados los datos se procederá a ejecutar las pruebas estadísticas correspondientes a los efectos de identificar diferencias estadísticamente significativas de los valores obtenidos de la evaluación entre los dos grupos de estudio.

3.5. Criterios a considerar en la elaboración de videos tutoriales

A los efectos de elaborar una síntesis de las principales recomendaciones para grabar videos educativos, se adaptaron y modificaron los consejos de Scipion (s/f), sumando aquellos que se consideraban de utilidad a partir de la experiencia de los investigadores.

Las recomendaciones se resumen en los siguientes ítems:

1. Asegúrate de entender las necesidades de tu target
2. Sé breve, la atención disminuye rápido
3. Regla de Marketing: la regla de Anuncia – Enseña – Anuncia
 - 3.1. Preséntate a la audiencia
 - 3.2. Explica de qué va a tratar tu vídeo y por qué es relevante
 - 3.3 Explica o desarrolla la temática
 - 3.4 Conclusión
 - 3.5 Llamada a la acción
4. Trabaja el título de tu vídeo
5. Prepárate antes de apretar el botón “Grabar”

Estas recomendaciones, sumadas a las sugerencias e instructivos de aplicaciones disponibles y accesibles en forma libre y gratuita para la grabación de videos educativos, conforman una caja de herramientas básicas al momento de crear videos como elementos constituyentes de los MOOC.

4. RECURSOS HUMANOS

El presente trabajo de investigación es desarrollado por un equipo conformado por dos docentes y un estudiante de la carrera de sistemas.

5. REFERENCIAS BIBLIOGRÁFICAS

Baldomero, M., Salmerón, J. L. y López, E. (2015). Comparativa entre instrumentos de evaluación de calidad de cursos MOOC: ADECUR vs Normas UNE 66181:2012. *RUSC. Universities and Knowledge Society Journal*, 12(1), págs. 131-145. doi <http://dx.doi.org/10.7238/rusc.v12i1.2258>

Cabero Almenara, Julio, Llorente Cejudo, M^a del Carmen y Vázquez Martínez, Ana Isabel. (2014). Las tipologías de MOOC: su diseño e implicaciones educativas, 18(1): 13-26. Recuperado de <http://www.ugr.es/~recfpro/rev181ART1.pdf>

Cheng, J.C.Y. (2014). An Exploratory Study of Emotional Afford -dance of a Massive Open Online Course. *European Journal of Open, Distance and e-Learning* 17(1), 43-55. Recuperado de <http://goo.gl/sJuxAh>

Delors, J. (1996). La educación encierra un tesoro. Madrid: Santillana.

Gobierno Vasco (2014). Comparación de diferentes plataformas de código abierto. Recuperado de: http://edx.asmoz.org/pdf/estudio_comparativo.pdf

Koutropoulos, A., & Hogue, R. J. (2012). How to succeed in a MOOC-Massive online open course. *eLearning Guild, October 08, 2012*. Retrieved from [http://www.cedma-europe.org/newsletter%20articles/eLearning%20Guild/How%20to%20Succeed%20in%20a%20MOOC%20-%20Massive%20Online%20Open%20Course%20\(Oct%2012\).pdf](http://www.cedma-europe.org/newsletter%20articles/eLearning%20Guild/How%20to%20Succeed%20in%20a%20MOOC%20-%20Massive%20Online%20Open%20Course%20(Oct%2012).pdf)

Mengual-Andrés, S., Roig Vila, R., y Lloret Catalá, C. (2015). Validación del Cuestionario de evaluación de la calidad de cursos virtuales adaptado a MOOC.RIED. *Revista Iberoamericana de Educación a Distancia*, 18 (2), 145-169. doi:<http://dx.doi.org/10.5944/ried.18.2.13664>.

Moya López, Mónica. Los MOOC/COMA: un nuevo reto educativo para el siglo XXI. Una metodología didáctica para el aprendizaje en línea. *Virtualis*, 4 n° 8, (2013): 84-103. Recuperado de <http://aplicaciones.ccm.itesm.mx/virtualis/index.php/virtualis/index>

Pereira, J., Sanz-Santamaría, S., Gutiérrez, J. (2014). Comparativa técnica y prospectiva de las principales plataformas MOOC de código abierto. *RED, Revista de Educación a Distancia*. Número 44. Número monográfico sobre “Buenas prácticas de Innovación Educativa: Artículos seleccionados del II Congreso Internacional sobre Aprendizaje, Innovación y Competitividad, CINAIC 2013”. 15 de noviembre de 2014. Recuperado de: <http://www.um.es/ead/red/44>

Scipion, F. (s/f). Material multimedia didáctico: ¿cómo hacer un vídeo educativo de primera para youtube? Recuperado de <https://www.lifestylealcuadrado.com/material-multimedia-didactico-como-hacer-un-video-educativo/>

Procesos de Diseño de Componentes para Espacios Virtuales de Trabajo Orientados a la Educación

Darío Rodríguez, Roberto García, Norberto Charczuk, Federico Ribeiro, Hernan Amatriain, Rodolfo Priano, Ramón García-Martínez

Laboratorio de Investigación y Desarrollo en Espacios Virtuales de Trabajo
Grupo de Ingeniería de Espacios Virtuales de Trabajo y Grupo Investigación en Sistemas de Información

Departamento de Desarrollo Productivo y Tecnológico. Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús. Argentina. Tel +54 11 5533 5600 Ext. 5194
dariorodriguez1977@gmail.com, rgm1960@yahoo.com

Resumen

La evolución de la calidad de las comunicaciones basadas en la tecnología de Internet es la base de la tendencia de desarrollo de los Espacios Virtuales de Trabajo (EVT). Como una familia emergente de aplicaciones a desarrollar, necesita de un abordaje ingenieril que defina los distintos procesos constructivos. Es así que, la Ingeniería de Espacios Virtuales de Trabajo (IEVT) entiende en los procesos y las metodologías utilizadas para: ordenar, controlar y gestionar la tarea de desarrollo de los EVTs y todos los artefactos software integrables a estos; así como la concepción de las metodologías y el establecimiento de buenas prácticas de uso de los mismos. En este proyecto se busca definir tres procesos: [a] Proceso de Diseño de Dispositivos Educativos Centrados en EVTs, [b] Proceso de Diseño de Objetos de Aprendizaje inteligentes Integrables a EVTs, y Proceso de Evaluación de EVTs Orientados a Educación.

Palabras clave: Diseño de Componentes, Procesos de Diseño, Espacios Virtuales de Trabajo Orientados a Educación.

Contexto

El proyecto que se presenta se articula con la Línea de Investigación Prioritaria “3. Desarrollos Informáticos” del Instituto de Economía, Producción y Trabajo en el marco de la Resolución del Consejo Superior N° 113/14 de la Universidad Nacional de Lanús, promoviendo

instrumentos metodológicos orientados a la mejora del sistema productivo de la industria del Software con foco en Espacios Virtuales de Trabajo. Por otra parte responde a la interpretación que de dicha Línea de Investigación se presenta en el Documento de Trabajo “Líneas de Investigación, Programas I+D+I, Grupos de I&D y Producción Científica” del Departamento de Desarrollo Productivo y Tecnológico cuando propone estructurar dentro de la Línea del Campo Epistemológico “Productividad en Informática” el Programa I+D+I en Ingeniería de Espacios Virtuales de Trabajo.

Introducción

Los Espacios Virtuales de Trabajo (EVT) están destinadas a facilitar la mediación en el interior de equipos cuyos miembros no están físicamente contiguos [Gibson y Cohen, 2003], y tienen que desarrollar un objeto conceptual (por ejemplo: investigación, desarrollo de proyectos software, artículos técnicos, informes, documentación de diseño de edificios, planes de negocio, planes de inversión corporativos, entre otros). El EVT debe satisfacer el requisito de mantener y documentar las diferentes versiones del objeto conceptual que está siendo desarrollado por el equipo de trabajo de colaboración; dejando constancia de la evolución del acuerdo entre los miembros del grupo de trabajo desde las especificaciones iniciales del objeto conceptual hasta su etapa final de desarrollo.

Mikropoulos y Natsis [2011], tras analizar una década de investigación con trabajo de campo en el área de los EVT de uso educativo, señalan entre otras, la tendencia de sistematizar la integración de nuevos artefactos a los EVT (como los objetos de aprendizaje), así como dotarlos de un marco metodológico para su uso en el ámbito educativo y la evaluación de su usabilidad por la comunidad educativa [Redfern y Naughton, 2002].

Un Objeto de Aprendizaje es "una colección de contenidos, ejercicios, y evaluaciones que son combinados sobre la base de un objetivo de aprendizaje simple" [Wayne Hodgins, 2002]. En el contexto de los EVT orientados a Educación, un Objeto de Aprendizaje es un conjunto de recursos digitales, autocontenible y reutilizable, con un propósito educativo y constituido por al menos tres componentes internos: contenidos, actividades de aprendizaje y elementos de contextualización [Vessey y Conger, 1994]. El Objeto de Aprendizaje debe tener una estructura de información externa (metadatos) que facilite su almacenamiento, identificación y recuperación [Tejada et al., 2001]. Brito [2010] sostiene que a partir de la masiva y vertiginosa irrupción de las Tecnologías de la Información y Comunicación, los recursos diseñados deben ser accesibles, portables y mantenibles, para dar respuestas asequibles a las necesidades actuales. Focalizando sobre estos aspectos, la filosofía de Objetos de Aprendizaje (OA) en conjunto a la de Patrones de Diseño de OA, se erigen como pilares fundamentales en el desarrollo de materiales tecno-pedagógico-comunicacionales libres; proponiendo un enfoque integral que posibilita emplear todo el potencial de las TIC y, por lo tanto, incrementar y extender los beneficios de aplicación en la Educación a Distancia. Un Objeto de Aprendizaje Inteligente es un Objeto de Aprendizaje que utiliza alguna tecnología de Sistemas Inteligentes [García-Martínez et al., 2003]. para enriquecer su comportamiento [Karampiperis y Sampson, 2004; Silveira et al., 2005; Riley, et al., 2009; Desmarais y Baker, 2012].

Las tecnologías digitales han reconfigurado significativamente la Identidad, la Intimidad y la Imaginación durante las últimas décadas" [Gardner y Davies, 2014], por lo que los cambios culturales que se han producido en los estudiantes a partir de la interacción con esta revolución digital plantea nuevos interrogantes. En tal sentido, habrá que repensar cuáles serán los mejores dispositivos y escenarios para el aprendizaje y cómo ampliar el alcance de las aulas para lograr las competencias esperadas en los estudiante ya que, según Gardner y Davies [2014], "los jóvenes de ahora no solo crecen rodeados de aplicaciones sino que además han llegado a entender el mundo como un conjunto de aplicaciones, a ver sus vidas como un conjunto de aplicaciones ordenadas". Del mismo modo se debe encontrar cómo potenciar nuestra enseñanza a través de los recursos tecnológicos disponibles en esta sociedad en red y cómo organizar y mediar pedagógica y eficientemente los contenidos de nuestros cursos. El docente debe incorporar competencias para crear o bien transformar sus cursos en espacios educativos mediados por tecnologías Web.

Las dificultades que se presentan para adaptar los cursos educativos a un escenario con TIC son complejas ya que "tanto las posibilidades que ofrecen las TIC para la enseñanza y el aprendizaje, como las normas, sugerencias y propuestas de uso pedagógico y didáctico de las mismas, son siempre e irremediamente reinterpretadas y reconstruidas por los usuarios, profesores y estudiantes, de acuerdo con los marcos culturales en los que se desenvuelven" [Monereo y Coll, 2008].

Desde el propio campo informático un equipo de la Universidad Jaime I – España propone una metodología para seleccionar tecnologías Web 2.0 para la docencia [Grangel et al., 2012], y en la Argentina la Dirección de Educación a Distancia de la Universidad Nacional de La Plata formula pautas de trabajo para propuestas educativas mediadas por tecnologías digitales [González et al., 2012].

Centrando el tema en los espacios web, Marqués Graells [1999] analiza las principales funciones

que pueden realizar en el ámbito educativo, presenta diversas tipologías y establece un modelo para la identificación y evaluación de estos espacios considerando diversos criterios de calidad.

Ferreira Szpiniak y Sanz presentan diferentes modelos de evaluación de EVEAs comparándolos, y revisan el concepto de usabilidad y de heurísticas de usabilidad [Ferreira Szpiniak y Sanz, 2007]; por otra parte, proponen modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje [Ferreira Szpiniak y Sanz, 2012].

El proyecto que se presenta se funda en los resultados de investigaciones desarrolladas por el Grupo de Ingeniería de Espacios Virtuales, los que se pueden visualizar en la siguiente página web:

<http://sistemas.unla.edu.ar/sistemas/gisi/GIEVT.html>

La propuesta se basa en los siguientes resultados preliminares del grupo de trabajo:

- [a] Se ha estudiado la utilización de espacios virtuales en la formación de investigadores [Rodríguez et al., 2009; 2010a; Rodríguez y García-Martínez, 2013a; 2013c; 2014a].
- [b] Se han sentado las bases para una Ingeniería de Espacios Virtuales de Trabajo [Rodríguez et al., 2009; 2010b; Rodríguez y García-Martínez, 2012a; 2012b; 2014b].
- [c] Se han propuesto una serie de formalismos de modelado de interacción entre los miembros del grupo dentro de un espacio virtual de trabajo colaborativo [Rodríguez y García-Martínez, 2014a] que se puede describir brevemente como: [a] Tablas Concepto-Categoría-Definición: se utiliza para representar los conocimientos fácticos del modelo conceptual de dinámica grupal. Un concepto puede ser de alguna de las siguientes categorías: actor, objeto ó interacción: [b] Procedimientos de Interacción: Describen interacciones compuestas entre los actores vinculadas al desarrollo de un objeto conceptual; [c] Diagramas de Interacción Grupal: Se utilizan para representar de manera integrada las

interacciones de todos los actores considerados en el proceso de modelado; [d] Diagramas de Secuencia de Dinámica Grupal: Se utilizan para expresar la dinámica grupal entre los actores en la línea de tiempo que impone la interacción; y [e] Diagrama de Desarrollo de Objetos Conceptuales: Formaliza las interacciones constructivas de un objeto conceptual desarrollado por los miembros del equipo de trabajo mediado por el espacio virtual.

- [d] Con base en los formalismos propuestos se han desarrollado procesos de formalización de los mecanismos de awareness [Herrera et al., 2013; 2014].
- [e] Se ha explorado el uso de la sociometría en espacios virtuales de trabajo [Charczuk et al., 2013; 2015].

Este proyecto se basa en los siguientes resultados preliminares del grupo de trabajo:

- [a] Se ha estudiado la utilización de espacios virtuales en la formación de investigadores [Rodríguez et al., 2009; 2010a; Rodríguez y García-Martínez, 2013a; 2013c; 2014a].
- [b] Se han propuesto formalismos de modelado de interacciones en grupos de trabajo mediados por espacios virtuales [Rodríguez et al., 2009; 2010b; Rodríguez y García-Martínez, 2012a; 2012b; 2014b].

Objetivos

La evolución de la calidad de las comunicaciones basadas en la tecnología de Internet es la base de la tendencia de desarrollo de los Espacios Virtuales de Trabajo (EVT). Como una familia emergente de aplicaciones a desarrollar, necesita de un abordaje ingenieril que defina los distintos procesos constructivos. En este proyecto se busca definir tres procesos: [a] Proceso de Diseño de Dispositivos Educativos Centrados en EVTs, [b] Proceso de Diseño de Objetos de Aprendizaje inteligentes Integrables a EVTs, y Proceso de Evaluación de EVTs Orientados a Educación.

Objetivo Específico 1: Desarrollar un Proceso de Diseño de Dispositivos Educativos Centrados en Espacios Virtuales de Trabajo

Objetivo Específico 2: Desarrollar un Proceso de Diseño de Objetos de Aprendizaje Inteligentes Integrables a Espacios Virtuales de Trabajo

Objetivo Específico 3: Desarrollar un Proceso de Evaluación de Espacios Virtuales de Trabajo Orientados a Educación

Resultados Esperados

Al concluir el proyecto se espera haber alcanzado los siguientes resultados:

- (i) Un Proceso de Diseño de Dispositivos Educativos integrable a etapas de producción de EVT's Educativos.
- (ii) Un Proceso de Diseño de Objetos de Aprendizaje Inteligentes integrable a etapas de producción de EVT's Educativos.
- (iii) Un Proceso de Evaluación de EVT's integrable a etapas de producción de EVT's Educativos.

Metodología

Para construir el conocimiento asociado al presente proyecto de investigación, se seguirá un enfoque de investigación clásico [Riveros y Rosas, 1985; Creswell, 2002] con énfasis en la producción de tecnologías [Sábato y Mackenzie, 1982]; identificando métodos, materiales y abordaje metodológico necesarios para desarrollar el proyecto:

[a] Métodos:

[a.1] Revisiones Sistemáticas:

Las revisiones sistemáticas [Argimón, 2004] de artículos científicos siguen un método explícito para resumir la información sobre determinado tema o problema. Se diferencia de las revisiones narrativas en que provienen de una pregunta estructurada y de un protocolo previamente realizado.

[a.2] Prototipado Evolutivo Experimental (Método de la Ingeniería):

El prototipado evolutivo experimental [Basili, 1993] consiste en desarrollar una solución inicial para un determina-

do problema, generando su refinamiento de manera evolutiva por prueba de aplicación de dicha solución a casos de estudio (problemáticas) de complejidad creciente. El proceso de refinamiento concluye al estabilizarse el prototipo en evolución.

[b] Materiales:

Para el desarrollo de los formalismos y procesos propuestos se utilizarán:

- Formalismos de modelado conceptual usuales en la Ingeniería de Software [Rumbaugh et al., 1999; Jacobson et al., 2013] y en la Ingeniería del Conocimiento [García-Martínez y Britos, 2004].
- Modelos de Proceso usuales en Ingeniería de Software [IEEE, 1997; ANSI/IEEE, 2007; Oktaba et al., 2007].

Formación de Recursos Humanos

En este proyecto participan 3 investigadores formados y se desarrollan 4 tesis de maestría. Por ser un proyecto que inicia en el año 2017 no hay antecedentes de formación de recursos humanos en el año previo.

Referencias

- ANSI/IEEE, (2007). *Draft IEEE Standard for software and system test documentation*. ANSI/IEEE Std P829-2007.
- Argimón J. (2004). *Métodos de Investigación Clínica y Epidemiológica*. Elsevier
- Basili, V. (1993). *The Experimental Paradigm in Software Engineering*. LNCS, Vol. 706. ISBN 978-3-540-57092-9.
- Charczuk, N., Rodríguez, D., García-Martínez, R. 2013. *Propuesta de Técnicas de Diagnóstico Sociométrico de Dinámicas Grupales para Utilizar en Ambientes de Trabajo Colaborativo*. Proc. VIII TEYET. ISBN 978-987-1676-04-0.
- Charczuk, N., Rodríguez, D., García-Martínez, R. 2015. *Participation Metrics within Virtual Collaborative Workspaces Oriented to Generation of Didactic Interventions*. Computer Science & Technology Series. Pág. 87-97. EDULP. ISBN 978-987-1985-71-5.
- Charum, V. (2007). *Modelo de Evaluación de Plataformas Tecnológicas*. Tesis de Magister en Telecomunicaciones. Instituto Tecnológico de Buenos Aires.
- Creswell, J. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Prentice Hall. ISBN 10: 01-3613-550-1.

- Ferreira Szpiniak, A., Sanz, C. (2007). *Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje*. Proc. XIII CACIC. Pág. 932-947. ISBN 950-656-109-3.
- Ferreira Szpiniak, A., Sanz, C. (2012). *MUa un modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje. Aplicación a un caso de estudio*. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología. TE&ET. ISSN, 9959, 94-103.
- García Martínez, R., Britos, P. (2004). *Ingeniería de Sistemas Expertos*. Nueva Librería. ISBN 987-1104-15-4.
- Gibson, C., Cohen, S. (2003). *Virtual Teams That Work: Creating Conditions for Virtual Team Effectiveness*. J Wiley.
- González, A. H., Esnaola, F., Martín, M. (2012). *Propuestas educativas mediadas por tecnologías digitales*. Dirección de Educación a Distancia, Innovación en el Aula y TIC. UNLP.
- Grangel, R., Campos, C., Rebollo, C., Remolar, I., Palomero, S. (2012). *Metodología para seleccionar tecnologías Web 2.0 para la docencia*. Jornadas de Enseñanza de la Informática (18es: 2012: Ciudad Real).
- Herrera, A., Rodríguez, D., García Martínez, R. (2014). *Topologías de Interacción para Espacios Virtuales de Trabajo Colaborativo*. Rev. Iberoamericana de Educación en Tecnología y Tecnología en Educación, 14: 74-86.
- Herrera, A., Rodríguez, D., García-Martínez, R. (2013). *Taxonomía de Mecanismos de Awareness*. Proc. XIX CACIC. Pág. 651-660. ISBN 978-987-23963-1-2.
- IEEE, (1997). *IEEE Standard for Developing Software Life Cycle Processes. IEEE Std 1074-1997* (Revisión of IEEE Std 1074-1995; Replaces IEEE Std 1074.1-1995)
- Jacobson, I., Ng, P. W., McMaha, P. E., & Jaramillo, C. M. Z. (2013). *La esencia de la ingeniería de software: El núcleo de Semat*. Revi Lat.Am. Ing. Soft., 1(3), 71-78.
- Marqués Graells, P. (1999). *Criterios para la clasificación y evaluación de espacios web de interés educativo*. Educar, (25): 095-111.
- Mikropoulos, T., Natsis, A. (2011). *Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999-2009)*. Computers & Education, 56(3); 769-780.
- Monereo, C., Coll, C., 2008, *Psicología de la educación virtual*, Ediciones Morata, Madrid.
- Oktaba, H., Garcia, F., Piattini, M., Ruiz, F., Pino, F., Alquicira, C. (2007). *Software Process Improvement: The Competisoft Project*. IEEE Computer, 40(10): 21-28.
- Redfern, S., & Naughton, N. (2002). *Collaborative Virtual Environments to Support Communication and Community in Internet-Based Distance Education*. Journal of Information Technology Education, 1(3).
- Riveros, H. y Rosas, L. (1985). *El Método Científico Aplicado a las Ciencias Experimentales*. Trillas. ISBN 96-8243-893-4.
- Rodríguez, D., Bertone, R. García-Martínez, R. (2010a). *Formación de Investigadores Mediada por Espacios Virtuales*. Fundamentación y Prueba de Concepto. Proc. V TEYET. Pág. 512-421. ISBN 978-987-1242-42-9.
- Rodríguez, D., Bertone, R., García-Martínez, R. (2009). *Consideraciones sobre el Uso de Espacios Virtuales en la Formación de Investigadores*. Rev- de Informática Educativa y Med. Audiovisuales, 6: 35-42. ISSN 1667-8338.
- Rodríguez, D., Bertone, R., García-Martínez, R. (2010b). *Collaborative Research Training Based on Virtual Spaces*. En Key Competencies in the Knowledge Society. Springer Verlag 344-353. ISBN 978-3-642-15377-8.
- Rodríguez, D., Bertone, R., Pollo-Cattaneo, F., García-Martínez, R. (2012a). *Modelo Colaborativo de Formación de Investigadores*. Proc. II JEIN. Pp. 183-191. ISSN 2313-9056.
- Rodríguez, D., Charczuk, N., Garbarini, R., García-Martínez, R. (2012b). *Trabajo Colaborativo basado en Espacios Virtuales*. Proc. II JEIN. Pág. 192-199. ISSN 2313-9056.
- Rodríguez, D., García Martínez, R. (2012a). *Modeling the Interactions in Virtual Spaces Oriented to Collaborative Work*. En Software Engineering: Methods, Modeling, and Teaching, Vol 2. Pág. 79-84. ISBN 978-612-4057-84-7.
- Rodríguez, D., García Martínez, R., Merlino, H., Charczuk, N., Lacabanne, M., Caracciolo, B., Iglesias, F. (2013). *Espacios Virtuales para Trabajo Colaborativo*. Proc. XV WICC Pág. 1116-1120. ISBN 978-9-872-81796-1.
- Rodríguez, D., García-Martínez, R. (2012c). *Modelado de Interacciones en Espacios Virtuales Dedicados a Trabajo Colaborativo*. XVIII CACIC. Pág. 589-598. ISBN 978-987-1648-34-4.
- Rodríguez, D., García-Martínez, R. (2013a). *Elementos de Análisis y Diseño para Espacios Virtuales para la Formación de Investigadores*. Revista Latinoamericana de Ingeniería de Software, 1(2): 45-56, ISSN 2314-2642.
- Rodríguez, D., García-Martínez, R. (2013b). *Propuesta de Proceso de Diseño de Espacios Virtuales de Trabajo Educativo Personalizables*. VIII TEYET. ISBN 978-987-1676-04-0.
- Rodríguez, D., García-Martínez, R. (2013c). *Viabilidad Tecnológica de Formación de Investigadores Mediante Espacios Virtuales*. 6to Seminario Internacional de Educación a Distancia. Trabajo No 237. Red de Universitaria de Educación a Distancia de Argentina (RUEDA). UN Cuyo. 10 al 12 de Octubre. Mendoza. Argentina.
- Rodríguez, D., García-Martínez, R. (2014b). *Modelado de Interacciones Aplicado a Diseño de Espacios Virtuales de Trabajo*. En "Ingeniería del Software e Ingeniería del Conocimiento: Dos Disciplinas Interrelacionadas". Pág. 377-393. ISBN 978-958-8815-31-2.
- Rodríguez, D., García-Martínez, R. (2014c). *Proposal of Design Process of Customizable Virtual Working Spaces*. En Modern Advances in Applied Intelligence. Springer Verlag 450-459. ISBN 978-3-319-07454-2.
- Rodríguez, D., Pollo-Cattaneo, F., Bertone, R., García-Martínez, R. 2010c. *Elementos para el Análisis y Diseño Conceptual de Espacios Virtuales de Trabajo Colaborativo Orientados a la Formación de Investigadores*. Proc. XVI CACIC. Pág. 364-373. ISBN 978-950-9474-49-9.
- Rodríguez, D., Ramón García-Martínez, R. (2014a). *A Proposal of Interaction Modelling Formalisms in Virtual Collaborative Work Spaces*. Lecture Notes on Software Engineering, 2(1): 76-80. ISSN 2301 3559.
- Rumbaugh, J., Jacobson, I., Booch, G. (1999). *The Unified Modeling Language, Reference Manual*. Addison Wesley, ISBN-10: 02-0130-998-X.
- Sabato J, Mackenzie M. (1982). *La Producción de Tecnología: Autónoma o Transnacional*. Instituto Latinoamericano de Estudios Transnacionales - Technology & Engineering. ISBN 9789684293489.

Producción Creativa de Material Educativo Digital para un Espacio Innovador de Aula Invertida

Leda Digi3n¹ y Marisa Digi3n²

¹Dpto. de Inform1tica. Facultad de Ciencias Exactas y Tecnolog1as.
Universidad Nacional de Santiago del Estero.
ldigion@unse.edu.ar

²1rea Acad1mica Matem1tica. Facultad de Ciencias Econ3micas.
Universidad Nacional. de Jujuy.
marisadigion@gmail.com

Resumen

Este trabajo presenta los t3picos generales de un proyecto de investigaci3n aplicada, para la producci3n de material educativo digital que se desea incorporar, como material curricular de ense1anza en el aula virtual de la asignatura “An1lisis Matem1tico”; incluida en los planes de estudio de las carreras de Contador P1blico, Administrador de Empresas y Licenciado en Econom1a Pol1tica de Facultad de Ciencias Econ3micas, Universidad Nacional de Jujuy.

La incorporaci3n de dicho material digital no pretende reemplazar a los recursos convencionales que se utilizan actualmente en la citada materia. La intenci3n actual es complementarlos, posibilitando la realizaci3n de cierto tipo de actividades que de manera tradicional, no son posibles visualizarlas. Se espera que la incorporaci3n de recursos tecnol3gicos otorgue a los aprendices (estudiantes en aula virtual), una idea real de los procesos de aproximaci3n que se requieren para la compresi3n de dicho tema. Con la complementaci3n descripta, y el uso de otros soportes educativos basados en herramientas de aprendizaje visual y recursos tecnol3gicos de interacci3n, insertos en un modelo actual

como el de aprendizaje inverso o aula invertida, se quiere contribuir a generar mayor aprendizaje significativo basado en pensamiento creativo.

Palabras clave: aula invertida, material educativo digital, creatividad, aprendizaje significativo.

Contexto

Este proyecto se sit1a en un espacio de colaboraci3n generado por la c1tedra Tecnolog1as Inform1ticas Avanzadas, del Dpto. de Inform1tica (FCEyT – UNSE) en el marco del proyecto “Implicancias conceptuales interdisciplinarias en la b1squeda de creatividad y criticidad para ambientes colaborativos y distribu3dos, soportados por computador y mediados por la gesti3n del conocimiento” (C3d. N3 23/C134 CICYT-FCEyT, UNSE).

Se busca aplicar los conocimientos adquiridos en ambientes de trabajo colaborativo y distribu3dos, en el curso de una asignatura de grado como la nombrada. Y como inter3s acad1mico, se quiere incorporar la innovaci3n tecnol3gica y pedag3gica para lograr aprendizaje significativo basado en pensamiento creativo y mediado por producci3n digital.

Introducción

Se publicó un informe en *KnowledgeWorks* [1] titulado "*Exploring the Education Workforce: New Roles for an Expanding Learning Ecosystem*", el cual tiene por objetivo explorar cómo los roles de los educadores pueden cambiar en la próxima década. El informe ofrece siete ejemplos de las futuras funciones del educador para "ayudar a los agentes educativos a imaginar qué tipo de nuevos roles educativos podrían contribuir a la creación de ecosistemas de aprendizaje flexibles y rigurosos, que permitan a los alumnos y los adultos que los apoyan a prosperar". Estos siete nuevos roles son: *The Learning Pathway Designer* (el Diseñador del Camino de Aprendizaje), *Competency Trackers* (Rastreadores de Competencias), *Pop-up Reality Producers* (Productores de Realidades Emergentes), *Social Innovation Portfolio Directors* (Directores de Portafolios Sociales de Innovación), *Learning Naturalist* (Naturalistas del Aprendizaje), *Micro-Credentialing Analysts* (Analistas de Micro-acreditaciones) y el *Data Steward* (Mayordomo de Datos). También este reporte señala que el cambio tecnológico acelerado, el incremento de personalización, la gran diversidad de modelos educativos y la proliferación de sistemas de datos, son ejemplos de tendencias que llevarán a más roles diversos que tendrá el educador.

En esta propuesta de investigación que se presenta, se considera que el soporte digital que debe disponer y aplicar el educador en el proceso de aprendizaje es vital, en todos los roles mencionados. Especialmente en el rol de "Diseñador de Camino de Aprendizaje", pues trabaja con los estudiantes para fijar objetivos, observar el progreso de estudio, y proporcionar modelos con una

secuencia de actividades alineadas con sus competencias.

Entonces, los materiales educativos digitales constituyen un producto que posee características como las técnicas, que pueden ser evaluadas mediante el uso de métricas adecuadas; y otras, que están relacionadas con la significación del aprendizaje, en un ámbito donde la cantidad de variables a tener en cuenta son muchas.

El uso de metodologías adecuadas de desarrollo en principio, puede solucionar el problema de la calidad, para obtener un material educativo digital libre de errores, o con una tasa tolerable de errores no encontrados ni detectados. Pero como cada producto digital desarrollado debe producir un aprendizaje de cierto tipo, o desarrollar una habilidad o capacidad en particular sumada a la calidad técnica, se necesita una evaluación desde el punto de vista de los aprendizajes que se pretende lograr; y por ende, requiere un modelo instructivo adecuado con soporte tecnológico válido.

Con la producción adecuada de material educativo digital específico, se quiere realizar una contribución a la enseñanza de la asignatura Análisis Matemático (Universidad Nacional de Jujuy). La incorporación de dicho material digital no pretende reemplazar a los recursos convencionales que se utilizan actualmente en la citada materia. La intención que se persigue es complementarlos, posibilitando la realización de cierto tipo de actividades que, de manera tradicional, no son posibles visualizarlas; y con un usuario que opere interactivamente con las herramientas TICs.

Al ser una asignatura que aborda contenidos de la matemática dinámica (Cálculo Diferencial e Integral), requiere de procesos que permitan observar las transformaciones que surgen de esa

dinámica. Ejemplos concretos son los que, precisamente, dieron origen a esta rama de la Matemática: la determinación de la pendiente de la recta tangente a la gráfica de una función, el cálculo del área de una región plana de lados no rectos y la determinación de la velocidad instantánea de un móvil. En este sentido, la incorporación de recursos tecnológicos otorgaría a los estudiantes, una idea real de los procesos de aproximación que se requieren para la comprensión de estos temas, lo cual redundará en el logro de aprendizajes más significativos. Con la complementación descripta, y el uso de otros soportes educativos basados en herramientas de aprendizaje visual y recursos tecnológicos de interacción, con modelo de aprendizaje invertido en aula virtual, se espera generar mayor aprendizaje significativo con pensamiento creativo y crítico.

Líneas de Investigación, Desarrollo e Innovación

Uno de los problemas que actualmente existen en la educación es no tener claro para qué, cómo y cuándo nos sirve el medio digital como un apoyo en la educación de grado. Existe una falta de modelos dirigidos a los profesores como creadores/usuarios del material digital, que ayuden al diseño y desarrollo de estos objetos educativos. De aquí se deriva que en la práctica, el material digital es desarrollado y utilizado sin aprovechar el potencial de sus capacidades de representación multimedial, simulación e interacción. De ahí la necesidad de contar con pautas que orienten al profesional docente en el desarrollo y uso de este tipo de recursos para el aprendizaje.

Por otro lado, se observa que los avances teóricos para el diseño de materiales educativos digitales se han

desarrollado por líneas separadas: el diseño instructivo por un lado, los diferentes tipos de diseño (diseño de información, diseño de experiencias, diseño de interfaz, diseño gráfico) por otro, y finalmente, la práctica por un tercer camino.

Con la intención de romper esta disociación y apoyar al docente de grado, se quiere generar un modelo para diseñar Materiales Educativos Digitales que integre los hallazgos del diseño instructivo, y de áreas del diseño multimedial; y se encuadre en la teoría constructivista, bajo las cuales las personas construyen el conocimiento basado en lo que saben, y el aprendizaje es así un proceso activo [2], y creativo en un espacio innovador de aprendizaje en aula invertida.

Se busca que el modelo pueda promover la maduración del razonamiento causal, el cual representa una de las actividades más básicas e importantes de los procesos cognitivos, ya que permite realizar otras funciones de más alto nivel: como el entender conceptos, hacer predicciones, entender implicaciones, realizar inferencias, explicar fenómenos y la resolución de problemas [3].

Una ventaja que señala la literatura [2], es que las aplicaciones educativas multimedia e hipermedia, presentan un considerable potencial como herramientas cognitivas, al mostrar, proponer y dar evidencia directa al aprendiz (estudiante en modalidad virtual), permitiéndole ver, observar, interpretar, reflexionar, buscar evidencias y ligar la experiencia adquirida con el conocimiento previo. Para ésto, en primer lugar se debe adecuar el modelo a las necesidades del docente que lo implementará, ya que si a éste le resulta difícil o poco práctico de usar, entonces simplemente no será una herramienta que

contribuya a la implementación de su práctica.

Por lo nombrado, se propone investigar en las siguientes áreas de estudio que se listan.

- Sobre diseño digital y recursos mediales

La selección de los contenidos, es uno de los problemas recurrentes en educación que comienzan con el planteo del docente de qué enseñar, para qué enseñar y cómo enseñar. En el análisis del “qué enseñar”, de acuerdo a los "principios básicos", ejes de todo el desarrollo, el docente debe seleccionar la información a presentar y transmitir, determinando los contenidos y también su organización que dependerá de las estrategias elegidas por el docente o su contexto [4].

- La interacción con la interfaz colaborativa

Según Jacob Nielsen (2000) [5] en el diseño de una interfaz es fundamental la gramática de la interacción, que contribuye no sólo a imponer una manera de leer sino, sobre todo, un modo de hacer. Dicha gramática determina y condiciona la utilidad, ya que se debe adaptar el diseño de la interfaz al tipo de interacción que se quiere proponer al usuario, sin un criterio único y lineal predeterminado.

- Sobre la planificación didáctica

En el contexto universitario se observan aún modelos educativos tradicionales que, en gran parte, hacen difícil la implementación de ambientes donde el uso de las TICs sea cada vez más ubicuo y diverso [6]. En este sentido resulta importante, como etapa inicial, fortalecer la capacitación docente, diagnosticar la realidad en la que estudiantes y docentes están inmersos y analizar sus potencialidades, para luego proyectar el diseño de nuevos ambientes de aprendizaje soportados por las TICs.

- Sobre el pensamiento creativo

La creatividad se define como la creación, identificación, planteamiento y solución divergente de un problema [7]. Creatividad es la capacidad de ver nuevas posibilidades y hacer algo al respecto. Cuando una persona va más allá del análisis de un problema e intenta poner en práctica una solución, se produce un cambio. La creatividad implica ver un problema, tener una idea, hacer algo sobre ella y tener resultados positivos. Habría que identificar si existen algunos otros rasgos de la personalidad creativa y tratar de ir creando un clima de relaciones y una estrategia de trabajo adecuada en el aula, para favorecer el desarrollo de estas cualidades en los alumnos, ya que ellas contribuirán a fomentar y a acrecentar su capacidad creativa.

- Sobre aprendizaje invertido

La Educación Basada en Competencias es la tendencia pedagógica más relevante en la actualidad. Le sigue muy de cerca el Aprendizaje Invertido y Gamificación [7].

El aprendizaje invertido es una técnica didáctica en el que la exposición de contenido se hace por medio de videos que pueden ser consultados en línea de manera libre, mientras el tiempo de aula se dedica a la discusión, resolución de problemas y actividades prácticas bajo la supervisión y asesoría del profesor.

En agosto de 2014, al menos setenta y dos profesores del Tecnológico de Monterrey, implementaron el modelo de Aprendizaje invertido, impactando alrededor de seis mil estudiantes de diferentes cursos y grados impartidos entre agosto 2013 y agosto 2014. En el Reporte *Edu Trends* edición “Aprendizaje Invertido”, realizado por el Observatorio de Innovación Educativa, se listan los resultados y experiencias de algunos de los profesores que han implementado este modelo en sus cursos [8].

Resultados y Objetivos

La finalidad de este proyecto, es desarrollar material educativo digital complementario al dictado tradicional de la asignatura Análisis Matemático (área específica); para generar aprendizaje significativo basado en pensamiento creativo, y aplicable en el modelo de clase inversa (aprendizaje invertido).

Como objetivos específicos, se espera:

- Desarrollar material educativo que permita facilitar a los alumnos, la recuperación y uso de información sobre la temática del Cálculo Diferencial.
- Valorar específicamente la introducción de recursos mediales precisos para una audiencia determinada.
- Estudiar el modelo de aprendizaje invertido, y los recursos pedagógicos y tecnológicos necesarios para su implementación.
- Aplicar el diseño instructivo para desarrollar un modelo de enseñanza invertido, en un entorno personalizado y virtual de aprendizaje (EVA).
- Observar el impacto que tiene el uso de la tecnología de imágenes e interacción de la aplicación, incorporadas a los procesos de aprendizaje de los estudiantes del grupo seleccionado.
- Validar por medio de un análisis estadístico, el uso y aplicación del material educativo digital en la enseñanza aprendizaje del grupo.
- Comparar formalmente la generación de aprendizaje en un grupo experimental y en un grupo control, utilizando como variable de análisis el material educativo digital, incorporado en el modelo instructivo válido, como el aprendizaje invertido.

Se plantea la siguiente hipótesis de trabajo. Para entender los procesos de cambio y sus efectos, así como las posibilidades que los sistemas de enseñanza aprendizaje conllevan los avances tecnológicos, hay que situarse en el marco de un espacio de innovación. Entonces:

“El uso de nuevas tecnologías de información y comunicación (TICs) en el ámbito de áreas específicas de la enseñanza de la Matemática, como el Cálculo Diferencial, provoca transformaciones positivas en la configuración del proceso pedagógico de este tipo de ciencia, a partir de la aplicación de un nuevo diseño instructivo del material educativo digital; y consecuentemente se producen cambios en los roles que han desempeñado los estudiantes y docentes, como nuevos tutores y aprendices interactuando en un entorno virtual de enseñanza aprendizaje”.

Sobre la metodología y plan de actividades, para el desarrollo de esta propuesta, se aplicarán los dos enfoques de investigación.

El enfoque cuantitativo se aplicará a partir de la recolección y el análisis de datos, contestando preguntas de investigación, y para probar la hipótesis establecida previamente. Con el enfoque cualitativo, se descubrirán y refinarán las preguntas, o se producirán nuevas preguntas, para continuar con la investigación.

Formación de Recursos Humanos

A partir de las líneas de investigación planteadas, se propone la participación de los siguientes recursos humanos:

- Docente Responsable del dictado de la asignatura de grado Tecnologías Informáticas Avanzadas (carreras de LSI- PUI -PI) y Tecnologías de la

Información y Comunicación en Educación II (carrera de PI- P.2014), Ing. Leda Digión.

- Alumnos de las asignaturas mencionadas, y que participan del Taller integrador “Aplicación de TICs en espacio innovador educativo virtual”.

- Alumnos avanzados de las carreras de grado LSI-PUI-PI, en preparación de anteproyecto de Trabajo Final.

- Colaboración del Docente Externo: Lic. Marisa Digión. Responsable del Equipo cátedra de la asignatura “Análisis Matemático” (UNJu).

Referencias

[1] Bolkan, J. Los 7 futuros roles del docente (2015). *Reporte Semanal de Innovación Educativa para Profesores*. Elaborado por el Observatorio de Innovación Educativa del Tecnológico de Monterrey.

Disponible:

<http://www.observatorioedu.com/blog/2015/9/1/los-7-futuros-roles-del-docente>

[2] Solomonidou, C. (2009). Constructivist design and evaluation of interactive educational software: a research-based approach and examples. *Open Education: The Journal for Open and Distance Education and Educational Technology*. N° 5 (1), pp. 6-24.

[3] Jonassen, DH. and Ionas IG. (2008). Designing effective supports for causal reasoning. *Education Technology Research Developers*. N° 56, pp. 2876-308.

[4] Cabero, J. y López, E. (2009). Descripción de un instrumento didáctico para el análisis de modelos y estrategias de enseñanza de cursos universitarios en red (A.D.E.C.U.R). *Pixel-Bit. Revista de Medios y Educación*. N° 34, pp. 13-30.

[5] Nielsen, J. (2000). Usabilidad. Diseño de sitios Web, Prentice Hall, Madrid.

[6] Digión, L. y Figueroa, S. (2015). A quality indices-based pedagogic proposal for training virtual tutors. 9th International Technology, Education and Development Conference. Madrid, España. *International Academy of Technology, Education and Development (IATED)*. Pp. 5723-5731. ISBN: 978-84-606-5763-7.

Disponible:

<http://library.iated.org/view/DIGION2015AQU>

[7] Penagos, J.C. and Aluni, R. (2000). Preguntas más frecuentes sobre creatividad. Universidad de las Américas-Pue.

Disponible:

http://homepage.mac.com/penagoscorzo/creatividad_2000/creatividad8.html.

[8] Radar de Innovación Educativa 2015 (2015). *Reporte Edu Trends*. Observatorio de Innovación Educativa. Tecnológico de Monterrey.

Disponible:

<http://bit.ly/ObservatorioGPlus>.

Realidad Aumentada, Realidad Virtual e Interacción Tangible para la Educación

María José Abásolo^{1,2}, Cecilia Sanz¹, Marcelo Naiouf¹, Armando De Giusti¹,
Graciela Santos³, María L. Castro³, María José Bouciguez³

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, csanz, mnaiouf, degiusti}@lidi.info.unlp.edu.ar

²Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

³Facultad de Ciencias Exactas - Universidad Nacional del Centro Pcia. Bs. As.
(UNICEN)

Resumen

La línea de investigación y desarrollo presentada consiste en estudiar, desarrollar y evaluar aplicaciones de realidad virtual, realidad aumentada y juegos, en particular para apoyar los procesos de enseñanza-aprendizaje. Uno de los principales objetivos es la formación de recursos humanos y fortalecimiento de la investigación mediante el trabajo intergrupar entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Realidad Virtual, Realidad Aumentada, Interacción Tangible, Enseñanza, Aprendizaje, TIC

Contexto

La línea de investigación Realidad Aumentada y Realidad Virtual forma parte del proyecto 11/F017 "Cómputo Paralelo de Altas Prestaciones. Fundamentos y Evaluación de rendimiento en HPC. Aplicaciones a Sistemas Inteligentes, Simulación y Tratamiento de Imágenes", acreditado por la Universidad Nacional de La Plata (UNLP) en el marco del Programa de Incentivos. Además en esta línea se lleva a cabo el proyecto "REFORTICCA Recursos para el Empoderamiento de FORMadores en TIC, Ciencias y Ambiente" aprobado como Proyecto de Innovación y Transferencia en Áreas

Prioritarias de la Provincia de Buenos Aires de la CICPBA, el cual se lleva a cabo entre los grupos de investigación III- LIDI (UNLP), EcienciaTEC y CIFICEN (UNICEN) de diferentes áreas involucradas: TIC, enseñanza, ciencias físicas, matemática y medio ambiente.

Introducción

En los últimos años han surgido nuevas tecnologías y paradigmas como la Realidad Aumentada (RA), la Realidad Virtual (RV), la Interacción Tangible (IT), que se presentan como herramientas que pueden apoyar los procesos de enseñanza-aprendizaje tanto en los diferentes niveles educativos de la educación formal como en educación no formal, y educación especial (Hernández Ortega et al., 2012).

Realidad Aumentada

La RA propone aumentar las capacidades de percepción humana, para visualizar la información digital - como texto, imágenes, vídeos o animaciones en 3D- directamente embebida en el mundo físico real (Manresa-Yee et al, 2011). Existen variedad de aplicaciones de RA que pueden encontrarse en la web para apoyar el proceso de enseñanza-aprendizaje en

áreas como biología, matemática, astronomía, anatomía, idiomas, ecología, etc.

El uso de RA puede combinarse con la geolocalización del participante, es decir, la localización de su ubicación en el globo terráqueo, en base a dispositivos físicos como GPS, brújulas, acelerómetros y giroscopios incorporados a celulares de última generación. Se provee contenidos relevantes del entorno -en forma de objetos 3D, imágenes, indicadores de puntos de interés (POIs) entre otros- dependiendo de la ubicación del usuario en un cierto lugar en el mapa terrestre. El desafío de aprender a incorporar la RA y la geolocalización en los contextos educativos, implica la voluntad de centrar el aprendizaje en la participación activa del estudiante, en sus intereses, en situaciones relevantes y directamente relacionadas con su vida real, lo cual supone un cambio en los planteamientos pedagógicos que exigen el diseño de nuevas propuestas metodológicas y el uso de recursos didácticos capaces de facilitar los nuevos procesos. Pero para ello, necesitamos docentes que conozcan y aprendan a utilizar estas herramientas, para luego mediar el contenido de su disciplina desde una mirada pedagógica – didáctica, pudiendo vislumbrar las posibilidades educativas de la RA y la geolocalización para enriquecer los procesos de enseñanza – aprendizaje; brindando escenarios para explorar, interactuar y relacionarse con su entorno, generando recursos y conocimientos de manera creativa y lúdica.

Realidad Virtual

Por su parte, la RV es un término que se aplica a experiencias visuales donde el participante se ve inmerso e interactúa en un ambiente o escena virtual 3D con diferentes grados de inmersión (Manresa- Yee et al, 2011). Los videojuegos y simulaciones de procesos y fenómenos se presentan como entornos de aprendizaje interactivo participativo que cautiva a un

jugador ofreciendo desafíos que requieren mayores niveles de dominio, que pueden conjugar la participación y diversión con el rigor y la resolución de nuevas situaciones (Aldrich, 2009) (Squire, 2008). Los videojuegos y simulaciones se emplean en la enseñanza de las ciencias. (Bouciguez, 2013) (Santos, 2016) describen algunos videojuegos educativos para enseñar Física y se caracterizan en función de la tecnología que involucra, los aportes y limitaciones al aprendizaje desde una perspectiva constructivista social. Esta caracterización puede brindar información útil sobre cómo utilizar un videojuego, qué requerimientos tecnológicos necesita y qué potencialidades puede aportar a una situación de enseñanza y aprendizaje.

Interacción Tangible

La IT permite, a través de la manipulación física de objetos superficies y espacios del entorno real, denominados representaciones tangibles, controlar sus contrapartes digitales. La IT ofrece oportunidades para el desarrollo de actividades educativas colaborativas, dado que permite que el grupo se concentre en el objetivo a cumplir y la tecnología no resulte un distractor (Artola, 2013).

Los jóvenes y niños, nativos en la era digital, tienen un acercamiento natural a las Tecnologías de Información y Comunicación (TIC) y apropian naturalmente tecnologías como la RA y RV, popularizadas recientemente por su uso en publicidad, revistas educativas y juegos. Sin embargo, los docentes quedan un paso atrás de sus alumnos en lo que respecta a apropiación de las nuevas tecnologías. Se detecta la necesidad de empoderar a los docentes de escuelas primarias y secundarias en la apropiación de TIC acercándoles aplicaciones y herramientas existentes, algunas de las cuales desarrolladas por nuestro grupo de investigación y otras disponibles como software abierto.

El actual contexto social plantea nuevos desafíos a los fines de educar en ciencia y tecnología a una población estudiantil diversa, para colaborar y desenvolverse como ciudadanos globales. Se apunta a una educación adecuada que privilegie la comprensión, la comunicación, tanto oral como escrita, la autonomía en el aprendizaje, la obtención, selección y análisis crítico de la información y la resolución de problemas en contexto. Esto supone modificar las estrategias de enseñanza para el logro de aprendizajes significativos y orientados a la acción, a partir de la participación activa del alumno en la resolución de actividades el desarrollo de estrategias de pensamiento.

Por otra parte es conocido que los docentes aún hacen uso escaso o insuficiente de la tecnología en la enseñanza, y en particular para la enseñanza de Matemática, Física o Ambiente. Se asume que esto se debe a dos razones de peso: por un lado, desconocimiento de las nuevas tecnologías y, por otro, la falta de formación para el diseño de propuestas de enseñanza que integren las TIC. A esto debe sumarse la existencia de pocos recursos tecnológicos educativos que se adecuen a las nuevas corrientes pedagógicas basadas en nociones como el “aprender haciendo”, “aprendizaje activo”, el aprendizaje basado en problemas, ambientes personales de aprendizaje (PLEs) etc. orientadas a promover el pensamiento crítico, la curiosidad, creatividad y co-construcción de conocimiento (Castañeda y Adell, 2013) (Adams Becker et al., 2016)

Se parte del supuesto que son los docentes quienes deben apropiarse de la tecnología y adquirir las competencias necesarias para enriquecer sus prácticas docentes con la tecnología para incidir en una mejora de la educación en Ciencias en el ámbito de la Provincia de Buenos Aires. Es de central importancia el papel del docente responsable del diseño de las propuestas pedagógicas, para guiar el proceso de enseñanza y aprendizaje, decidir qué

herramienta será más apropiada, así como anticipar dificultades y obstáculos de aprendizaje.

El objetivo general de la temática de investigación es brindar a docentes de educación primaria y secundaria recursos educativos y herramientas TIC, basadas en RA, RV e IT, para ser usadas como soporte del proceso de enseñanza-aprendizaje de ciencias físicas y matemáticas y medio ambiente. En particular se orienta a colaborar con el empoderamiento del docente ofreciendo cursos de capacitación sobre el uso de las TIC en la enseñanza, poniendo a disposición recursos que integren los conocimientos disciplinares, didácticos y tecnológicos, y posibilitando el intercambio de experiencias en el aula que incorporen las TIC realizadas en colaboración con los docentes.

Líneas de investigación y desarrollo

- RA aplicada a la educación
- Simulaciones y Videojuegos educativos para la enseñanza de Matemática y Física
- IT aplicada a la educación

Resultados y Objetivos

- Realización de una encuesta y cuestionarios online a una muestra representativa de docentes de nivel primario y secundario sobre sus conocimientos y usos de tecnologías, y el uso que hacen de la tecnología en el aula (Coll, 2008).
- Ofertar cursos semipresenciales en los que se les ofrecerán conocimientos sobre las tecnologías y los posibles usos educativos mediante una plataforma EVEA de uso libre complementada con talleres presenciales.

- Realización de prácticas de RA utilizando una herramienta de autor de desarrollo propio denominada AuthorAR (Moralejo et al., 2013) que permite la creación de escenas aumentadas basadas en el reconocimiento de marcadores. Dado que la herramienta fue diseñada como una herramienta para el maestro, la misma incorpora plantillas específicas para crear actividades educativas.
- Realización de prácticas de IT mediante una mesa interactiva portable, para lo cual se prevé la utilización de una herramienta de autor de desarrollo propio denominada TIES (Artola et al., 2015) la cual incorpora plantillas específicas para crear actividades educativas.
- Utilización de modelado y simulación en Matemática y en Física mediante GeoGebra (Martinovic, 2014) (Sezen Yüksel y Çıldır, 2015). Estudiar las posibilidades de aplicación de las TIC dadas a conocer a los docentes analizando propuestas educativas con TIC implementadas en aula por los docentes.
- Creación de un repositorio de recursos educativos para la enseñanza de ciencias físicas, matemáticas y medio ambiente
- Crear una comunidad virtual de aprendizaje colaborativo que permita a los docentes que emprenden el aprendizaje de las TIC mediante los materiales didácticos generados puedan compartir su experiencia y evacuar dudas. Esto incluye la elaboración de un repositorio de experiencias áulicas de propuestas educativas con TIC.

Formación de recursos humanos

La formación de recursos humanos es prioritaria en esta línea, y por esto se implementó la carrera de postgrado mencionada en la sección anterior. En el marco de los proyectos conjuntos mencionados se ha podido contar con la visita de profesores de otras universidades nacionales y extranjeras (UNLP, UNS, UNICEN y UIB), con experiencia y formación en los temas propuestos, los cuales impartieron cursos de posgrado.

En la actualidad hay en curso diferentes tesis de postgrado en el marco de esta línea de investigación:

- Mario Vincenzi. “La Realidad Aumentada en la educación. Vigencia, proyecciones y límites” (tesis de especialización en curso).
- Natalia Encina. “Evaluación de browsers de realidad aumentada para apoyar procesos de enseñanza -aprendizaje” (tesis de especialización en curso).
- Wilma Gavilanes “Metodología para la evaluación del impacto de experiencias con Realidad aumentada en educación superior” (tesis de doctorado en curso).
- María José Bouciguez “Ambientes virtuales altamente interactivos basados en videojuegos y simulaciones para la educación en ciencias” (tesis de doctorado en curso)

Referencias

Aldrich, C. (2009). Virtual worlds, simulations, and games for education: A unifying view. *Innovate* 5 (5), http://www.innovateonline.info/pdf/vol5_issue_5/Virtual_Worlds,_Simulations,_and_Games_for_Education- A_Unifying_View.pdf

Artola, Verónica; Sanz, Cecilia; Moralejo, Lucrecia; Pesado, Patricia; Baldasarri, Sandra (2015) Herramienta de autor para la creación de actividades educativas basadas en Interacción Tangible. XIII Workshop de Tecnología Informática Aplicada en Educación
 - CACIC 2015. Junín, Proceeding del

Congreso. ISBN: 978-987-3724-37-4 Buenos Aires, Argentina. Octubre 2015.

Artola, Verónica (2013) Interacción tangible en aplicaciones educativas. *Diseño e implementación de un prototipo basado en este paradigma de interacción orientado al aprendizaje colaborativo*

Tesis de grado Facultad de Informática UNLP
<http://sedici.unlp.edu.ar/handle/10915/46826>

Adams Becker, S., Cummins, M., Davis, A., and Yuhnke, B. (2016). 2016 NMC Technology Outlook for Australian Tertiary Education: A Horizon Project Regional Report. Austin, Texas: The New Media Consortium

[Bou13] María José Bouciguez, Graciela Santos, María José Abásolo (2013) *Potencialidad de los videojuegos en el aprendizaje de Física*. Actas de WEFA 2013 I Workshop de Enseñanza de Física en Argentina, Universidad Nacional del Centro de la Pcia. de Bs. As., ISBN 978-950-658-342-2.

Castañeda, L., Adell, J. (eds.) (2013) Entornos personales de aprendizaje: Claves para el ecosistema educativo en la red. Alcoy, Marfil
<http://www.um.es/ple/libro/>

Coll, C. Muri, M. T. y Onrubia, J. (2008). Análisis de los usos reales de las TIC en contextos educativos formales: una aproximación socio-cultural. En Revista Electrónica de Investigación Educativa, 10(1). Consultado el 28 de agosto de 2012, en:
<http://redie.uabc.mx/vol10no1/contenido-coll2.html>

Hernández Ortega, José; Pennesi Fruscio, Massimo; Sobrino López, Diego y Vázquez Gutiérrez, Azucena (2012) *Tendencias Emergentes en Educación con TIC*. Asociación Espiral, Educación y Tecnología, ISBN: 978-84-616-0448-7

Manresa-Yee, C.; Abásolo, M.J.; Mas Sansó, R.; Vénere, M. (2011) *Realidad Virtual, Realidad Aumentada e Interfaces Basadas en Vision*. XV Escuela Internacional de Informática, XVII Congreso Argentino de Ciencia de la Computación CACIC 2011. Editorial EDULP, ISBN 978-950-34-0765-3

Martinovic, D., Karadag, Z., & McDougall, D. (Eds.) (2014). Proceedings of the Fifth North American GeoGebra Conference, GeoGebra-NA 2014, November 21-22, 2014, Toronto, ON: University of Toronto.

Moralejo L., Sanz C., Pesado P., Baldassarri S. (2013) AuthorAR: Authoring Tool For Building Educational Activities Based On Augmented Reality 2013 International Conference on Collaboration Technologies and Systems (CTS 2013) San Diego, Estados Unidos. Mayo de 2013 Proceeding del Congreso. Páginas 377-381. ISBN: 978-1-4673-6404-1

Santos G. (2016) Videojuegos y estrategias para enseñar física. Cap. 3 del libro: "Pasaporte a la enseñanza de las ciencias. La modelización como eje organizador para la construcción de significados", Consuelo Escudero y Silvia Stipich (comp.). Noveduc libros del Centro de Publicaciones Educativas y Material Didáctico S.R.L. En proceso de edición. 2016. ISBN 978-987-538-482-8

Sezen Yüksel & Çıldır. (2015). The Impacts of Dynamic Geometry Software on Graphing Abilities of Prospective Physics Teachers: GeoGebra Sample. Eurasian J. Phys. & Chem. Educ. 7(1): 46-61.

[Squ08] Squire, K.D. (2008) *Game-based learning: An emerging paradigm for learning*. Performance Improvement Quarterly, 21 (2), 7-36.
<http://www3.interscience.wiley.com/journal/120835177/issue>

Repositorios Digitales con Contenidos Orientados a las Necesidades de Instituciones Educativas de la Región

Antonio Castro Lechtaler^{1,2}, Fernanda Carmona¹, José Texier¹, Fernando Emmanuel Frati¹, Alberto Riba¹, Marisa Gagliardi¹, Emiliano Peressini¹, Jusmeidy Zambrano¹, Germán Kurt Grin²

¹Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina; ²CISTIC/FCE - Universidad de Buenos Aires, Córdoba 2122, CABA

antonio.castro@fce.uba.ar, {fbcarmona, jtexier, fefrati, ariba}@undec.edu.ar,
marisagagliardi76@gmail.com, emiliano.peressini@gmail.com, jzambrano@undec.edu.ar,
german.grin@gmail.com

RESUMEN

Las políticas públicas como los programas “Conectar Igualdad” y “Joaquín V. González”, han alcanzado una amplia cobertura en la entrega de computadoras portátiles y han contribuido en cambios educativos y sociales en la Argentina, pero su impacto no ha sido el esperado por diversas causas en las Instituciones Educativas (IE). Adicionalmente, se aprobó la Ley 26.899 de Repositorios Institucionales (RI) y la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares. Esta línea de investigación propone una solución a las diferentes necesidades escolares presentes en las IE de la región, relacionadas con problemas en el acceso a bibliografía básica y técnica, para contribuir, sin costo alguno, en la formación del alumno y en la capacitación del RRHH de las mismas. Los Repositorios Institucionales (RI) pueden brindar un espacio de contención a obras de referencia, contenidos multimediales, televisivos, libros, imágenes, videos, manuales áulicos y otros recursos educativos adecuados a estos niveles iniciales. La implementación de un RI vinculado al requerimiento de otros espacios académicos constituye una oportunidad de articulación entre todos los niveles educativos, mediante el aporte tecnológico que caracterizan el espacio universitario y respaldado por el marco normativo vigente.

En este sentido, tanto el Centro de Investigación en Sistemas, Tecnologías de la Información y las Telecomunicaciones de la Universidad de Buenos Aires, como el Departamento de Ciencias Básicas de la Universidad Nacional de Chilecito han buscado articular sus recursos y producciones científicas – tecnológicas, garantizando un conjunto de trabajos acordes a las líneas de investigación de ambas instituciones. Dentro de esas líneas de trabajo se encuentran el desarrollo de Repositorios Institucionales (RI), Objetos de Aprendizaje (OA) y el movimiento de Acceso Abierto (AA); tanto en lo que respecta a los aspectos técnicos vinculados a ellos, como a su utilización y uso como productos útiles en la enseñanza primaria y secundaria.

Palabras clave: repositorios institucionales, objetos de aprendizaje, material educativo digital, acceso abierto.

CONTEXTO

En el marco del Plan de Fortalecimiento de la Investigación Científica, el Desarrollo Tecnológico y la Innovación en las Universidades Nacionales (Ac. Pl. N° 676/08 y N° 687/09) y del Reglamento PDTSCIN (Ac. Pl. N° 901/14), el Consejo Interuniversitario Nacional y el CONICET convocan, en el año 2014, a Institu-

ciones Universitarias Públicas a la presentación de Proyectos de Desarrollo Tecnológico y Social (PDTs) para la adjudicación de subsidios.

La convocatoria exigía que dentro del equipo de investigación del proyecto, la dirección fuese realizada a través de un Grupo Responsable, que debía incluir a docentes-investigadores de al menos dos Universidades, preferentemente de distintos CPRES (Consejos Regionales de Planificación de la Educación Superior). Dado que existía una coincidencia en las líneas de Investigación de ambas Universidades, la UBA presentó un proyecto conjuntamente con la UNDEC, el cual dentro del total de Proyectos presentados recibió el número de orden 407.

Evaluados los proyectos por pertinencia y calidad, el CIN, por Resolución CE N° 1055/15 Buenos Aires, del 20 de mayo de 2015, aprobó para ser financiados 68 proyectos, resultando el Proyecto Repositorios Digitales con Contenidos Orientados a las Necesidades de Escuelas Rurales (ER) Primarias y Secundarias dentro de ellos. Este PDTs obtuvo el Orden de Mérito 48 entre más de 600 trabajos, habiendo sido calificado con 99 puntos sobre 100, y por tal causa se le asignó la totalidad del monto del subsidio solicitado que fue de \$ 180.000,00.

Esta línea de I/D/I corresponde al diseño, desarrollo e implementación de proyectos que fortalecerán la investigación y las diferentes labores relacionadas con la gestión del conocimiento en cuanto a la visibilidad web y la preservación de la mayoría de la producción académica y científica tanto de la UNdeC como de la FCE de la UBA y otras Instituciones Educativas de nivel primario y secundario, fortaleciendo la articulación entre todos los niveles educativos. Esta línea de investigación se encuentra alineada con los núcleos estratégicos del Plan Nacional CyT 2020 ya que comprende el desarrollo de bases de

datos dirigidas a la generación de sistemas con fines socio-productivos inclusivos, al mejoramiento de las comunicaciones entre las comunidades y a la creación de Centros de Innovación Inclusiva.

En lo que respecta a la UNdeC cabe mencionar que tiene otros proyectos activos como el proyecto “Red para la creación y publicación de objetos virtuales de aprendizajes de calidad en Repositorios Institucionales” que corresponde a la convocatoria “Redes Internacionales 9”, promovida por la SPU, período de ejecución 2016-2017; un proyecto presentado en la convocatoria 2013-2014 del programa “Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica” de la Secretaría de Ciencia y Tecnología en la UNdeC cuyo título es “Fortalecimiento y visibilidad web de la producción científica de la UNdeC”.

Los proyectos del programa de Redes Internacionales permiten generar los espacios para discutir, exponer y articular, en diferentes ámbitos nacionales e internacionales, el trabajo y los resultados obtenidos, mientras que el proyecto correspondiente al programa de Financiamiento para el Estímulo y Desarrollo de la Investigación Científica y Tecnológica sirve de base para la búsqueda de trabajos científicos que puedan colaborar y responder a las necesidades de las IE.

1. INTRODUCCIÓN

La incorporación de las Tecnologías de la Información y Comunicación (TIC) en la Educación ha permitido extender los ambientes de enseñanza y aprendizaje, para así poder desarrollar ambientes virtuales colaborativos e interactivos, con el uso de diferentes materiales educativos digitales, que permitan el logro de los objetivos. Es por ello, que en el ámbito educativo se ha impuesto un nuevo concepto que busca la reutilización, permanencia, interoperabilidad, accesibilidad y compatibilidad de recursos digitales para el desarrollo de cur-

sos y programas de formación en línea a través de la Web, los llamados Objetos de Aprendizaje (OA). Lo que distingue a un OA de un material educativo digital es la introducción de información autodescriptiva, expresada a través de los metadatos, conjunto de atributos o elementos que permiten describir al objeto, es decir son entidades generalmente entregadas a través de Internet y diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos [5, 6].

En los últimos años, el desarrollo de Repositorios Institucionales (RI) de acceso abierto ha sido un tema prioritario en las políticas de educación, ciencia y técnica de muchos países, y en particular, en las universidades públicas de Argentina, aprobándose la Ley 26899 de Repositorios Institucionales (RI), y en los otros niveles educativos las Bibliotecas Escolares con la aprobación de la Ley 26.917 sobre el Sistema Nacional de Bibliotecas Escolares y Unidades de Información Educativa.

Los RI tienen como propósito recopilar, catalogar, gestionar, acceder, difundir y preservar información, permitiendo entre otras cosas el acceso libre y gratuito a todos los recursos que los conforman [1], [2]. Estos están constituidos por un conjunto de archivos digitales en representación de productos científicos y académicos que pueden ser accedidos por los usuarios [1]. En este sentido, un RI puede integrar a los OA en los denominados Repositorios de Objetos de Aprendizaje (ROA), bibliotecas digitales especializadas, orientados a facilitar la búsqueda y recuperación de los OA de manera que puedan ser utilizados en diversos ambientes educativos [11].

La filosofía del Acceso Abierto tiene como fin asegurar el acceso libre y abierto a la producción científica y académica, es decir, garantiza el acceso a través de Internet sin que los derechos del copyright sean una barrera [3], [4]. Es así que los RI

materializan el objetivo del OA porque la información que se deposita es una producción que tiene como propósito ser accesible sin restricciones, y preservada digitalmente.

El RI permitirá contar en la región con material de estudio generado por instituciones de nivel primario y secundario, contribuyendo a la articulación de todos los niveles. El compartir los trabajos, obras, documentos, libros, trabajos prácticos, entre otros, generados por los actores, permitirá mejorar la calidad de la educación que recibirán los estudiantes y aumentar la eficiencia de la investigación. Esa información podrá ser utilizada para explorar nuevas hipótesis, conocimientos previos o nuevos, validar métodos de estudio, técnicas de análisis o estrategias de aprendizaje utilizadas en las distintas IE bajo el principio de acceso a la información libre y gratuita, es decir, sin restricciones legales, técnicas y de acceso.

A través de los Centros de Innovación Inclusiva creados, todos los integrantes de la comunidad podrán constituirse no sólo en consumidores de saberes, sino en creadores de conocimiento. Aunque la brecha digital se está reduciendo gracias a las políticas de inclusión nacionales y provinciales, aún existen sectores sociales con dificultades para acceder a la bibliografía básica y técnica necesaria para la formación. Estos sectores serán los principales beneficiarios del RI, ya que con el apoyo directo y trabajo conjunto de las IE también se tendrá un impacto indirecto en la comunidad de la que forman parte. Serán los mismos estudiantes que adquieran las competencias necesarias para superar el alfabetismo digital y son quienes contribuirán desde sus hogares, escuelas o cualquier otro lugar para replicar estos resultados, sobre todo fomentando el hábito de la lectura/escritura en situaciones donde el acceso físico al material bibliográfico suele ser escaso. Esto dará lugar a nuevos proyectos orientados a esos espacios de la

comunidad, dirigidos precisamente a superar sus dificultades de inclusión digital.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este trabajo incluye las siguientes líneas de investigación:

- Repositorios institucionales y bibliotecas digitales.
- Objetos de aprendizaje.
- Gestión de la información y el conocimiento.
- Sistemas de información web y bases de datos.
- Índices bibliométricos.
- Interoperabilidad.
- Preservación digital.
- Recuperación de la información.

3. OBJETIVOS

El desarrollo de esta línea de investigación está permitiendo cumplir con los siguientes objetivos:

- Comprender los problemas y las tendencias asociadas a la producción abierta del conocimiento, el uso y la difusión.
- Analizar la visibilidad web de las IE intervinientes.
- Analizar los recursos educativos existentes que cumplan con las necesidades educativas de las IE intervinientes.
- Fomentar la formación de habilidades y conocimientos relacionados con los RI y los OA.
- Entender y valorar el movimiento mundial de Acceso Abierto y señalar las vías de su materialización.
- Implementar Repositorios de prueba para visualizar los diferentes recursos (educativos, académicos y científicos) que se produzcan en las IE de la región.
- Depositar y preservar los OA generados por los diferentes actores de las IE.

- Definir e implementar políticas de recopilación, distribución y mantenimiento para el funcionamiento del RI.
- Establecer principios y normas para que permitan la implementación de los Centros de Innovación Inclusiva [7, 8].
- Contar en la región con el primer RI con material educativo digital generado por IE de nivel primario y secundario.

4. RESULTADOS OBTENIDOS/ESPERADOS

- Relevamiento de los diferentes recursos educativos producidos y a producir por las IE, con especial énfasis en las escuelas rurales de montaña, e incluyendo además, a todas las instituciones educativas de la ciudad de Chilecito y su zona de influencia, siguiendo los lineamientos del proyecto PDTS, anteriormente nombrando.
- Curso a nivel de postgrado llamado “Gestión de la Información Académica y Científica”.
- Se desarrolló un proyecto de asignaturas sobre las Bibliotecas Digitales para presentarse en dos universidades [9, 10] y otras que lo puedan requerir.
- Se implementó un repositorio de prueba que está gestionando los recursos educativos y objetos de aprendizaje generados dentro del proyecto PDTS.
- Se han desarrollado talleres en las IE de nivel primario y secundario sobre herramientas de ofimática, de producción colaborativa de documentos, de búsqueda y selección de información en la web.
- Se desarrollaran en las IE de nivel primario y secundario talleres para la formación de habilidades y conocimientos relacionados con los RI y OA y producción de OA.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está formado por investigadores categorizados y otros en formación del CISTIC/FCE y de la Carrera de Licenciatura en Sistemas de Información de las Organizaciones Universidad de Buenos Aires por una parte; y por la otra, de investigadores doctorados y categorizados pertenecientes a de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC, especializados bibliotecas digitales, repositorios institucionales, objetos de aprendizaje, desarrollo de software, cómputo paralelo y tecnología grid; participando también alumnos avanzados de grado de ambas Universidades.

Por lo expresado, se cuenta con los recursos humanos con las habilidades y la formación académica en las diversas áreas de la propuesta, asegurando la concreción de esta línea de trabajo.

Los integrantes son docentes de las asignaturas las siguientes asignaturas:

En la Universidad de Buenos Aires: Teoría de los Lenguajes y Sistemas Operativos, Construcción de Aplicaciones Informáticas, y Sistemas de Datos.

En la Universidad Nacional de Chilecito: Programación I, Sistemas I, Arquitecturas Paralelas, Teoría de la Computación y Herramientas de Ingeniería de Software.

6. BIBLIOGRAFÍA

- [1]. J. Texier, “Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior,” presented at the 11th Latin American and Caribbean Conference for Engineering and Technology - 2013, Cancun, México, 2013, p. 9.
- [2]. M. De Giusti, N. Oviedo, A. Lira, A. Sobrado, J. Martínez, and A. Pinto, “SEDICI – Desafíos y experiencias en la vida de un repositorio digital,” *RENATA*, vol. 1, no. 2, pp. 16–33, Aug. 2011.
- [3]. P. Suber, “Ensuring open access for publicly funded research,” *BMJ*, vol. 345, 2012.
- [4]. D. Torres-Salinas, N. Robinson-García, and A. Cabezas-Clavijo, “Compartir los datos de investigación en ciencia: introducción al data sharing,” *Profesional de la Información*, vol. 21, no. 2, pp. 173–184, 2012.
- [5]. D. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE, 2001.
- [6]. E. Morales, F. García, A. Barron, A. Berlanga, and C. López, “Propuesta de Evaluación de Objetos de Aprendizaje,” presented at the Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE), Barcelona, 2005.
- [7]. La Innovación Inclusiva: <http://www.mincyt.gob.ar/noticias/una-jornada-para-pensar-sobre-la-innovacion-inclusiva-12550>.
- [8]. La Innovación al Servicio de la Inclusión Social: <http://www.mincyt.gob.ar/noticias/la-innovacion-al-servicio-de-la-inclusion-social-10731>
- [9]. Texier, J., Zambrano, J., & Carmona, F. B. (2016). Las Bibliotecas Digitales en el Currículum de las carreras de

- Ciencias de la Computación: una propuesta posible para Argentina y Venezuela. In XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016).
- [10]. Texier, J. (2016). Los repositorios institucionales y su importancia en las Universidades Venezolanas. Revista Scitus. Venezuela.
- [11]. Paulsson, F., Naeve, A. Establishing technical quality criteria for Learning Objects. Proceedings of To be published in the proceedings of eChallenges 2006, Barcelona, Spain

Repositorios Digitales de Acceso Abierto en la Facultad de Tecnología y Ciencias Aplicadas (UNCA)

María Vanesa Doria¹, Ivanna Maricruz Lazarte¹, Carola Victoria Flores¹, Juan Pablo Moreno, María Isabel Korzeniewski¹, María Carolina Haustein¹, Oscar Antonio Arias²

¹ Departamento de Informática, Facultad de Tecnología y Ciencias Aplicadas. Universidad Nacional de Catamarca, Argentina

² Instituto de Informática, Facultad de Tecnología y Ciencias Aplicadas. Universidad Nacional de Catamarca, Argentina

Maximio Victoria 55, Catamarca, Argentina. Tel: +54 (383) 4435112

vanesadoria@gmail.com, ilazarte@tecno.unca.edu.ar, carolaflores@tecno.unca.edu.ar, jpmoreno@tecno.unca.edu.ar, marisak@tecno.unca.edu.ar, carolina.haustein@gmail.com, oarias@tecno.unca.edu.ar

Resumen

Uno de los pilares de las Universidades Nacionales del país es su interacción con el medio, donde se pretende volcar la producción científica que se desarrolla dentro del ámbito académico a la sociedad.

La Facultad de Tecnología y Cs. Aplicadas de la UNCA no es ajena al crecimiento de repositorios digitales de acceso abierto a los recursos académicos y de investigación en las instituciones universitarias de acuerdo a la Ley N° 26.899 "Creación de Repositorios Digitales Institucionales de Acceso Abierto, Propios o Compartidos".

Este trabajo tiene como propósito mostrar el marco de la línea de I/D Repositorios Digitales de Acceso Abierto que sirve de base para la construcción, implementación y puesta en marcha del repositorio institucional de la Facultad de Tecnología y Cs. Aplicadas y además presenta un prototipo de repositorio temático para la disciplina Ingeniería de

Software desarrollado para el Dpto. Informática de la facultad mencionada.

Palabras clave: repositorios digitales, acceso abierto, repositorios institucionales, repositorios temáticos

Contexto

El Laboratorio de Tecnología de la Información y las Comunicaciones (LaTICs) de la Facultad de Tecnología y Ciencias Aplicadas (FTyCA) de la Universidad Nacional de Catamarca (UNCA), entre sus funciones, organiza actividades para dar continuidad a la I/D que surgen de tesis de postgrado y trabajos finales. En este marco, continúa con la línea de I/D de la tesis de postgrado denominada "Modelo de un Repositorio en Ingeniería de Software para la Facultad de Tecnología y Cs. Aplicadas de la UNCA" de la Mgtr. Maria Vanesa Doria, y que se presentará en las próximas convocatorias de Proyectos de

Investigación y Desarrollo 2017 de la FTCYA de la UNCA.

Introducción

La Real Academia Española (RAE) define un repositorio como “Lugar donde se guarda algo”. Esta definición se aplicó al léxico específico de la informática para designar como Repositorios Digitales (RD) de Acceso Abierto (AA), a un lugar donde se guardan documentos científicos y/o académicos siguiendo la filosofía del AA. Esta filosofía propone la publicación y difusión de la literatura académica y científica a través de Internet de manera libre y gratuita permitiendo a los usuarios la lectura, descarga, copia, distribución, impresión, búsqueda o enlace a los textos completos de los artículos y otras producciones académicas y científicas, sin otras barreras económicas, legales o técnicas que las que suponga Internet en sí misma (Budapest, 2002).

Los RD de AA se clasifican en:

- Repositorios institucionales, que contienen la producción de una organización (universidad, instituto, sociedades científicas).
- Repositorios temáticos, los cuales están referidos al depósito de producciones vinculadas a una disciplina. Según Lynch (2003), abrieron camino en el ciclo de comunicación científica, lo que permitió a las comunidades académicas identificar el potencial de esta alternativa de acceso a la literatura científica.

En la Argentina existen políticas nacionales para el desarrollo de repositorios institucionales a través del Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación (MINCYT), donde se creó el Sistema Nacional de Repositorios Digitales (SNRD). Este se encuentra en el marco jurídico de la Ley N° 26.899 "Creación de Repositorios Digitales Institucionales de Acceso Abierto, Propios o Compartidos" que establece que la producción científico-tecnológica financiada con fondos

públicos debe estar disponible en RD de AA y además, estas producciones deben ser depositadas mediante el proceso de autoarchivo en un plazo no mayor a 6 (seis) meses una vez publicada o aprobada la producción.

Los RD de AA, en el ámbito universitario, constituyen la clave para el almacenamiento, administración, localización y recuperación de documentos digitales de carácter científico-tecnológico y académico.

En el ámbito universitario Argentino se pueden mencionar las siguientes universidades que se han adherido a la ley mediante la creación de RD de AA:

- Universidad Nacional de Cuyo,
- Universidad de Buenos Aires,
- Universidad Nacional de Rosario,
- Universidad Nacional de La Plata,
- Universidad Nacional del Litoral,
- Universidad Nacional de Córdoba,
- Universidad Nacional de Salta,
- Universidad Nacional de Mar del Plata,
- Universidad Nacional de Tucumán, entre otras.

En las implementaciones de estas universidades en algunos casos se utiliza el término repositorios y en otros bibliotecas digitales, ya que se consideran sinónimos, según el contexto en el que se lo utilice, pero la diferencia fundamental es que los repositorios alojan recursos científicos, académicos y administrativos permitiendo la interoperabilidad con otros repositorios y la “cosecha de metadatos”, este concepto fue definido en el Open Archives Initiative Protocol for Metadata Harvesting (OAI-MHP, 2017), sin entrar en detalles técnicos, son acciones que básicamente puede realizar un proveedor de metadatos y que por tanto pueden ser solicitadas por un recolector de metadatos y prestar servicios de valor agregado, mientras las bibliotecas digitales tienen sus propios sistemas de estándares para recolectar información y no tiene como misión principal ser interoperable con otros sistemas. (Texier, 2013)

En la región latinoamericana se han aprobado legislaciones similares en Perú y

México, en apoyo de la creación de repositorios institucionales para la difusión de la producción científica (Bongiovani, 2014).

Atendiendo la relevancia del contexto presentado, la línea de I/D se enfoca en la construcción, implementación y puesta en marcha del Repositorio Institucional para la FTyCA de la UNCA (RITEYCA) y la implementación y puesta en marcha del Repositorio Temático en Ingeniería de Software (RIS).

RITEYCA comprenderá un conjunto de documentos de textos, teniendo como propósitos el de recolectar, organizar, preservar y difundir toda la producción científica- tecnológica y académica generada en la FTyCA.

RIS es un repositorio temático que alberga la producción científica-académica del Dpto. Informática de la FTyCA de la UNCA.

Los repositorios mencionados están basados en los estándares, políticas y protocolos comunes del SNRD del MINCyT de la República Argentina. Esto permitirá a futuro iniciar las gestiones para integrar RITEYCA y RIS al SNRD, lo que garantizará su interoperabilidad con los repositorios miembros del sistema.

Líneas de Investigación y Desarrollo

En la línea de I/D presentada en este trabajo se llevarán a cabo actividades relacionadas a las siguientes temáticas:

- Directrices de Creación de RD de AA: examinar las directrices elaboradas por el SNRD y definición de las políticas específicas para la FTyCA.
- Construcción de los RD de AA mencionados siguiendo la “Metodología para el Análisis, Diseño y Desarrollo de Servicios de Información Digital” propuesta por Navarro (2006). Esta metodología establece cinco etapas:
 - Etapa de planificación.
 - Etapa de diseño conceptual y lógico.

- Etapa de desarrollo.
- Etapa de mantenimiento.
- Etapa de evaluación.

Resultados Obtenidos/Esperados

El resultado obtenido es un prototipo del RIS.

Las producciones que se depositan en cada una de las colecciones del RIS deben ser un post-print, es decir, haber pasado por una instancia de evaluación. Las colecciones disponibles en el repositorio son:

- Artículos: Contiene artículos de revistas científicas y tienen vinculación con la temática del repositorio.
- Documentos de Eventos: Se incluyen los documentos resultantes de las presentaciones en jornadas, congresos, conferencias, etc.
- Tesis: Reúne Tesis de grado, Tesis de maestría, Tesis de Doctorado generados por alumnos y docentes.

Para la interoperabilidad del repositorio se tienen en cuenta las recomendaciones del protocolo Open Archives Initiative Protocol for Metadata Harvesting (OAI-MHP, 2017), y las Directrices del SNRD (Directrices, 2015). Tanto el protocolo como las directrices recomiendan el modelo de metadatos Dublin Core sin cualificar, por lo tanto, las producciones depositadas en el RIS se describen con este modelo.

En cuanto a Derechos de Autor, es importante definir qué derechos se ceden y cómo se lo realiza. En el RIS, excepto donde se diga explícitamente lo contrario, cada documento se publica bajo la Licencia Creative Commons 4.0 Internacional: CC BY-NC-SA. Es decir, no se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.

El depósito de documentos se realiza mediante el proceso de autoarchivo que consiste en una serie de simples pasos: 1) el

autor de una producción se loguea con su usuario y contraseña en el RIS; 2) elige la colección en la que desea depositar; 3) completa un formulario web con los datos de la producción digital que archiva, adjuntando el documento (artículo, documento de trabajo, libro, tesis, etc.). Este proceso luego pasa por una etapa de control por los gestores de contenido antes de ser publicados en el repositorio.

Actualmente se está trabajando en la construcción del repositorio institucional RITEYCA. De acuerdo a la metodología mencionada anteriormente, se están llevando a cabo las actividades referidas a la Etapa de Desarrollo.

Con la implementación de las primeras versiones de los repositorios RITEYCA y RIS se espera reunir, registrar, divulgar, preservar y dar acceso a las producciones científicas-tecnológicas y académicas de la FTyCA de la UNCA y brindar un servicio de búsqueda a texto completo a todos los interesados, facilitando la labor de docencia, investigación o estudio.

Formación de Recursos Humanos

El equipo de trabajo se encuentra conformado por la Comisión de Repositorios Digitales de Acceso Abierto perteneciente a LaTICs, según Resolución de la FTyCA N° 293/2016.

La estructura de esta comisión es la siguiente:

Coordinadores

- Mgtr. María Vanesa DORIA
- Mgtr. Carola Victoria FLORES
- Lic. Juan Pablo MORENO

Integrantes:

- Esp. María Isabel KORZENIEWSKI
- Lic. María Carolina HAUSTEIN
- Dra. Ivanna Maricruz LAZARTE
- Técnico Oscar Antonio ARIAS

La producción principal de esta línea de I/D es la Tesis de Postgrado “Modelo de un Repositorio en Ingeniería de Software para la Facultad de Tecnología y Cs. Aplicadas de la UNCA” (Doria, 2016).

Los seis docentes que integran el proyecto están categorizados en el Programa de Incentivos de la Secretaria de Políticas Universitarias (SPU) perteneciente al Ministerio de Educación y Deportes de la Nación Argentina. El integrante Técnico es No docente y pertenece al Instituto de Informática de la FTyCA.

Por lo antes expuesto, se cuenta con recursos humanos con habilidades y formación académica y técnica que permiten asegurar la concreción de la línea de I/D presentada.

Bibliografía

- Bongiovani, Paola C., Guarnieri, Griselda, Babini, Dominique, & López, Fernando A. (2014). Acceso Abierto en la Universidad Nacional de Rosario: Necesidades y prácticas de los docentes/investigadores. *Información, cultura y sociedad*, (30), 13-33. Recuperado en 17 de marzo de 2017, de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-17402014000100002&lng=es&tlng=es. [Consulta: 26 de febrero de 2016]
- Budapest Open Access Initiative. (2002) [online]. Disponible: <<http://www.soros.org/openaccess/read.shtml>> [Consulta: 26 de febrero de 2016]
- Directrices del SNRD, [online]. Disponible: http://repositorios.mincyt.gob.ar/pdfs/Directrices_SNRD_2015.pdf [Consulta: 26 de febrero de 2016]
- Doria, María Vanesa; Del Prado, Ana María, Haustein, María Carolina; *Repositorios digitales de acceso abierto* (2014) Editorial académica Española ISBN 978-3-659-00343-1

Doria, Maria Vanesa; Acceso Abierto y Repositorios Temáticos (2013) Editorial académica Española ISBN 978-3-659-06908-6

Doria, Maria Vanesa; Flores, Carola Victoria; Montejano, Germán Antonio; Modelo de Repositorio Temático en Ingeniería de Software: Un caso de éxito en la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca (2016) Editorial académica Española 978-3-8417- 5497-4

Doria, Maria Vanesa; “Modelo de un Repositorio en Ingeniería de Software para la Facultad de Tecnología y Cs. Aplicadas de la UNCA”; Editorial Científica Universitaria de la UNCA. Edición electrónica; ISBN 978-987-661-215-9

Ley 26.899, Link de acceso: <<http://www.infoleg.gob.ar/infolegInternet/anexos/220000-224999/223459/norma.htm>>, [Consulta: 26 de febrero de 2016]

Lynch, C; Institutional repositories: essential infrastructure for scholarship in the digital age. ARL No. 226: 1-7 (2003). [online] Disponible:

<http://scholarship.utm.edu/21/1/Lynch,_IRs.pdf> [Consulta: 26 de febrero de 2016]

Navarro, E., Planificación, diseño y desarrollo de servicios de información digital. En: Tramullas J, Garrido P (eds.). Software libre para servicios de información digital. Madrid: Pearson Prentice-Hall; ISBN 978-84-8322-299-7, 2006. p. 23-43

Sistema Nacional de Repositorio Digitales (SNRD), [online] Disponible: <<http://repositorios.mincyt.gob.ar/index.php>>, [Consulta: 26 de febrero de 2016]

OAI-PMH, [online] Disponible: <https://www.openarchives.org/pmh/> [Consulta: 26 de Marzo de 2017]

Texier, J (2013) Los repositorios institucionales y las bibliotecas digitales: una somera revisión bibliográfica y su relación en la educación superior. Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI'2013) "Innovation in Engineering, Technology and Education for

Competitiveness and Prosperity" August 14 - 16, 2013 Cancun, Mexico. [online] Disponible en: <http://www.laccei.org/LACCEI2013-Cancun/RefereedPapers/RP042.pdf> [Consulta: 26 de Marzo de 2017]

Resolviendo Numéricamente Ecuaciones Diferenciales Ordinarias: Diseño de un Software Educativo

Ascheri M. E., Pizarro R., Astudillo G., García P., Culla M. E., Pauletti C.

Facultad de Ciencias Exactas y Naturales. Universidad Nacional de la Pampa

Avda. Uruguay 151 - Santa Rosa - La Pampa - 02954-425166

mavacheri@gmail.com rubenpizarro71@gmail.com

RESUMEN

En el marco del curso Cálculo Numérico que se dicta para varias carreras de la Facultad de Ciencias Exactas y Naturales de la UNLPam, se está trabajando en un proyecto de investigación sobre la elaboración de un software educativo para la resolución numérica y gráfica de integrales y de ecuaciones diferenciales ordinarias, usando herramientas gratuitas existentes en la Web. Este software se ha desarrollado en diferentes etapas, permitiendo en la actualidad resolver temas relacionados con ecuaciones no lineales, interpolación, ajuste de curvas, integrales y ecuaciones diferenciales.

Los objetivos que se pretenden alcanzar son:

1. Generar el contexto educativo adecuado a los contenidos a desarrollar y a los objetivos propuestos en la asignatura.
2. Proporcionar una herramienta tecnológica para afianzar los conceptos teóricos y la aplicación de los métodos numéricos a través de la visualización gráfica del funcionamiento de los mismos.
3. Poner a disposición de los estudiantes una herramienta de acceso libre y disponible en la Web.

En este año, se finalizará con el proyecto obteniendo como resultado del mismo la posibilidad de disponer de una versión final del mencionado software y analizar cómo su utilización influye en el proceso de enseñanza y de aprendizaje de los temas desarrollados.

Palabras clave: software educativo, integración numérica, ecuaciones diferenciales ordinarias.

CONTEXTO

Este Proyecto de Investigación, está acreditado y financiado por la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa. Surge como continuidad de otros proyectos ya finalizados relativos a la elaboración de un software educativo para la resolución de temas de cálculo numérico.

Está relacionado con otros proyectos del Departamento de Matemática del cual este grupo de investigación es parte, en los que se abordan contenidos relacionados a la matemática, su enseñanza y la inclusión de tecnologías en el proceso de enseñanza y de aprendizaje.

INTRODUCCIÓN

La experiencia obtenida a partir de la

estrategia metodológica empleada en un Proyecto de Investigación anterior (Ascheri *et al.*, 2008, 2009, 2010, 2011), motiva a que se continúe con el desarrollo de software educativo que incluya otras temáticas de Cálculo Numérico para lograr una integración curricular y complementar los resultados obtenidos hasta el momento. Esta herramienta tiene como funcionalidad pretendida, guiar el aprendizaje como soporte del proceso de enseñanza. Actualmente, se ha finalizado con la etapa que se refiere a la elaboración del software educativo que contemple el tema resolución numérica y gráfica de integrales (Ascheri, *et al.*, 2014 y 2015). Dicho software se encuentra disponible en el sitio <http://secanu.exactas.unlpam.edu.ar/>



Figura 1: Software educativo de elaboración propia

Se ha comenzado a trabajar en el desarrollo del software que permita la resolución numérica y gráfica de ecuaciones diferenciales ordinarias. Dicho contenido está incluido en el curso de Cálculo Numérico que se dicta para las carreras de Profesorado en Matemática (3° Año), Licenciatura en Física (3° Año) e Ingeniería Civil (2° Año) de nuestra Facultad.

Según Alemán de Sánchez (1999), es importante que el software contemple no solamente las prácticas, sino que proporcione al estudiante ayuda en la solución de los problemas y brinde una

retroinformación completa, sin limitarse a indicar que se ha cometido un error, sino brindando información acerca del tipo de error. Este y otros aspectos serán considerados e implementados en los diversos ejemplos que se desarrollaran en el software.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN.

Las líneas de investigación que se abordan con el presente proyecto son las siguientes:

- Enseñanza de matemática, visualización gráfica de diferentes métodos numéricos.
- Inclusión de tecnologías en el proceso de enseñanza y de aprendizaje

3. RESULTADOS Y OBJETIVOS

Se ha desarrollado y puesto a punto el software educativo correspondiente a la *resolución numérica y gráfica de integrales*. Se propone ahora el logro del siguiente objetivo:

- ✓ Desarrollar el software educativo diseñando las etapas necesarias para lograr el producto, para la *resolución numérica y gráfica de ecuaciones diferenciales ordinarias*.

Para ello, se han planificado las siguientes tareas:

1. Hacer una búsqueda y análisis preliminar de elementos existentes en el campo de la Ingeniería de Software y en el de las Ciencias de la Educación.
2. Utilizar estos elementos, incorporándoles aspectos educativos relativos a la temática involucrada.
3. Experimentar y hacer un análisis y validación de este software.
4. Actualizar, de acuerdo a los resultados obtenidos, la aplicación web de la

asignatura "Cálculo Numérico" dependiente de la Facultad de Ciencias Exactas y Naturales de la UNLPam.

Para obtener los datos que permitirán realizar un análisis y validación de esta segunda fase del software educativo elaborado y obtener conclusiones, se utilizarán principalmente la observación, las encuestas y los resultados obtenidos por los estudiantes en las diferentes evaluaciones parciales y finales. Estos resultados constituirán el aporte de una metodología de investigación cuantitativa.

Se desarrollará una estrategia de triangulación que permitirá la coexistencia de la investigación cuantitativa y de la cualitativa (Erickson, 1999). Esta estrategia definida como la combinación de metodologías para el estudio de un mismo fenómeno, considera a los métodos cuantitativos y cualitativos como campos complementarios. Por medio de esta combinación de metodologías, se espera encontrar, con la misma metodología, posibles falencias o detectar contradicciones que no se puedan identificar por algunas de las herramientas (observación, encuestas, datos y demás) en forma aislada. En muchos casos, son necesarios ambos tipos de datos (cualitativos y cuantitativos), para una mutua verificación y de forma suplementaria (Vasilachis de Gialdino, 2006).

Un software educativo debe poner énfasis en lograr aprendizajes significativos, brindando posibilidades de vincular los nuevos conceptos con los que se tienen adquiridos y estableciendo relaciones no arbitrarias entre ellos. Debe tener en cuenta los objetivos de las unidades, además de los contenidos y sus relaciones. Así como complementarse con las herramientas existentes para facilitar la comprensión de las temáticas involucradas. También, lograr

un incremento de la motivación y facilitar el desarrollo de las destrezas. Asimismo, se lo debe incluir en actividades de evaluación, tanto para el estudiante como para el docente.

Para la elaboración de esta parte del software educativo se utilizarán distintas técnicas, tales como el uso de gráficas por computadoras y animación interactiva para ilustrar y presentar procesos y algoritmos (Alemán de Sánchez, 1999; Cataldi, 1999; Di Battista *et al.*, 1999; Galvis Panqueva, 1992; Marquès, 1996; Price *et al.*, 1998; Rivera Porto, 1997, entre otros).

La visualización de la conducta dinámica de los algoritmos presenta importantes beneficios educativos:

- Logran un incremento de la motivación.
- Facilitan el desarrollo de destrezas.
- Asisten en el desarrollo de habilidades analíticas.
- Ofrecen un buen soporte al docente.
- Permiten la exploración de las peculiaridades de un algoritmo, explorando de forma interactiva.

Las herramientas informáticas y las técnicas numéricas que usaremos son, respectivamente, la computadora y los métodos numéricos clásicos de resolución de ecuaciones diferenciales ordinarias (Ascheri y Pizarro, 2007; Chapra y Canale, 2007; Gerald y Wheatley, 2000; Mathews y Fink, 2000; Nakamura, 1997), tales como los métodos de Taylor, Euler, Heun, Runge-Kutta, pasos múltiples, entre otros.

El software diseñado brinda la posibilidad que durante el desarrollo del curso de Cálculo Numérico los docentes y estudiantes puedan en sus clases teóricas y prácticas visualizar el comportamiento de los distintos métodos mencionados y probar diferentes ejemplos con diversos conjunto de valores iniciales, haciendo variar el comportamiento del método (simulación).

El estudiante podrá, fuera del horario de clases, experimentar las diferentes alternativas del software ya que el mismo se encuentra disponible en línea. Debemos considerar que el mencionado software además se puede utilizar en la evaluación, lo que permite desarrollar, proponer y concretar evaluaciones que de otra forma no se podrían implementar. Incluyendo, así, la aplicación en los distintos procesos de la enseñanza y del aprendizaje.

FORMACIÓN DE RECURSOS HUMANOS

Se espera continuar con un grupo de trabajo considerando:

- La motivación de los integrantes a participar de jornadas y eventos científicos.
- La contribución al desarrollo de capacidades y habilidades intelectuales en jóvenes estudiantes que participen o se relacionen con el Proyecto.
- La posibilidad de interesar a otros docentes y estudiantes en participar en actividades de investigación.

A partir de la presentación de los resultados de proyectos relacionados con el actual se han adquirido compromisos tendientes a realizar intercambios de las producciones que surjan, logrando así ampliar la población a la que estará dirigida la implementación del software educativo y que no se restrinja sólo a la cátedra de Cálculo Numérico.

REFERENCIAS

Alemán de Sánchez, A. *La enseñanza de la matemática asistida por computadora.* (1999). En línea: <http://www.utp.ac.pa/articulos/enseñarmatemática.html>.

Ascheri, M. E.; Pizarro, R. *Libro de Texto para Estudiantes Universitarios: CÁLCULO NUMÉRICO.* EdUNLPam. (2007).

Software educativo para la resolución numérica y gráfica de temas de cálculo numérico. Ascheri M. E., Pizarro R. A., García P., Astudillo G. J., Culla M. E.. *Un software educativo con herramientas libres y acceso web para temas de cálculo numérico: primer prototipo.* Memorias de la II REPEM, La Pampa, Argentina. ISBN 978-950-863-103-9. (2008).

Ascheri M. E., Pizarro R. A., García P., Astudillo G. J., Culla M. E. *Software educativo para temas de cálculo numérico: primera etapa de evaluación.* Memorias XI Workshop de Investigadores en Ciencias de la Computación, WICC 2009, San Juan, ISBN: 978-950-605-570-7. (2009)..

Ascheri, M. E., Pizarro, R. A., Astudillo, G. J., García, P., & Culla, M. E. *Software educativo para temas de cálculo numérico.* Memorias XII Workshop de Investigadores en Ciencias de la Computación. ISBN: 978-950-34-0652-6. (2010).

Ascheri, M. E., Pizarro, R. A., Astudillo, G. J., García, P., & Culla, M. E. *Software educativo para temas de cálculo numérico.* Memorias XII Workshop de Investigadores en Ciencias de la Computación. ISBN: 978-950-673-892-1. (2011).

Ascheri, M. E.; Pizarro, R. A.; Astudillo, G.; García, P.; Culla, M. E. *Avances en la Elaboración del Software Educativo para la Resolución Numérica y Gráfica de Integrales,* Memorias de la V REPEM, Vol. 5, ISSN N° 2362-5716, pp. 322-330. (2014).

Ascheri, M. E.; Pizarro, R. A.; Astudillo, G.; García, P.; Culla, M. E.; Pauletti, C., *Utilización de un software educativo para la resolución numérica y gráfica de integrales,* Memorias del XVII Workshop de Investigadores en Ciencias de la Computación, ISBN N° 978-987-633-134-0, pp. 1-5. (2015).

- Cataldi, Z.; Lage, F.; Pessacq, R.; García Martínez, R.** *Revisión de Marcos Teóricos Educativos para el Diseño y Uso de Programas Didácticos.* (1999). En línea: www.itba.edu.ar/capis/webcapis/RGMITBA/comunicacionesrgm/c-icie99-revisionde%20marcosteoriciseducativos.pdf
- Chapra, S.; Canale, R.** *Métodos Numéricos para Ingenieros.* Mc Graw Hill/Interamericana de España, S. A. U. (2007).
- Di Battista, P.; Eades, G.; Tamassia, R.; Tollis, I.** *Graph Drawing: Algorithms for the Visualization of Graphs.* Prentice Hall. (1999).
- Erickson, F.** *Métodos cualitativos de investigación sobre la enseñanza.* Paidós. (1999).
- Galvis Panqueva, A.** *Ingeniería de Software Educativo.* Ediciones Unidades (1992).
- Gerald, C. y Wheatley, P.** *Análisis Numérico con aplicaciones.* México: Pearson Educación. (2000).
- Marquès, P.** *El software educativo.* Universidad Autónoma de Barcelona. (1996). En línea: http://www.lmi.ub.es/te/any96/marques_software
- Mathews, J.; Fink, K.** *Métodos Numéricos con MATLAB.* Prentice Hall. (2000).
- Nakamura, S.** *Análisis Numérico y Visualización Gráfica con MATLAB.* México: Pearson Educación. (1997).
- Price, B.; Beacker, R.; Small, I.** *An Introduction to Software Visualization, Software Visualization.* MIT Press. (1998).
- Rivera Porto, E.,** *Aprendizaje Asistido por Computadora. Diseño y Realización.* (1997). En línea: <http://www.geocities.com/eriverap/libros/Aprend-comp/apen1.html>
- Vasilachis de Gialdino, I.** *Estrategias de Investigación Cualitativa.* Gedisa, S. A. (2006).

Sub-redes de Gestión y sus Herramientas como Apoyo en un Departamento Pedagógico Universitario

María V. Godoy, Sonia I. Mariño, Pedro L. Alfonzo, Marcelo Caceres, Romina Alderete

Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste.

Argentina, 9 de Julio 1449, 3400 Corrientes, Argentina,

mvgg2001@yahoo.com, simarinio@yahoo.com, plalfonzo@hotmail.com

Resumen

La Educación Superior ofrece a la gestión de proyectos, una diversidad de ámbitos de aplicación de administración de información. Los artefactos de las TIC se constituyen en una potencial herramienta para mediar procesos de Gestión del Conocimiento. El trabajo expone la generación de Sub-redes de gestión destinadas a un Departamento Pedagógico Universitario, así como los artefactos software que las sustentan y median actividades de administración de datos orientadas a la toma de decisiones.

Palabras clave: Educación Superior, Gestión en la Información, Programas Nacionales

Contexto

La línea de investigación presentada en este documento se desarrolla en el Área de Ingeniería Web y en el marco de un Componente del Programa PROMINF (2013-2016) de la FaCENA, UNNE.

Introducción

En la sociedad actual, el conocimiento es una variable determinante del desempeño de las organizaciones, su gestión es clave. Las instituciones de Educación Superior (ES) deben tratar con este paradigma y aplicar las potencialidades que ofrecen las tecnologías para su desarrollo. En este sentido, Ruiz Tapia et al. (2016, p. 29) definen cómo las TIC configuran nuevas relaciones comunicacionales en la sociedad, en las universidades y en las actividades que ellas involucran.

El Proyecto de Mejoramiento de la Enseñanza en Carreras de Licenciatura en Ciencias de la Computación, Licenciatura en

Sistemas/Sistemas de Información/Análisis de Sistemas, Licenciatura en Informática, Ingeniería en Computación e Ingeniería en Sistemas de Información/Informática (PROMINF), atiende diversos aspectos para el mejoramiento de la carrera, tales como: Apoyo Mejoramiento Gestión Académica; Actividades Interinstitucionales; Recursos Humanos Académicos; Infraestructura; Equipamiento; Bibliografía

Uno de los componentes aborda la conformación de una Red de Gestión Integral (RG). Ésta permitirá intercambiar información entre diversas unidades académicas del país. Se diseñará a tal efecto un sistema de RG que contendrá información de: Proyectos Finales de Carrera, Prácticas Profesionales Supervisadas, Tutorías y Gestión Departamental (casos a elegir), lo que constituirá un insumo para el monitoreo de la carrera y una fuente de centralización de documentación.

La Red de Gestión de RRHH propuesta se enmarca en lo expuesto, dado el potencial que ofrecen las herramientas TIC en procesos de administración transformando los datos en información valiosa. Se propone desde su inicio trabajar en colaboración con la Universidad Nacional de Misiones y la Universidad Nacional del Centro, a fin de que los módulos sean transferidos entre las mencionadas Universidades.

Líneas de investigación y desarrollo

La ejecución del presente proyecto se define a través de las siguientes líneas de I+D

- Línea 1 de I+D: Desarrollo Económico – Social Regional Sustentable: Se definen sub-redes de conocimiento y las herramientas que dan soporte a las

mismas que aportarán al desarrollo interno del departamento en cuanto a gestión de información, así como relaciones cooperativas entre otros departamentos de Universidades regionales. Lo expuesto contribuirá a la posibilidad de contar con datos, para favorecer los procesos de aseguramiento de la calidad a los que son periódicamente sometidas las carreras universitarias

- Línea 2 de I+D orientada a la Educación. En estos ámbitos pueden establecerse procesos y herramientas que den sustento a actividades académicas o administrativas. En este caso se aborda la segunda y, se enmarca en la ejecución de un Componente del Programa PROMINF.
- Línea 3 –Métodos ágiles. La generación de las herramientas que soportan las sub-redes de conocimiento se diseñan aplicando una metodología ágil orientada a procesos en Educación Superior.

A partir de las líneas de I+D expuestas se sostiene la importancia de su tratamiento para el Departamento y la Carrera.

Resultados obtenidos/esperados

Como se expresó en PROMINF (2013) la propuesta se sustenta en “las TICs, cuyas herramientas informáticas ponen de manifiesto el potencial y la efectividad de las Redes operando en tres ámbitos, “las 3Cs”, de forma relacionada:

- Comunicación (ayudan a poner en común conocimientos).
- Comunidad (ayudan a formar e integrar comunidades).
- Cooperación (ayudan a hacer cosas juntos).

La conceptualización de las **Sub-Redes de Conocimiento** se fundamenta en la normativa existente, datos registrados en planillas electrónicas, documentos y el conocimiento tácito de los recursos humanos que intervienen en cada uno de los procesos administrativos susceptibles de ser gestionados por las misas.

Además, este proceso “reúne varias etapas de desarrollo y varios actores (...) forman grafos

complejos con jerarquías bien definidas, comunicaciones formales e informales, que contienen información pública, privada y crítica en una distribución longitudinal, transversal y sectorial dentro de la unidad académica” (PROMINF, 2013))

Por ello, se han logrado al momento los siguientes objetivos:

- Profundización de métodos ágiles y su aplicación en la definición de soluciones para la gestión departamental
- Estudio y utilización del entorno de programación o framework SIU-TOBA
- Definición de una base de datos como repositorio de datos para la posterior explotación de la información
- Definición de requerimientos de tres sub-redes de conocimiento, así como la conexión entre ellas.
- Diseño y desarrollo de tres sub-redes de conocimiento, que aportan a la gestión departamental, y las herramientas informáticas que la soportan.
- Elaboración y presentación de trabajos en congresos y los informes técnicos ante autoridades de la FACENA.

Formación de recursos humanos

- Se están formando dos graduados en esta línea de I+D que favorecerá el desarrollo integral de artefactos software de apoyo a la gestión departamental. Particularmente en el desarrollo web en el mencionado framework, que permitirá fortalecer los sujetos que con estos conocimientos aportan a desarrollos de gestión departamental con miras a administración académica.
- Integrantes del Área de Ingeniería Web fortalecen su visión y formación en desarrollos tecnológicos para la gestión departamental con miras a la transferencia de conocimiento y tecnología a otros ámbitos de Educación Superior.

- Se realizan Actividades de formación académica, previendo el dictado de cursos de especialización en la temática.
- Se promueve la participación en actividades de difusión académica científica, conferencias, publicaciones en congresos y revistas nacionales e internacionales.

Bibliografía

Agile Project Management AdvisoryService: ExecutiveUpdate, 5(20), 1–4. Disponible en: <http://www.torak.com/site/files/Lessons Learned From The First Scrum by Dr.Jeff Sutherland.pdf>

CARBALLO, R. (2006). Innovación y gestión del conocimiento. Ed. Díaz de Santos.

DAVENPORT, T. y PRUSAK, L. (1998). Working Knowledge. Harvard Bussines School Press, Boston, Massachusett

DEEMER, P., BENEFIELD, G., LARMAN, C. y VODDE, B (2009). Información Básica de ScrumtheScrum Primer Version 1.1. Scrum Training Institute. Traducción de Leo Antoli. Agile, Spain. Disponible en: http://www.goodagile.com/scrumprimer /scrumprimer_es.pdf.

GARCÍA PEÑALVO, F. J.; GARCÍA DE FIGUEROLA, C., MERLO VEGA, J. A. (2010). Open Knowledge Management in Higher Education. Disponible en: <http://hdl.handle.net/10366/121874>.

FROST, A. (2014). A Synthesis of Knowledge Management Failure Factor. Disponible en: <http://www.knowledge, management, tools.net>.

LEY 26.427. Creación del sistema de pasantías educativas en el marco del sistema educativo nacional. Publicación en Boletín Oficial: 22/12/2008

NÚÑEZ PAULA I. A. y NÚÑEZ GOVÍN Y. (2005). Propuesta de clasificación de las herramientas , software para la gestión del conocimiento. Acimed 13(2). (En línea). Disponible en:

http://bvs.sld.cu/revistas/aci/vol13_2_05 /aci03_05.htm.

PALACIO, J. Y RUATA, C. (2009). Scrum Manager: Gestión de proyectos. Disponible en: http://www.scrummanager.net/files/sm_ proyecto.pdf

PASSONI, L. (2005). Gestión del conocimiento: una aplicación en departamentos académicos. Tesis, Disponible en: http://www.accessmylibrary.com/coms2 /summary_0286, 31232685_ITM.

PEREZ, O. (2011). Cuatro enfoques metodológicos para el desarrollo de SoftwareRUP – MSF – XP – SCRUM, Inventum No. 10 Facultad de Ingeniería UNIMINUTO. Junio de 2011 64 , ISSN 1909 – 2520; p. 64, 78

PÉREZ LINDO A., RUIZ MORENO, L. VARELA, C., GROSSO, F., CAMÓS, C., TROTTINI, A, M, BURKE, M. L. y DARIN, S. (2005). Gestión del conocimiento. Un nuevo enfoque aplicable a las organizaciones y la universidad. Grupo Editorial Norma. Bs. As.

PÉREZ LINDO, A. (2007). De la revolución cognitiva a la gestión del conocimiento. Documento presentado en el curso de postgrado “Gestión del Conocimiento”. Doctorado de Ciencias Cognitivas. Facultad de Humanidades. Universidad Nacional del Nordeste, Inédito.

PÉREZ LINDO, A. (2007b). El Gobierno inteligente en la Republica Universitaria. De Platón a la Gestión del Conocimiento. Documento presentado en el curso de postgrado “Gestión del Conocimiento”. Doctorado de Ciencias Cognitivas. Facultad de Humanidades. Universidad Nacional del Nordeste, Inédito.

POLANYI, M. (1983). The Tacit Dimension. Gloucester, Massachussets, 1, 25.

RUIZ TAPIA, J. A., MARTÍNEZ, M., SÁNCHEZ, M. DE LA L. (2016). El impacto de las TICs en la calidad de la educación superior, Revista de

- Investigación en Ciencias Contables y Administrativas. Vol. 1(1), pp. 28, 44.
- PROMINF (2013). UNNE, UNICEN, UNAM. Redes para la Planificación, Ejecución y Control de Actividades Académicas. Programa PROMINF.
- Reglamento para el Régimen de Adscripciones, Resolución N 1286 /13, FACENA, UNNE
- SCHWABER, K. (1995). Scrum Development Process, in OOPSLA Business Object Design and Implementation Workshop, J. Sutherland, D. Patel, C. Casanave, J. Miller, and G. Hollowell, Eds. London: Springer.
- SUTHERLAND, J., VIKTOROV, A. y BLOUNT, J. (2007). Distributed Scrum: Agile Project Management with Outsourced Development Teams, hicss, pp.274a, 40th Annual Hawaii International Conference on System Sciences (HICSS'07).
- SUTHERLAND, J. (2004). Agile development: Lessons learned from the first scrum. Cutter
- SUTHERLAND, J., Y SCHWABER, K. (2011). The Scrum Papers : Nut , Bolts , and Origins of an Agile Framework. Extraído de <http://jeffsutherland.com/ScrumPapers.pdf>

Tecnología e Innovación en Ambientes de Aprendizaje: Desarrollo y Gestión

Stella Maris Massa¹, Adolfo Spinelli², Felipe Evans³, Jorge Petrillo⁴, Carlos Rico⁵

Fac. de Ingeniería, Universidad Nacional de Mar del Plata, Argentina

(7600) Juan B. Justo 4302 +54-223-4816600

smassa@fi.mdp.edu.ar¹, spinelliadolfo@gmail.com², evansfelipe@gmail.com³, petrillo@mdp.edu.ar⁴,
crico@crico.com.ar⁵

Resumen

Se presenta un proyecto I+D+T, (Investigación, Desarrollo y Transferencia) bianual (2016-2017).

El proyecto aborda la problemática de la inclusión de las tecnologías de la información y de la comunicación (TIC) en ambientes de aprendizaje desde una visión innovadora.

En los países en vías de desarrollo se concibe cada día más a la universidad como un agente clave para el desarrollo de potencialidades, recursos y valores para la propia sociedad.

Por otra parte, en la actualidad la tecnología desarrolla medios innovadores para llegar a mayor cantidad de personas denominándolas "Tecnologías emergentes". Entre ellas se encuentran los dispositivos móviles, la realidad aumentada, la gamificación, el m-learning, el flipped classroom (clases invertidas), los videojuegos y los cursos masivos en línea (MOOC).

En este escenario y con la mirada en nuestro contexto educativo es que este proyecto se focaliza en ambientes de aprendizaje que incluyan tecnologías emergentes y en la producción, difusión y preservación de Recursos Educativos Abiertos que favorezcan el desarrollo de competencias de formación profesional. El objetivo es indagar espacios innovadores de interacción, modelos de representación de la información y de gestión de contenidos de ambientes de aprendizaje.

Palabras clave: Ambientes de aprendizaje, Innovación, Tecnología.

Contexto

El proyecto I+D+T que se presenta está enmarcado en el Grupo de Investigación en Tecnologías Interactivas (GTI) radicado en el Departamento de Informática de la Facultad de Ingeniería de UNMDP. El grupo de trabajo está integrado por especialistas en Informática, Educación y Ciencias Sociales. Esta pluralidad ha contribuido al desarrollo de las líneas de investigación del Grupo GTI y en particular en las propuestas para el proyecto que se describe en este artículo.

Introducción

El presente proyecto aborda la problemática de la inclusión de las tecnologías de la información y de la comunicación (TIC) en ambientes de aprendizaje desde una visión innovadora.

En ese sentido, Medina [1] señala que las decisiones metodológicas de cambio en ambientes de aprendizaje deberían sustentarse en principios como los siguientes: transferencia, problematización de lo aprendido o a aprender, prospectiva, toma de decisiones, complejidad, autenticidad, realismo analógico, incertidumbre, integración del saber y hacer y multiculturalismo.

Del mismo modo, el cambio educativo, según Fullan [2], es un proceso dinámico que implica la interacción de variables en el tiempo. El autor agrupa los factores necesarios para implementar transformaciones en educación en tres grandes categorías: en primer lugar las características de la innovación; en segundo lugar las características locales tales como ubicación, nexos geográficos, gestión admi-

nistrativa, autoridades y/o directriz de sectores y, finalmente, las características externas, relacionadas con las políticas gubernamentales o las iniciativas apoyadas por agencias internacionales.

En la actualidad la tecnología ha desarrollado mejores medios innovadores para llegar a mayor cantidad de personas denominándolas “Tecnologías emergentes” [3]. Ramírez Montoya [4] las denomina nuevas tecnologías disruptivas e innovaciones basadas en la ciencia con la capacidad de crear una industria nueva o la de transformar una existente. Comprenden las telecomunicaciones, los dispositivos móviles, la realidad aumentada, dando lugar a la gamificación, el m-learning, el flipped classroom (clases invertidas) y los cursos masivos en línea (MOOC), entre otros [5].

En lo que respecta a las tecnologías emergentes se debe mencionar el Movimiento Educativo Abierto que es un espacio en el que se dan a conocer las nuevas tendencias y los nuevos procesos que están surgiendo, además de mostrar la forma de utilizarlos, reutilizarlos y de integrarlos a los procesos educativos [6]. Particularmente los Recursos Educativos Abiertos (REA) representan una gran variedad de contenido digital, incluyendo cursos completos, materiales de cursos, módulos, libros de texto, vídeos, tests, simulaciones, videojuegos y cualquier otro medio de transmisión del conocimiento en formato digital.

Para compartir efectivamente los REA se crean los Repositorios, que se constituyen como un reservorio indexado de los mismos [7].

En este escenario y con la mirada en nuestro contexto educativo es que este proyecto se focaliza en ambientes de aprendizaje que incluyan tecnologías emergentes y en la producción, difusión y preservación de REA que favorezcan el desarrollo de competencias de formación profesional.

Desde esta perspectiva y tal como señalan [8], [9] y [10], entre otros, es que estos “nuevos ambientes de aprendizaje” se pueden enriquecer con la incorporación de las tecnologías para apoyar el desarrollo de competen-

cias.

En esa línea, autores como Beck & Wade [11] señalan que los videojuegos están ayudando a los jugadores a adquirir habilidades necesarias en los empleos actuales y futuros. Ha cambiado dramáticamente la forma en que miembros de la generación actual ven el mundo de negocios, el modo en que piensan acerca del trabajo, riesgos, éxitos, y qué esperan de sí mismos.

Entre los numerosos tipos de videojuegos que existen, son de especial interés en este proyecto los “juegos serios” (Serious Games). Los Serious Games son aplicaciones interactivas creadas con una intencionalidad educativa, que proponen la explotación de la jugabilidad como experiencia del jugador. Presentan a los jugadores retos y misiones que implican tomas de decisiones, resolución de problemas, búsqueda de información selectiva, cálculos, desarrollo de la creatividad y la imaginación, etc., logrando el efecto inmersivo en el juego, como una prolongación de la experiencia vital del usuario [12].

Conceptualmente, los Serious Games pueden ser considerados una iniciativa que se concentra en el uso de los principios de diseño de juegos para otros fines no meramente lúdicos, por ejemplo, capacitación, publicidad, simulación, o educación [13].

Es entonces que la situación generada por el uso masivo de las TIC ha provocado un replanteamiento de las metodologías de enseñanza y aprendizaje y está influyendo con claridad para que se creen nuevos escenarios formativos. Cabero [14] ya indicaba que esos nuevos escenarios formativos tendrían las siguientes características: tecnológicos/ mediáticos, amigables, flexibles, individualizados, colaborativos, activos, interactivos/dinámicos, deslocalizados espacialmente de la información, pluripersonales, y pluridimensionales/multiétnicos.

Es importante destacar que los problemas que se nos presentan en la actualidad, en líneas generales, no son tecnológicos. Proviene en saber qué hacer, cómo hacerlo, para quién y por qué hacerlo.

Urge pues potenciar la investigación didáctico-educativa y pensar en estrategias y soluciones pedagógicas que nos oriente hacia dónde dirigir el cambio en el desafío de los nuevos escenarios formativos que se nos presentan en la actualidad.

Líneas de Investigación

Las principales líneas de investigación que se abordan en este proyecto son:

- Entornos formativos potenciados por tecnologías interactivas.
- Serious Games y Gamificación.
- Repositorios de recursos educativos abiertos.

Resultados y Objetivos

En relación al proyecto de I+D+T, se plantearon los siguientes objetivos específicos:

- Explorar y analizar iniciativas educativas innovadoras de integración de la tecnología en ambientes de aprendizaje.
- Diseñar, implementar y evaluar ambientes de aprendizaje enriquecidos con tecnologías emergentes que contribuyan al desarrollo de competencias.
- Validar los principales criterios y procedimientos de evaluación apropiados para ambientes de aprendizaje enriquecidos con tecnología que permitan verificar si los estudiantes han desarrollado las competencias asignadas.
- Generar soluciones de software (REA) basadas en el uso adecuado de las tecnologías emergentes en procesos de aprendizaje.
- Implementar un prototipo de Repositorio para la gestión y preservación de los REA indexado por competencias enmarcado en una estrategia institucional de integración de sistemas y convergencia de servicios en el entorno digital.

A la fecha de presentación de este artículo se ha transitado el primer año del proyecto. A continuación se describen los resultados y acciones más relevantes.

- Se construyó un Modelo de proceso de desarrollo para Serious Games (MPDSG). El mismo está descrito en [15].
- En consonancia con el Modelo MPDSG se está desarrollando el Serious Game “Power Down the Zombies”. El objetivo educativo del videojuego es mejorar la toma de decisiones relativas al uso racional, eficiente y consciente de la energía; como así también poder evaluar los impactos medioambientales y sociales de los usos tecnológicos de la energía y reflexionar críticamente sobre el uso que debe hacerse de los recursos naturales. Está destinado a estudiantes de 4to año de escuelas secundarias. Los procesos de Elicitación y Especificación fueron descritos en [16] y [17].
- Como aportes a la producción de REA se ha impulsado una iniciativa de producción de Videojuegos Educativos con la participación de estudiantes de ingeniería. Estas actividades resultan ser además una instancia de aprendizaje y desarrollo de competencias profesionales: contribución a la generación de desarrollos tecnológicos y/o innovaciones tecnológicas, colaboración, planeación de proyectos, comunicación, toma de responsabilidades y roles dentro del equipo de trabajo y manejo del tiempo [18].
- Se dictó el taller “Videojuegos y aprendizaje” en varias instancias a docentes municipales de Mar del Plata y en el XI Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2016). Los destinatarios fueron docentes y estudiantes involucrados con las TIC. La descripción completa de esta actividad puede verse en [19].
- En el mes de agosto de 2016 organizamos el 1er Festival Creativo MDQ. El Festival estaba conformado por un conjunto de eventos en donde el eje es la industria del

videojuego y la educación. Este primer festival tuvo como objetivo principal pensar de forma distinta la labor docente, enfocándonos en el juego y en particular el videojuego. Se conformó como un espacio para crear redes entre educadores, desarrolladores, estudiantes, artistas, músicos, y otras personas que tienen relación de una u otra forma en la industria del gaming. El Festival estuvo coorganizado por nuestro Grupo de Investigación GTI y el Grupo de Investigación Humanidades Digitales - Tecnología - Gamificación 2.0" (HUM 2.0) de la Facultad de Humanidades.

- Se desarrollaron experiencias áulicas de inclusión de videojuegos en el aula con ingresos a la Facultad de Ingeniería y en varios cursos de una escuela de nivel medio. Se diseñaron e implementaron secuencias didácticas con el objeto de incorporar el uso de un videojuego como eje central de una dinámica de aprendizaje [20] y [21].

Formación de recursos humanos

El equipo de trabajo del proyecto está integrado por especialistas en Ingeniería de Software y Tecnología Informática aplicada a la Educación. En lo disciplinar participan varios profesionales de la Ingeniería y profesores de Ciencias Básicas.

Varios integrantes acreditan estudios de posgrado: una tesis doctoral: Doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP [22] y dos tesis de maestría en Tecnología Informática Aplicada en Educación [23], [24].

Integrantes del proyecto se encuentran desarrollando dos tesis de posgrado de la UNLP en el marco del proyecto de investigación, correspondientes a la Maestría en Ingeniería de Software y una del Doctorado en Educación de la Universidad Nacional de Rosario.

Se realizaron además numerosas actividades de transferencia referidas a la gestión y asesoramiento y capacitación en los temas relativos

al proyecto instaurando esta línea de investigación en la Facultad de Ingeniería con la posibilidad de formar nuevos recursos humanos (becarios y tesistas).

Referencias Bibliográficas

- [1] Medina, A. (2001). Los métodos en la enseñanza universitaria. En A. García-Valcárcel, (Comp.) *Didáctica universitaria* (pp. 155-198). Madrid: La Muralla.
- [2] Fullan, M. (2007). *Las fuerzas del cambio con creces*. Madrid: Akal.
- [3] Veletsianos, G. (2011). Designing opportunities for transformation with emerging technologies. *Educational Technology*, 51(2), 6.
- [4] Ramírez Montoya, M. (2014). *Tecnologías emergentes en el movimiento educativo abierto*. Monterrey, México: Tecnológico de Monterrey.
- [5] Gros, B. (2015). La caída de los muros del conocimiento en la sociedad digital y las pedagogías emergentes. *Education in the Knowledge Society*, 16(1), 58-68. <http://dx.doi.org/10.14201/eks20151615868>
- [6] Ramírez Montoya, M. y García-Peñalvo, F.J., 2015. *Movimiento Educativo Abierto*. *Virtualis* 6, 12, 1-13.
- [7] Ferreras-Fernández, T., & Merlo-Vega, J. A. (2015). Repositorios de acceso abierto: un nuevo modelo de comunicación científica. *La Revista de la Sociedad ORL CLCR en el repositorio Gredos. Rev. Soc. Otorrinolaringol. Castilla Leon Cantab. La Rioja*, 6(12), 94 -113.
- [8] Cohen, L., Manion, L., y Morrison, K. (2004). *A Guide to Teaching Practice* (5ta ed.). Londres: Routledge.
- [9] Johnson, L., Adams Becker, S., Gago, D. Garcia, E. y Martín, S. (2013). *NMC Perspectivas Tecnológicas: Educación Superior en América Latina 2013-2018. Un Análisis Regional del Informe Horizon del NMC*. Austin, Texas: The New Media Consortium www.nmc.org/pdf/2013-technology-outlook-latin-america-ES.pdf.
- [10] Baeten, M., Dochy, F., Struyven, K., Parmentier, E. y Vanderbruggen, A. (2016). Student-centred learning environments: An investigation into student teacher's instructional preferences and

approaches to learning. *Learning Environments Research*. 19(1), 43-62. doi: 10.1007/s10984-015-9190-5

[11] Beck, J. & Wade, M. (2006). *The Kids are alright: how the Gamer Generation is Changing the Workplace*. Boston: Harvard Business School Press.

[12] del Moral Pérez, M.E. (2013). *Advergaming & edutainment: fórmulas creativas para aprender jugando*. Ponencia inaugural del Congreso Internacional de Videojuegos y Educación (CIVE, 2013). Cáceres, España.

[13] Iuppa, N. & Borst, T. (2006). *Story and simulations for serious games: Tales from the trenches*. United States: Focal Press, Taylor & Francis Group.

[14] Cabero, J. (2007). *Las necesidades de las TIC en el ámbito educativo: oportunidades, riesgos y necesidades*. *Tecnología y Comunicación Educativas* Año 21, No. 45.

[15] Evans, F., Spinelli, A.; Zapirain, E., Massa, S. M. Soriano, F. (2016). *Proceso de desarrollo de serious games. Diseño centrado en el usuario, jugabilidad e inmersión*. III Congreso Argentina de Ingeniería (CADI 2016), IX Congreso Argentino de Enseñanza de la Ingeniería (CAEDI 2016). Capítulo 1. Universidad Tecnológica Nacional - Facultad Regional Resistencia. Universidad Nacional del Nordeste. Chaco, Argentina. ISBN 978-950-42-0173-1.

[16] Spinelli, A., Massa, S. M., Evans, F. (2016). *El proceso de creación de un videojuego como herramienta para la toma de decisiones en el uso eficiente de la energía*. II Congreso Argentino de Energías Sustentables. Facultad Regional Bahía Blanca, Universidad Tecnológica Nacional. Universidad Nacional del Sur. Universidad Nacional del Comahue. Bahía Blanca. ISBN 978-987-1896-62-2

[17] Spinelli, A.; Massa, S. M., Zapirain, E. (2016). *La construcción narrativa de un Serious Game*. IV Congreso Internacional Videojuego y Educación (CIVE 2016). Universidad de VIGO, VIGO, España.

[18] Spinelli, A.; Massa S. M.; Zapirain, E.; Kühn, F.; Rico, C. (2016). *Elicitación de requerimientos para un Serious Game*. II Jornadas Argentinas de Tecnología, Innovación y Creatividad. Universidad Caece. Mar del Plata. Argentina.

[19] Hinojal, H., Morcela, A., Petrillo, J. (2016). *Evidencias de aplicabilidad de videojuegos comerciales en educación*. II Jornadas Argentinas de Tecnología, Innovación y Creatividad. Universidad Caece. Mar del Plata. Argentina.

[20] Farías, Y., Fornasier, M., Moro, L., Morcela, A. (2016). *Aprender energía a partir de un videojuego*. III Jornadas de Investigación en Educación. Facultad de Humanidades. UNMDP.

[21] Morcela, A., Farias, Y. (2016). *Aprendizaje en contexto 3.0*. II Jornadas Argentinas de Tecnología, Innovación y Creatividad. Universidad Caece. Mar del Plata. Argentina.

[22] Massa, S. M. (2013). *Objetos de aprendizaje: Metodología de desarrollo y Evaluación de la calidad*. Tesis Doctoral. Facultad de Informática. UNLP. En Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/26207>

[23] Bacino, G. (2015). *Aula extendida en la educación superior en Ingeniería. Una propuesta de aplicación en el área tecnológica básica de electrotecnia*. Tesis de Maestría. Facultad de Informática. UNLP. Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/45476>

[24] Revuelta, M. (2016). *Laboratorio Remoto en un Entorno Virtual de Enseñanza Aprendizaje*. Tesis de Maestría. Facultad de Informática. UNLP. Repositorio Institucional de la UNLP. <http://sedici.unlp.edu.ar/handle/10915/55888>

Tecnologías de Código Abierto para la Enseñanza de la Ingeniería: Herramientas de Vigilancia Tecnológica

Guagliano Miguel, Tornillo Julián, Carroso Lucas. Pavlicevic Juan.
 Facultad de Ingeniería - Universidad Nacional de Lomas de Zamora
 Instituto de Investigaciones de Tecnología y Educación (IIT&E)
 Camino de Cintura y Avda. Juan XXIII –Llavallol
 Teléfono: 011 – 4282-7880

institutoiite@gmail.com; ing.guaglianom@gmail.com; julianeloytornillo@gmail.com;
lucas.carroso@gmail.com; jpavlicevic@gmail.com

Resumen

En estos tiempos, la Universidad se ha convertido en una organización compleja y, por ello, debe definir estrategias orientadas a mejorar los procesos de decisiones en búsqueda de nuevas oportunidades y detección de amenazas de manera anticipada y así generar procesos de transformación de los datos en información y de ésta en conocimiento. En este sentido, la implementación de herramientas de Vigilancia Tecnológica y la Inteligencia Competitiva (VTeIC) en Universidades permite a dichas Unidades Académicas acceder de manera eficiente a información de calidad respecto a las fronteras del conocimiento regional, nacional e internacional de una tecnología o sector; y además optimizar los recursos de investigación.

El presente artículo tiene como finalidad exponer los pasos metodológicos para, mediante las herramientas de VTeIC, conocer el estado del arte de las aplicaciones de enseñanza en el nivel universitario en ingeniería, con el fin de llegar a la adopción y adaptación de softwares de código abierto aplicables a la enseñanza.

A partir de este estudio, se espera proveer a la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora (FI-UNLZ) alternativas en softwares de código abierto, que contribuyan a elevar la calidad educativa en el entorno virtual de la institución.

Palabras clave: Tecnología, Código Abierto, Vigilancia Tecnológica, Inteligencia Competitiva, Enseñanza, Ingeniería, Universidad

Contexto

El Instituto de Investigaciones en Tecnología y Educación (IIT&E) de LA FI-UNLZ trabaja desde el año 2006 en actividades de Investigación y Desarrollo orientadas a contribuir al mejoramiento de la educación formal y profesional, mediante la incorporación de tecnologías de información y comunicación (TIC) en los procesos de gestión, enseñanza y aprendizaje. En este sentido, el crecimiento ha sido acelerado, creciente e ininterrumpido.

En el año 2016, en el marco del Programa “Enseñanza de la ingeniería: Desarrollo y Evaluación de modelos, estrategias y tecnologías para mejorar los indicadores académicos y la eficiencia organizacional” se pone en marcha el Proyecto “Desarrollar un Programa de Vigilancia e Inteligencia Estratégica (VeIE) en la FI-UNLZ orientado a mejorar las políticas estratégicas que impacten en los procesos de educación, investigación y transferencia en el ámbito de la enseñanza superior”, del cual se desprende esta publicación.

Actualmente el IIT&E desarrolla sus actividades como Centro Asociado a la Comisión de Investigaciones Científicas (CIC), el cual depende del Ministerio de Ciencia Tecnología e Innovación de la Provincia de Buenos Aires. Esto impacta

positivamente en los resultados que obtiene el instituto ya que favorece a la investigación, el desarrollo y la transferencia de los resultados a la región de influencia de la Universidad.

Introducción

De acuerdo a Minnaard et al. (2015), el volumen importante de información que hoy en día presenta un fácil acceso a la misma y su velocidad de tránsito, plantea retos importantes a la gestión diaria en todas las organizaciones. Entendiéndose organización, como *“formaciones complejas y plurales compuestas por individuos y grupos que constituyen un sistema de roles coordinados mediante un sistema de autoridad y comunicación articulado por un sistema de significados compartidos”*. Por lo tanto, las estrategias orientadas a mejorar los procesos de decisiones en búsqueda de nuevas oportunidades y detección de amenazas de manera anticipada, tienen que estar alineadas y dentro de la visión de los procesos de transformación de los datos en información y de ésta en conocimiento con valor agregado.

De acuerdo con la norma UNE 166006:2011 (2011), la Vigilancia Tecnológica es una herramienta fundamental en el marco de los sistemas de gestión de I+D+i y se define como *“el proceso organizado, selectivo y sistemático, para captar información del exterior y de la propia organización sobre ciencia y tecnología, seleccionarla, analizarla, difundirla y comunicarla, para convertirla en conocimiento con el fin de tomar decisiones con menor riesgo y poder anticiparse a los cambios”*. La Inteligencia Competitiva comprende, además, el análisis, la interpretación y comunicación de información de valor estratégico acerca del ambiente de negocios, de los competidores y de la propia organización, que se transmite a los responsables de la toma de decisiones, como elemento de apoyo para ajustar el rumbo y marcar posibles caminos de evolución, de interés para la organización, haciendo foco en el proceso de obtención, análisis, interpretación y difusión de información de valor estratégico sobre la

industria y los competidores (Gibbons y Prescott, 1996).

La utilización de las tecnologías de la información y la comunicación (TIC) a nivel mundial, ha surgido y adquirido un rol central en los últimos años, dando lugar a la aparición de nuevas temáticas tales como la Vigilancia e Inteligencia (Escorsa, P. Maspons, R., 2001).

La gestión de estas temáticas plantea el desafío de lograr anticiparse a las innovaciones tecnológicas, lo que obliga a las organizaciones a estar monitoreando permanentemente acerca de las limitaciones y las nuevas oportunidades que pueden coexistir en su entorno respecto de la evolución de la ciencia y la tecnología en términos globales y la dinámica del entorno que la acompaña.

Así, la Vigilancia Tecnológica es una herramienta que ayuda a comprender y explicar la evolución de la tecnología y permite a las instituciones y organizaciones anticiparse a los efectos negativos que sobre su actividad pueden tener, y aprovechar las oportunidades que se pueden llegar a presentar, ayudando a la identificación de los escenarios más probables y al estudio del impacto previsible sobre la actividad de la empresa que dichas tecnologías emergentes pueden provocar (Guagliano M, 2014).

El software es cada vez más el gran intermediario entre la información y la inteligencia humana. (Hernández, J. M. 2005). En particular, los softwares libres se basan en la cooperación y la transparencia, y garantizan una serie de libertades a los usuarios que son de especial interés para los trabajos que se dan en espacios colaborativos.

Como indica Hernández (2005), *el software libre no tiene prácticamente coste de licencia.* Es decir que suponen un costo mucho menor a los softwares privados o propietarios. Este ahorro resulta atractivo para las organizaciones en general, y las universidades en particular, ya que brinda la posibilidad de destinar esta porción del presupuesto para otros fines, tales como infraestructura IT u

otros recursos que favorezcan el aprendizaje mediante estos softwares.

Para Stallman (2005), *un software libre es un asunto de libertad, no de precio* (Stallman R., 2004), e indica que para que un software sea libre, el usuario debe gozar de las siguientes cuatro libertades:

1. *La Libertad Cero es la libertad de ejecutar el programa con cualquier propósito, de la forma que quieras.*
2. *La Libertad Uno es la libertad de ayudarte a ti mismo cambiando el programa para que se ajuste a tus necesidades.*
3. *La Libertad Dos es la libertad de ayudar al prójimo distribuyendo copias del programa.*
4. *Y la Libertad Tres es la libertad de ayudar a construir tu comunidad publicando una versión mejorada de modo que los otros puedan beneficiarse de tu trabajo.*

De esta manera, el software libre les otorga la libertad a los usuarios para ejecutarlo, copiarlo, modificarlo y distribuirlo; ya sea con o sin modificaciones.

El término “software libre” (en inglés “free software”) ha provocado confusiones debido a la doble aceptación que en inglés tiene la palabra “free” (puede significar tanto gratuito como libre). Es por ello que, como indica HERNÁNDEZ, en Estados Unidos se comenzó a emplear principalmente el término “código abierto” (en inglés “Open Source”) a fin de evitar dar la percepción de que es un recurso totalmente gratuito pero manteniendo el valor central del mismo ya que el código fuente está disponible.

Por otra parte, la universidad tiene una responsabilidad y unos intereses mucho más amplios que una empresa o un particular cuando adquiere tecnologías de la información (García A. y Cuello R., 2007). En este sentido, la universidad tiene la obligación de utilizar el software – en este caso de código abierto- de manera

responsable y para fines que contribuyan a la comunidad o a la sociedad en su conjunto.

García y Cuello también indican que la Universidad debe exigir 3 condiciones al software a incorporar:

- Debe operar en la lengua propia del lugar en que se va a utilizar (localización).
- Debe poder garantizar el acceso a la información en todo momento, en el presente y en el futuro (perennidad).
- No ha de permitir que personas no autorizadas tengan acceso a los datos confidenciales de los particulares o a información reservada (seguridad).

De acuerdo con lo que concluye Amatriain, *en un tema de tanta trascendencia, no podemos permitir que las instituciones educativas y las universidades esperen a subir en el vagón de cola. Hemos de tomar posición y situarnos en primera línea de esta revolución* (Amatriain X., 2004). Es por ello que resulta de gran importancia que las universidades sean uno de los principales usuarios de softwares de código abierto, ya sea para incorporarlos a los procesos de enseñanza, como también para los proyectos de investigación.

Tal como se ha mencionado, la Universidad es una organización compleja y resulta indispensable anticiparse a los escenarios más probables en términos de educación, investigación y transferencia.

Teniendo en cuenta todo lo mencionado, la Facultad de Ingeniería se vio en la necesidad de contar con nuevas soluciones de herramienta de softwares de código abierto que contribuyan a elevar la calidad educativa en el entorno virtual de la institución. Es por eso, que resulta muy importante establecer una metodología que permita de manera eficaz y eficiente alcanzar los objetivos del presente trabajo.

Líneas de Investigación, Desarrollo e Innovación

El presente artículo presenta los pasos de la metodología que permitirá, a través de la Vigilancia Tecnológica e Inteligencia Competitiva, poder identificar el Estado del arte de las aplicaciones de enseñanza en el nivel universitario en ingeniería, con el fin de llegar a la adopción y adaptación de softwares de código abierto aplicables a la enseñanza.

El ámbito de la experimentación es el IITE - perteneciente a la FI UNLZ- donde funciona un gabinete que cuenta con equipo informático y mobiliario adecuado, que permite el desarrollo de todas las tareas.

Se facilitará el desarrollo del trabajo con la implementación de herramientas de vigilancia e inteligencia, como pueden ser: buscadores especializados, metabuscadores, multibuscadores, plataformas de vigilancia, bases de datos de acceso libre, entre otros.

El trabajo se inscribe en un diseño metodológico de características experimentales ya que se exploran las potencialidades de los softwares propuestos.

De acuerdo a lo antes mencionado, se propone:

1. Relevar las necesidades y requerimientos de las materias de las carreras de la FI-UNLZ mediante una encuesta ad-hoc.
2. Construir las sentencias de búsqueda que permitan aplicar las herramientas de Vigilancia Tecnológica e Inteligencia Competitiva.
3. Mediante técnicas de Vigilancia Tecnológica relevar los softwares de código abierto volcando los resultados obtenidos en una planilla indicando fortalezas y debilidades.
4. Evaluar algunos de los softwares recolectados mediante técnicas y herramientas de Inteligencia Competitiva, volcando los resultados obtenidos en una planilla indicando fortalezas y debilidades.

Resultados y Objetivos

La metodología propuesta y su implementación, son un factor clave para mejorar e incrementar los niveles de calidad educativos en la institución estudiada.

Los objetivos específicos del trabajo son:

1. Determinar los aspectos metodológicos y pedagógicos fundamentales de las materias de las carreras de la FI-UNLZ.
2. Identificar las palabras clave y construir las sentencias de búsqueda para aplicar las herramientas de Vigilancia Tecnológica e Inteligencia Estratégica.
3. Relevar los softwares de código abierto mediante Vigilancia Tecnológica.
4. Participar en la evaluación y el análisis de los softwares recolectados mediante técnicas y herramientas de Inteligencia Competitiva.

Formación de Recursos Humanos

La realización del trabajo implica formar y consolidar un equipo de trabajo multidisciplinario para llevar a cabo las etapas de la metodología propuesta.

El Proyecto “Desarrollar un Programa de Vigilancia e Inteligencia Estratégica (VeIE) en la FI-UNLZ orientado a mejorar las políticas estratégicas que impacten en los procesos de educación, investigación y transferencia en el ámbito de la enseñanza superior” se encuentra bajo la dirección Mg. Ing. Juan Pavlicevic y la co-dirección de Dr. Marta Comoglio. Además, integran el equipo de investigación 2 doctorandos y un becario que es estudiante de la carrera de Ingeniería Industrial.

El programa al que pertenece cuenta con docentes investigadores, personal técnico, personal de apoyo y becarios (Becas CIN y Becas LomasCYT). Todos los Recursos Humanos tienen una relación constante mediante actividades de docencia o

transferencia. Asimismo, se puede garantizar que este proyecto favorecerá a todas las actividades académicas que se realizan en la institución y también a las actividades de Gestión Académica.

Referencias

- Amatriain, X. (2004). El software libre en la educación: guía para su justificación e implementación. Actas de las III Jornadas de Software Libre. Escuela politécnica Superior de Ingeniería, Universitat politécnica de Catalunya, URL: <http://portal.jornadespl.org/biblioteca/iii-jornades>.
- Asociación Española de Normalización y Certificación AENOR (2011). Gestión de la I+D+i: Sistema de vigilancia tecnológica e inteligencia competitiva. Norma UNE 166006. Madrid, España.
- Escorsa, P. and Maspons R. (2001). De la Vigilancia Tecnológica a la Inteligencia Competitiva. España, FT-Prentice Hall, Pearson.
- García, A. M. D., & Cuello, R. O. (2007). La promoción del uso del software libre por parte de las universidades. Revista de Educación a Distancia, (17).
- Gibbons, PT and Prescott, JE. (1996). Parallel competitive intelligence processes in organizations. International Journal of Technology, Special Issue on Informal Information Flow Management, 11, pp 162-178.
- Guagliano M (2014). “Desarrollo Metodológico para la Generación de Productos de Vigilancia Tecnológica e Inteligencia Estratégica del Sector Autopartista”. Trabajo final de Tesis de Especialización en Gestión de la Innovación, UNLZ Argentina.
- Hernández, J. M. (2005). Software libre: técnicamente viable, económicamente sostenible y socialmente justo. Infonomía.
- Minnaard et al. (2015). La información y la tecnología para la toma de decisiones: aplicación Data WareHouse en instituciones universitarias. In XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015).

- Stallman, R. (2004). Software libre para una sociedad libre. Madrid: Traficantes de Sueños, 2004.

Tecnologías Innovadoras como Mediadoras de Procesos de Enseñanza y Aprendizaje. Exploración de Herramientas de Realidad Aumentada

Cecilia Sanz²; Tatiana Gibelli¹; Edith Lovos¹; Paula Suárez¹; Álvaro Saldivia; Sergio Condó; Romina Cariaga; María Luján Colueque, Verónica Cuevas³

¹Centro Interdisciplinario en Derechos, Inclusión y Sociedad (CIEDIS)
Sede Atlántica, Universidad Nacional de Río Negro,

² Investigador asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Instituto de Investigación en Informática LIDI (III-LIDI), Facultad de Informática, Universidad Nacional de la Plata

³ Centro Universitario Zona Atlántica (CURZA)
Universidad Nacional del Comahue

csanz@lidi.info.unlp.edu.ar - {tgibelli, elovos, psuarez, scondo, asaldivia, rcariaga, mcolueque}@unrn.edu.ar -vcuevas1976@gmail.com

Resumen

En este trabajo se presentan objetivos y avances preliminares, así como metas de trabajo futuras de un proyecto de investigación acreditado por la Universidad Nacional de Río Negro (UNRN) e iniciado en agosto de 2016. Este proyecto propone un trabajo de investigación-acción, para analizar y conocer el impacto de diferentes tecnologías digitales en contextos educativos específicos. Se busca innovar tanto en metodologías de inclusión como en el tipo de tecnología a utilizar, con el fin de analizar sus posibilidades y barreras en los procesos de enseñanza y aprendizaje. Se indagan tecnologías emergentes como el caso de Realidad aumentada con la intención de diseñar desarrollos y metodologías innovadoras de aplicación en el escenario educativo.

Palabras clave: TIC, mediación, educación, realidad aumentada.

Contexto

Esta investigación se desarrolla en el marco de un proyecto bianual (2016-2018) acreditado y financiado por la Universidad Nacional de Río Negro, titulado: “La mediación de las tecnologías de la información y la comunicación en procesos educativos. Innovaciones para mejorar los procesos de enseñanza y aprendizaje”. El mismo surge como continuación de un proyecto previo (2014-2016) denominado “Tecnologías de la Información y la Comunicación en la Enseñanza y Aprendizaje en Nivel Superior. Habilidades de Autorregulación del Aprendizaje y Trabajo Colaborativo”. Éste resultó fundamental para la conformación de un equipo de investigación en la sede atlántica de la UNRN que aborde la temática de la incorporación de las TIC en los procesos de enseñanza y aprendizaje. En especial nos centramos en ese proyecto en el nivel universitario priorizando el potencial de las TIC para el desarrollo de las

habilidades de autorregulación de los aprendizajes y trabajo colaborativo. Este nuevo proyecto busca explorar tecnologías más emergentes como la realidad aumentada y su uso en diferentes áreas y niveles educativos.

Objetivos

El objetivo principal de este proyecto, es investigar metodologías y estrategias innovadoras que favorezcan procesos educativos mediados por Tecnologías de la Información y de la Comunicación.

Los objetivos específicos son:

- Explorar marcos teóricos que permitan interpretar distintos contextos educativos mediados por TIC.
- Evaluar la mediación de TIC en situaciones educativas concretas.
- Proponer innovaciones tecnológicas y metodológicas para necesidades educativas específicas, en particular, relacionadas con la realidad aumentada y el uso de dispositivos móviles.
- Promover el desarrollo de prácticas docentes innovadoras con apoyo de recursos informáticos.

Se espera que las metodologías y estrategias investigadas permitan potenciar los procesos de enseñanza y aprendizaje y el desarrollo de habilidades (pensamiento crítico, trabajo en equipo, comunicación, etc) necesarias para desarrollarse en la sociedad del conocimiento.

Resultados Preliminares y Esperados

Se inició el proyecto con una primera instancia de revisión bibliográfica y

exploración de la temática por parte de todos los integrantes del equipo de investigación.

Se comenzó con una revisión de teorías que fundamentan el uso de tecnología digital en educación, centrando la atención en la Teoría Cognitiva del Aprendizaje Multimedia (TCAM). Esta teoría, fundada por Richard E. Mayer de la Universidad de California incorpora algunos elementos de las teorías clásicas del procesamiento de información como la Teoría de la Carga Cognitiva de Sweller, la Teoría de la Codificación Dual de Paivio, el Modelo de Memoria de Trabajo de Baddeley y el organigrama para representar la memoria y los procesos cognitivos de Atkinson y Shiffrin y los aplica principalmente a la multimedia (Mayer, 2005). A partir de ella se proponen principios de aplicación que pueden utilizarse para mejorar el diseño de materiales multimediales de acuerdo a la manera como la mente humana está configurada para procesar la información. En general, estos principios se refieren a considerar la manera como se presenta la información gráfica, escrita y auditiva (Andrade-Lotero, 2012). Entendemos que este es un campo que debe seguir profundizando pues las investigaciones que prueban nuevos diseños multimedia han encontrado varias limitaciones referidas, principalmente, a la medición de la carga cognitiva, a la experiencia y conocimiento previo de los sujetos, a la motivación y a otros aspectos afectivos frente al contenido y formato de la instrucción (Artino, 2008).

En relación a las tecnologías incipientes a explorar se decidió centrar la investigación en realidad aumentada (RA). La RA se caracteriza por: (a) una combinación de objetos virtuales y reales en un escenario real, (b) usuarios interactuando en tiempo real y (c) una alineación entre los objetos reales y

virtuales (Azuma et al., 2001). Respecto a su potencialidad en el plano educativo, permite incorporar multimedia al proceso enseñanza aprendizaje, innovar en la práctica docente y promover el diseño de materiales educativos atendiendo a los requerimientos didácticos (Avendaño, 2012). Cabero y Barroso (2015), señalan que para su incorporación en los procesos de enseñanza y aprendizaje es necesario tener en cuenta ciertos principios, entre ellos: el diseño de los entornos debe ser flexible de manera que su inclusión en el espacio de enseñanza y aprendizaje, no se convierta en un problema técnico; la producción de materiales debe tener en cuenta múltiples plataformas y soportes y se hace necesaria la formación de los docentes en competencias didácticas que les permitan incluir las TIC en sus prácticas de enseñanza, posibilitando a través de la RA la creación de “escenografías educativas enriquecedoras” más allá de las cuestiones tecnológicas.

Como experiencia piloto, se llevó adelante una implementación y posterior análisis de una experiencia didáctica usando RA con alumnos de primer año de la Lic. en Sistemas (UNRN) cohorte 2016. La experiencia consistió en el uso de la herramienta EPRA (Salazar Mesía, et. al; 2015) para la apropiación y manejo de conceptos de básicos de algoritmia. Se observaron en esta experiencia tanto aspecto motivacionales y de satisfacción como de rendimiento de los estudiantes.

Por otra parte, se realizó una revisión de experiencias con RA para la enseñanza de la matemática. Bujak y colaboradores (2013) analizan el uso de la RA en matemática desde desde tres perspectivas: física, cognitiva y contextual, presentando para cada una de ellas ejemplos de aplicaciones existentes de RA así como pautas para el diseño de futuras experiencias considerando cuestiones

prácticas y tecnológicas. En la dimensión física destacan que las interacciones naturales con objetos físicos fomentan la comprensión de los conceptos en contexto. En la dimensión cognitiva mencionan que la alineación espacio temporal de la información puede servir como andamiaje en la comprensión de los conceptos abstractos. Finalmente en la dimensión contextual plantean las posibilidades que brinda la RA para el aprendizaje colaborativo en entornos no tradicionales generado experiencias de aprendizaje significativas. Asimismo diversas experiencias que utilizan RA para el aprendizaje en matemática destacan el aumento en la motivación del alumno respecto a otras aplicaciones TIC convencionales. Otro aspecto considerado para aplicar RA en matemática es la usabilidad del aplicativo por parte del alumno, es decir, si es de fácil manejo, intuitivo, diseño con diversos grados de dificultad, etc.

Asimismo se inició una revisión de herramientas que permiten el diseño de actividades con RA por geolocalización usando dispositivos móviles, sin necesidad de contar con conocimientos de programación. Se llevó adelante un análisis comparativo entre dos herramientas que posibilitan el diseño de juegos móviles educativos a través de la RA, permitiendo la creación de actividades en las que existe un recorrido y toma relevancia el posicionamiento del jugador. Los resultados del mismo han sido aprobados para su presentación al Congreso Iberoamericano de Educación y Sociedad (CIEDUC 2017). En base a esta revisión se ha planificado una experiencia didáctica utilizando estas herramientas en una asignatura del primer año de la Lic. en Comunicación Social para el primer cuatrimestre de 2017.

Nos proponemos como objetivo a futuro el diseño, implementación y

evaluación de propuestas educativas con RA para la enseñanza y aprendizaje en particular en el ámbito de la UNRN. En pos de este objetivo se comenzaron a delinear algunas propuestas siguiendo una Guía para el diseño de actividades educativas con RA, desarrollada por la Dra. Sanz, en el marco del proyecto de investigación “Tecnologías para sistemas de software distribuidos. Calidad en sistemas y procesos. Escenarios educativos mediados por TICs” del III-LIDI (UNLP).

Con el propósito de ir planificando la recolección de datos durante la implementación de las propuestas se realizó una revisión de variables e instrumentos que suelen utilizarse para el análisis de experiencias educativas con tecnologías digitales como la realidad aumentada. Uno de los aspectos que se analiza en este tipo de investigaciones es la motivación que es considerada la fuente de energía responsable de que los estudiantes decidan hacer un esfuerzo. Suele utilizarse para evaluar esta variable el modelo ARCS (Keller, 2010) que considera cuatro factores motivacionales: la Atención, Relevancia, Confianza y Satisfacción. Uno de los instrumentos que consideramos de utilidad para la recolección de datos es para esta variable es el cuestionario IMMS tipo Likert con 36 ítems basado en el modelo ARCS según la propuesta de Di Serio, Ibáñez y Delgado (2013), que resulta relevante porque ha sido validado y utilizado en varios estudios de investigación utilizando la tecnología como un factor de motivación en el aprendizaje. Otra de las variables a las que suele atenderse en el diseño de materiales educativos digitales es la usabilidad. Este es un concepto amplio. Bevan (1995) define las componentes de usabilidad: Eficacia (grado de exactitud con que se realizan las tareas cumpliendo los objetivos para

los que está diseñado), Eficiencia (rapidez para realizar las tareas para las que ha sido diseñado) y Satisfacción (cumplimiento de las expectativa para mantener la motivación del usuario). Existen diferentes cuestionarios para medir la usabilidad a partir de estos componentes, en particular se indagó en el utilizado por Sanchez Riera (2013) en su tesis doctoral.

A partir de las experiencias a desarrollar durante este año se prevé la realización de informes técnicos y publicaciones que den cuenta de las metodologías diseñadas, sus fortalezas y debilidades.

Formación de Recursos Humanos

El proyecto de investigación es dirigido por una docente externa a la Universidad Nacional de Río Negro, la Dra Cecilia Sanz, docente de la Facultad de Informática de la Universidad Nacional de la Plata. El equipo de investigadores está formado por: dos docentes con dedicación completa y cinco docentes concurrentes de la UNRN, una docente externa a la institución y dos alumnos de la carrera Lic. en Sistemas de la UNRN.

En cuanto a la formación profesional de los integrantes, la dirección externa con experiencia y antecedentes relevantes en la temática del proyecto constituye un aporte importante a la formación del equipo local. En este sentido está previsto el dictado por parte de la directora del proyecto de un curso de postgrado sobre el tema realidad aumentada, que se llevará adelante en la Sede Atlántica de la UNRN durante el primer cuatrimestre de 2017. El curso está destinado a todos los integrantes del proyecto siendo abierto a otros interesados. Asimismo los integrantes del proyecto realizan otros

cursos de postgrado y actualización vinculadas a la temática.

Por otra parte, se participa de la formación nuevos investigadores. La directora y la codirectora participan en la dirección de tesis de posgrado vinculadas a la temática, y se ha presentado una becaria al programa de becas CIN convocatoria 2016. Como resultado de este proyecto se espera crear posibilidades para la dirección conjunta de otras tesis.

Referencias

- Andrade-Lotero, L. A. (2012). Teoría de la carga cognitiva, diseño multimedia y aprendizaje: un estado del arte. *Magis. Revista Internacional de Investigación en Educación*, 5(10).
- Artino, A. R. (2008). Cognitive Load Theory and the Role of Learner Experience: An Abbreviated Review for Educational Practitioners. *Association for the Advancement of Computing In Education Journal, AACE Journal*, 16 (4), 425-439.
- Avendaño, V. y Domínguez, L. A. (2012). Realidad aumentada: Una exploración al escenario de la virtualidad educativa. Madrid: Editorial Académica Española.
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6), 34-47.
- Bujak, K. R., Radu, I., Catrambone, R., Macintyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*, 68, 536-544.
- Cabero, J. y Barroso, J. (2015). Realidad Aumentada: posibilidades educativas. En Ruiz-Palmero, J., Sánchez-Rodríguez, J. y Sánchez-Rivas, E. (Edit.). *Innovaciones con tecnologías emergentes*. Málaga: Universidad de Málaga.
- Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Di Serio, A., Ibáñez, B. y Delgado, C. (2013). Impact of an Augmented Reality System on Students Motivation for a Visual Art Course. *Computers & Education*, 68, 586-596.
- Keller, J. M. (2010). Motivational design for learning and performance. *Science & Business Media*. 227–286, Springer.
- Salazar Mesía, N., Gorga, G., & Sanz, C. V. (2015). EPRA: Herramienta para la Enseñanza de conceptos básicos de programación utilizando realidad aumentada. In *X Congreso sobre Tecnología en Educación & Educación en Tecnología (TE & ET)(Corrientes, 2015)*.
- Sánchez Riera, A. (2013). Evaluación de la tecnología de realidad aumentada móvil en entornos educativos del ámbito de la arquitectura y la edificación. Tesis doctoral. Universidad Politécnica de Catalunya. Barcelona.

TIC Modelos y Objetos de Enseñanza su Aplicación en Carreras Tecnológicas de Nivel Superior

Hidalgo Sebastián^{1,2}, Ivañe Flavia Camila², Moret Pablo²

1 - Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora.
Instituto de Investigaciones en Tecnología y Educación (IIT&E)
Camino de Cintura y Av. Juan XXIII – Llavallol.
Teléfono: 011-4282-7880

2 - Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora.
Laboratorio de Modelado y Simulación
Camino de Cintura y Av. Juan XXIII – Llavallol.
Teléfono: 011-4282-7880

Resumen

Se buscará mostrar el vínculo existente entre el uso de tecnologías informáticas y de cómputo para el estudio de casos que los estudiantes pueden hallar en la vida real, para comprender cómo el modelado de situaciones, manejando este tipo de herramientas, puede permitir el aprendizaje en la carrera y dar lugar al conocimiento de dichas herramientas para su uso en la vida profesional.

En el ámbito de la Logística abundan desarrollos de software a medida y especializados, los cuales son utilizados en el ámbito empresarial, pero que resultan de carácter específicos, no aplicables al ámbito académico. En el laboratorio de Simulación de la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora, se cuenta con herramientas de simulación versátiles, que permiten ser adaptadas al diseño de casos de simulación de procesos productivos, a su vez se ha llevado a cabo un relevamiento exhaustivo de las necesidades del empleo de estas herramientas en las distintas cátedras de la facultad de ingeniería, del cual se detectó la necesidad de diseñar un caso de estudio para la cátedra de Logística, adaptando las

herramientas disponibles y permitiendo el empleo de las mismas para el entendimiento y la aplicación de los elementos percibidos en el entorno real desde un ambiente simulado, que permite detectar problemas, solucionarlos y medirlos, en poco tiempo y sin comprometer recursos de elevado valor.

Palabras Clave: Gestión Logística, Simulación de procesos industriales, Optimización de recursos.

Contexto

Para contextualizar el contenido de este trabajo debemos comprender a qué nos referimos cuando hablamos de Simular un entorno industrial, asociado a procesos logísticos y por qué buscamos la bajada al entorno académico. En principio una simulación será todo aquel modelo extraído de la realidad que permite la observación de variables en un determinado tiempo acotado y que permite la medición de otras variables que serán el out-put de nuestro modelo. Estas variables de salida deberán ser representativas de lo que se quiere medir, directa o indirectamente y su grado de error será tanto mayor en virtud de la complejidad del modelo y la expertiz del desarrollador de dicho modelo. Una simulación de un

proceso logístico englobará la anterior definición a un proceso en el cual intervengan variables propias de dichos entornos, entendiendo a la administración logística como el estudio y asignación de recursos para garantizar, entre otras cosas, la respuesta a los requerimientos de abastecimiento dentro de una cadena suministro en cantidad, calidad y tiempo establecidos. Llevado a un entorno académico requerirá adaptar las herramientas disponibles para que estudiantes de carreras técnicas puedan comprender aspectos de la realidad de la temática tratada, definir las variables de un modelo y obtener resultados concluyentes en un lapso de tiempo adaptado a los plazos normales para el estudio de dicha temática a lo largo de la carrera cursada. Dentro del material consultado en torno a trabajos de carácter semejante podemos citar el trabajo de Juan Carlos Romero Tinjaca¹: “Cuando una simulación se convierte en una herramienta de educación directa, sirve de gran ayuda a los estudiantes pues se establece una relación entre los procesos reales y un proceso abstracto ya que se puede obtener resultados sin producir ningún inconveniente y ayudará a la toma de decisiones ante el proceso que se simula.”

Introducción

La simulación de procesos industriales requiere de una metodología procedimental para poder ser llevada a cabo, que requerirá de un análisis inicial de la necesidad de efectuar dicha simulación y cuáles son los

resultados buscados, en otras palabras, surge de la necesidad de medir algo contrastable con una posible realidad y extraer conclusiones, normalmente, información documentada validando la tarea realizada. Este proceso puede ir acompañado de un entorno virtual para que el usuario pueda hallar similitud con la realidad, por ejemplo, un escenario donde se pueda representar el lay-out de un proceso industrial determinado, un modelo de este procedimiento se detalla en la *Figura 1*. Dentro del entorno de aprendizaje, el contacto con herramientas que se valgan de la manipulación de la información para obtener datos relevantes es de gran importancia para el aprendizaje de la interpretación de la información, a la vez, que permite al alumno entrar en contacto con la mencionada metodología de trabajo. En este punto podemos citar el trabajo de Rosa Imelda García Chi³, en donde se afirma que “El aprendizaje basado en competencias que se establece en los programas de estudio de los Institutos Tecnológicos permite el diseño de estrategias de aprendizaje, las cuales se incorporan en la Instrumentación didáctica de la planeación académica que realiza el docente, donde se incorpora el uso de tecnologías de información, que promueve en los estudiantes una motivación e involucramiento en el aprendizaje significativo de la Simulación de Procesos y Servicios, así

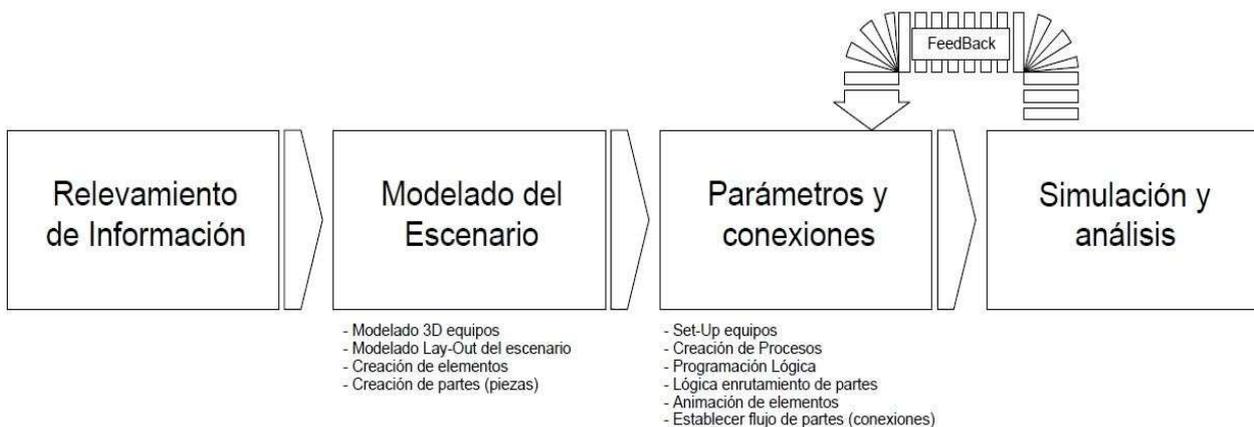


Figura 1 –Método de desarrollo de modelos de simulación de procesos industriales - Fuente: Elaboración propia

como en el desarrollo de habilidades instrumentales como lo es el uso de una herramientas de software para la toma de decisiones”.

Para el desarrollo de este trabajo, partimos de la necesidad de hallar una disciplina que se ajustara a la posibilidad de elaborar sistemas interconectados que permitieran la medición de parámetros, a su vez, hallar tal disciplina que tuviera poco desarrollo en el ámbito académico. Los procesos logísticos cumplen, en definitiva con la lógica de los procesos industriales, de una manera especial, pero representables en modelos computarizados, dado que, podemos observar que:

- ✓ Existen elementos de entrada que serán “transformados” a nuevos elementos con un valor agregado. Encaja en la definición de proceso.
- ✓ Observamos existencias a lo largo de los procesos que pueden cuantificarse y optimizarse.
- ✓ Existen recursos comprometidos para poder lograr los procesos satisfactoriamente. (humanos, maquinaria, energía, etc.)
- ✓ Rigen los principios de la economía, dado que si hay recursos comprometidos, estos no son ilimitados.
- ✓ Existen transporte de materiales que pueden representarse en escenarios computarizados.
- ✓ Pueden medirse tiempos de procesos.
- ✓ Puede determinarse la calidad de los procesos a partir de las métricas de tiempo, volumen, recursos utilizados y cumplimiento de objetivos.

Realizado este relevamiento se procedió a la realización de la investigación para la adaptación de las herramientas de simulación en el entorno académico a dicha disciplina.

Líneas de Investigación y Desarrollo

La investigación desarrollado parte de dos hipótesis principales que definen el camino a seguir para el relevamiento de la información disponible, las necesidades percibidas y el posterior desarrollo del caso siguiendo las directivas antes expuestas. Por lo tanto podemos establecer que:

- ✓ El campo de los procesos logísticos ha sido poco explorado en términos de la simulación industrial aplicado al entorno académico propiamente dicho.
- ✓ El uso de tales herramientas en el entorno de la educación permitirá formar competencias y brindar un pensamiento pragmático al alumno que interactúa con un modelo semejante.

En relación a nuestra segunda hipótesis podemos citar el trabajo de G.M. Rodrigo Andrés⁴: “De las diferentes técnicas didácticas la simulación permite representar o imitar procesos o productos buscando la creación de escenarios que faciliten el aprendizaje, el desarrollo o fortalecimiento de competencias de los estudiantes o personal que desempeña procesos de planeación, ejecución y control de procesos logísticos. En la simulación de procesos logísticos puede utilizarse laboratorios, infraestructura o Tecnologías de Información o software de modelamiento de simulación discreta.”

En primer lugar la detección y análisis de herramientas ya existentes del entorno académico vinculadas a la temática en cuestión permitirán focalizar los esfuerzos de la investigación en vía de ampliar el campo de los recursos existentes para el desarrollo académico de los alumnos. En segundo lugar tales recursos deberán estar preparados para que exista un valor agregado en la formación del pensamiento crítico.

Se puede afirmar entonces que estamos frente a una investigación que implica el

relevamiento de elementos del entorno profesional ya existentes, su complejidad y aplicabilidad al entorno académico, una posterior contrastación con los recursos dentro del mismo entorno donde se quiere aplicar los elementos relevados y un proceso de adaptación o correspondencia entre lo observado y las necesidades percibidas en conjunción con docentes de las cátedras que son alcanzadas por la investigación en cuestión. Finalmente el producto de la investigación será evaluado en forma experimental e incrementado su valor con las sucesivas experiencias prácticas que realizadas en el ámbito académico.

Resultados y Objetivos

Los objetivos perseguidos dentro del marco del siguiente trabajo serán:

- ✓ Relevar los medios y herramientas de aprendizaje práctico dentro del entorno académico vinculados al uso de tecnologías de la información y herramientas de cómputo asociadas al marco de procesos logísticos.
- ✓ Detectar elementos del ámbito profesional que puedan adaptarse a un entorno de aprendizaje.
- ✓ Vincular estos elementos a las herramientas disponibles y generar medios de aprendizaje práctico que permitan incrementar los saberes desde un punto de vista pragmático.
- ✓ Evaluar y documentar los resultados obtenidos.

Los resultados serán medidos y estudiados como el valor percibido dentro del entorno académico de los elementos de aprendizaje generados.

Formación de Recursos Humanos

El presente trabajo se enmarca en el proyecto de Investigación “Innovaciones en los Procesos de Enseñanza en Carreras

Científico Tecnológicas: Impacto en la percepción de la calidad, en los Indicadores Académicos y en las competencias de Egreso”, que se integra en el programa “Enseñanza de la Ingeniería. Desarrollo y Evaluación de Modelos Estrategias y Tecnologías para Mejorar los indicadores Académicos y la Eficiencia Organizacional, junto a otros tres proyectos. Los mismos cuentan con aval institucional y participación en el Programa de Incentivos de la Secretaría de Políticas Universitarias. En el ámbito del programa participan, 22 docentes investigadores, de los cuales, 6 son doctorandos¹, 6 becarios CIN y 3 alumnos avanzados de la carrera de Ingeniería, de los cuales 2 Becas CIN y un alumno pasante se encuentran afectados concretamente al proyecto en el que se inscriben las actividades que se presentan.

Referencias

1. Juan Carlos Romero Tinjaca. (2010). Propuesta de diseño de una herramienta en simulación virtual en almacenamiento logístico. Corporación Universitaria Minuto de Dios, Facultad de Ingeniería. SOACHA. [Consultado el 12/09/16]
2. Eduin Dionisio Contreras Castañeda, Julián David Silva Rodríguez. (2014). Logística inversa usando simulación en la recolección de envases de plaguicidas: estado del arte. Revista Ingeniería Industrial, 33-50, 18. [Consultado el 19/09/16]
3. Rosa Imelda García Chi, Arturo Eguia Alvarez, Gloria Emilde Izaguirre Cardenas. Uso de la herramienta de software Promodel como estrategia didáctica en el aprendizaje basado en competencias de simulación de procesos y servicios. [Consultado el 19/09/16]

¹ Doctorado de Ingeniería FIUNLZ

4. G.M. Rodrigo Andrés, Z.M.Abdul, C.A. José Alejandro. (12 Junio 2014). Utilización de simulación discreta como estrategia de aprendizaje de logística empresarial. Teckne, Vol. 1, Pág. 13. [Consultado el 22/09/16]
5. Lina Marcela González Maya, Manuela Rodríguez Gómez. (2009). Juegos y Ejercicios prácticos para las materias del área de gestión de la producción y logística en ingeniería de Producción. Departamento de Ingeniería de Producción, Escuela de Ingeniería, Universidad EAFIT, Medellín. [Consultado el 29/09/16]
6. Pablo Moret, Claudia Iravedra, Leandro Rodriguez. (2014). Modelos y Objetos de enseñanza, diseño y simulación: Su aplicación en carreras tecnológicas del nivel superior. Instituto de investigación en tecnología y educación IIT&E. Facultad de Ingeniería. Universidad Nacional de Lomas de Zamora. [Consultado el 03/09/16]

Uzi: Máquina Virtual sobre Plataforma Arduino para Robótica Educativa

Ricardo Moran, Gonzalo Zabala, Matías Teragni, Sebastián Blanco

Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática
Universidad Abierta Interamericana

Av. Montes de Oca 745, Ciudad Autónoma de Buenos Aires, República Argentina
(+54 11) 4301-5323; 4301-5240; 4301-5248

{Ricardo.Moran, Gonzalo.Zabala, Matias.Teragni, Sebastian.Blanco}@uai.edu.ar

Resumen

El objetivo de este proyecto es el desarrollo de una máquina virtual para Arduino que ejecute un set de instrucciones definido especialmente para facilitar la enseñanza de robótica en las escuelas.

Palabras clave: Physical Etoys, Arduino, máquina virtual, lenguaje de programación, robótica educativa

Contexto

El presente proyecto será radicado en el Centro de Altos Estudios en Tecnología Informática (CAETI), dependiente de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana. El mismo se encuentra inserto en la línea de investigación “Sociedad del conocimiento y Tecnologías aplicadas a la Educación”. El financiamiento está dado por la misma Universidad Abierta Interamericana

Introducción

En los últimos años, la aparición de kits de robótica orientados a usuarios no

expertos fomentó el desarrollo de un conjunto significativo de proyectos educativos usando robots en diferentes niveles de educación, desde jardín de infantes hasta educación de grado. La plataforma de hardware Arduino, siendo abierta y de bajo costo, se ha popularizado muy rápidamente a nivel mundial para proyectos de este estilo.

La plataforma Arduino provee un entorno de desarrollo simplificado (basado en el lenguaje de programación C++) en el que muchos conceptos avanzados de programación de microcontroladores están “escondidos” al usuario. Sin embargo, este entorno es todavía demasiado complejo para algunos de los usuarios menos experimentados, sobre todo los niños más pequeños. Esto limita la complejidad de los problemas que estos usuarios pueden atacar.

Por estas razones, y aprovechando el carácter abierto de Arduino, han surgido múltiples intentos de proveer un entorno de programación más adecuado para principiantes. Uno de estos intentos es Physical Etoys, un ambiente de programación visual para chicos diseñado para proveer mecanismos de comunicación con distintas plataformas de hardware, incluyendo Arduino. Usando Physical Etoys, un alumno puede

programar robots de manera completamente gráfica, simplemente arrastrando y soltando instrucciones en la pantalla.

Sin embargo, Physical Etoys tiene un problema: dado que los programas se ejecutan en la computadora y las órdenes son transmitidas hacia el robot, Physical Etoys requiere que los robots estén continuamente conectados a la computadora para poder funcionar. Si bien la interactividad que este modelo de programación provee es muy importante (sobre todo para los alumnos que recién están comenzando), la falta de autonomía es un limitante muy importante. Por un lado, la comunicación constante con la computadora no está permitida en algunas competencias. Por otro lado, la cantidad de proyectos que pueden realizarse de esta forma es limitada (por ejemplo, quedan fuera algunas actividades de robótica situada que requieran la toma de decisiones en un entorno complejo y dinámico). Otros entornos de programación similares a Physical Etoys presentan el mismo problema (por ejemplo, Scratch 4 Arduino).

Una posible solución sería compilar los programas generados visualmente a un código nativo que pueda ejecutar el robot sin necesidad de una conexión con la computadora. Esta opción es la que eligieron otros entornos de programación (por ejemplo, Minibloq). Sin embargo, esta solución tiene sus propios problemas. Separar el “tiempo de compilación” del “tiempo de ejecución” imposibilita la forma de trabajo interactiva que caracteriza a ambientes como Physical Etoys.

Líneas de Investigación, Desarrollo e Innovación

No existe un entorno de programación visual para robótica que resuelva estos problemas, por lo tanto, creemos necesario un enfoque diferente. Los siguientes requerimientos deben cumplirse:

1. La ejecución de los programas debe hacerse directamente en el Arduino sin necesidad de interacción con la computadora (a excepción de la transmisión del programa).
2. En caso que el Arduino estuviera conectado con la computadora, todas las posibilidades de interacción que ofrece Physical Etoys deben mantenerse.
3. Para el usuario final debe ser transparente si el arduino funciona en modo “autónomo” o “interactivo”.

Otros requerimientos importantes (aunque no esenciales):

4. El entorno de programación debe ofrecer mecanismos de abstracción que permitan ocultar los detalles de algunos conceptos avanzados de programación de microcontroladores (por ejemplo: timers, interrupciones, modelos de concurrencia, etc.) Estos conceptos pueden luego ser introducidos a un ritmo compatible con las necesidades de los alumnos, pero desconocerlos no debería ser un limitante.
5. El entorno de ejecución debe ofrecer mecanismos para encontrar y arreglar errores (debugging). Particularmente, debe ser posible detener en cualquier

momento la ejecución, observar el estado interno del programa, y continuar ejecutando paso a paso de forma interactiva.

Nuestra propuesta para cumplir con estos requisitos es la utilización de una máquina virtual que ejecute en el Arduino un set de instrucciones sencillo diseñado especialmente para robótica educativa. Esta máquina virtual será transmitida (compilada) al Arduino una sola vez, y a partir de ese momento las herramientas de desarrollo diseñadas especialmente para este propósito podrán comunicarse con la máquina virtual para transmitir instrucciones individuales o programas enteros a través del puerto serie.

Esta máquina virtual, además de ejecutar los programas, será la encargada de abstraer todo mecanismo de comunicación y de detectar cuando el Arduino está conectado con la computadora, en cuyo caso podrá comunicarse con la misma en forma interactiva (ya sea enviando información de los sensores y estado interno del motor de ejecución, o recibiendo comandos que modifiquen el estado de los actuadores).

Si bien no se conocen entornos que cumplan con los requisitos antes planteados, el desarrollo de máquinas virtuales y lenguajes de programación de alto nivel para microcontroladores como Arduino no es nuevo. La mayoría de los desarrollos relevados se basan en lenguajes de programación de propósito general como Java, Scheme, o Python.

Resultados y Objetivos

Como objetivos específicos de este proyecto podemos enumerar los siguientes:

1. Definición de un set de instrucciones específico para robótica educativa (y particularmente, para Arduino)
2. Implementación de una máquina virtual simple que ejecute este set de instrucciones y permita al mismo tiempo monitorear el estado del robot desde la computadora.
3. Implementación de un lenguaje de programación que permita evaluar el funcionamiento de la máquina virtual.
4. Implementación de herramientas de desarrollo: compilador, inspector, debugger, etc.
5. Integración con Physical Etoys.

Formación de Recursos Humanos

El equipo de trabajo está conformado por un investigador adjunto del Centro de Altos Estudios en Tecnología Informática (CAETI) quien ejerce el rol de director del proyecto, dos doctorandos, y un ayudante alumno de la Facultad de Tecnología Informática de la Universidad Abierta Interamericana.

Bibliografía

- J.E. Smith, R. Nair: “Virtual Machines: Versatile Platforms for Systems and Processes” in The Morgan Kaufmann Series in Computer Architecture and Design Series, Morgan Kaufmann Publishers (2005)
- C. Schlegel, “Communication patterns as key towards component interoperability,” in Software Engineering for Experimental Robotics (Series STAR, vol. 30), D. Brugali, Ed. Berlin, Heidelberg: Springer-Verlag, , pp. 183–210.

- Smartsoftware (2007)
- Ando, N., Suehiro, T., Kitagaki, K., Kotoku, T., Yoon, W.K.: RT-middleware: Distributed component middleware for RT (robot technology). In: International Conference on Intelligent Robots and Systems 2005 (IROS 2005), pp. 3933–3938 (2005)
 - Brooks, A., Kaupp, T., Makarenko, A., Oreback, A., Williams, S.: Towards component-based robotics. In: Proc. of 2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'05), pp. 163–168. Alberta, Canada (2005)
 - Christian Schlegel, Thomas Häßler, Alex Lotz and Andreas Steck. Robotic Software Systems: From Code-Driven to Model-Driven Designs. In procs. Of ICAR 2009. International Conference on Advanced Robotics. IEEE Press (2009)
 - Bert Freudenberg, Yoshiki Ohshima, and Scott Wallace, "Etoys for one laptop per child," in 7th International Conference on Creating, Connecting and Collaborating through Computing - C5 2009, Kyoto, 2009, pp. 57-64.
 - R. Rahul, A. Whitchurch, and M. Rao, "An open source graphical robot programming environment in introductory programming curriculum for undergraduates," in 2014 IEEE International Conference on MOOCs, Innovation and Technology in Education, IEEE MITE 2014, Patiala, 2014, pp. 96-100.
 - A. Rao, "The application of LeJOS, Lego Mindstorms Robotics, in an LMS environment to teach children Java programming and technology at an early age," in 5th IEEE Integrated STEM Education Conference, ISEC 2015, 2015, pp. 121-122.
 - M. Elizabeth and C. Hull, "Occam-A programming language for multiprocessor systems," *Computer Languages*, vol. 12, no. 1, pp. 27-37, 1987.
 - C. L. Jacobsen, M. C. Jadud, O. Kilic, and A. T. Sampson, "Concurrent event-driven programming in occam- π for the Arduino," *Concurrent Systems Engineering Series*, vol. 68, pp. 177-193, 2011.
 - Ryan Suchocki and Sara Kalvala, "Microscheme: Functional programming for the Arduino," in *Scheme and Functional Programming Workshop*, Washington, D.C., 2014, pp. 21-29.

Tesis Doctorales

Evaluación de Rendimiento y Eficiencia Energética de Sistemas Heterogéneos para Bioinformática

Enzo Rucci

Directores por UNLP: Armando De Giusti y Marcelo Naiouf

Directores por UCM: Carlos García Sanchez y Guillermo Botella Juan

Fecha de exposición: 10/03/2016

Facultad de Informática, Universidad Nacional de La Plata, Argentina

erucci@lidi.info.unlp.edu.ar

1. Introducción

En primer lugar, se presenta la motivación de esta tesis. Luego se enuncian los objetivos y la metodología a emplear. Finalmente, se describen los alcances y limitaciones de esta investigación.

1.1. Motivación

En un mundo con recursos energéticos limitados y una constante demanda por mayor poder computacional, el problema del consumo energético se presenta como uno de los mayores obstáculos para el diseño de sistemas que sean capaces de alcanzar la escala de los Exa_ops. Por lo tanto, la comunidad científica está en la búsqueda de diferentes maneras de mejorar la eficiencia energética de los sistemas de cómputo de altas prestaciones (HPC, por sus siglas en inglés) [1].

Una estrategia reciente para incrementar el poder computacional de los sistemas HPC consiste en incorporarles aceleradores y coprocesadores, como pueden ser las unidades de procesamiento gráfico (GPU, por sus siglas en inglés) de AMD y NVIDIA o los coprocesadores Xeon Phi de Intel. Más allá de que estos dispositivos no comparten las mismas características, la estrategia se basa en utilizarlos para las secciones de código que requieren cómputo intensivo y dejar las tareas livianas a la unidad central de procesamiento (CPU, por sus siglas en inglés), como la entrada/salida (E/S) y las comunicaciones. Estos sistemas híbridos que emplean diferentes recursos de procesamiento se denominan arquitecturas heterogéneas y son capaces de obtener picos de rendimiento muy superiores a los de las CPUs. En los últimos años, esta estrategia se ha ido consolidando por una capacidad adicional de los aceleradores y coprocesadores: el incremento en el poder de cómputo se logra al mismo tiempo que se limita el consumo de potencia, lo que permite obtener mejores cocientes FLOPS/Watt. El ranking TOP500 [2] sirve para dar evidencia a esta tendencia mientras que el Green500 [3] la refuerza. En noviembre de 2010, ambos rankings contaban con 17 sistemas con aceleradores. Sin embargo, cinco años más tarde, este número se incrementó a 104 para el TOP500 y a 103 para el Green500.

Uno de los primeros aceleradores en ser incorporados a los sistemas de cómputo de altas prestaciones fueron las GPUs. En sus inicios estos dispositivos fueron pensados para el procesamiento gráfico y sus diseños estuvieron orientados a este tipo de aplicaciones. En la última década, se introdujeron algunas modificaciones a las arquitecturas de estos dispositivos y se desarrollaron librerías que permiten evitar el uso de primitivas de gráficos para programarlos. Estas modificaciones posibilitaron una amplia extensión de su uso. La clave del poder computacional de las GPUs se encuentra en su arquitectura: cientos o miles de núcleos pequeños y simples que ejecutan instrucciones en orden y operan en grupo como un procesador vectorial sumado a una memoria principal con alto ancho de banda la convierten en un arquitectura masivamente paralela capaz de lograr un pico de rendimiento muy por encima al de las CPUs. Sin embargo, los algoritmos deben ser codificados de forma tal que reflejen su arquitectura para lograr alto rendimiento. En ese sentido, tanto la arquitectura de las GPUs como sus modelos de programación difieren sustancialmente de los correspondientes a las CPUs, que es donde la comunidad HPC se ha focalizado tradicionalmente. En consecuencia, el programador se ve obligado a aprender conocimientos específicos de las arquitecturas y de su programación, lo que representa una tarea ardua.

Una alternativa a las GPUs son los coprocesadores Xeon Phi, los cuales fueron desarrollados por Intel para aplicaciones que requieren de cómputo de altas prestaciones. Se diseñaron como un complemento de las CPUs aunque pueden operar en forma individual, lo que los diferencia de las placas gráficas. Su arquitectura se basa en la de los procesadores Xeon pero con un mayor número de núcleos y de capacidades vectoriales más amplias. Al compartir los mismos modelos de programación que los procesadores Xeon, no resulta necesario aprender un nuevo lenguaje o herramienta para programarlos, con lo cual se reduce el tiempo requerido para desarrollar o portar una aplicación en este tipo de dispositivos. Por haber sido desarrollado recientemente, su utilidad y rendimiento son evaluados en la actualidad en diferentes áreas de aplicación, como pueden ser álgebra lineal, predicción del clima o diferentes tipos de simulación [4].

Las *Field Programmable Gate Arrays (FPGAs)* son otro tipo de acelerador basado en circuitos integrados reconfigurables. Aunque tradicionalmente fueron utilizadas para el procesamiento digital de señales, su uso ha crecido en diferentes áreas y probablemente se expanda aun más considerando que Intel compró recientemente Altera, una de las principales empresas fabricantes de FPGAs [5]. HPC es una de las áreas que se ha interesado en estos aceleradores debido principalmente a la evolución en su capacidad de cómputo y a su bajo consumo energético. Si bien tanto la frecuencia del reloj como el pico de rendimiento suelen ser más bajos que los correspondientes a las CPUs y a las GPUs, la capacidad de configurar el hardware para que se adapte al problema específico a resolver le da la posibilidad a las FPGAs de obtener mejores rendimientos, además de ser en general más eficientes desde el punto de vista energético [6]. La programación de estas placas se realiza a través de lenguajes de descripción de hardware (HDL, por sus siglas en inglés), como pueden ser Verilog o VHDL. Desafortunadamente, los HDLs son engorrosos, propensos a errores y requieren tener que mantener una explícita noción del tiempo, lo que implica una complejidad adicional [7]. En la actualidad, las principales empresas fabricantes de FPGAs (Altera y Xilinx) trabajan en el desarrollo de herramientas que permitan disminuir el esfuerzo de programación de estos dispositivos; en particular, a través de OpenCL, lo que resulta más familiar para los programadores HPC.

Más allá del tipo de acelerador utilizado, la programación de sistemas heterogéneos representa un verdadero desafío. Para lograr aplicaciones de alto rendimiento, los programadores deben enfrentar una serie de dificultades como pueden ser: estudiar características específicas de cada arquitectura, contemplar aplicaciones con múltiples niveles de paralelismo y aplicar técnicas de programación y optimización particulares para cada una de ellas, lograr una distribución del trabajo y un balance de carga entre los diferentes dispositivos de procesamiento que permita obtener buen rendimiento y afrontar el surgimiento de nuevos lenguajes y modelos de programación junto a la ausencia de un estándar y de herramientas afianzadas para este tipo de sistemas.

Una de las áreas que se ve afectada por los problemas actuales de los sistemas HPC es la bioinformática, ya que cuenta con un número creciente de aplicaciones que requieren de cómputo de altas prestaciones para alcanzar tiempos de respuesta aceptables. Esta situación se debe principalmente al crecimiento exponencial que ha experimentado la información biológica en los últimos años [8].

Una operación fundamental en cualquier investigación biológica es el alineamiento de secuencias, la cual consiste en comparar dos o más secuencias biológicas, como pueden ser las de ADN o las de proteínas. El propósito de esta operación es detectar qué regiones dentro de una secuencia comparten una historia evolutiva común. El alineamiento de secuencias resulta esencial en el análisis filogenético, el perfilado de enfermedades genéticas, la identificación y cuantificación de regiones conservadas o unidades funcionales, y el perfilado y predicción de secuencias ancestrales. También resulta importante en el desarrollo de nuevas drogas y en la investigación forense criminal, y es por ello que actualmente la comunidad científica realiza un gran esfuerzo en este campo de la bioinformática [9].

El algoritmo de Smith-Waterman (SW) [10] es un método popular para el alineamiento local de secuencias que ha sido utilizado como base para otros algoritmos posteriores y como patrón con el cual comparar otras técnicas de alineamiento. Sin embargo, su complejidad computacional es cuadrática, lo que lo hace inviable para procesar grandes conjuntos de datos biológicos. En la actualidad, se emplean diversas heurísticas que permiten reducir el tiempo de ejecución a costo de una pérdida en la sensibilidad de los resultados. Por fortuna, el proceso de alineamiento exhibe cierto paralelismo inherente que puede ser explotado para reducir el costo computacional de SW sin perder precisión.

De manera de procesar el creciente volumen de información biológica con tiempos de respuesta aceptables, es necesario desarrollar nuevas herramientas computacionales que sean capaces de acelerar primitivas claves y algoritmos elementales, especialmente aquellos que son costosos computacionalmente como es el caso del algoritmo SW. En ese sentido, el empleo de paralelismo se vuelve fundamental y para ello resulta indispensable contar con una evaluación integral de los diferentes sistemas HPC que considere el rendimiento alcanzable y el consumo energético asociado.

1.2. Objetivos y metodología

El objetivo general de esta tesis consiste en evaluar el rendimiento y la eficiencia energética de sistemas para cómputo de altas prestaciones al acelerar el alineamiento de secuencias biológicas mediante el algoritmo SW.

Los objetivos específicos son los siguientes:

- Describir y comparar las propuestas científicas para el alineamiento de secuencias biológicas utilizando sistemas para cómputo de altas prestaciones con el objeto de caracterizarlas.

- Identificar los factores que inciden en el rendimiento y la eficiencia energética de los sistemas para cómputo de altas prestaciones al acelerar el alineamiento de secuencias.
- Diseñar y desarrollar soluciones algorítmicas para aquellos sistemas de cómputo de altas prestaciones que no tengan implementaciones disponibles o, en caso contrario, que superen a las existentes en cuanto a aceleración.
- Medir y analizar el rendimiento y la eficiencia energética de las herramientas seleccionadas y de las propuestas desarrolladas.

Esta investigación doctoral se plantea como un estudio proyectivo que pretende dar solución a la necesidad de contar con implementaciones aceleradas del alineamiento de secuencias mediante el algoritmo SW en sistemas de cómputo de altas prestaciones junto a evaluaciones de rendimiento que además consideren la eficiencia energética.

1.3. Alcances y limitaciones

Con respecto al caso de estudio seleccionado, se decidió focalizar en el alineamiento de proteínas en lugar del de ADN por dos razones. En primer lugar, porque resulta ventajoso desde el punto de vista biológico en la mayoría de los casos [11]. En segundo lugar, el alineamiento de secuencias de proteínas es más complejo debido a un alfabeto más grande y al uso de matrices de sustitución, por lo que resolver el alineamiento de secuencias de ADN resulta relativamente simple una vez resuelto el primero.

En relación a la explotación de sistemas heterogéneos, las GPUs son el acelerador dominante en la comunidad de HPC al día de hoy por su alto rendimiento y bajo costo de adquisición y existe vasta investigación científica sobre el uso de esta clase de acelerador para el alineamiento de secuencias como se describe en la Sección 2.2.2. Por ese motivo se decidió priorizar el desarrollo de nuevas soluciones algorítmicas para sistemas heterogéneos basados en coprocesadores Xeon Phi y basados en FPGAs. Los coprocesadores Xeon Phi fueron introducidos recientemente y, al inicio de esta tesis, no existían implementaciones disponibles para el alineamiento de secuencias. En cuanto a las FPGAs, el estudio de su aplicación para el procesamiento de secuencias biológicas sí cuenta con antecedentes aunque estas implementaciones fueron desarrolladas con HDLs tradicionales y en su mayoría poseen una o más limitaciones que restringen su uso en el mundo real (ver Sección 2.2.2). El reciente desarrollo de herramientas basadas en OpenCL por parte de las principales empresas fabricantes de FPGAs representa uno de los motivos del estudio de esta clase de acelerador para el alineamiento de secuencias.

2. Alineamiento de secuencias biológicas de alto rendimiento

Esta sección se centra en la fuerte relación entre el alineamiento de secuencias biológicas y el HPC.

2.1. Algoritmo SW para alineamiento de secuencias biológicas

El algoritmo SW computa el alineamiento local óptimo entre dos secuencias siguiendo un enfoque matricial y se puede dividir en dos fases: (1) Cómputo de la matriz de similitud (o también llamada matriz de alineamiento), para obtener el puntaje de alineamiento óptimo; y (2) Retroceso, para obtener el alineamiento óptimo.

1. *Cómputo de la matriz de similitud*: dadas dos secuencias $q = q_1q_2q_3 \dots q_m$ y $d = d_1d_2d_3 \dots d_n$, se construye una matriz de similitud H de $(m+1) \times (n+1)$ elementos. Las relaciones de recurrencia para completar la matriz, con las modificaciones de Gotoh para admitir un modelo de penalización de gap por afinidad [12], se detallan a continuación:

$$H_{i,j} = \max \begin{cases} 0 \\ H_{i-1,j-1} + SM(q_i, d_j) \\ E_{i,j} \\ F_{i,j} \end{cases} \quad \begin{aligned} E_{i,j} &= \max \begin{cases} H_{i,j-1} - G_{oe} \\ E_{i,j-1} - G_e \end{cases} \\ F_{i,j} &= \max \begin{cases} H_{i-1,j} - G_{oe} \\ F_{i-1,j} - G_e \end{cases} \end{aligned}$$

Los residuos de la secuencia q , usualmente llamada secuencia de consulta, etiquetan las $_$ las mientras que los residuos de la secuencia d , usualmente llamada secuencia de la base de datos, $Hi;j$ representa el puntaje de alineamiento para las subsecuencias de q y d terminadas en los residuos qi y dj , respectivamente. $Ei;j$ y $Fi;j$ son los puntajes de los alineamientos de las mismas subsecuencias de q y d pero terminadas en un gap en q y en d , respectivamente. SM es la matriz de sustitución que define los puntajes de sustitución entre todos los pares de residuos. Matrices de sustitución usuales en el alineamiento de secuencias de proteínas son las familias BLOSUM o PAM. G_{oe} es la suma de las penalizaciones por inserción y extensión de gap, mientras que G_e es la penalización

por extensión de gap. $H_{i,j}$, $E_{i,j}$ y $F_{i,j}$ son inicializadas en 0 cuando $i = 0$ o $j = 0$. El puntaje de alineamiento óptimo S es el máximo puntaje de alineamiento de la matriz H .

2. *Retroceso*: el alineamiento óptimo se calcula realizando un proceso de retroceso desde la posición donde se encuentra el puntaje de alineamiento óptimo S hasta llegar a una posición donde el valor sea 0, siendo éste punto el inicio del alineamiento.

El algoritmo SW tiene complejidad temporal cuadrática. El puntaje de alineamiento óptimo se puede calcular linealmente en cuanto a espacio. Sin embargo, para obtener el alineamiento óptimo entre las secuencias resulta necesario almacenar la matriz H completamente, por lo que su complejidad espacial se vuelve cuadrática.

Resulta importante notar que ninguna celda de la matriz puede tener valor inferior a 0. Adicionalmente, existe un orden estricto en el cómputo de la matriz H debido a las dependencias de datos inherentes al problema. Toda celda en la matriz H tiene dependencias sobre otras tres celdas: la que está a la izquierda, la que está por encima y la anterior en la diagonal.

2.2. Aceleración del algoritmo SW

Esta sección se enfoca en el estado del arte de la aceleración del algoritmo SW.

2.2.1. Dependencias de datos y paralelismo

Aunque las dependencias mencionadas en la Sección 2.1 restringen las formas en que se puede computar la matriz de similitud, el algoritmo SW posee cierto paralelismo inherente que puede ser explotado para reducir su costo computacional. En general, las implementaciones suelen emplear alguno de los dos siguiente esquemas:

- El esquema intra-tarea se basa en acelerar el cómputo de un único alineamiento por vez. Las implementaciones que adoptan este esquema suelen computar las celdas de cada antidiagonal de la matriz en paralelo, aprovechando que estos cálculos son independientes entre sí. También es posible computar varias celdas de una fila o de una columna simultáneamente; sin embargo, como este enfoque ignora las dependencias de datos antes mencionadas, se requiere de un ajuste posterior de los valores de las celdas para mantener la corrección del algoritmo.
- El esquema inter-tarea se basa en acelerar el cómputo de múltiples alineamientos al mismo tiempo. La mayor ventaja de este enfoque es la independencia en el cálculo de cada alineamiento.

Ambos enfoques han sido explorados por la comunidad científica en una amplia variedad de trabajos y dispositivos, describiéndose los más destacados en las siguientes secciones.

2.2.2. Implementaciones existentes

Para poder evaluar el desempeño de una implementación en forma independiente de las secuencias de consulta y de las bases de datos utilizadas para las diferentes pruebas al igual que de la arquitectura de soporte, se emplea la métrica de rendimiento Cell Updates Per Second (CUPS). Dada una secuencia de consulta Q y una base de datos D , el valor de GCUPS (mil millones de CUPS) se calcula como $\frac{jQj}{jDj \cdot t_{109}}$, donde jQj es el número de total de residuos en la secuencia de consulta, jDj es el número total de residuos en la base de datos y t es el tiempo de ejecución en segundos requerido para computar las matrices de similitud [13]. A continuación, se describen las implementaciones más destacadas de acuerdo a la arquitectura de soporte empleada.

Implementaciones en CPU. Las implementaciones en CPU se han valido de instrucciones vectoriales para acelerar el cómputo de las matrices de similitud. En general, se pueden identificar diferentes enfoques para la utilización de vectorización en el cómputo de las matrices de similitud, los cuales se ilustran en la Figura 1. La mayoría de los esfuerzos se han concentrado en el enfoque intra-tarea. Sin embargo, es el enfoque inter-tarea el que ha producido los mejores resultados: en el año 2011, Rognes presentó SWIPE, una implementación paralela que explota las instrucciones SSSE3 bajo el enfoque inter-tarea [14]. SWIPE mostró rendimientos superiores a sus implementaciones contemporáneas por lo que se la considera la implementación SW más rápida para el conjunto de instrucciones SSE.

Implementaciones en GPU. Entre las propuestas basadas en GPU, se destaca CUDASW++ [13] y sus versiones posteriores [15, 16]. En su última versión, CUDASW++ combina cómputo concurrente en la CPU y en la GPU empleando el esquema inter-tarea en ambos casos. También se caracteriza por una cuidadosa organización de los accesos de memoria para aprovechar acceso coalescente y por la explotación de memoria

constante para las secuencias de consulta y de memoria compartida para la matriz de sustitución. En la actualidad, se la considera la implementación SW más rápida para sistemas basados en GPUs compatibles con CUDA.

Implementaciones en FPGA. En general, las implementaciones en FPGA se basan en crear bloques básicos que sean capaces de computar el valor de una celda de una matriz en cada ciclo de reloj. Luego, múltiples instancias de estos bloques son combinados a la vez para formar arreglos sistólicos que puedan procesar grandes cantidades de datos en paralelo. Desafortunadamente se hace difícil realizar una comparación analítica que resulte justa debido a diferentes causas [17, 18]: (1) existe una amplia variedad en tipos de FPGAs y cada una implementa sus circuitos de forma diferente, lo que dificulta una comparación directa; (2) existen muy pocas implementaciones que sean completas desde el punto de vista funcional, mientras que las que reportan los mejores resultados suelen contemplar pruebas sintéticas con el objetivo de mostrar la viabilidad del empleo de esta clase de aceleradores aunque su uso potencial en el mundo real es limitado; y (3) falta de documentación en los detalles de implementación que resultan claves para la comparación.

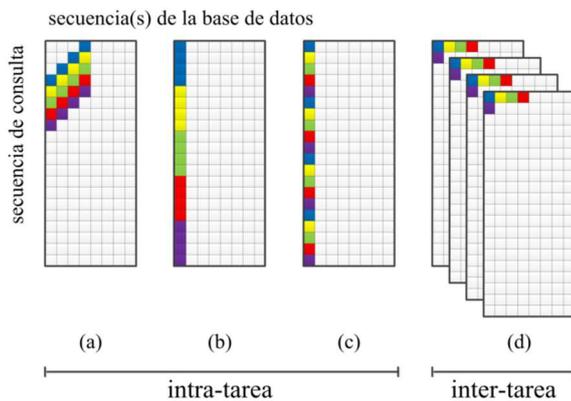


Figura 1: Enfoques para paralelización SIMD del algoritmo Smith-Waterman (adaptado de [14]). (a) Vectorización en las anti-diagonales. (b) Vectorización en la secuencia de consulta. (c) Vectorización por franjas en la secuencia de consulta. (d) Vectorización en múltiples secuencias de la base de datos.

Implementaciones en Xeon Phi. En el año 2014, Liu y Schmidt presentaron la herramienta SWAPHI [19] basada en el modelo `o_oad` de OpenMP. SWAPHI soporta los esquemas de paralelismo intra-tarea e inter-tarea, además de diferentes técnicas para la obtención de los puntajes de sustitución. XSW [20] es una herramienta alternativa publicada en el año 2014. Al igual que SWAPHI, XSW emplea el esquema inter-tarea. Sin embargo, a diferencia de SWAPHI, XSW se ejecuta en modo nativo con Pthreads como modelo de programación. Una versión extendida de XSW, conocida como XSW 2.0, fue desarrollada posteriormente [21]. Esta implementación adopta el modelo `o_oad` para combinar cómputo concurrente en la CPU.

3. Contribuciones y resultados

A continuación se describen las tres principales contribuciones de esta tesis, las cuales se encuentran sustentadas por los siguientes trabajos publicados en congresos internacionales [22, 23], revistas indexadas [24, 25] y capítulos de libros [26].

Optimización de SW para sistemas heterogéneos basados en aceleradores Xeon Phi: SWIMM

La primera contribución de esta tesis consiste en la introducción de la herramienta SWIMM para búsquedas de similitud¹ en sistemas heterogéneos basados en aceleradores Xeon Phi, la cual trabaja con bases de datos biológicas reales y permite configurar los parámetros de ejecución. Además, admite tres modos de ejecución: (1) cómputo en el host, (2) cómputo en los coprocesadores y (3) cómputo concurrente en el host y los coprocesadores. Mediante esta herramienta, se estudió y evaluó la viabilidad de esta clase de sistemas heterogéneos para el alineamiento de secuencias de proteínas mediante el algoritmo SW. Para beneficio de la comunidad científica, SWIMM se encuentra disponible en un repositorio web público².

En primer lugar, considerando la compatibilidad de código entre los procesadores Xeon y los coprocesadores Xeon Phi, se desarrolló una implementación común para ambos dispositivos siguiendo el enfoque inter-tarea. La propuesta explota dos niveles de paralelismo: a nivel de hilo mediante OpenMP y a nivel de datos usando vectorización guiada. Luego, se desarrollaron versiones alternativas para cada dispositivo que reemplazan la vectorización guiada por vectorización manual a través de funciones intrínsecas. En todos los casos se estudiaron y aplicaron diferentes técnicas de programación y optimización a las distintas implementaciones. Finalmente, utilizando los códigos desarrollados para ambos dispositivos se implementó una versión híbrida capaz de computar alineamientos en el Xeon y en el Xeon Phi de forma simultánea. Esta implementación admite dos modos de distribución de la carga de trabajo (estática y dinámica).

Posteriormente, se seleccionaron secuencias de consulta y bases de datos de proteínas representativas para la comunidad científica que sirvieran como benchmarks. En particular, se eligieron 20 secuencias de consulta de variada longitud y las bases de datos de datos Swiss-Prot y Environmental NR, las cuales fueron utilizadas conjuntamente en investigaciones similares anteriormente [14, 16, 19, 20]. Además, se seleccionaron otras implementaciones de referencia con las cuales poder realizar comparaciones: SWIPE para CPU, SWAPHI para Xeon Phi, XSW 2.0 para CPU-Xeon Phi y CUDASW++ para CPU-GPU. =

Se probaron las distintas implementaciones sobre dos arquitecturas heterogéneas diferentes. A continuación se detallan los principales resultados y conclusiones del análisis experimental:

- Para lograr implementaciones SW de alto rendimiento tanto en los Xeon como en los Xeon Phi, se adaptó el código con el `_n` de:
 - Explotar dos niveles de paralelismo: usar más hilos no resulta suficiente; se debe explotar paralelismo de datos también. En ese sentido, la vectorización guiada ofrece portabilidad de código y un menor costo de programación. Sin embargo, la mejora de rendimiento que se puede lograr es pequeña comparada a la de la vectorización manual.
 - Aprovechar el hecho de que los puntajes de similitud no requieren de un amplio rango de representación. Mediante la explotación de tipos de dato enteros de bajo rango es posible aumentar el paralelismo a nivel de datos (más elementos por vector). En particular, se mostró que resulta conveniente computar los alineamientos usando enteros de 8 bits en primera instancia y utilizar progresivamente representaciones más amplias cuando sea necesario.
 - Emplear técnicas de procesamientos por bloques que mejoren la localidad de datos, especialmente en el Xeon Phi considerando que su memoria caché es más pequeña.
 - Lograr una carga de trabajo balanceada entre el Xeon y el Xeon Phi en versiones híbridas. Una diferencia de potencia computacional (en GCUPS) muy grande entre ambos dispositivos puede llevar a un serio desbalance de carga, lo que se traslada a una reducción del rendimiento.
- Con respecto a los GCUPS, la implementación con intrínsecas SSE logró 118 GCUPS en el Xeon E5-2670 con 32 hilos mientras que su equivalente con AVX2 alcanzó 355 GCUPS en el Xeon E5-2695 v3 con 56 hilos. Aun así, la implementación con KNC sólo obtuvo un máximo de 50 GCUPS para 228 hilos. El pobre desempeño del Xeon Phi se debe a la ausencia de capacidades vectoriales para explotar enteros de menos de 32 bit.
- En cuanto a la comparación con otras implementaciones SW, la versión SSE de SWIMM se mostró equiparable con SWIPE mientras que la versión AVX2 logró superarlo ampliamente alcanzando diferencias de hasta 1.4. Cabe destacar que, hasta donde llega el conocimiento del tesista, esta es la primera implementación SW para el conjunto de instrucciones AVX2. Por otra parte, la versión KNC de SWIMM se mostró competitiva con SWAPHI, siendo capaz de superarlo para secuencias medianas y largas. Complementariamente, la versión híbrida de SWIMM con distribución dinámica (SSE+KNC) superó notablemente a su alternativa XSW 2.0, con mayores diferencias para secuencias de consulta largas. Por último, en el sistema basado en Xeon E5-2695 v3, SWIMM (versión AVX2+KNC) logró mejores rendimientos que CUDASW++ 3.0 utilizando los recursos del host y una GPU NVIDIA K20c. Resulta importante mencionar que la próxima generación de coprocesadores Xeon Phi (Knights Landing) unificará su conjunto de instrucciones con la generación actual de procesadores Xeon (Skylake) adoptando las extensiones AVX-512 de 512 bits. Como este conjunto de instrucciones sí posee capacidades para explotar enteros de bajo rango, se espera un incremento notable en el rendimiento de ambos dispositivos usando la metodología propuesta.

Optimización de SW para sistemas heterogéneos basados en aceleradores FPGA: OSWALD

La segunda contribución de esta tesis consiste en la presentación de la herramienta OSWALD para búsquedas biológicas en sistemas heterogéneos basados en aceleradores FPGA, la cual trabaja con bases de datos biológicas reales y permite configurar los parámetros de ejecución. Mediante esta herramienta, se estudió y evaluó la viabilidad de esta clase de sistemas heterogéneos para el alineamiento de secuencias de proteínas mediante el algoritmo SW. A diferencia de las implementaciones disponibles en la literatura, se decidió explorar los beneficios de utilizar una tecnología innovadora, como lo es OpenCL en el ámbito de las FPGAs, en lugar de HDLs tradicionales como VHDL o Verilog. Al igual que la herramienta anterior, OSWALD se encuentra disponible en un repositorio web público para beneficio de la comunidad científica³.

Como primer paso, se desarrolló una implementación para una única FPGA. En cuanto a la paralelización de los alineamientos se optó por el esquema inter-tarea. Para el desarrollo del algoritmo, se estudiaron y aplicaron diferentes técnicas de programación y optimización. Luego, se evaluó el rendimiento y el consumo de recursos de las diferentes versiones desarrolladas de forma de encontrar la configuración más beneficiosa. A continuación, se extendió la implementación mono-FPGA para que sea capaz de soportar ejecución en múltiples aceleradores al mismo tiempo. Finalmente, utilizando el código Xeon de SWIMM se implementó una versión híbrida capaz de computar alineamientos en la CPU y en las FPGAs de forma simultánea.

Empleando las secuencias de consulta y las bases de datos de proteínas utilizadas anteriormente, se probó la implementación multi-FPGA y la implementación híbrida sobre dos arquitecturas heterogéneas diferentes. También se probaron otras implementaciones de referencia con las cuales comparar las propias: SWIMM para sistemas heterogéneos basados en Xeon Phi y CUDASW++ 3.0 para los basados en GPUs compatibles con CUDA. Lamentablemente no fue posible realizar una comparación con otras implementaciones FPGA debido a la falta de disponibilidad del código fuente. Si bien un análisis teórico es posible, las significativas diferencias entre las distintas implementaciones impiden una comparación justa. A continuación se detallan los principales resultados y conclusiones del análisis experimental:

- De la exploración de los diferentes kernels SW se puede extraer:
 - Al computar las matrices de alineamiento empleando técnicas de procesamientos por bloques no sólo se mejora la localidad de datos sino que también se reduce los requerimientos de memoria del kernel, lo que favorece el uso de memoria privada de baja latencia.
 - Al explotar un mayor grado de paralelismo de datos a través de la vectorización de operaciones se logra incrementar el rendimiento en forma notable a expensas de un aumento moderado en el consumo de recursos.
 - Al explotar tipos de dato enteros de bajo rango no sólo se reduce el consumo de recursos sino que también se mejora el rendimiento. De esta forma se aprovecha el hecho de que los puntajes de similitud no requieren de un amplio rango de representación. En particular, se mostró que resulta conveniente computar los alineamientos usando enteros de 8 bits en la FPGA y recomputar en el host cuando sea necesario.
 - Lograr una carga balanceada entre los diferentes dispositivos de procesamientos
 - es un factor clave para obtener alto rendimiento. En la implementación multi-FPGA, ésto se logra en la etapa de preprocesamiento al dividir la base de datos en número de chunks que sea múltiplo del número de FPGAs y procurando que los chunks contengan aproximadamente la misma cantidad de residuos. Por su parte, la implementación híbrida requiere de un esquema más complejo por las diferentes capacidades computacionales que pueden tener el host y los aceleradores. Al estimar la potencia de cómputo de cada dispositivo de procesamiento, no sólo se logra una carga de trabajo balanceada con un overhead mínimo sino también una implementación general para cualquier sistema heterogéneo de esta clase.
- Con respecto a los GCUPS, la implementación mono-FPGA logró 58.4 GCUPS mientras que la implementación multi-FPGA con dos aceleradores obtuvo 114 GCUPS, lo que representa una escalabilidad casi lineal. Por su parte, la implementación híbrida alcanzó 178.9 y 401.1 GCUPS en los sistemas basados basados en Xeon E5-2670 y en Xeon E5-2695 v3, respectivamente.

³OSWALD se encuentra disponible online en <https://github.com/enzorucci/OSWALD>

- En cuanto a la comparación con otras implementaciones SW, OSWALD y SWIMM mostraron comportamientos similares en los dos sistemas heterogéneos utilizados. El rendimiento de OSWALD estuvo siempre por encima del de SWIMM (incluso para secuencias cortas), logrando mayores diferencias a medida que la carga de trabajo aumentaba. Por su parte, CUDASW++ 3.0 superó a todas las demás implementaciones en el sistema basado en Xeon E5-2670, principalmente por el gran poder computacional de la GPU NVIDIA K20c. En contraste, presentó el peor rendimiento en el sistema basado en Xeon E5-2695 v3 debido a su incapacidad de aprovechar las extensiones AVX2.

El costo de programación y la falta de portabilidad en los códigos FPGA han limitado tradicionalmente su aplicación a las búsquedas SW. OSWALD es una implementación portable, completamente funcional y general para acelerar búsquedas de similitud en sistemas heterogéneos basados en FPGA que demostró ser competitivo con otras implementaciones SW de referencia.

Eficiencia energética de SW en sistemas heterogéneos

Como en la actualidad la comunidad HPC no sólo está interesada en mejorar el rendimiento sino también en reducir el consumo energético, la tercera contribución de esta tesis consiste en el estudio y evaluación de la eficiencia energética de diferentes sistemas heterogéneos para el alineamiento de secuencias de proteínas mediante el algoritmo SW. Para ello, no se realizaron pruebas especiales sino que se monitorizó el consumo de potencia por parte de los diferentes sistemas en los experimentos realizados anteriormente.

En primer lugar, se analizó el consumo de potencia de las implementaciones híbridas de SWIMM y de OSWALD en forma individual. Posteriormente, se analizó comparativamente el rendimiento y el consumo de potencia obtenidos en las diferentes arquitecturas heterogéneas utilizadas. A continuación se detallan los principales resultados y conclusiones del análisis experimental:

- Algunas de las técnicas que, además de incrementar los GCUPS, mejoraron la eficiencia energética son:
 - Explotar mayor paralelismo de datos enteros de bajo rango, como se evidenció en el sistema basado en Xeon E5-2695 v3 (AVX2) comparado al basado en Xeon E5-2670 (SSE).
 - Usar Hyper-Threading. El uso de esta tecnología provocó un incremento mayor en el rendimiento que el correspondiente al consumo de potencia en ambas arquitecturas.
 - Lograr una carga de trabajo equilibrada entre host y aceleradores. De esta manera, no sólo se mejoraron las prestaciones sino que también se redujo el consumo energético.
- En cuanto a la computación basada únicamente en aceleradores (sin ejecución simultánea del host):
 - La FPGA puede no ser la mejor opción desde el punto de vista del rendimiento. Sin embargo, sí lo puede ser desde la perspectiva del consumo de potencia; su bajo consumo puede ser de utilidad cuando éste aspecto sea la mayor prioridad.
 - El Xeon Phi no representa una buena elección tanto desde la óptica del rendimiento como del consumo de potencia.
- A través de SWIMM, las CPUs presentaron un buen balance entre rendimiento y consumo. Sus prestaciones mejoraron significativamente al explotar las extensiones AVX2. Respecto a la combinación con aceleradores:
 - El uso de arquitecturas heterogéneas basadas en coprocesadores Xeon Phi no fue una buena opción en términos del consumo de potencia. La integración del Xeon Phi al cómputo de los alineamientos proveyó poca mejora de rendimiento al host y decrementó su cociente GCUPS/Watt. Se espera que la incorporación futura de las extensiones AVX-512 cambie esta situación.
 - En el sistema basado en Xeon E5-2670, CUDASW++ 3.0 mejoró los GCUPS y el cociente GCUPS/Watt de SWIMM (versión Xeon). Sin embargo, no fue capaz de repetir sus prestaciones en el sistema basado en Xeon E5-2695 v3 debido a que CUDASW++ 3.0 sólo explota las instrucciones SSE. Por ese motivo, la combinación de CPUs con GPUs puede ser una buena alternativa en sistemas que puedan explotar las instrucciones SSE, pero que no dispongan de las extensiones AVX2. El desarrollo de una herramienta que dé soporte a las extensiones AVX2 permitiría aumentar la eficiencia energética de esta clase de sistemas.
 - La computación híbrida basada en CPU-FPGA se ubicó como la opción más eficiente energéticamente teniendo en cuenta que OSWALD obtuvo los cocientes GCUPS/Watt más altos en

ambos sistemas. Aun más, la FPGA fue el único acelerador que mejoró las tasas GCUPS/Watt del host en ambos sistemas.

Aunque existen otras evaluaciones de eficiencia energética en el contexto de SW [27, 28], éstas poseen uno o más limitaciones que reducen su utilidad en el mundo real. En esta evaluación, se han empleado secuencias de consulta y bases de datos representativas de la comunidad bioinformática, potentes arquitecturas orientadas a HPC y eficientes implementaciones que han demostrado ser las más rápidas de su clase. En base a estas condiciones y teniendo en cuenta el rol de SW en la bioinformática, se espera que esta evaluación le resulte útil a cualquier centro de investigación y desarrollo de este área al momento de seleccionar un sistema HPC para sus aplicaciones de acuerdo con las prioridades que posea.

4. Conclusiones

De forma de procesar el creciente volumen de información biológica con tiempos de respuesta aceptables, resulta necesario desarrollar nuevas herramientas computacionales que sean capaces de acelerar primitivas claves y algoritmos elementales eficientemente en términos de rendimiento y consumo energético. Por ese motivo, esta tesis se ha planteado como objetivo general evaluar el rendimiento y la eficiencia energética de sistemas para cómputo de altas prestaciones al acelerar el alineamiento de secuencias biológicas mediante el método de Smith-Waterman.

En primer lugar, se estudiaron las posibles formas de paralelizar el algoritmo SW y se describieron las implementaciones existentes sobre diferentes plataformas de procesamiento: CPU, GPU, FPGA y Xeon Phi. El análisis para cada dispositivo incluyó la evolución temporal de sus implementaciones así como también una descripción detallada de los aportes, las limitaciones, los resultados y la experimentación realizada en cada uno de los trabajos. Este análisis fue posible gracias al estudio de las diferentes arquitecturas y modelos de programación además del de los distintos algoritmos para alineamiento de secuencias biológicas llevado al inicio de la tesis.

El siguiente paso consistió en desarrollar nuevas soluciones algorítmicas para sistemas heterogéneos. Como se mencionó en la Sección 1.3, se decidió priorizar el desarrollo de nuevas soluciones algorítmicas para sistemas heterogéneos basados en Xeon Phi y basados en FPGA. Como punto de partida, se desarrollaron y optimizaron implementaciones para las CPUs y los Xeon Phi en forma individual antes de combinarlos en una implementación híbrida. La integración de estas implementaciones dio como resultado la herramienta SWIMM. A partir del análisis inicial del estado del arte, se identificó a SWIPE como la herramienta más rápida para búsquedas de similitud en CPU. Mediante los experimentos realizados, se mostró que la versión SSE de SWIMM es equiparable con SWIPE mientras que la versión AVX2 logró superarlo ampliamente alcanzando diferencias de hasta 1.4_. Cabe destacar que, hasta donde llega el conocimiento del tesista, esta es la primera implementación SW para el conjunto de instrucciones AVX2. En cuanto a los Xeon Phi, al inicio de esta investigación no existían implementaciones disponibles para este coprocesador. Sin embargo, en el año 2014 aparecieron SWAPHI y XSW 2.0. Mientras que SWAPHI sólo explota el coprocesador, XSW 2.0 es capaz de aprovechar la potencia del host en simultáneo. A través de los experimentos realizados, se mostró que la versión KNC de SWIMM resulta competitiva con SWAPHI, siendo capaz de superarlo para secuencias medianas y largas. Complementariamente, la versión híbrida de SWIMM con distribución dinámica (SSE+KNC) superó notablemente a su alternativa XSW 2.0, logrando aceleraciones de hasta 3.9_.

El siguiente paso en el trabajo experimental fue el desarrollo de soluciones algorítmicas para sistemas heterogéneos basados en aceleradores FPGA. A diferencia de las implementaciones disponibles en la literatura, se decidió explorar los beneficios de utilizar una tecnología innovadora, como lo es OpenCL en el ámbito de las FPGAs. En primer lugar, se exploró el rendimiento y el consumo de recursos de diferentes kernels de forma de encontrar la configuración más beneficiosa. Posteriormente, se desarrolló una versión híbrida empleando el código de SWIMM, siendo compilada junto a las anteriores en la herramienta OSWALD. Hasta donde llega el conocimiento del tesista, ésta la primera implementación utilizando OpenCL sobre FPGAs para búsquedas de similitud. Además, en base al análisis inicial del estado del arte, es una de las pocas implementaciones que es completamente funcional y general para esta clase de sistemas, además de facilitar la portabilidad por el empleo de OpenCL. Desafortunadamente, la ausencia del código fuente impide una comparación con otras implementaciones basadas en FPGA y aunque un análisis teórico es posible, los diferentes dispositivos, tecnologías y enfoques empleados no sólo complican una comparación directa sino que también dificultan hacerla en forma justa.

El último paso del trabajo experimental consistió en relacionar el rendimiento alcanzado por los diferentes sistemas empleados con el correspondiente consumo de potencia. De acuerdo al análisis realizado, se puede afirmar que los sistemas basados únicamente en CPU son capaces de obtener un buen balance entre rendimiento y consumo de potencia a partir de la explotación de multihilado e instrucciones vectoriales, destacándose aquellos que disponen de las extensiones AVX2. Por otra parte, la incorporación de aceleradores al cómputo del host demostró mejorar el rendimiento global del sistema en todos los casos, aunque la proporción de mejora varió de acuerdo al acelerador elegido. Desafortunadamente no ocurrió lo mismo en cuanto a la eficiencia energética. Los Xeon Phi no representan una buena opción para este tipo de aplicación. La ausencia de capacidades vectoriales para enteros de bajo rango es la causa principal del pobre rendimiento de este coprocesador, lo que provoca que el incremento en el consumo energético sea mucho mayor que la ganancia de rendimiento. En sentido opuesto, las GPUs sí pueden explotar este tipo de datos y es por ello que se puede aprovechar su gran poder computacional para acelerar los alineamientos. Aunque el aumento en el consumo de potencia es elevado, la mejora lograda por el uso de esta clase de aceleradores resulta superior, lo que se traduce en una mayor eficiencia energética. En el caso de las FPGAs, su capacidad de reconfiguración hace posible adaptar el hardware a los requerimientos del algoritmo SW. Si bien el rendimiento alcanzable de estos aceleradores puede ser inferior al correspondiente a las GPUs, su bajo consumo de potencia lleva a mayores tasas de eficiencia energética. Por último, a diferencia de otras evaluaciones de rendimiento y eficiencia energética en el contexto de SW, en esta tesis se han empleado secuencias de consulta y bases de datos representativas de la comunidad bioinformática, potentes arquitecturas orientadas a HPC y eficientes implementaciones que han demostrado ser las más rápidas de su clase. Por ese motivo se considera que la evaluación presentada puede ser de mayor utilidad en el mundo real. De acuerdo a los resultados obtenidos y a las contribuciones realizadas, se espera que esta tesis aporte a una mayor adopción de SW por parte de la comunidad bioinformática y a un procesamiento más eficiente de los alineamientos de secuencias biológicas en términos de rendimiento y consumo energético.

5. Líneas de trabajo futuro

Algunas de las líneas de trabajo futuro que se desprenden de esta tesis pueden ser:

- Desarrollar soluciones algorítmicas para las nuevas generaciones de procesadores Xeon y Xeon Phi. Como se mencionó anteriormente, Intel ha anunciado que unificará el conjunto de instrucciones en las próximas generaciones de los Xeon y los Xeon Phi al integrar las extensiones AVX-512 en ambos dispositivos. Este conjunto de instrucciones dispone de capacidades para explotar enteros de bajo rango, lo que representa una de los principales factores para obtener alto rendimiento. Teniendo en cuenta los resultados obtenidos con la metodología empleada, se espera un incremento notable en el rendimiento de ambos dispositivos.
- Explorar la explotación de dispositivos no convencionales en HPC para la aceleración del alineamiento de secuencias, como pueden ser los DSPs o las arquitecturas basadas en procesadores ARM. Si bien estos dispositivos poseen capacidades computacionales limitadas, su bajo consumo energético los vuelve atractivos desde el punto de vista de la eficiencia energética.
- Extender la evaluación de rendimiento y eficiencia energética a otros dominios. Así como el algoritmo SW resulta representativo de la bioinformática, interesa extender el análisis realizado a aplicaciones que resulten características de otras áreas.

Referencias

- [1] W. Feng, X. Feng, and R. Ge, *Green supercomputing comes of age*, IT Professional, vol. 10, no. 1, pp. 17–23, 2008.
- [2] Top500. [Online]. Available: www.top5000.org
- [3] Green500. [Online]. Available: www.green500.org
- [4] *Front matter*, in High Performance Parallelism Pearls, J. R. Je_ers, Ed. Boston: M. Kaufmann, 2015, pp. i–ii.
- [5] (2015, 06) Intel to acquire altera. HPC Wire. [Online]. Available: http://www.hpcwire.com/o_-the-wire/intel-to-acquire-altera/
- [6] M. Vestias and H. Neto, *Trends of CPU, GPU and FPGA for high-performance computing*, in Field Programmable Logic and Applications (FPL), 2014 24th International Conference on, Sept 2014, pp. 6.
- [7] S. O. Settle, *High-performance Dynamic Programming on FPGAs with OpenCL*, 2014.
- [8] National Center for Biotechnology Information. [Online]. Available: <http://www.ncbi.nlm.nih.gov/>
- [9] B. Schmidt, *Bioinformatics: High Performance Parallel Computer Architectures*, B. Schmidt, Ed. CRC Press, 2010.

- [10] T. F. Smith and M. S. Waterman, *Identification of common molecular subsequences*, *Journal of Molecular Biology*, vol. 147, no. 1, pp. 195_197, March 1981.
- [11] A. Siegel. (2011, July) What are the advantages/disadvantages of using Protein Alignment vs DNA Alignment? [Online]. Available: <https://www.ocf.berkeley.edu/~asiegel/posts/?p=14>
- [12] O. Gotoh, *An improved algorithm for matching biological sequences*, in *J. of Mol. Biol.*, vol. 162, 1981, pp. 705_708.
- [13] Y. Liu, D. L. Maskell, and B. Schmidt, *CUDASW++: optimizing Smith-Waterman sequence database searches for CUDA-enabled graphics processing units*, *BMC Research Notes*, vol. 2:73, 2009.
- [14] T. Rognes, *Faster Smith-Waterman database searches with inter-sequence SIMD parallelization*, *BMC Bioinformatics*, vol. 12:221, 2011.
- [15] Y. Liu, B. Schmidt, and D. Maskell, *CUDASW++2.0: enhanced Smith-Waterman protein database search on CUDA-enabled GPUs based on SIMT and virtualized SIMD abstractions*, *BMC Research Notes*, vol. 3, no. 1, p. 93, 2010.
- [16] Y. Liu, A. Wirawan, and B. Schmidt, *CUDASW++ 3.0: accelerating Smith-Waterman protein database search by coupling CPU and GPU SIMD instructions*, vol. 14:117, 2013.
- [17] L. Hasan and Z. Al-Ars, *Computational Biology and Applied Bioinformatics*. InTech, 2011, ch. 9, pp. 187_202.
- [18] S. Dydel and P. Bala, *Large Scale Protein Sequence Alignment Using FPGA Reprogrammable Logic Devices*, in *Field Programmable Logic and Application*, ser. LNCS. Springer, 2004, vol. 3203, pp. 23_32.
- [19] Y. Liu and B. Schmidt, *SWAPHI: Smith-Waterman protein database search on Xeon Phi coprocessors*, in *25th IEEE International Conference on Application-specific Systems, Architectures and Processors (ASAP 2014)*, 2014.
- [20] L. Wang, Y. Chan, X. Duan, H. Lan, X. Meng, and W. Liu, *XSW: Accelerating Biological Database Search on Xeon Phi*, in *Fourth International Workshop on Accelerators and Hybrid Exascale Systems 2014*, 2014.
- [21] __. (2014) XSW 2.0: A fast Smith-Waterman Algorithm Implementation on Intel Xeon Phi Coprocessors. [Online]. Available: <http://sdu-hpcl.github.io/XSW/>
- [22] E. Rucci, A. De Giusti, M. Naiouf, G. Botella, C. Garcia, and M. Prieto-Matias, *Smith-Waterman algorithm on heterogeneous systems: A case study*, in *2014 IEEE International Conference on Cluster Computing (CLUSTER)*, Sep. 2014, pp. 323_330.
- [23] E. Rucci, C. García, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias, *Smith-Waterman Protein Search with OpenCL on an FPGA*, in *2015 IEEE Trustcom/BigDataSE/ISPA*, vol. 3, Aug. 2015, pp. 208_213.
- [24] __, *An energy-aware performance analysis of SWIMM: Smith-Waterman implementation on Intel's Multicore and Manycore architectures*, *Concurrency and Computation: Practice and Experience*, vol. 27, no. 18, pp. 5517_5537, 2015.
- [25] __, *OSWALD OpenCL Smith-Waterman on Altera's FPGA for Large Protein Databases*, *International Journal of High Performance Computing Applications*, Jun. 2016.
- [26] __, *State-of-the-Art in Smith-Waterman Protein Database Search on HPC Platforms*, in *Big Data Analytics in Genomics*, K.-C. Wong, Ed. Springer, Oct. 2016, pp. 197_223, doi: 10.1007/978-3-319-41279-5_6.
- [27] K. Benkrid, A. Akoglu, C. Ling, Y. Song, Y. Liu, and X. Tian, *High Performance Biological Pairwise Sequence Alignment: FPGA Versus GPU Versus Cell BE Versus GPP*, *Int. J. Recon_g. Comput.*, vol. 2012, pp. 7:7_7:7, 2012.
- [28] D. Zou, Y. Dou, and F. Xia, *Optimization schemes and performance evaluation of Smith-Waterman algorithm on CPU, GPU and FPGA*, *Concurr. Comput.*, vol. 24, no. 14, pp. 1625_1644, 2012.

MADÉ-mlearn: Marco para el Análisis, Diseño y Evaluación de M-learning

Susana Isabel Herrera, Doctora en Ciencias Informáticas

Directora: Dra. Cecilia Verónica Sanz - Codirectora: MSc. Marta Cristina Fénnema Facultad de Informática de la Universidad Nacional de La Plata

Fecha de exposición: 29 de Septiembre de 2016 – Disponible en: <http://hdl.handle.net/10915/55492>

1. Motivación y metodología

Esta tesis presenta una investigación realizada en el campo de las tecnologías informáticas aplicadas a la Educación. Aborda como situación problemática la inclusión de las tecnologías móviles en el aprendizaje, fenómeno denominado como *mobile-learning* o *m-learning*. Los interrogantes que motivaron la investigación fueron:

¿Cómo contribuir a la mejora del aprendizaje utilizando tecnologías móviles?

¿En qué consiste el *m-learning* y cómo se vincula con otras formas de aprendizaje mediado por TIC? ¿Cuáles son las teorías que fundamentan el *m-learning*?

¿Cuáles son los aspectos que deberían ser considerados al analizar y/o evaluar una experiencia de *m-learning*?

¿Qué variables y aspectos se deben considerar al diseñar experiencias de *m-learning* de una manera sistémica?

El objeto de estudio lo constituyen las experiencias de *m-learning*. La investigación desarrollada fue cuali-cuantitativa (enfoque mixto).

Se trata de una investigación cualitativa porque:

- La investigación se inició a partir de interrogantes, sin seguir un proceso estrictamente definido (aunque sí planificado), que fue modificándose durante sus 4 años de desarrollo;
- No se definieron hipótesis estrictas, el proceso fue guiado por los objetivos definidos;
- En la mayor parte de la investigación se procedió de manera inductiva, comenzando con una indagación en el mundo social, observando qué es lo que ocurre, recogiendo datos a partir de entrevistas o cuestionarios y, luego, generalizando;
- En algunos casos se utilizaron métodos de recolección de datos no estandarizados (entrevistas y cuestionarios abiertos) obteniendo perspectivas y puntos de vistas de los involucrados;
- El investigador principal formó parte del fenómeno estudiado, dado que participó mediante diferentes roles en las experiencias que formaron parte de la investigación.

Al mismo tiempo, se trata de una investigación cuantitativa principalmente porque:

- Se realizaron estudios cuantitativos estadísticos para avalar y aceptar la publicación de los hallazgos preliminares relacionados con el *m-learning*; por lo tanto, algunas recolecciones de datos se hicieron mediante instrumentos específicamente diseñados que permitieron realizar un análisis cuantitativo de la información;

Por otra parte, es una investigación aplicada ya que los resultados obtenidos son conocimientos que se aplican inmediatamente en la mejora del aprendizaje mediado por TIC; es decir, en la solución de un problema inmediato del mundo real.

Las variables principales involucradas en la investigación son:

- Aspectos a considerar en el análisis, diseño y evaluación de experiencias de *m-learning*: pedagógicos, tecnológicos, culturales, de interacción;
- Efectividad del marco de análisis, diseño y evaluación.

La descripción completa de este apartado se encuentra en el Capítulo 1 de la tesis.

2. Aporte de la tesis a la disciplina

Los principales aportes de esta tesis son: a) un marco teórico sobre *m-learning* y Computación Móvil, b) *MADE-mlearn*, un marco conceptual que guía el análisis, diseño y evaluación de experiencias de *m-learning*, y c) *MADE-mlearn App*, una aplicación que permite realizar la evaluación de experiencias de *m-learning*, usando *MADE-mlearn*, en forma automatizada. Además, en el contexto de la tesis fue desarrollada la aplicación móvil educativa y colaborativa *Educ-Mobile*. Los resultados parciales de la tesis fueron publicados, con referato nacional e internacional, en:

- World Engineering Education Forum (WEEF 2015). Mobile Technologies in Engineering Education. Autores: Herrera, S. (expositor), Fennema, M., Morales, M., Palavecino, R., Goldar, E., Zuain, S. Ed. IEEE. ISBN 978-1-4799-8707-8. Florencia. Septiembre, 2015.
- IX Congreso Tecnología en Educación y Educación en Tecnología. Presentación de DEMO: Educ-Mobile, Juego educativo colaborativo para *m-learning*. Autores: Herrera, S. I., Najar, P., Morales, M.I., Sanz, C., Fennema, M.C. ISBN 978-987-28186-0-9. Chilecito, La Rioja. 2014.
- XX Congreso Argentino de Ciencias de la Computación (CACIC 2014). Eficiencia de aplicaciones móviles según su arquitectura. Autores: Najar, P., Ledesma, E., Rocabado, S., Herrera, S., Palavecino, R. ISBN 978-987-3806-05-6. Buenos Aires. Octubre, 2014.
- International Conference on Collaboration Technologies and Systems (CTS). Collaborative *m-learning* practice using Educ-Mobile. Autores: Herrera, S. (expositor), Sanz, C. Ed. IEEE. ISBN 978-1-4799-5157-4. Minneapolis. Mayo, 2014.
- Libro: Aprendizaje Basado en Dispositivos Móviles. Autores: Herrera, S., Morales, M., Fennema, M., Sanz, C. Santiago del Estero: EDUNSE - Editorial de la UNSE. Pag. 196. ISBN 978-987-1676-18-7. Santiago del Estero. 2014.
- II Congreso Argentino de Ingeniería. CADI 2014. Estrategias de *m-learning* para la enseñanza de la Matemática en Carreras de Ingeniería. Autores: Herrera, S.I., Morales, M.I. y Fennema, M. C. San Miguel de Tucumán. Octubre, 2014.
- Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología. *MADE- mlearn*: un marco para el análisis, diseño y evaluación de experiencias de *m-learning* en el nivel de postgrado. Autores: Herrera, S., Sanz, C., Fennema, M. ISSN 1850- 9959. La Plata: Universidad Nacional de La Plata. 2013. vol. n°10. Pp. 7 - 15.
- VIII Congreso Tecnología en Educación y Educación en Tecnología (TE&ET 2013). *M-learning* en Zonas de Recursos Limitados. Autores: Rocabado, S., Herrera, S., Morales, M., Estelles, C. ISBN 978-987-1676-04-0. Santiago del Estero. Junio, 2013.

2.1. Marco teórico

Para la elaboración del marco teórico se hizo el siguiente recorrido: se presentaron conceptualizaciones, teorías y características del *m-learning*; se revisaron antecedentes de experiencias y de marcos para el análisis de experiencias y, finalmente, se describieron aspectos de Computación Móvil importantes para el desarrollo y uso de aplicaciones.

Respecto al *m-learning*, se planteó su relación con otras modalidades de aprendizajes del *continuum* “educación presencial–educación a distancia” (Zangara, 2014). Se presentaron diferentes definiciones y conceptualizaciones de *m-learning*, propuestos por diversos autores (Woodill, 2011;

Sanchez Prieto et al., 2013; Traxler, 2011; Pachler et al., 2010; Seipold y Pachler, 2011; Sharples et al., 2007). Se obtuvo una definición propia, a la cual responde todo el planteo de la tesis:

El m-learning es el proceso de adquirir conocimiento mediante una relación dialógica entre el entorno y las personas y/o las personas entre sí, a través de una mediación con tecnología móvil, tanto en contextos de aprendizaje formales como no formales e informales; involucra en el alumno competencias tecnológicas para manipular los dispositivos móviles, competencias relacionadas con el aprendizaje autónomo y con la capacidad de interacción y comunicación.

Se analizaron las teorías que sustentan al *m-learning*: Teoría de la Actividad (TA) (Engeström, 1987), Marco Conversacional de Laurillard (2002), modelo dialéctico de Taylor-Sharples (Sharples et al., 2007), enfoque socio-cultural de Pachler et al. (2010). Este último estudia el *m-learning* desde diferentes aspectos: educativo, social, cultural, mediático, tecnológico y semiótico; es consistente con la definición planteada.

Se determinaron las principales características del *m-learning*: el ecosistema, los modos de interacción, los enfoques para la implementación de prácticas y el *m-learning* colaborativo, la relación con la vida cotidiana y cuestiones generales de aprendizaje que son importantes para el diseño de prácticas de *m-learning*. Sobre estas características se resalta lo siguiente:

- En relación al ecosistema, se describieron cada uno de sus principales componentes propuestos por Woodill (2011): dispositivos, infraestructura, conceptos, contenidos, plataforma y herramientas; cuando se estudian experiencias de *m-learning*, se contextualizan cada uno de estos componentes.
- Respecto a los enfoques de interacción, fueron caracterizados los tres modos alternativos: a) recuperación de información, b) recopilación y análisis de información, c) comunicación, interacción y colaboración en redes. Luego, se trató el *m-learning* colaborativo, resultado del modo de interacción c).
- En relación a los enfoques para la implementación de prácticas de *m-learning* (Pachler et al., 2012), se abordó tanto el enfoque top-down como el bottom-up.
- Se plantearon cuatro parámetros que deben ser considerados para balancear las tensiones que surgen del uso de las tecnologías móviles para atender las demandas de las instituciones educativas, por un lado, y las competencias de la vida cotidiana, por otro. Esos parámetros son: configuración del aprendizaje, relación con los recursos de aprendizaje, énfasis institucional en la experticia, y modos de representación.

Se identificaron fortalezas y debilidades del *m-learning* en relación a la educación de adultos y a la educación superior.

Respecto a los antecedentes, se revisaron experiencias de *m-learning* para adultos; destacando la importancia de que los adultos aprendan a través de su vinculación con su contexto y en particular, objetos de su vida cotidiana. Se abordaron experiencias en educación superior, donde más se utilizaron celulares hasta la actualidad (Cukierman et al., 2007; Thomas et al., 2015; Deroche et al., 2015). Se presentaron experiencias realizadas en Argentina por el equipo de investigación de Computación Móvil de la Universidad Nacional de Santiago del Estero (UNSE) (Herrera et al., 2012). Se describieron experiencias de *m-learning* desarrolladas en educación primaria y secundaria, destacando las experiencias desarrolladas en Argentina por el equipo ya mencionado de UNSE (Herrera et al., 2014), y por investigadores de la Universidad Nacional de La Plata (Llitas, 2015). Se analizaron experiencias de *m-learning* colaborativo (Organista y Serrano, 2011; Nouri et al., 2011; Nouri et al., 2012; Dewitt et al., 2014), ya que esta tesis considera que el aprendizaje

colaborativo contribuye a la mejora del aprendizaje actual. Asimismo, se presentaron los marcos para el análisis de *m-learning* que fueron desarrollados por otros autores (Park, 2011; Navarro et al., 2015; Pachler et al., 2012), mencionando sus principales fortalezas y debilidades. Se destaca la visión ecológica del marco de Pachler (2012).

Tanto los aspectos teóricos como los antecedentes de *m-learning*, fueron considerados tanto en el proceso de construcción de *MADE-mlearn* como en el diseño de la experiencia “Práctica de *m-learning* colaborativo usando *Educ-Mobile*”.

Respecto a Computación Móvil, en el marco teórico se presentó su conceptualización según diversos autores (B’Far, 2004; Pernici, 2006; Talukder et al., 2010), y las características principales. Se determinaron los conceptos básicos que deben ser considerados al proponer y evaluar experiencias de *m-learning*. Dichos conceptos hacen referencia tanto a tecnologías móviles blandas (sistemas operativos, aplicaciones móviles, herramientas de desarrollo) como a duras (redes inalámbricas, redes móviles, dispositivos).

Se hizo hincapié en aspectos relacionados con el desarrollo de aplicaciones móviles (Challiol, 2012): Arquitectura de la Aplicación, Modelo de Contexto, Modelo de Usuario, Representación del Espacio, Dinámica de la Aplicación, Representación de los Puntos de Interés, Modelo de Dominio y Mecanismos de Sensado. Se presentó la evolución de las redes de telefonía móvil (Talukder et al., 2010; Rocabado, 2014), que permiten ejecutar aplicaciones en línea, en movimiento, cada vez más ricas y complejas. Se revisaron las herramientas que se utilizan actualmente para desarrollar aplicaciones de calidad. Esto es importante ya que, la implementación de experiencias de *m-learning* involucra, generalmente, el uso de aplicaciones móviles. Sin duda, uno de los principales aspectos considerados fue el estudio de arquitecturas alternativas para el desarrollo de aplicaciones y el análisis de la eficiencia de éstas.

Los aspectos considerados se tuvieron en cuenta para elaborar *MADE-mlearn*, para el diseño de la experiencia de *m-learning* colaborativo y para el desarrollo de *Educ-Mobile*.

El marco teórico se encuentra plasmado en los capítulos 2, 3 y 4 de la tesis.

2.2. *MADE-mlearn*

El Marco para el Análisis, Diseño y Evaluación de experiencias de *m-learning* (*MADE-mlearn*) incorpora la concepción del *m-learning* como una modalidad de aprendizaje mediado por dispositivos y redes móviles (Zangara, 2014) y, principalmente, como una modalidad que permite el aprendizaje situado -en cualquier momento y lugar- (Traxler, 2011), en contextos propios -formales, no formales e informales- y que propone considerar aspectos de la vida cotidiana de las personas en procesos educativos (Sharples et al., 2007; Pachler et al., 2010; Seipold y Pachler, 2011), con diversos tipos de interacción (Woodill, 2011), bajo diversos paradigmas pedagógicos y diferentes tipos de actividades didácticas.

MADE-mlearn fue elaborado desde un enfoque ecológico o sistémico. Desde el punto de vista estructural, está compuesto por un conjunto de subcaracterísticas pedagógicas, tecnológicas y socio-culturales; las cuales fueron seleccionadas en base a: a) estudios exploratorios teóricos y de antecedentes; y b) un relevamiento de campo realizado en la región NorOeste de Argentina (NOA), presentados en Herrera y Fennema (2011). Estas subcaracterísticas están agrupadas en características, categorías y ejes (ver Fig. 1).

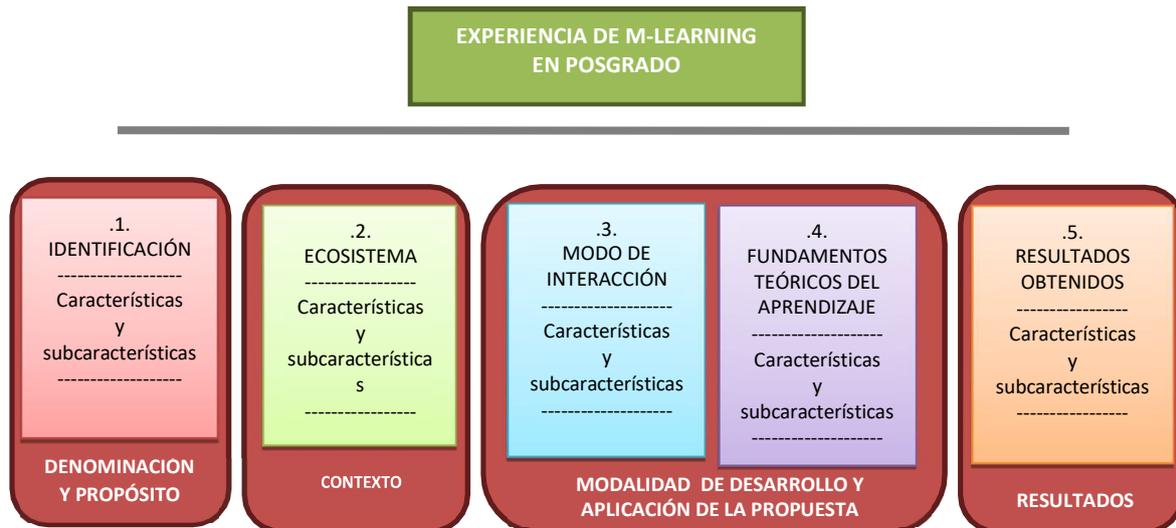


Figura 1. Ejes y categorías del *MADE-mlearn*.

Desde lo procedimental, *MADE-mlearn* constituye una guía para el análisis y evaluación de experiencias de *m-learning* que ya fueron implementadas; así como también es una guía para el diseño de nuevas experiencias de *m-learning* que se desean implementar. En la tesis se define cómo se deben realizar los procesos de análisis, diseño y evaluación de experiencias, cuando se usa este marco. El proceso de **análisis** de experiencias de *m-learning* consiste básicamente en un proceso de análisis y síntesis; se analizan cada una de las subcaracterísticas y se sintetizan en las características, categorías y ejes. El proceso de análisis propuesto es sencillo. Los pasos a seguir para el análisis de experiencias con *MADE-mlearn* son: 1. Definición del equipo de análisis, 2. Revisión global y distribución de trabajo, 3. Análisis secuencial de las subcaracterísticas, y 4. Elaboración de síntesis. En cada una de estas etapas, se ejecutan actividades en función de entradas y se obtienen productos en la salida. Al terminar el proceso, se obtiene un Documento Final de Análisis de la experiencia de *m-learning*.

Por otro lado, los pasos a seguir para el **diseño** de nuevas experiencias con *MADE-mlearn* son: 1. Definición del equipo de diseño, 2. Revisión global de *MADE-mlearn* y definición del modo de trabajo, 3. Diseño de la experiencia según subcaracterísticas, y 4. Elaboración de especificación final. Se diseñó una herramienta de documentación y trabajo colaborativo, tanto para el análisis como para el diseño, que consiste en una planilla de cálculo sencilla que contiene todas las subcaracterísticas sistematizadas.

Finalmente, a efectos de que *MADE-mlearn* pueda ser utilizado para evaluar una experiencia de *m-learning*, se realizó un proceso de operacionalización de sus categorías, características y subcaracterísticas. Esta operacionalización consistió en traducir estos aspectos en valores cuantitativos y cualitativos, para poder viabilizar dicha evaluación.

La operacionalización permite que cualquier evaluador con conocimientos en Educación mediada por TICs, sea experto o no en *m-learning*, pueda asignar valores a las subcaracterísticas y, a partir de allí, a través de la aplicación de fórmulas automáticas, pueda obtener valores cuantitativos que indiquen el grado de pertinencia y calidad de la información de cada característica, luego de las categorías y, finalmente, pueda obtener una valoración general del planteamiento y especificación de una determinada experiencia de *m-learning*.

En general, el criterio adoptado es:

La calidad de la especificación de una experiencia de m-learning es más alta cuanto mayor sea la cantidad de subcaracterísticas de MADE-mlearn que se encuentren documentadas y mayor sea la completitud y pertinencia de la información documentada.

Al finalizar el proceso de evaluación con *MADE-mlearn*, un evaluador podrá obtener como resultado un nivel alto, medio o bajo de calidad de la especificación de la experiencia de *m-learning*, según la documentación provista.

El proceso de **evaluación** con *MADE-mlearn* se basa en la estrategia GOCAME, desarrollada por el grupo de investigación GIDIS WEB de la Universidad Nacional de La Pampa (Lew et al., 2011). Esta estrategia se aplica a la medición de la calidad del software, desde el punto de vista de la calidad del producto. Tiene un amplio alcance dado que propone diferentes etapas que se deben llevar a cabo, como ser: definición de requerimientos de calidad, diseño de la medición, etc. Fue aplicada en acabadas oportunidades, un ejemplo, es la medición de la eficiencia de aplicaciones móviles según su arquitectura (Najar et al., 2014). En este caso, no se trata de evaluar calidad del software sino de evaluar la calidad de la especificación realizada en relación a una experiencia de *m-learning*. Es decir, el objeto a ser evaluado es diferente. Sin embargo, se toma como base a la estrategia GOCAME dado que evalúa la calidad en función de la medición de atributos, características y subcaracterísticas; aspectos similares a los considerados por *MADE-mlearn*.

Las etapas del proceso propuesto son: 1. Revisión de la documentación, 2. Valoración cuantitativa y cualitativa de las subcaracterísticas, 3. Valoración cuantitativa de las características, 4. Valoración cuantitativa de las categorías, y 5. Valoración de la calidad de la especificación de la experiencia de *m-learning*.

La descripción completa de *MADE-mlearn* se encuentra en el Capítulo 5 de la tesis.

2.3. MADE-mlearn App

En la tesis se desarrolló una aplicación web, denominada *MADE-mlearn App*, que permite realizar la evaluación en forma automatizada. Esto facilita y hace viable el proceso de evaluación, dada la gran cantidad de subcaracterísticas de *MADE-mlearn*.

MADE-mlearn App es una aplicación web que ha sido desarrollada utilizando un método de desarrollo ágil: Programación Extrema (o XP, del inglés *eXtreme Programming*). Está disponible en www.made-mlearn.com.ar.

Se definieron las historias de usuario y, a partir de éstas, se inició el diseño de la aplicación, trabajando con las tarjetas de programación.

Como resultado del diseño, se obtuvo un modelo de datos que contiene las tablas correspondientes a *ejes*, *categorías*, *características* y *subcaracterísticas*. Estas últimas, están relacionadas con sus correspondientes *indicadores*. También es importante observar que un *usuario* que pertenece a una *institución*, puede tener el perfil de administrador o de evaluador. Además, existen *experiencias* que son evaluadas por los *usuarios* evaluadores; y el resultado final se relaciona con la tabla *Resultados*. El modelo está implementado en una Base de Datos en *SQL Server*.

La descripción de *MADE-mlearn App* se encuentra al final del Capítulo 5 de la tesis.

3. Validación de la efectividad de MADE-mlearn

MADE-mlearn ha sido validado tanto en el análisis y evaluación de experiencias ya implementadas, como así también en el diseño de nuevas experiencias.

La validación del análisis se hizo en base a: a) la aplicación de *MADE-mlearn* completo (y el proceso definido) para analizar dos experiencias de *m-learning* en educación superior; b) la aplicación parcial de *MADE-learn* para el análisis de 12 experiencias de *m-learning* colaborativo, realizada por Gallo Vargas (2014).

La validación del diseño de nuevas experiencias con *MADE-mlearn* fue un proceso que se extendió aproximadamente dos años; involucró el desarrollo de la aplicación móvil *Educ-Mobile*. Se diseñó e implementó la experiencia denominada “Práctica de aprendizaje colaborativo usando *Educ-Mobile*”. Para adquirir robustez en los resultados de la validación, fue necesario replicar la práctica en otros cursos (adaptando la práctica usando *MADE-mlearn*). En la Figura 2 se muestra el proceso completo de validación del diseño con *MADE-mlearn*.



Figura 2. Validación del Diseño de experiencias de *m-learning* usando *MADE-mlearn*.

La experiencia “Práctica de aprendizaje colaborativo usando *Educ-Mobile*” fue diseñada e implementada en el curso de posgrado Enseñanza de la Tecnología de la Especialización en Enseñanza en Tecnologías de la UNSE. Una vez implementada fueron evaluados los objetivos de aprendizaje (mediante una evaluación en línea que calificaba a los alumnos del 0 al 100) y la motivación de los alumnos (mediante una encuesta implementada en *GoogleDrive*).

En cuanto a *Educ-Mobile*, es un juego educativo sincrónico colaborativo basado en posicionamiento para prácticas de *m-learning*. Es una aplicación Android que se adapta para el aprendizaje de diferentes contenidos en todos los espacios curriculares de la educación media y superior. Su objetivo consiste en que los estudiantes logren resolver en forma colaborativa, interactuando con sus dispositivos móviles, consignas relacionadas con los contenidos del área en cuestión.

La aplicación cuenta con tres juegos grupales (ver figuras 3 y 4). Cada grupo debe jugarlos de manera secuencial, hasta cumplir con todos o hasta que se cumpla el tiempo máximo (2 horas). Cada integrante de grupo debe seguir diversos recorridos, resolviendo consignas individuales que le permiten acceder a otras grupales que se resuelven en forma cooperativa y colaborativa. La resolución de cada consigna individual y grupal va otorgando puntajes, que se van acumulando.

Al ingresar, la aplicación va habilitando cada juego y los jugadores tienen que ir a las distintas estaciones donde se encuentran códigos QR que deben leer. La aplicación provee al alumno un croquis que indica la ubicación de cada estación. El código QR permite que la aplicación identifique si el jugador se encuentra en la estación asignada y automáticamente muestra la consigna a resolver.



Figura 3. Pantalla principal de Juegos.

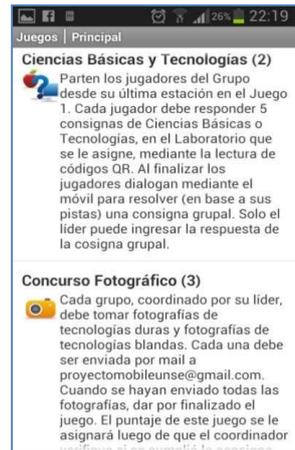


Figura 4. Pantalla principal de juegos.

Las figuras 5 a 7 muestran ejemplo de una consigna individual.

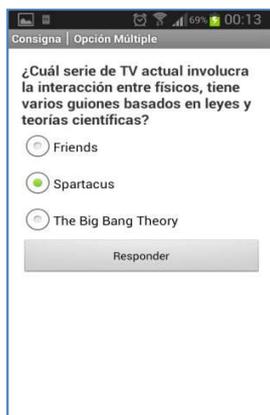


Figura 5. Consignas de opción múltiple.

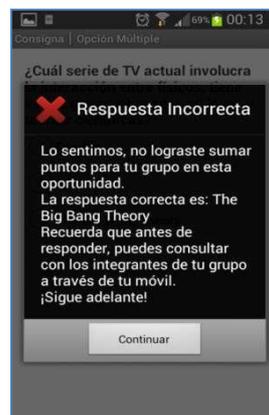


Figura 6. Mensaje de respuesta incorrecta.



Figura 7. Mensaje de respuesta correcta.

Para responder, los jugadores pueden consultar al resto del grupo o a quien deseen a través de su dispositivo (usando chat, SMS, llamadas, etc.). Además de las consignas individuales, cada equipo recibe pistas para resolver acertijos que deben descubrir colaborando entre sus integrantes. Cada grupo cuenta con un líder, que guía la comunicación del grupo y, además, es el responsable de ingresar las respuestas de las consignas grupales que, cuando están bien resueltas, duplica el puntaje obtenido por el grupo en forma individual. Los jugadores del mismo equipo no se encuentran físicamente mientras se desarrolla el juego. La aplicación informa el puntaje que van obteniendo, de manera individual y grupal. Las consignas están relacionadas con las temáticas mencionadas.

Cada juego (1 a 3) involucra tres tipos de consignas: opción múltiple (selección de la respuesta correcta), ingreso de texto y multimedia.

El grupo ganador se define en función del puntaje total y del tiempo de duración de la jugada, siguiendo una fórmula, que prioriza ampliamente el puntaje obtenido sobre el tiempo empleado.

Información más detallada está disponible en www.educ-mobile.com.ar.

Finalmente, respecto a la validación de la evaluación usando *MADE-mlearn*, ésta se hizo aplicando la técnica de juicio a expertos, usada en diversas investigaciones de educación y TIC (Maldonado Mahauad, 2015; Cabero y Llorente, 2013). Se seleccionaron ocho expertos, quienes evaluaron la experiencia "Práctica de aprendizaje colaborativo con *Educ-mobile*". Para la evaluación utilizaron la *MADE-mlearn App*. Luego de que los expertos realizaron la evaluación con la aplicación, se les solicitó que respondan cuatro preguntas abiertas sobre la evaluación usando *MADE-mlearn*.

En la validación del análisis, como así también del diseño y de la evaluación, se realizaron análisis de resultados cuantitativos. Estos permitieron detectar fortalezas y debilidades de *MADE-mlearn*. A partir de ellas, se abordaron a las siguientes conclusiones:

- *MADE-mlearn* constituye un marco que guía efectivamente el análisis, diseño y evaluación de experiencias de *m-learning*; y, además, permite definir correctamente los requisitos iniciales de aplicaciones móviles de *m-learning* involucradas en las experiencias.
- *MADE-mlearn* constituye un verdadero aporte al área Tecnología Informática aplicada en Educación, ya que avanza respecto de los marcos existentes. Es decir:
 - constituye un marco sustentado en sólidas teorías basadas en la interacción entre alumnos y con el docente y en la mediación de tecnologías móviles de uso cotidiano, que permiten crear nuevos contextos situados de aprendizaje;
 - da soporte no sólo al análisis sino también al diseño y evaluación de cualquier tipo de experiencias de *m-learning*, sea o no de Educación a Distancia, en contextos latinoamericanos (gracias a que fue construido apoyado en un ecosistema de *m-learning* definido previamente).

La validación de la efectividad de *MADE-mlearn* en el análisis, diseño y evaluación de experiencias de *m-learning*, se encuentra en los capítulos 6, 7 y 8, respectivamente.

4. Posibles líneas de investigación futura

A partir del proceso de validación de *MADE-mlearn*, en el cual se identificaron fortalezas y debilidades, se definieron como posibles líneas de investigación futura:

- Estudio de la suficiencia y claridad de las subcaracterísticas de *MADE-mlearn*, mediante “juicio de expertos” para profundizar el trabajo ya realizado;
- Optimización del proceso de Evaluación con *MADE-mlearn* (mejora de la escala de Likert utilizada al momento);
- Optimización de *MADE-mlearn App*: agregar evaluaciones personalizadas y ampliar su aplicación al análisis y diseño
- Creación de un repositorio digital abierto con las experiencias de *m-learning* analizadas usando *MADE-mlearn*; para a aumentar la replicabilidad de las experiencias y promover la escalabilidad de las mismas.

Algunas de estas líneas ya se están abordando desde el año 2017 en el marco del proyecto de investigación de Computación Móvil de la UNSE, codirigido por la tesista.

Referencias

- B'Far, R. (2004). *Mobile Computing Principle. Designing and Developing Mobile Applications with UML and XML*. Cambridge University Press. ISBN 9780521817332.
- Cabero Almenara, J. y Llorente Cejudo, M. (2013). *La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información y comunicación (TIC)*. Eduweb, 7(2).
- Challiol, C. (2012). *Apuntes de Curso de Posgrado sobre Computación Móvil*. Curso dictado en Universidad Nacional de Santiago del Estero. Secretaría de Extensión de la Facultad de Ciencias Exactas y Tecnologías.
- Cukierman, U., Gonzalez, A., Ibáñez, E., Iglesias, L., Palmieri, J., Rozenhauz, J., Santangelo, H., Sanz, C. y Zangara, A. (2007). *Una experiencia de uso de celulares en un curso de articulación escuela media y universidad en modalidad a distancia*. VirtualEduca, Brasil.
- Deroche, A., Acosta, M., Vegega, C., Bernal Tomadoni, L., Straccia, L., Pytel, P. y Pollo Cattaneo, M. (2015). *Propuesta de Desarrollo de Aplicación Móvil para la Evaluación Dinámica en Asignatura de Grado en Ingeniería en Sistemas de Información*. XVII WICC. ISBN 978-987-633-134-0. Salta.

- DeWitt, D., Siraj, S. y Alias, N. (2014). *Collaborative mLearning: A Module for Learning Secondary School Science*. Educational Technology & Society, 17 (1), 89–101.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki, Finland: Orienta-Konultit.
- Gallo Vargas, A. S. (2014). *Dispositivos móviles en actividades educativas colaborativas: análisis y recopilación de experiencias*. Trabajo integrador de la Especialización en Tecnología Informática aplicada a la Educación. FI, UNLP.
- Herrera, S., Fennema, M. C. y Sanz, C. V. (2012). *Estrategias de m-learning para la formación de posgrado*. En Anales del VII TEYET. Pergamino. Recuperado el 15 de diciembre de 2013, de: <http://hdl.handle.net/10915/18313>.
- Herrera, S., Morales, M.I., Fennema, M. C. y Sanz, C.V. (2014). *Aprendizaje basado en dispositivos móviles. Experiencias en la Universidad Nacional de Santiago del Estero*. ISBN 978-987-1676-18-7. EDUNSE. Sgo. del Estero.
- Herrera, S.I. y Fennema, M.C. (2011). *Tecnologías Móviles Aplicadas a la Educación Superior*. Congreso Argentino de Ciencias de la Computación. La Plata.
- Laurillard, D. (2002). *Rethinking university teaching. A conversational framework for the Effective Use for Educational Technology*. Ed. Routledge. 2nd Edition. New York. ISBN 0415256798.
- Lew, P., Olsina, L., Becker, P. y Zhang, L. (2011). *An integrated strategy to systematically understand and manage quality in use for Web App*. Requirements Engineering. 17(4), 299-330.
- Llitas, A. B. (2015). *Un enfoque de modelado de Actividades Educativas Posicionadas que contemplan elementos concretos*. Tesis de Maestría. Facultad de Informática. Universidad Nacional de La Plata.
- Maldonado Mahauad, J. (2015). *Desarrollo de un Marco de Análisis para la Selección de Metodologías de Diseño de Objetos de Aprendizaje (OA) basado en criterios de calidad para contextos educativos específicos*. Tesis de Maestría. La Plata. Recuperado el 19 de diciembre de 2015, de: <http://hdl.handle.net/10915/45063>.
- Najar, P., Ledesma, E., Rocabado, S., Herrera, S. y Palavecino, R. (2014). *Eficiencia de aplicaciones móviles según su arquitectura*. CACIC 2014. ISBN 978-987-3806-05-6. La Matanza, Buenos Aires.
- Navarro, C., Molina, A., Redondo, M. y Juárez-Ramírez, M. (2015). *Framework para Evaluar Sistemas M-learning: Un Enfoque Tecnológico y Pedagógico*. VAEP-RITA Vol. 3, Núm. 1. ISSN 1932-8540. Pp. 38-45.
- Nouri, J. (2012). *A theoretical grounding of learning mathematics in authentic real-world contexts supported by mobile technology*. IADIS Mobile Learning, pp. 35-42. ISBN 978-972-8939-66-3. 2012.
- Nouri, J., Cerratto-Pargman, T., Eliasson, J. y Ramberg, R. (2011). *Exploring the Challenges of Supporting Collaborative Mobile Learning*. International Journal of Mobile and Blended Learning, Volume 3 Issue 4, pp. 54-69. IGI Publishing Hershey, USA. ISSN: 1941-8647.
- Organista-Sandoval, J. y Serrano-Santoyo, A. (2011). *Implementación de una actividad educativa en modo colaborativo con apoyo de smartphones: una experiencia universitaria*. Edutec: Revista electrónica de tecnología educativa. ISSN 1135- 9250, N° 36, Pp. 3-19.
- Pachler, N., Bachmair, B. y Cook, J. (2010). *Mobile learning: structures, agency, practices*. Ed. Springer. Nueva York.
- Pachler, N., Seipold, J. y Bachmair, B. (2012). *Mobile Learning: Some Theoretical and Practical Considerations*. In The "My Mobile" Handbook. Guidelines and Scenarios for Mobile Learning in Adult Education, K. Friedrich, M. Ranieri, N. Pachler, and P. de Theux, Eds. 2012, pp. 11–16.
- Park, Y. (2011). *A Pedagogical Framework for Mobile Learning: Categorizing Educational Applications of Mobile Technologies into Four Types*. International Review of Research in Open and Distance Learning. ISSN 1492-3831. Vol. 12.2. Virginia Tech, USA.
- Pernici, B. (2006). *Mobile Information Systems. Infrastructure and Design for Adaptivity and Flexibility*. Germany: Springer-Verlag.
- Rocabado, S. H. (2014). *Caso de estudio de comunicaciones seguras sobre redes móviles ad-hoc*. Tesis de finalización de Maestría en Redes de Datos. Facultad de Informática, Universidad Nacional de La Plata.
- Sánchez Prieto, J. C., Olmos Migueláñez, S. y García Peñalvo, F. J. (2013). *Mobile Learning: Tendencias and Lines of Research*. TEEM '13, Salamanca, Spain. ACM 978-1-4503-2345-1/13/11.
- Seipold, J. y Pachler, N. (2011). *Evaluating Mobile Learning Practice Towards a framework for analysis of user-generated contexts with reference to the socio-cultural ecology of mobile learning*. Zeitschrift für Theorie und Praxis der Medienbildung ISSN 1424-3636. Themenheft Nr. 19: Mobile Learning in Widening Contexts: Concepts and Cases.
- Sharples, M., Taylor, J. y Vavoula, G. (2007). *Theory of learning for the mobile age*. In Andrews, R., & Haythornthwaite, C. (Eds.) The SAGE Handbook of E-learning Research. London: Sage, pp. 221-47.
- Talukder, A., Ahmed, H. y Yavagal, R. (2010). *Mobile Computing: Technology, Applications, and Service Creation*. Second Edition. McGraw-Hill Professional.
- Thomas, P., Cristina, F., Dapoto, S. y Pesado, P. (2015). *Dispositivos Móviles: Desarrollo de Aplicaciones Orientadas a Educación*. XVII Workshop de Investigadores en Ciencias de la Computación (WICC). ISBN 978-987-633-134-0.
- Traxler, J. (2011). *Mobile Learning: Starting in the Right Place, Going in the Right Direction?* International Journal of Mobile and Blended Learning, 3(2), 57-67, April-June 2011.
- Woodill, G. (2011). *The mobile learning edge*. Ed. Mc Graw Hill. ISBN: 978-0-07-173984-9.
- Zangara, A. (2014). *Apostillas sobre los conceptos básicos de educación a distancia o... una brújula en el mundo de la virtualidad*. Maestría en Educación a Distancia. Fac. de Informática de la Universidad Nacional de La Plata. La Plata.

Resumen Tesis Doctoral

Título:

Plataforma de Desarrollo de Laboratorios Remotos de Redes de Sensores Inalámbricos basados en Cloud Computing

Autor: Dr. Ing. Pablo Daniel Godoy

Director: Dr. Carlos García Garino

Co-Director: Dr. Ricardo Cayssials

Universidad donde se desarrolló la tesis: Universidad de Mendoza.

Doctorado aprobado por Resolución Ministerial nº 8/93, y acreditado por CONEAU por Resolución 210/11.

Fecha de defensa y aprobación: 4/8/2016

Resumen

Las redes de sensores inalámbricos, o WSN por sus siglas en inglés, son sistemas formados por un gran número de dispositivos denominados nodos. Estos nodos poseen capacidad de procesamiento de datos, memoria, capacidad de comunicación inalámbrica y sensores de diferentes tipos. En algunos casos, estos nodos tienen la capacidad de actuar sobre el medio ambiente que los rodea. El desarrollo de estos dispositivos ha sido posible gracias a los avances en el campo de los MEMS (Sistemas Micro Electromecánicos), que han permitido el desarrollo de componentes electrónicos de pequeñas dimensiones y bajo costo.

Las WSN poseen actualmente un gran número de aplicaciones, por ejemplo: monitoreo del medio ambiente, agricultura, monitoreo de la salud, automatización de procesos y fábricas, automatización de edificios, aplicaciones militares, etc. Se espera que el número de aplicaciones crezca en los próximos años, como también el número de WSN y tipos de nodos. Frente a los métodos tradicionales de medición empleados en estas aplicaciones (instrumentos y estaciones meteorológicas), las WSN permiten una mayor densidad espacial de puntos de medición. Además permiten el acceso en tiempo real a los datos y monitoreos de larga duración.

Los laboratorios remotos son sistemas que

permiten el acceso de manera remota a laboratorios de distinta naturaleza. Generalmente el acceso es a través de Internet, aunque puede ser a través de redes LAN de organizaciones o universidades. Existen en la actualidad laboratorios remotos de distinto tipo, por ejemplo, laboratorios remotos de física, de circuitos analógicos, de circuitos digitales, de robótica, de redes de computadoras, de WSN, etc (Marianetti, 2006; Steyn and Hancke, 2011). En general, son sistemas complejos, que incluyen un número considerable de componentes y de recursos tecnológicos.

Los laboratorios remotos constituyen una valiosa herramienta en los campos de la investigación científica, desarrollo de aplicaciones, formación y entrenamiento de usuarios y tareas de enseñanza. Estos permiten a los usuarios acceder a laboratorios con un gran número de equipos o instrumentos, o con equipamiento de costo elevado, que pueden estar ubicados en distintos lugares del mundo, al cual el acceso sería difícil o imposible por otros medios. Por ejemplo, existen laboratorios remotos de redes de computadoras con miles de nodos, que los usuarios difícilmente podrían montar localmente, pero que pueden ser accedidos de manera remota.

Los laboratorios remotos pueden clasificarse en dos grandes grupos: orientados a tareas de investigación científica o desarrollo de aplicaciones y orientados a tareas de

formación y entrenamiento de usuarios. En el caso de laboratorios remotos orientados a investigación científica, el usuario posee un nivel de conocimientos muy elevado, y puede llevar a cabo los experimentos sin ninguna restricción. Un laboratorio remoto orientado a tareas de formación y entrenamiento de usuarios está destinado a usuarios con variados niveles de conocimientos, e incluye componentes que facilitan el acceso y la configuración de experimentos por parte de los usuarios. En el caso de un laboratorio remoto destinado a formación y entrenamiento de usuarios sin conocimientos elevados, los experimentos que los usuarios pueden realizar están predefinidos y restringidos, de modo que no tengan que configurar completamente el experimento. Además, para el caso de tareas de docencia, los profesores o entrenadores necesitan poder realizar el seguimiento de los estudiantes, con lo cual surge la necesidad de componentes adicionales que se encarguen de esta tarea (Jona and Uttal, 2013).

Motivación

Existe un número considerable de publicaciones que reportan aplicaciones de las WSN en el campo de la agricultura. Por este motivo, en la provincia de Mendoza y la región, las WSN pueden llegar a convertirse en una valiosa herramienta en este campo. Pueden ser de utilidad en el estudio de diferentes fenómenos meteorológicos que afectan los cultivos, como ser heladas o eventos peligrosos para los cultivos, por ejemplo plagas. Prueba de ello es la existencia de una tesis de doctorado recientemente aprobada que emplea las WSN como herramienta en el estudio y predicción de heladas (Iacono, 2015). Las WSN pueden ser también de utilidad para el estudio del efecto de diferentes factores ambientales sobre la calidad de las frutas obtenidas, como también del efecto de los diferentes procesos de traslado, almacenamiento, procesamiento, etc. Además, ingenieros agrónomos o agricultores necesitan de datos espacial y temporalmente distribuidos para desarrollar sus tareas. Estos datos permiten aplicar de

manera correcta técnicas de cultivo (deshoje, estrés hídrico, etc) y maximizar sus beneficios. Para potenciar el uso de las WSN, de manera de permitir que éstas sean aplicadas ampliamente en la provincia y el país, nace la necesidad de contar con personal capacitado en el uso de las mismas. Por tanto, surge la necesidad de contar con herramientas que asistan en el proceso de formación y entrenamiento de los usuarios en el uso de las WSN.

El estudio del estado del arte de los laboratorios remotos de WSN permite ver que todos éstos han sido diseñados con propósitos de investigación científica. Éstos permiten a los usuarios acceder sin restricciones a los nodos de las WSN. Los usuarios pueden configurar el sistema operativo, los protocolos de comunicación, cargar componentes de software en los nodos, y realizar cualquier tipo de experimentos. Para realizar estas tareas, los usuarios necesitan contar con conocimientos elevados sobre WSN y programación.

Sin embargo, un usuario sin conocimientos avanzados necesitará realizar inicialmente experimentos simples, y luego, a medida que adquiera mayores conocimientos, podrá realizar experimentos más complejos. Por lo tanto, para este usuario será muy difícil realizar sus primeros experimentos y prácticas con WSN empleando los laboratorios de WSN disponibles actualmente. Se concluye que es de gran utilidad el estudio y desarrollo de laboratorios remotos de WSN destinado a usuarios que deseen formarse y entrenarse en el uso de las WSN, de fácil uso, destinados a usuarios sin conocimientos avanzados de WSN ni de programación. Por este motivo, esta tesis propone un modelo conceptual para laboratorios remotos de WSN que cumple estas características, y una implementación de dicho modelo.

Los mecanismos utilizados para permitir el acceso remoto a los equipos bajo prueba empleados por distintos laboratorios remotos son similares, diferenciándose fundamentalmente por el sistema bajo prueba. Por este motivo, las conclusiones y aportes que puedan realizarse en la implementación

de un laboratorio remoto de WSN, pueden ser útiles también para otro tipo de laboratorios remotos.

Pensando en los trabajos futuros que continuarán el trabajo presentado en esta tesis, se espera que el trabajo realizado sea el punto de partida de una línea de investigación más amplia y de largo plazo, destinada a estudiar la aplicación de laboratorios remotos y TIC en educación y desarrollo de aplicaciones. Un laboratorio remoto destinado a tareas de educación requerirá de herramientas especiales que permitan al docente realizar tareas de seguimiento y evaluación, no presentes en los laboratorios remotos de WSN existentes (Jona and Uttal, 2013). Esta tesis propone la implementación de módulos que sean de ayuda cuando este laboratorio remoto sea aplicado a tareas de docencia.

Como motivación adicional, se menciona el hecho de que los laboratorios remotos de WSN existentes actualmente emplean nodos sensores no disponibles en el mercado local. Por lo tanto, es de interés el desarrollo de un laboratorio remoto que emplee nodos sensores disponibles en el mercado local.

La tesis

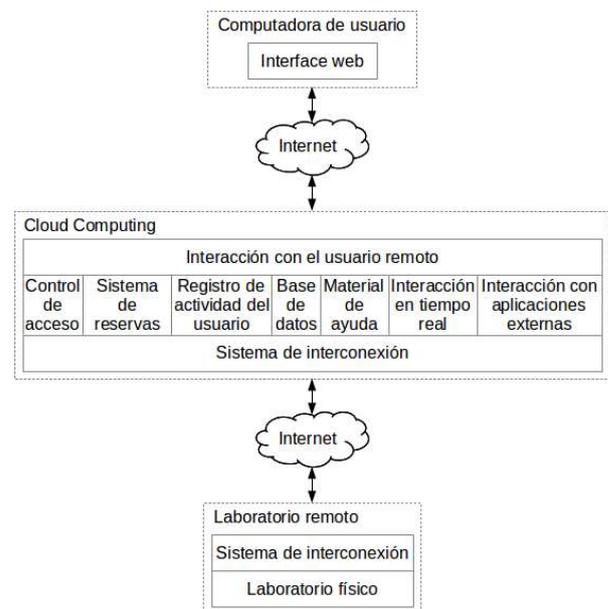
En esta tesis se propone un enfoque para la implementación de laboratorios remotos de WSN para ser empleados por usuarios que deseen formarse, entrenarse o experimentar en programación y funcionamiento de las WSN, que pueda ser empleado por usuarios con pocos conocimientos sobre WSN o programación, como también para la realización de experimentos o desarrollos. La hipótesis principal de esta tesis doctoral es: “Se pueden desarrollar laboratorios remotos de WSN, que puedan ser utilizados por usuarios que deseen formarse, entrenarse o experimentar en programación y funcionamiento de las WSN, o en desarrollo de aplicaciones que utilicen WSN, mediante el adecuado diseño e implementación de sus módulos de gestión”.

Para probar la hipótesis, se ha propuesto y desarrollado un modelo para laboratorios remotos de WSN que cumpla los requisitos planteados en la hipótesis. Como

componentes distintivos del modelo para laboratorios remotos de WSN propuesto respecto a los existentes, este modelo emplea una interface web que permite realizar experimentos de manera fácil y rápida, de manera que pueda ser empleado para poner a prueba conceptos básicos de las WSN, sin necesidad de poseer conocimientos avanzados en WSN o programación. Se ha realizado también una implementación y puesta en funcionamiento de un prototipo, empleando nodos de WSN disponibles en el mercado local.

Pensando en futuras aplicaciones del laboratorio remoto propuesto a tareas de educación, se ha planteado la incorporación de un módulo para seguimiento de los usuarios, basado en el almacenamiento de la actividad que realizan. Esta información debe almacenarse de manera segura y confiable, de modo que pueda ser utilizada para seguimiento y evaluación de la actividad que los usuarios realizan.

La siguiente figura muestra un esquema del modelo propuesto. Dicho modelo fue implementado para la realización de experimentos.



Experimentos realizados

Se realizaron experimentos de prueba de concepto, que son experimentos que los usuarios pueden realizar en el laboratorio

remoto. Se presentan estos experimentos en el formato en el cual los usuarios los realizarían, y una realización del experimento.

Estos experimentos incluyen:

- Tiempo de respuesta de la WSN
- 5.1.2 Potencia de RF recibida
- 5.1.3 Tiempo de respuesta en función de la ocupación del canal de comunicaciones

Además, se realizaron experimentos para evaluar las latencias introducidas por el laboratorio remoto, y se compara la latencia con y sin Cloud.

Aportes a la disciplina

Esta tesis presenta seis aportes en el campo de los laboratorios remotos de WSN y de laboratorios remotos en general, que se listan a continuación. Los dos primeros aportes son los aportes principales de esta tesis relacionados con la hipótesis enunciada en la sección. Los cuatro restantes, si bien no surgen directamente de la hipótesis planteada, también constituyen contribuciones al campo de los laboratorios remotos.

1) Una arquitectura para laboratorios remotos de WSN orientados a usuarios que deseen formarse, entrenarse o experimentar en programación y funcionamiento de las WSN, o en desarrollo de aplicaciones que utilicen WSN: Se han documentado en la bibliografía laboratorios remotos de WSN de diferentes tipos, pero ningún laboratorio remoto de WSN con la arquitectura propuesta en este trabajo. Esta arquitectura es el aporte principal de la tesis.

2) Una propuesta para el despliegue de laboratorios remotos empleando servicios de Cloud Computing: Se propone un nuevo modelo de implementación de laboratorios remotos no documentado en la bibliografía, el despliegue de varios componentes propios de un laboratorio remoto empleando servicios de Cloud Computing. Como se muestra en esta tesis, el empleo de Cloud Computing brinda ventajas propias de esta tecnología a los laboratorios remotos.

3) Una propuesta para interconectar dispositivos o equipos electrónicos al Cloud: Se ha propuesto un sistema de interconexión

entre los equipos del laboratorio físico y los módulos de gestión desplegados en el Cloud no documentado en la bibliografía, basado en una aplicación cliente-servidor. Dicho sistema podría ser utilizado para conectar cualquier dispositivo o equipo electrónico a componentes de software desplegados en el Cloud.

4) Construcción de una interface web didáctica: Esta interface permite a los usuarios programar la WSN de manera fácil y didáctica, de modo que los usuarios no necesiten conocimientos avanzados de programación ni WSN.

5) Una nueva clasificación de los laboratorios remotos: Se ha esbozado en la literatura la existencia de laboratorios remotos destinados a tareas de investigación científica, desarrollo o educación, pero ningún autor ha propuesto una clasificación formal ni ha realizado una comparación exhaustiva. Por este motivo, se propone en esta tesis una clasificación de los laboratorios remotos, distinguiendo entre laboratorios remotos destinados a investigación científica o desarrollo, y destinados a formación, educación y entrenamiento de usuarios.

6) Construcción de un prototipo que emplea nodos disponibles en el mercado local: Existe un número grande de nodos sensores comerciales disponibles en el mercado. Los laboratorios remotos de WSN construidos típicamente emplean nodos de la familia TelosB, que permiten a los usuarios programar completamente los nodos. Sin embargo, estos nodos no están disponibles en el mercado local, ni son los más adecuados para usuarios que aún no poseen los conocimientos de programación necesarios para programar este tipo de nodo. Por ello, el prototipo emplea nodos de fácil programación y disponibles en el mercado local.

5. Trabajos futuros

Los trabajos futuros se dividen en dos: posibles mejoras al laboratorio remoto desplegado y profundizar la investigación en cuanto a la integración de laboratorios remotos al Cloud.

5.1 Posibles mejoras

Entre los trabajos futuros se incluyen algunas ideas para mejorar o agregar funcionalidades al laboratorio remoto, que han surgido durante el proceso de implementación del modelo propuesto, durante la realización de experimentos, y la lectura de recientes publicaciones (Orduna et al., 2015). Estas mejoras posibles se describen a continuación.

5.2 Profundizar la investigación en cuanto a la integración de laboratorios remotos al Cloud

Del estudio del estado del arte y los trabajos relacionados, pudo observarse que la integración de laboratorios remotos con tecnologías de Cloud Computing es un campo con pocas y muy recientes publicaciones. Sin embargo, como pudo verificarse en esta tesis, el paradigma de Cloud Computing puede brindar características beneficiosas a los laboratorios remotos. Por otro lado, también se llegó a la conclusión de que en este campo hay varios problemas por resolver, como proponer y analizar distintas arquitecturas para la interconexión de los equipos bajo prueba con los componentes del laboratorio remoto implementados en el Cloud, disminuir las latencias introducidas, etc. En este trabajo se propuso un mecanismo para lograr la interconexión, pero pueden proponerse y estudiarse otros mecanismos y otros esquemas de interconexión.

5.3 Estudio de los aspectos pedagógicos para su aplicación a tareas de docencia

Como se mencionó anteriormente en las limitaciones, el laboratorio remoto ha sido diseñado de manera que una de sus aplicaciones futuras sea en tareas de docencia. Pero previo a emplear y presentar el laboratorio remoto como apto para esta aplicación, es necesario realizar un estudio de su aplicabilidad a tareas de docencia desde el punto de vista pedagógico.

Se propone como trabajo futuro realizar este

análisis, partiendo de un estudio del estado del arte sobre condiciones a cumplir y métodos para evaluar la aptitud de un laboratorio remoto, o herramienta tecnológica en general, para ser aplicado a docencia. Dicho estudio del estado del arte permitirá conocer las mejoras a implementar y los experimentos a realizar para poder aplicar el laboratorio remoto a tareas de docencia.

5.4 Aplicación del laboratorio remoto a tareas de docencia

Una vez realizado el trabajo futuro mencionado en el punto, se espera que el laboratorio remoto implementado pueda emplearse como una herramienta de enseñanza en la carrera de licenciatura en ciencias de la computación de la Universidad Nacional de Cuyo, la cual cuenta con una materia relacionada con las WSN. Además, luego de estudiar la aplicación del laboratorio remoto a otro tipo de equipos bajo prueba, se espera que el laboratorio remoto pueda ser empleado para su aplicación en enseñanza y docencia de otro tipo de equipo bajo prueba.

5.5 Implementación de un modelo simulado para interactuar con el laboratorio remoto

Las WSN tienen un gran número de aplicaciones, algunas de las cuales fueron descritas en la sección. Cada aplicación tiene características de funcionamiento diferentes en cuanto a aspectos como: tipos de datos, frecuencias de muestreo de datos, requisitos respecto a la forma de sensar datos, tipos de posibles fallas, tiempos de procesamiento de datos, acciones a tomar según los datos sensados, etc. Esto hace imposible que un laboratorio remoto de WSN pueda adaptarse a todas las aplicaciones posibles. Por otro lado, de acuerdo al estudio del estado del arte realizado en los capítulos [cha:Antecedentes] y [cha:RelatedWork], puede afirmarse que casi todos los laboratorios remotos de WSN presentados en la bibliografía están diseñados para realizar investigación sobre la programación y desarrollo de software para las WSN, y no sobre el comportamiento de las

WSN en distintas aplicaciones.

Se propone implementar diferentes tipos de modelos que interactúen con el laboratorio remoto a través de aplicaciones creadas para tal fin. Con este propósito fue incorporado en el modelo de capas del laboratorio remoto el módulo de interacción con aplicaciones externas. Este módulo permite la interacción con aplicaciones que pueden realizar diferentes acciones en función del comportamiento de modelos simulados. Estos modelos estarán implementados en software y serán ejecutados en el Cloud, para aprovechar una de las características constructivas más importantes del laboratorio remoto, el uso de servicios de Cloud Computing, que provee grandes capacidades de cómputo y almacenamiento de datos.

Referencias

- Jona, K. and Uttal, D. (2013). Don't forget the teacher: New tools to support broader adoption of remote labs. In Remote Engineering and Virtual Instrumentation (REV), 2013 10th International Conference on, pages 1–2.
- Marianetti, O. (2006). Laboratorios remotos, un aporte para su diseño y gestión. Universidad de Mendoza, Facultad de Ingeniería, Maestría en Teleinformática.
- Orduna, P., Gomez-Goiri, A., Rodriguez-Gil, L., Diego, J., Lopez-de Ipina, D., and Garcia-Zubia, J. (2015). wCloud: Automatic generation of WebLab-Deusto deployments in the Cloud. In Remote Engineering and Virtual Instrumentation (REV), 2015 12th International Conference on, pages 223–229.
- Steyn, L. P. and Hancke, G. P. (2011). A survey of wireless sensor network testbeds. In IEEE Africon'11.