Tecnologías Emergentes y Avances de la Computación en Mexico



SMCC


LUCHAR PARA LOGRAR
LOGRAR PARA DAR
UNIVERSIDAD AUTONOMA DE CHIHUAHUA


SMCC


UNIVERSIDAD TECNOLOGICA DE CHIHUAHUA
PENSAR Y CREAR
PARA SERVIR
UTCH

Tecnologías Emergentes y Avances de la Computación en México

Esta obra ha sido editada por la Universidad Autónoma de Chihuahua.

© Departamento Editorial, UACH
Ciudad Universitaria
Chihuahua, Chih., México

Editores:
Karina Caro
Karina Figueroa

Coordinación de contenidos:
Fernando Martínez

El presente monográfico se publica bajo una licencia *Creative Commons* del tipo: Reconocimiento – Compartir Igual



ISBN: 978-607-9424-94-7



Primera edición: Noviembre, 2016.

Editado en Chihuahua, México.







ENCUENTRO NACIONAL DE CIENCIAS DE LA COMPUTACIÓN

_____ENC 2016 _____

14, 15 y 16 de Noviembre, 2016

Chihuahua, Chihuahua, México

El ENC 2016 fue organizado por:

Sociedad Mexicana de Ciencias de la Computación, SMCC

Universidad Tecnológica de Chihuahua, UTCH

Universidad Autónoma de Chihuahua, UACH

Libro Electrónico, disponible en línea con acceso abierto en: www.smcc.mx

Asociaciones que apoyaron:

Asociación Mexicana de Interacción Humano-Computadora (AMEXIHC)

Sociedad Mexicana de Ciencias de la Computación (SMCC)

Editores:

Karina Caro
Karina Figueroa



Sociedad Mexicana de Ciencias de la Computación

http://www.smcc.mx

Comité Organizador

Karina Caro (CICESE, Organizador General)

Karina Figueroa (UMSNH, Organizador General)

Marcela D. Rodríguez (UABC, Presidenta SMCC)

Daniel León (UTCH)

Dynhora Ramírez (UTCH)

Fernando Martínez (UACH)

Organizadores de Talleres:

Soft Computing Applications:

Ángel G. Andrade (UABC)

Anabel Martínez (UPP)

Guillermo Galaviz (UABC)

Oscar Montiel (CITEDI)

Roberto Sepúlveda (CITEDI)

Computación Clínica e Informática Médica:

Ana I. Martínez García (CICESE)

Marcela D. Rodríguez (UABC)

Fernando Arámbula Cosío (CCADET, UNAM)

Alicia Martínez Rebollar (CENIDET)

Leo Joskowicz, (UNIVERSIDAD HEBREA DE JERUSALÉN, ISRAEL)

Inteligencia Artificial: Métodos y aplicaciones:

Félix Fernando González Navarro (UABC)

Juan Carlos Nieves Sánchez (UMEA UNIVERSITY)

Fernando Orduña Cabrera (ITSC)

Jorge Eduardo Ibarra Esquer (UABC)

Knowledge Discovery and Cloud Computing Applications:

Víctor Jesús Sosa Sosa (CINVESTAV)

Maricela Claudia Bravo Contreras (UAM)

José Alejandro Reyes Ortiz (UAM)

Network Systems and Protocols:

J. Antonio García Macías (CICESE)

Rolando Menchaca Méndez (CIC-IPN)

Organizadores de Consorcio de Posgrado:

Jéssica Beltrán Márquez (ITSON)

Vania Carolina Álvarez Olivas (UACH)

Organizador de Pláticas Invitadas:

Marcela D. Rodríguez (UABC)

Organizadores de Tutoriales:

J. Antonio García Macías (CICESE)

Luis A. Castro (ITSON)

Organizadores de Sesión de Carteles:

Pedro César Santana Mancilla (UCOL)

Juan Pablo García Vázquez (UABC)

Pablo Pancardo García (UJAT)

Organizadores de Feria de Posgrado:

Marco Antonio Ramírez Salinas (CIC-IPN)

Santiago Jaime Reyes Herrera (CIC-IPN)

Ernesto Urquidi Martínez (CIC-IPN)

Elvia Cruz Morales (CIC-IPN)

Comité Organizador Local

Daniel León (UTCH, Organizador General Local)

Dynhora Ramírez (UTCH, Organizador General Local)

Fernando Martínez (UACH)

Raymundo Cornejo (CONACYT-UACH)

Ricardo Guevara (UTCH)

Magdalena Campos (UTCH)

Andrés Pérez (UTCH)

Emilio Loya (UTCH)

Daniel Vázquez (UTCH)

Luis González (UACH)

Vania Álvarez (UACH)

Claudia Barraza (UACH)

Héctor Valenzuela (UACH)

Benjamín Cervantes (UTCH)

Marisol de León (UTCH)

Fátima López (UTCH)

Susana Bueno (UTCH)

Isela Ramos (UTCH),

Blanca González (UTCH)

Raquel Hernández (UTCH)

Rosario Vázquez (UTCH)

Melissa Baca (UTCH)

Cristina Barba (UTCH)

Eva Pérez (UTCH)

Alfonso Barroso (UTCH)

Addai Guerrero (UTCH)

Alejandro Enríquez (UTCH)

Jorge Campos (UTCH)

Eli Gómez (UTCH)

Jorge Chaparro (UTCH)

Ivonne Siqueiros (UTCH)

Roberto Aguilar (UTCH)

Evelyn Hinojos (UTCH)

Silvia Benites (UTCH)

Jorge Salomón (UTCH)

Francisco Cortes (UTCH)

Teresita García (UTCH)

Oscar Beltrán (UTCH)

Marco Orozco (UTCH)

Milton Batres (UTCH)

Dafnis Villagrán (UTCH)

Ángel Almanza (UTCH)

Roy Terrazas (UTCH)

Liliana Rivas (UTCH)

María del Mar Flores (UTCH)

Comité de Programa

Ángel G. Andrade	UABC
Fernando Arambula Cosío	CCADET – UNAM
Armando Arce	ITSON
Jessica Beltrán	CICESE
Maricela Bravo	UAM
Larisa Burtseva	UABC
Karina Caro	CICESE
José Carlos Carrasco	ITESM
Mario Chacón	ITCH
Edgar Chávez	CICESE
Laura Cruz	ITCM
Karina Figueroa	UMSNH
Carlos Flores	UCOL
Brenda L. Flores Rios	UABC
Salvador Flores-Frias	UABC
Guillermo Galaviz	UABC
Carlos Eric Galván Tejada	ITESM
J. Antonio García Macías	CICESE
Juan Pablo García Vázquez	UABC

Laura Gaytán	UCOL
Gabriel González	CENIDET
Félix F. González Navarro	UABC
Everardo Gutiérrez	UABC
Jorge Eduardo Ibarra-Esquer	UABC
Valeria Javalera	ITESCA
Roberto López Avitia	UABC
Iván Lopez-Arevalo	CINVESTAV
Gabriel A. Lopez-Morteo	UABC
Jaime Lopez-Veyna	SUPER COMPUTER CENTER, BARCELONA
Anabel Martínez	UPP
Leonardo Martínez	UAM
Ana I. Martínez García	CICESE
Fernando Martínez	UACH
Israel Marck Martínez	CICESE
Rolando Menchaca-Méndez	CIC
Hugo Mitre	CIMAT
José Mitre	IMP
Luis G. Montané-Jimenez	UV
Azucena Montes	CENIDET

Oscar Montiel	CITEDI
Rene Navarro	UNISON
Juan Carlos Nieves Sánchez	UMEA UNIVERSITY
Arturo Olvera-López	BUAP
Fernando Orduña Cabrera	ITESCA
Hugo Pablo	UAM
Pablo Pancardo García	UJAT
David Pinto	BUAP
Alberto Ramírez-Trevino	ITESCA
Nelson Rangel	ITCM
José Alejandro Reyes	UAM
Ana Rios	UAT
Jorge Rodas	UACJ
José Rodríguez	CINVESTAV
Marcela Rodríguez	UABC
Cesar E. Rose Gómez	ITH
Pedro C. Santana Mancilla	UCOL
Roberto Sepúlveda	CITEDI
Leonid Sheremetov	IMP
Víctor Jesús Sosa	CINVESTAV

Eric Sadit Tellez Avila	INFOTEC - CONACYT
Edgar Tello	UAT
Ma-Angelica Vargas-Acosta	UABC
Pilar Vidriales	IMP
Luis Vizcarra	UABC
Abigail Zamora-Hernández	UV
Vania Álvarez	UACH

Revisores Adicionales

Blanco Vega, Ricardo

Cruz-Sandoval, Dagoberto

Esquivel, Karime

Guerrero, Esteban

Moctezuma, Daniela

Navarro, Rene

Pacheco, Alberto

Pacheco, Nancy

Perez Castro, Cynthia Beatriz

Ramirez, Alberto

Ramirez-Alonso, Graciela

Soto-Mendoza, Valeria

Villavicencio, Jesus Rigoberto

Prefacio

Las diferentes ramas de la Computación no dejan de ser disciplinas en evolución que congregan a investigadores y practicantes en todo el mundo. La Sociedad Mexicana de Ciencias de la Computación (SMCC) tiene el propósito de fomentar la colaboración entre la comunidad de investigadores mexicanos, promoviendo sus líneas de investigación y dando a conocer sus contribuciones. Para lograr lo anterior, organiza el Encuentro Nacional de Ciencias de la Computación (ENC), también conocido en otras ediciones como Congreso Internacional de Ciencias de la Computación. Así, el ENC se distingue por ser el evento que anualmente organiza la SMCC. Este año el ENC se realizó del 14-16 de Noviembre en Chihuahua, Chihuahua., fungiendo como Instituciones sede: la Universidad Tecnológica de Chihuahua (UTCH) y la Universidad Autónoma de Chihuahua (UACH).

Con el objetivo de identificar e integrar a las comunidades científicas de distintas áreas de la Computación con presencia en México, el ENC'2016 consistió de artículos presentados en diferentes Talleres Temáticos, Consorcio de Posgrado y Sesión de Carteles, cuyo llamado se definió por los intereses de las comunidades activas en el país. De esta forma, el ENC quedó conformado principalmente por 5 talleres: Inteligencia Artificial: métodos y aplicaciones, Network Systems and Protocols, Computación Clínica e Informática Médica, Soft Computing Applications y Knowledge Discovery and Cloud Computing Applications. Adicionalmente, se contó con cuatro Ponencias Magistrales: Dr. Édgar Chávez (CICESE), Dr. Luis A. Pineda (IIMAS-UNAM), Dr. Gregorio Toscano (CINVESTAV-Tamaulipas) y Dr. Jason Nawyn (MIT). Además, el ENC´2016 se conformó de la Feria de Posgrado en Tecnologías de Información y Comunicaciones (TICs), Paneles de Discusión y Tutoriales, a los cuales asistieron estudiantes de educación media superior, superior y posgrado. La publicación y edición de este libro digital, con ISBN 978-607-9424-94-7, corre a cargo de la Universidad Autónoma de Chihuahua. Este contiene 56 trabajos de un total de 92, que fueron seleccionados por los respectivos comités de programas.

Enfatizamos nuestro agradecimiento a todos los miembros de la SMCC pertenecientes al CITEDI, CICESE, UMSNH, CENIDET, INFOTEC, ITESM, ITSON,

UABC, quienes con el apoyo de sus instituciones, fue posible realizar este evento. Agradecemos en particular a la Asociación Mexicana de Interacción Humano Computadora (AMEXIHC) por su gran colaboración y apoyo en las gestiones administrativas del ENC. Queremos también agradecer a los organizadores de Talleres, Consorcio de Posgrado, Feria de Posgrado, Tutoriales, Sesión de Carteles y a los Comités de Revisión, por su esfuerzo y tiempo invertido para asegurar la calidad de los trabajos publicados en este libro.

Agradecemos encarecidamente a la Universidad Tecnológica Chihuahua (UTCH), y a la Facultad de Ingeniería de la Universidad Autónoma de Chihuahua (UACH), por el apoyo institucional, recursos aportados, seguimiento y facilidades para difundir en los medios de comunicación el ENC'2016; especialmente: al C. Dr. Benjamín Palacios Perches, Rector de la Universidad Tecnológica de Chihuahua y al C. M.I. Javier González Cantú, Director de la Facultad de Ingeniería de la Universidad Autónoma de Chihuahua.

De la misma manera, agradecemos a las direcciones administrativas y académicas de la UTCH, así como a los departamentos de Planeación, Extensión Universitaria, Finanzas y Vinculación de ambas dependencias educativas, UTCH y UACH que brindaron su apoyo en las gestiones administrativas y de logística para realizar el ENC'2016 en sus instalaciones y por fomentar la participación de los estudiantes. Particularmente agradecemos al M.G.T.I. Andrés Pérez García (UTCH) Magdalena Campos (UTCH), Lic. Emilio Loya Miranda (UTCH), Lic. Ricardo Guevara (UTCH), Daniel Vázquez (UTCH), M.I. David Maloof Flores (UACH), Dr. Gilberto Wenglas Lara, M.I. Ana Lucía Corral Flores (UACH), y Lic. Heriberto Ramírez Luján, por todas las facilidades brindadas.

Agradecemos a la Lic. Ivonne M. Barriga Márquez directora de Turismo de gobierno del estado y a la Lic. Norma Rico, de área Chihuahua Buro de convenciones, por todas las facilidades para poder difundir el estado en este evento, así como a las instituciones y empresas que apoyaron como patrocinadores: Dirección de Turismo de Gobierno del Estado de Chihuahua, DataZone, NU Technology, Impresos León, AG Comercializadora E Industrializadora, Distribuidora JC FERRECONSTRUYE, SAMSARA, El Almacén Papelería, RAM Servicios Tecnológicos En Comunicación S DE RL, COMPUPLUS, COPY.

Agradecemos infinitamente al M.I. Daniel León y a la M.I. Dynhora Ramírez, quienes presidieron el Comité Local y al Dr. Fernando Martínez, miembro organizador, por todo el apoyo brindado en la organización de este evento. De igual forma, agradecemos a todos los profesores y estudiantes voluntarios de ambas instituciones que formaron parte de este Comité. A todos ellos un reconocimiento por su disposición incondicional durante la organización del ENC y por la hospitalidad brindada. Finalmente, agradecemos a todos los asistentes al ENC'2016 por el interés en participar y contribuir a que este evento sea el que caracterice a la comunidad científica en Ciencias de la Computación.

Noviembre del 2016.

Karina Caro
Karina Figueroa
Coordinadores del ENC'2016

Tabla de Contenido

Consorcio de Posgrado

Diseño de un Protocolo de Red para la Animación de Marionetas Digitales por Fusión de Sensores RGBD 1
Elisabet González, Alberto Pacheco

Implementación y Evaluación de Algoritmos de Localización Libres de Distancia 3
José Alfredo López Urrea, Erica Ruiz, Adolfo Espinoza Ruiz, Joaquin Cortez

Design of a Remote Monitoring System of Water Environmental Variables 6
Joaquin Mass Sánchez, Erica Ruiz, Adolfo Espinoza Ruiz, Armando García Berumen

Bit-Error Rate Evaluation of Hybrid MIMO LSTBC Receivers Using Limited-Precision QR Decomposition 8
Jorge Robles, Joaquin Cortez, Erica Ruiz, Adolfo Espinoza

Sistema de Visión para Conocer el Campo de Velocidades en un Fluido mediante el Trazado de Partículas. 10
Monica Valenzuela-Delgado, Wendy Flores-Fuentes, Miguel Bravo-Zanoguera, Alejandro S. Ortiz-Perez, Moises Rivas-Lopez, Felix F. Gonzalez-Navarro, Oleg Sergiyenko

Sincronización de Labios en Marionetas Digitales 12
Angelica Yañez, Alberto Pacheco, Rafael Vázquez

Sesión de Carteles

Estadística y Estudio Léxico-Sintáctico del Habla de la Lengua Yaqui para el Desarrollo de Herramientas de TI 15
José Manuel Barba Soto, Fernando Orduña Cabrera, Valeria Javalera Rincón

Sistema de Consulta de Datos Semánticos para Dispositivos Móviles 17
Frida Cardenas, Ana B. Rios-Alvarado, Alan Díaz Manríquez, Edgar Tello Leal

Predicción de la Precipitación Pluvial mediante Algorítmos Genéticos 20
Lorena Carmona, Daniel Cervantes-Guerrero, Andrés Martinez, Carlos Eric Galván Tejada, Jorge I. Galván-Tejada

Análisis y Desarrollo de Algoritmos Genéticos para la Alineación de Cadenas de ADN 22
Edith A. Escarrega Payan, Eduardo Vázquez-Santacruz

Implementación de Mapping Personalizado para Determinar la Relevancia en Elasticseach 24
Joyce Selene Anaid Lozano Aguilar, Alejandra García Hernández, Alejandro Mauricio González

Detección de Accidentes de Ciclismo Usando el Sensor de Aceleración: una Propuesta Preliminar Utilizando Redes Neuronales 26
Jonathan Martínez De Casas, Carlos Eric Galván Tejada, Jorge Issac Galván Tejada, José Maria Celaya Padilla

The Aggregative K-means Clustering Algorithm 28
Fernando Orduña-Cabrera, Valeria Javalera-Rincón, Alberto Ramirez-Treviño

Desarrollo Herramientas Virtuales de Terapia física y Rehabilitación mediante Intel RealSense: Un Enfoque Preliminar 30
María L. Ortiz-Pasillas, Cinthia B. Juárez Serrano, Kenia Miranda, Carlos E. Galván-Tejada, Jorge I. Galván-Tejada

Videojuego Serio para Promover la ráctica y Entendimiento de Conocimientos Básicos de Química 32
Jorge Alberto Sánchez Amaya, Carlos Eric Galván Tejada, Luis Miguel Rada, Jorge I. Galván-Tejada, Jose Maria Celaya Padilla, Juan Ruben Delgado Contreras

Reconocimiento y Ubicación de Rugas Palatinas mediante Redes Neuronales Artificiales y Visión por Computadora 34
Israel Velázquez-Almazán, Eduardo Vázquez-Santacruz, Ivet Gil-Chavarría

Búsqueda de Modelo Univariado en Características de Datos Descriptores de Imágenes para Cáncer de Mama con Diagnóstico Asistido por Computadora 36
Laura Zanella, Eric Galván, Jorge Galván, José Celaya

Taller: Soft Computing Applications

An Optimization based Approach for Maximizing the Information Content of Keypoints Detected on a Digital Image 39
Fernando Fausto, Erik Cuevas, Marco Pérez-Cisneros

Solución del Problema PATS con el Algoritmo de Recocido Simulado 45
Felipe Leza Álvarez, Israel Marck Martínez Pérez

On-line Path Planning for Mobile Robots using Parallel Evolutionary Artificial Potential Field 52
Ulises Orozco-Rosas, Oscar Montiel, Roberto Sepúlveda

Application of Genetic Algorithm for solving Fairness based Resource Allocation Problem in Mobile Systems 59
Adolfo Reyna-Orta, Ángel G. Andrade, Guillermo Galaviz

Optimización de Ubicaciones de Vehículos Aéreos no Tripulados para Proveer Comunicación en Situaciones de Desastre 64
Gabriela Lizeth Rodríguez Cortes, Anabel Martínez Vargas, Oscar Montiel Ross

Analysis of Taxi Destination Prediction Using a Neural Network 68
Normando Zubia, Luis-Carlos González-Gurrola

Taller: Computación Clínica e Informática Médica

LaborCheck: Sistema de Monitoreo Automático de Variables Usuario-Computadora y de Interacción Social 73
Misael Camaños, Wendy Sánchez, Alicia Martínez, Manuel Mejía-Lavalle

Interacción entre Adultos Mayores con Demencia y Robots de Compañia: Un Estudio Cuantitativo 80
Dagoberto Cruz-Sandoval, Luis González, Jesus Favela

Propuesta de Adquisición de Datos Para el Problema Inverso Electrocardiográfico 85
Esteban Herrera, Rafael Lemuz López, Carlos Guillén Galván, Ángel Ramos del Olmo

Implementación de un Modelo de Detección de Aislamiento Social en Adultos Mayores ... 90
Virginia Ortiz, Wilfrido Campos, Alicia Martinez, Hugo Estrada

Uso de Red Neuronal para Estimación del Estrés en Personas Jóvenes en Actividades Sedentarias ... 97
Jorge H. Robles, Cesar R. Encinas, Jesus G. Zarate, Joaquin Cortez, Jessica Beltrán, Luis A. Castro

Una Propuesta de Ontología de Desórdenes de Conducta en Infantes ... 101
Andrea Tamborrell Hernández, Hugo Lasserre Chávez, Maria Somodevilla García, Ivo Pineda Torres

Designing Visualization Tools to Support Older Adults Care Process ... 106
Amanda Tapia, Jessica Beltran, Valeria Soto-Mendoza, Karina Caro

Taller: Inteligencia Artificial: Métodos y Aplicaciones

Aplicación de un Algoritmo de Clasificación para Generar un Mapa de Radio para Localización en Interiores ... 112
Isabel Ambriz, Pablo García, Ángel Andrade, Luis Vizcarra

A Machine Learning Pipeline to Automatically Identify and Classify Roadway Surface Disruptions ... 117
Ezra Aragón, Manuel Ricardo Carlos Loya, Luis-Carlos González-Gurrola, Hugo Escalante

Automatic Characterization of Vertically Aligned Nanotubes by Image Analysis ... 122
Jesús Caro, Félix Fernando González, Oscar Peréz, Mario Alberto Curiel, Benjamín Valdez

Aceleración en GPU de un Algoritmo de Evolución Diferencial para el Problema de Inferencia en Redes Reguladoras de Genes ... 126
Angel Flores Olivas, Luis-Carlos González-Gurrola, Eduardo Rodriguez-Tello

Consistent Transformations of Knowledge Bases in ASP Part I ... 130
Juan Acosta

Silla de Ruedas Autónoma e Inteligente que Avanza a Lado del Acompañante ... 138
Luis A Jara, Fernando Orduña, Alberto Ramirez

Multi-objective Optimization for Local Descriptors in Image Mosaicing ... 143
Cynthia Beatriz Perez Castro, Daniel Mendoza

Captura Inteligente de Imágenes de Actas Electorales mediante Dispositivos Móviles ... 147
Alexis Plata, Victor Morales

SVM to Infer Drivers Distracted Behavior: a Feasibility Experiment ... 151
Marcela D. Rodríguez, Antonio Ordorica, José M. Ramírez, Luis A. Castro, Jessica Beltran, Ángel G. Andrade

Uso del Lenguaje Natural para la Búsqueda de Respuestas en un Sistema de Gestión de Conocimiento Académico 155
Cesar Enrique Rose-Gomez, María Trinidad Serna-Encinas, Oscar Mario Rodríguez-Elías

Taller: Knowledge Discovery and Cloud Computing Applications

An Approach for the Extraction and Visualization of Semantic Knowledge in Private Cloud Storage 162
Leyla Alvarez-Medina, Ivan Lopez-Arevalo, José Luis González, Ana B. Rios-Alvarado

Web Services for Controling User's Environments Through Ontologies 167
Jose A. Bernal, Jose Rodriguez

An ABC-based Approach for Web Services Clustering 172
Maricela Bravo, Roman Mora, Roberto Alcantara, Ivan Flores

Addressing Clarity and Coherence during Ontology Construction 177
Maricela Bravo, José A. Reyes-Ortiz, Roberto Alcántara, José Rodriguez

Semantic Analyzer for Spanish, Using Ontologies 183
Alma Delia Cuevas Rasgado, Yedid Erandini Niño Membrillo, Farid Garcia Lamont

Traducer of Letters of the Mexican Sign Language to Text Using Leap Motion 191
Gil Diaz, José Guadalupe Rodríguez

An Ontological Model for Business Process Domain 196
Tania Guerrero, Edgar Tello, Luis de La Cruz

Text Mining System for Information Extraction and Classification of Scientific Documents 200
Ana B. Rios-Alvarado, Edgar Tello Leal, Fabian A. Ramirez-Tovar

SEXWISE: An IBM Watson-Powered Mobile Application to Promote Sexual Education .. 205
Miguel Royo-Leon, Laura M. Rodriguez, Paola Gallardo, Timothy Gonzalez, Rosalyn Negron, Diego Aguirre, Olac Fuentes

VM-into-PM: A Ball-into-bins Scheme for Virtual Machine Allocation in Private Cloud ... 211
Víctor J. Sosa-Sosa, Jose Luis González Compeán, Mario Gómez-Rodríguez, Edwing Aldana-Bobadilla

Dos Aproximaciones para Expandir Consultas en un SRIB usando Sinónimos 216
Ana Laura Lezama Sánchez, Mireya Tovar Vidal, Darnes Vilariño Ayala

Algoritmos de Aprendizaje Supervisados para el Análisis de Sentimientos en el Dominio de Restaurantes y Laptops 222
Karen Leticia Vazquez Flores, Mireya Tovar Vidal, David Pinto

Taller: Network Systems and Protocols

An Agent based Framework for Multi-sink Wireless Sensor Networks 229
Miriam Carlos-Mancilla, Ernesto López-Mellado, Mario Siller

Cifrado Caótico Parcial para Comunicaciones de Video en Tiempo Real 235
Rogelio Hasimoto-Beltran, Marcos Calderon-Calderon, Salvador Ruiz-Correa

Implementación de Canal Seguro para la Protección de imágenes de Actas de Escrutinio . . 241
Oscar Ruiz Hernández, Víctor Morales Rocha

A Li-Fi Based Application with an Android Device 245
Sergio Sandoval-Reyes, Victor Hugo Herver-Segura

Consorcio de Posgrado



Organizadores:
Jéssica Beltrán Márquez (ITSON)
Vania Carolina Álvarez Olivas (UACH)

Diseño de un Protocolo de Red para la Animación de Marionetas Digitales por Fusión de Sensores RGBD

Elisabet González Juárez
División de Estudios de Posgrado e Investigación
Instituto Tecnológico de Chihuahua
Chihuahua, México
+52(614)2012000
eli.jvn@gmail.com

Alberto Pacheco González
División de Estudios de Posgrado e Investigación
Instituto Tecnológico de Chihuahua
Chihuahua, México
+52(614)2012000
apacheco@itch.edu.mx

RESUMEN

Este trabajo de investigación aborda el diseño de un protocolo de red para generar animaciones vectoriales en base a la captura de movimiento en tiempo real usando múltiples sensores RGBD. El protocolo forma parte de la arquitectura de la plataforma FSA-MD, usando una topología apropiada para reducir la latencia del flujo de datos. Se emplea un servidor dedicado para adquirir los datos derivados de los sensores RGBD provenientes de los clientes de captura, para ser fusionados y transmitidos a los clientes heterogéneos para visualizar animaciones. Las pruebas preliminares muestran que el protocolo WebRTC es una opción viable porque permite adaptar una topología estrella multipunto de tiempo real. Se espera que esta tecnología, aunada al diseño del protocolo de red específico para la aplicación de fusión de datos, sea validado de manera efectiva en etapas posteriores con clientes basados en distintas plataformas y dispositivos móviles.

Conceptos CSS

• **Computing methodologies → Artificial intelligence → Computer vision → Image and video acquisition → Motion capture**

Palabras clave

Captura de movimiento; Protocolos de red de tiempo real; Marionetas digitales; Fusión de datos en redes de sensores

1. INTRODUCCIÓN

Tomando como referencia la arquitectura de la plataforma de fusión de sensores RGBD para la animación de marionetas digitales (FSA-MD) publicada en [5], se propone el diseño e implementación de un protocolo de red capaz de integrar y fusionar múltiples flujos de datos multimedia capturados desde diferentes sensores RGBD en tiempo real. Con esto se busca manipular una marioneta digital por medio de uno o más sensores RGBD, y así mismo dar soporte a múltiples clientes multimedia en diversas plataformas para visualizar dichas marionetas. Los sensores RGBD obtienen información de color y profundidad y permiten llevar a cabo la captura de movimiento, esto es, muestrear y registrar movimientos [3]. En los últimos años se han desarolllado diversas aplicaciones con estos sensores, algunas de las cuales se mencionan en [5]. Una de esas aplicaciones es la animación en tiempo real de marionetas digitales, las cuales son personajes animados que se utilizan en la producción de narrativas digitales [4]. Este tipo de aplicaciones requiere la integración y fusión de datos de varios sensores RGBD. En este trabajo se propone utilizar un protocolo de red de tiempo real basado en WebRTC para la fusión de datos, para transferir no un flujo de cuadros de video sino un flujo de vectores de las características extraídas de la red heterogénea de sensores RGBD.

El objetivo de esta investigación es diseñar un protocolo de red dentro de un sistema basado en la plataforma FSA-MD que permita manipular en tiempo real una marioneta digital a través de una red heterogénea de sensores RGBD, comunicados mediante dicho protocolo de red. Así también, se busca validar el protocolo de red mediante pruebas piloto utilizando diversos clientes multimedia para la reproducción de animaciones digitales en múltiples plataformas.

2. TRABAJOS RELACIONADOS

Existen algunos trabajos para el control de marionetas digitales 3D en los cuales se crean avatares sobre ambientes virtuales [1]. Por otro lado, algunos trabajos se han centrado en animaciones 2D [4]. Uno de estos trabajos se desarrolló en el I. T. de Chihuahua en donde se creó un sistema compuesto por un sensor RGBD y un reproductor de animación de marionetas [4]. Existen otros trabajos que utilizan varios sensores RGBD y para comunicarlos utilizan un protocolo de red [2]. Para esto hay distintos tipos de protocolos, siendo los protocolos de tiempo real los más adecuados para la transmisión de flujos multimedia, tal es el caso del trabajo presentado en [7]. Como parte de estos protocolos se encuentra WebRTC el cual es un protocolo diseñado para aplicaciones de transmisión multimedia en tiempo real, especialmente para sistemas de videoconferencias. Esta tecnología ha sido muy utilizada en los últimos años para diversas aplicaciones [6]. A pesar de que existen varios trabajos relacionados con el tema de investigación, no se ha encontrado alguno que haga una fusión e integración de datos obtenidos por una red heterogénea de sensores RGBD para aplicaciones de tiempo real.

3. METODOLOGÍA

A continuación se muestran las diversas etapas del presente proyecto de investigación de tesis:

1) Definición del problema y revisión de literatura. Se establecieron las hipótesis, se hizo una selección, clasificación, análisis y síntesis de la información consultada y se diseñó la plataforma FSA-MD para identificar los puntos a resolver.

2) Análisis, diseño e implementación del módulo de captura de movimiento y reproductores de animación de marionetas. Los módulos de captura o clientes de captura (CC) adquieren la información generada por los sensores RGBD para después identificar los puntos del esqueleto definido por cada sensor para generar un vector de posiciones de cada articulación. Para los reproductores de animación de marionetas o clientes de visualización (CV) cada marioneta debe tener definido cuáles son los puntos de rotación de las articulaciones (pivotes) así como su rango de movimiento. Una vez definido esto, los CV deben leer la secuencia de movimientos adquirida por los CC.

3) Análisis, diseño e implementación de APIs y servicios web. Las funciones que describen los movimientos de las marionetas digitales se definen en una API, la cual debe implementarse posteriormente en los CC y CV.

4) Análisis, diseño e implementación del protocolo de tiempo real para la fusión de sensores RGBD. Por medio del protocolo de flujo de captura, el servidor de fusión de datos recibe los flujos multimedia provenientes de la red heterogénea de sensores RGBD, los cuales son previamente adquiridos y procesados por los CC. Posteriormente se realiza la fusión e integración de dichos flujos para envíar los paquetes de información por medio del protocolo de flujo de cuadros de animación a los CV.

5) Pruebas preliminares y finales. Se han hecho pruebas piloto con alumnos de algunas instituciones educativas para producir narrativas digitales utilizando las marionetas digitales. También, es necesario realizar pruebas para evaluar el rendimiento y calidad del protocolo y el servidor de fusión de sensores RGBD.

4. RESULTADOS LOGRADOS

Se ha concluido con la definición del problema (etapa 1), obteniendo el diseño de la plataforma FSA-MD que consta de 3 componentes principales: módulos de captura de movimiento, un servidor para la fusión de datos RGBD y reproductores de animación de marionetas digitales. Para los CC y CV (etapa 2) se definió la estructura de las marionetas digitales y se desarrollaron 2 métodos para la elaboración de una marioneta digital, logrando automatizar algunas partes del proceso reduciendo los tiempos de elaboración. En cuanto al diseño de los CC y CV se hizo la implementación de uno de ellos en Processing. En cuanto al diseño de APIs (etapa 3), se definió una API para los movimientos del cuerpo de la marioneta, la cual se implementó en JavaScript para el Live Editor Avatar Demo[1]. Con este editor interactivo fue posible probar el funcionamiento de las marionetas digitales detectando pivotes faltantes, límites incorrectos para el rango de movimiento de las articulaciones y defectos en los movimientos definidos en las funciones del API. Finalmente se hizo la elección del protocolo a utilizar para el servidor de fusión de datos (etapa 4). Se eligió el *framework* WebRTC que está enfocado en la transmisión multimedia en tiempo real entre navegadores web y actualmente es utilizado en aplicaciones para conferencias y videoconferencias.

5. RESULTADOS ESPERADOS

Se busca implementar un servidor web en una tarjeta Raspberry Pi para realizar pruebas con WebRTC. En este servidor se espera sea posible instalar un certificado de seguridad para poder utilizar WebRTC el cual requiere de un servidor con este tipo de seguridad. El protoclo basado en WebRTC se implementará bajo una topología tipo estrella y un servidor dedicado o MCU (*Multipoint Control Unit*), el propósito de este tipo de arreglo es identificar a cada participante de la conferencia como CC o CV. El mayor trabajo se centrará en el MCU, el cual deberá asignar un identificador a cada miembro de la conferencia, determinar el tipo de cliente que es y adquirir y fusionar los datos. Las pruebas de este arreglo se harán primeramente con texto y después se trabajará con audio como fuente multimedia para resolver los detalles de codificación y evaluar la calidad de transmisión del protocolo.

6. CONCLUSIONES

Los resultados preliminares de este trabajo parecen indicar que es viable utilizar WebRTC para el diseño del protocolo propuesto en este trabajo. Además, la versatilidad que tiene WebRTC para soportar distintas plataformas y su enfoque en transmisión multimedia en tiempo real genera un gran aporte al diseño del protocolo de red ya que son dos elementos importante que este trabajo de investigación busca cubrir. Al trabajar con WebRTC se ha encontrado que solo se puede utilizar en ambientes seguros, por lo cual es necesario instalar certificados de seguridad en los servidores web. A pesar de esto, se ve posible adaptar WebRTC en el protocolo de red para lograr la animación de marionetas digitales mediante fusión de sensores RGBD. El reto para la etapa final de este proyecto consistirá hacer la adaptación de WebRTC a la propuesta de diseño del protocolo de red, esperando sintetizar el flujo de *frames* de animación en tiempo real tomando como base un innovador y eficiente formato basado en gráficos vectoriales.

7. REFERENCIAS

[1] Bholsithi, W., Wongwaen, N. and Sinthanayothin, C. 2014. 3D Avatar Developments in Real Time and Accuracy Assessments. In *2014 International Computer Science and Engineering Conference* (July 30 - August 1, 2014). ICSEC 2014. 85-90.

[2] Cassola, F., Paredes, H. and Fonseca, B. 2014. Online-Gym: Multiuser Virtual Gymnasium Using RINIONS and Multiple Kinect Devices. In *2014 6th International Conference on Games and Virtual Worlds for Serious Applications* (September 09 - 12, 2014). VS-GAMES 2014. 25-28.

[3] Ijjina, E. P. and Mohan, C. K. 2015. Human Action Recognition based on Motion Capture Information using Fuzzy Convolution Neural Networks. In *2015 Eighth International Conference on Advances in Pattern Recognition* (January 04 - 07, 2015). ICAPR 2015.

[4] Pacheco, A., Ramírez, M. and Guzmán, C. L. 2015. Marionetas Digitales: Tecnología Emergente para Narrar Historias con Personajes Animados Mediante la Captura Digital de Movimiento. In *Tecnologías Emergentes en la Educación*. Pearson. Distrito Federal, 161-175.

[5] Pacheco, A., Robledo, I. and Gonzalez, E. 2016. Plataforma de Fusión de Sensores RGB-D para la Animación de Marionetas Digitales. *Revista de Tecnologías Emergentes en la Educación*. 1, 1 (Jun. 2016), 101-106.

[6] Wenzel, M. and Meinel, C. 2016. Full-Body WebRTC Video Conferencing in a Web-Based Real Time Collaboration System. In *Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Desing* (May 04 - 06, 2016). CSCWD 2016. 334-338.

[7] Zhen-ping, F. and Kang, B. 2015. Analysis and implementation of streaming media system based on RTP and MPEG-4. In *2015 4th International Conference on Computer Science and Network Technology* (December 19 - 20, 2015). ICCSNT 2015. 1286-1289.

[1] http://podcast.itch.edu.mx/live-editor/

Implementación y Evaluación de Algoritmos de Localización Libres de Distancia

José López U.
Instituto Tecnológico de Sonora
Sonora México
2nd line of address
jlopezur@gmail.com

Erica Ruiz I.
Instituto Tecnológico de Sonora
Sonora México
2nd line of address
erica.ruiz@itson.mx

Adolfo Espinoza R.
Instituto Tecnológico de Sonora
Sonora México
adolfo.espinoza@itson.edu.mx

Joaquín Cortez G.
Instituto Tecnológico de Sonora
Sonora México
joaquin.cortez@itson.edu.mx

RESUMEN
Una gran gama de aplicaciones han sido diseñadas para las redes inalámbricas de sensores y actuadores (WSAN). Con el paso del tiempo, ha surgido la necesidad de localizar la fuente de información de una WSAN para la solución inmediata de diferentes problemas. Gracias a esta problemática, han surgido diferentes técnicas de localización de bajo costo y sin la necesidad de hardware adicional. Estas técnicas de localización obtienen errores de localización muy atractivos en simulación, sin embargo, surge la necesidad de observar el desempeño de estas técnicas en un escenario real exterior para observar si es factible la implementación de estos en casos reales. Este documento describe el análisis realizado sobre el canal de transmisión para la implementación de un sistema de localización, así como la el sistema de localización propuesto en hardware para evaluar el desempeño de algoritmos de localización en escenarios reales.

Keywords
WSN, WSAN, LOCALIZACIÓN, RSSI.

1. INTRODUCCIÓN
Las WSAN están compuestas de sensores y actores distribuidos en un área geográfica de interés. Los sensores son dispositivos de poca capacidad de procesamiento, muy bajo consumo energético y bajo costo, responsables de monitorizar el ambiente físico, mientras que los actores ejecutan una tarea acorde a los datos colectados y reportados por los sensores durante un evento [1]. En WSAN es muy importante conocer la posición de los nodos sensores, ya que permite determinar la ubicación geográfica de un evento y responder oportunamente a éste, además de que facilita el enrutamiento a través de la red y reduce el consumo de energía de los nodos. La localización precisa de los sensores es un requerimiento crítico en una amplia variedad de aplicaciones en WSAN [2], [3] tales como rastreo de animales, el cual permite analizar el comportamiento animal y las interacciones con otras especies; logística, donde los sensores reportan su posición cuando requieren ser encontrados; en salud, donde la ubicación de equipo médico y empleados es necesario en casos de emergencias y en el monitoreo de la evolución espacial de un fenómeno extraordinario entre otros [4]. A la estimación de la posición de un nodo de interés (NOI), se le conoce como localización.

La localización en WSANs se describe acorde a un sistema de coordenadas, definido por nodos de referencias con posiciones conocidas [5]. De manera general en la literatura los algoritmos de localización se clasifican en basados en distancia y libres de distancia (o basados en conectividad). Los primeros asumen que la fuerza de la señal decae conforme aumenta la distancia, de esta manera, las mediciones de la fuerza de la señal se convierten en distancias estimadas y éstas a su vez se emplean para estimar la posición del NOI. Estas técnicas presentan resultados muy precisos, sin embargo, requieren hardware especializado lo que las hace más costosas en redes grandes [6]. Si no es posible estimar la distancia entre nodos o si ésta tiene errores muy grandes se emplean los algoritmos libres de distancia los cuales utilizan información de conectividad. Estos algoritmos asumen que el rango de transmisión es constante o que la distribución de los nodos es uniforme y conocida a priori. Esto significa que el desempeño depende de la diferencia entre los valores esperados y reales del rango de transmisión y la distribución de los nodos [7]. Estas técnicas disminuyen su precisión respecto a las basadas en distancia pero su implementación es sencilla y de bajo costo.

2. TRABAJO RELACIONADO
La gran ventaja de la utilización de la estimación de distancia basada en RSS (Received Signal Strength) ha provocado que la comunidad investigadora se enfoque en su investigación. En el estado del arte existen múltiples investigaciones que demuestran la gran ventaja de la utilización de estas técnicas, sin embargo, también existen estudios que muestran la inestabilidad de la potencia de la señal recibida en escenarios reales. Estudios como [8] realizan experimentos con 2 nodos sensores Crossbow en un escenario ideal al minimizar los factores que pueden causar variaciones en los niveles de potencia transmitida como pueden ser: obstáculos, redes existentes sobre la misma frecuencia, el movimiento de personas dentro del área de despliegue y la falta de línea de vista entre los nodos. Los datos mostrados por los autores exhiben un comportamiento deseable en escenarios de simulación, pero al trasladar el experimento a escenarios reales, se puede observar cómo esta métrica varia de forma inesperada. Estos comportamientos pueden ser causados por scattering o shadowing de la señal, haciendo así, que la potencia percibida por el dispositivo no concuerde con la simulación, por esta razón, los autores establecen que no es viable la utilización de esta métrica en sistemas de localización comerciales. Otro estudio de gran interés es [9], en el cual realizan pruebas con 256 nodos sensores Texas Instrument en un escenario interior, donde se tomaron medidas de potencia para la estimación de la distancia y se comparan contra un escenario de simulación, los datos obtenidos en esta investigación, son favorables para la utilización de esta

métrica por la estabilidad observada en los experimentos. Gracias a esto, se establece la necesidad de observar el comportamiento de los algoritmos de localización en escenarios reales con módulos XBee por la sencillez y el bajo costo de estos dispositivos, y determinar los escenarios en los cuáles es posible aplicar técnicas de RSS.

3. OBJETIVO DE LA INVESTIGACIÓN

Por medio de esta investigación se pretende observar el comportamiento de la estimación de la distancia por medio de RSS en un escenario real, con el fin de observar la fiabilidad de la métrica RSS como estimador de distancia para la base de utilización en algoritmos de localización libres de distancia. Por lo que este trabajo tiene por objetivo desarrollar un sistema de localización en hardware, con la finalidad de evaluar el desempeño de algoritmos de localización y determinar cuál de estos presenta mejor desempeño bajo diferentes escenarios de evaluación. El sistema de localización propuesto consta de cuatro nodos anclas y un nodo móvil basados en tecnología Arduino y trasceivers XBee, con el cual se evaluarán en escenarios reales, en ambientes exteriores diferentes algoritmos de localización a un salto, libre de distancia, publicados en la literatura.

4. METODOLOGÍA

Como parte de la metodología empleada en este trabajo primeramente se realiza una investigación del estado del arte, acerca de los sistemas de localización físicos propuestos a la fecha así como la evaluación de algoritmos de localización. Se define el alance del proyecto, particularmente se seleccionan los algoritmos de localización a evaluar así como los escenarios de prueba. Enseguida se evalúan dichos algoritmos en escenarios ideales a través de simulaciones en la plataforma MATLAB. Se realiza un análisis estadístico, a través de la metodología 2K factorial [11] para determinar el impacto que tienen diversos factores como: el modelo de propagación, niveles de ruido, densidad de nodos, y relación de nodos ancla por densidad de nodos, en el desempeño de los algoritmos en cuestión. Posteriormente se implementa en hardware el sistema de localización propuesto, para ello se mide el patrón de radiación de los nodos transceptores, se implementa un protocolo de comunicación para sincronizar la transmisión entre los nodos de referencia y el nodo de interés, se implementa la red de nodos y se programan los algoritmos en los sistemas arduino y se configuran los transceivers Xbee. Como siguiente etapa se evalúa el desempeño de los algoritmos de localización seleccionados en el sistema propuesto, para lo cual es necesario definir la cama de pruebas, realizar las mediciones de los parámetros de desempeño y analizar los resultados. Finalmente se documenta a detalle el procedimiento y resultados de la investigación.

5. ESTADO DE LA INVESTIGACIÓN

Para la correcta estimación de la posición del nodo desconocido, primero se observa el comportamiento de RSS en el área a caracterizar. Esto es importante debido a que se hace una representación física del canal de transmisión, lo cual permite tener conocimiento previo del comportamiento del canal a utilizar. Los parámetros observados en este experimento fueron el patrón de radiación y la isotropía del canal. Las lecturas de RSS se realizaron en una red sensores implementada en las instalaciones del Instituto Tecnológico de Sonora. Esta red cuenta con cinco dispositivos; cuatro nodos ancla y un nodo móvil, encargado de recibir las señales de los cuatro nodos ancla. Estas mediciones fueron tomadas en un grid de 10x10, con una separación de un metro entre cada punto, dando un total de 121 mediciones. Para utilizar la métrica de RSS como un estimador de distancia, se asume que la radiación de la antena es isotrópica, es decir, que trasmite la señal a una potencia igual a todas las direcciones. Esta propiedad de isotropía, fue estudiada por medio de un experimento donde se colocó el nodo transmisor en el centro de una circunferencia con radio de 1 metro mientras que el receptor, era trasladado alrededor de esta circunferencia con una separación de 10° entre medición. Los resultados de este experimento se observan en la Figura 1 y 2.

6. RESULTADOS ESPERADOS

Con los resultados obtenidos del patrón de radiación y la isotropía de las antenas de los módulos Xbee, se pretende implementar un sistema de localización que sea capaz de estimar la distancia a partir de las lecturas de RSS obtenidas en tiempo real. Estos resultados servirán para observar la fiabilidad de la utilización de la métrica RSS en escenarios reales con la finalidad de determinar si esta métrica funciona como base para la estimación de la distancia en un sistema de localización comercial.

7. CONCLUSIONES Y TRABAJO FUTURO

El canal de transmisión observado en esta investigación proporciona la base necesaria para el despliegue de un sistema de localización, como trabajo futuro, se pretende programar diferentes algoritmos de localización libres de distancia en dispositivos Arduino y XBee debido a su bajo costo y capacidad de procesamiento necesario para los mismos. Es importante observar el comportamiento de los dispositivos utilizados en el canal de transmisión que se pretende utilizar pues, aunque los desarrolladores de los dispositivos afirmen que trabajan de cierta forma, las redes inalámbricas suelen depender directamente del medio donde se utilicen.

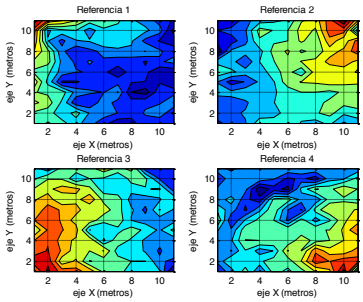


Figura 1. Patrón de radiación obtenido.

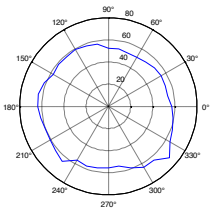


Figura 2. Isotropía del canal de transmisión.

8. REFERENCIAS

[1] D. Vassis, G. Kormentzas, and C. Skianis, “Performance evaluation of single and multi-channel actor to actor communication for wireless sensor actor networks,” Ad Hoc Networks, vol. 4, no. 4, pp. 487–498, 2006.

[2] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, “Locating the nodes: cooperative localization in wireless sensor networks,” IEEE Signal Process. Mag., vol. 22, no. 4, pp. 54–69, 2005.

[3] G. Mao, B. Fidan, and B. Anderson, "Wireless sensor network localization techniques," Comput. Networks, vol. 51, no. 10, pp. 2529–2553, 2007.

[4] L. Castro and J. Favela, "Contiuous Tracking of User Location in WLANs Using Recurrent Neural Networks," in ENC 2005, 2005, pp. 174–181.

[5] V. H. Perez-Gonzalez, D. Munoz-Rodriguez, C. Vargas-Rosales, and R. Torres-Villegas, "Relational position location in ad-hoc networks," Ad Hoc Networks, vol. 24, no. PA, pp. 20–28, 2015.

[6] A. Savvides, C.-C. Han, and M. B. Srivastava, "Dynamic Fine-Grained Localization in Ad-Hoc Wireless Sensor Networks," Cent. Embed. Netw. Sens., 2001
B. J. Dil, P. J. M. Havinga, B. J. Dil@utwente, Nl, and P. J. M. U. Havinga@ewi, "COM-LOC: A Distributed Range-Free Localization Algorithm in Wireless Networks.

[7] Parameswaran, Ambili Thottam, Mohammad Iftekhar Husain, Shambhu Upadhyaya, y others. "Is rssi a reliable parameter in sensor localization algorithms: An experimental study". En Field Failure Data Analysis Workshop (F2DA09), 5, 2009.

[8] Heurtefeux, Karel, y Fabrice Valois. "Is RSSI a Good Choice for Localization in Wireless Sensor Network?", 732–39. IEEE, 2012. doi:10.1109/AINA.2012.19

[9] Heurtefeux, Karel, y Fabrice Valois. "Is RSSI a Good Choice for Localization in Wireless Sensor Network?", 732–39. IEEE, 2012. doi:10.1109/AINA.2012.19.

[10] Francisco J. Martinez, Chai Keong Toh, Juan-Carlos Cano, Carlos T. Calafate, Pietro Manzoni Determining the Representative Factors Affecting Warning Message Dissemination in VANETs Wireless Personal Communication. Springer Science 2011. doi:10.1007/s11277-011-0379-3.

Design of a Remote Monitoring System of Water Environmental Variables

Joaquin Mass S., Erica Ruiz I., Adolfo Espinoza R., Armando Garcia B.

joaquin.mass@itson.edu.mx, erica.ruiz@itson.mx, adolfo.espinoza@itson.edu.mx,
armando.garcia@itson.edu.mx
Instituto Tecnológico de Sonora, Sonora México

ABSTRACT

Data acquisition has always been essential for the human being, and Wireless Sensor Networks have helped with this issue. Marine environments, though, are usually a hostile place for this kind of job due to corrosion, high tides, wind, among other problems sensor networks deployed on solid ground do not suffer. The port of Manzanillo is a really busy one, so it could potentially harm the bay's environment because of its high activity. Therefore, the main goal of this project is to develop a Wireless Sensor Networks suited for monitoring various water parameters in real-time, in order to prevent a real catastrophe from occurring and damaging permanently the area's environment.

Keywords: Wireless Sensor Networks, Monitoring, Water environmental.

1. PROBLEM STATEMENT

Water environmental monitoring has been a field of interest for several years; being one of the principal reasons studying how the ecosystem is affected by industrial and touristic development, urbanism, as well as overexploitation of marine species. The port of Manzanillo Colima is considered the port with the largest volume of trading operations in México.It is internationally recognized for having operating routes to Japan, Hong Kong, Australia, New Zealand, United States, and Canada. Besides, this port is capable of receiving 14-meter and 80,000 DWT (Deadweight tonnage) vessels [1], generating potential risk conditions for ecosystem in the region (chemical spills, for example). The University of Colima employs a Coastal Observation System (COS) with the goal of measuring and analyzing oceanic currents in the zone. This system uses drifters which are deployed on the bay.

However, the fact that these drifters cannot communicate with each other or with some kind of data-acquisition point, is a huge limitation; therefore, the drifter's system is unconscious of its environment. Also, the collected data cannot be analyzed immediately due to not having a communication network to centralize data. Finally, the COS is required to be able to easily extend to obtain other kinds of physical variables such as temperature, pH, salinity, dissolved oxygen, etc. Each drift has a GPS module attached. To collect data, each buoy is deployed manually every morning. During the day, university students look out for the buoys from land using binoculars. At the end of the day, these drifters are localized and collected, and then taken to a laboratory to extract the data. Nowadays, with the progress of electronics, wireless communication, and software systems, is possible to design a low-cost COS that is able to measure different variables or execute various studies while being monitored and controlled remotely in real time. Studies have been made on data acquisition for environmental monitoring to predict parameters [2], and terrestrial monitoring to prevent forest fires [3]. However, marine applications have some unique challenges [4]:

- The marine environment is rough (corrosive).
- Buoys deployment must take into consideration the movement caused by tide, sea waves, wind, etc.
- Energy consumption is high due to the necessity of covering long distances, added signal losses.
- Instrumentation cost is higher in comparison with terrestrial WSN's.

In regards of development of this kind of applications, some groups are running studies about data acquisition devices for wireless networks on marine environments. For example, [5] uses magneto elastic sensors to measure pH, and [6] also presents a sensor network trying to reduce costs.

Other relevant projects could be:

- The pilot project for environmental observation of Columbia River, integrating a real-time sensor network, a data-addressing system, and advanced numerical models [8].
- Envisense Center's project that supervises movement of sediment on Briksdalsbreen glacier, in Norway, using wireless probes buried deep on the ice [8].
- OCEANSENSE project approaches the localization problem on monitoring the ocean surface proposing a mobile scheme called perpendicular intersection [9].

The objective for this work is to develop a technological platform able to carried out and manage studies of oceanographic monitoring on coastal zones. This platform must be:

- *Configurable*, in order to let monitoring of various variables (using different kind of sensors) and executing varied studies.
- *Scalable*. Supporting a variable number of nodes and applications.
- *Operating in real time*. This will let monitoring and management at any moment while executing remotely.
- *Inexpensive*. Design must take into consideration the components cost as well as energy consumption, so that the system may provide high autonomy levels.

2. METHODOLOGY

This project consists of the following general phases:

- *Bibliographic research*. During this stage, the functioning of the COS will be studied deeply. Also, an analysis will be made regarding different projects currently ongoing on coastal zones across the world.
- *Requirements analysis*. In this phase, a study will be made in order to identify specific requirements that the COS has regarding the most relevant physical variables to monitor, as well as communication resources for information. At this stage, different technical requirements could be identified to complement the current COS that the University of Colima has through development of a technological

platform for communication, processing, and presentation of data.

- *Scenario definition*. The monitoring platform scenario will be defined, helping to identify the following relevant points:
  - Characterization of data to acquire.
  - Characteristics of monitoring requirements, associated to existing technologies.
  - Conditions on which the equipment will be exposed, in order to design a node case for data acquisition.
  - Definition and interactions of data treatment software and presentation for the runway device, nodes, and cloud-data processing and storage area.
  - Requirements for communication protocols (transport, network, MAC, and physical layers) associated to the application's scenario.
- *Analysis of information technologies*. In this stage, different IT's will be compared to help sustain the choice of technology to implement and ensure a trustworthy communication between the drifters and the information processing center.
- *Development of the data-harvesting system*. The COS's prototype node will be proposed. Development boards and acquisition devices will be selected in order to simplify later on, and generate schematics and PCB's. Regarding software, it's considered the implementation of the data-harvesting system and driver developing to support different devices used in physical variable acquisition and communication system with selected technologies.
- *Design and implementation of the communication protocol*. Firstly, a systematic revision will be made about communication protocols used for data transmission (transport layer, network layer, and media access protocols) applied to wireless sensor networks on water environments, analyzing pros and cons. Then, the best suited protocol for the defined scenario will be developed and implemented inside the data-harvesting system.
- *Development of end-user software*. A visual system is proposed to display the acquired data. The objective here is the processing and displaying of data for a better understanding.
- *Design and development of evaluation tests*. At the end of the integration process, the evaluation will focus on field tests which will let analyze the COS's prototype performance on a real-world environment. These metrics will be considered: transmission reliability (delivery success rate), lost-packets percentage, energy consumption, highest transmission range on harsh conditions for the node, etc.

3. EXPECTED RESULTS

At the end of this project, it is expected to reach the following results:

- A data transfer protocol where the following stages are clearly recognized: initialization, where clusters are formed along its respective node members and clusterheads; route discovery point-to-point between clusterheads and the sink node on land; clusterhead update, where the leader of each cluster is interchanged for energy-saving purposes; and maintenance, where a new cluster is formed in case of a disconnection between clusterhead and sink node.
- Hardware implementation of such protocol, preferably on RaspBerry Pi or Arduino Due boards, and XBee modules.
- End-user graphical interface for comfortable data reading.

4. PRELIMINARY RESULTS

At the time of writing this paper, some of these phases have been completed. Bibliographic research on different databases has been done, as well as a general analysis of requirements, where temperature and ocean currents are selected as relevant. Data to acquire, buoys conditions on water, and requirements for the communication protocols have been defined too.

For us, this consortium represents a great opportunity to strengthen or create bonds between the universities and participants which collaborate in the project. Students and teachers alike will expand their knowledge and experiences while contributing to the prevention of ecological disasters in the region. And finally, let's not forget the technological advance and logbooks that will be inherited to next generations of researchers.

5. REFERENCES

[1] Al-Haija, Q., Al-Qadeeb, H., & Al-Lwaimi, A. (2013). Case Study: Monitoring of AIR Quality in King Faisal University Using a Microcontroller and WSN. *Procedia Computer Science*, 21, 517-521.

[2] Bayo, A., Antolin, D., Medrano, N., Calvo, B., & Celma, S. (2010). 1. Early Detection and Monitoring of Forest Fire with a Wireless Sensor Network System. *Proceedings Eurosensors XXIV*. Linz.

[3] Albaladejo, C., Sánchez, P., Iborra, A., Soto, F., López, J. A., & Torres, R. (2010). Wireless Sensor Networks for Oceanographic Monitoring: A Systematic Review. *Sensors*, 6948-6968.

[4] Ong, K. G., Yang, X., Mukherjee, N., Wang, H., Surender, S., & Grimes, C. A. (2004). A Wireless Sensor Network for Long-term Monitoring of Aquatic Environments: Design and Implementation. *Sensor Letters*, 2 (1), 48-57.

[5] Dunkels, A., Marie Feeney, L., Grönvall, B., & Voigt, T. (2004). An integrated approach to developing sensor network solutions. In Proceedings of the *2nd International Workshop on Sensor and Actor Network Protocols and Applications* (SANPA'04).

[6] Dang, T., Frolov, S., Bulusu, N., Feng, W.-C., & Bapista, A. (2007). Near Optimal Sensor Selection in the Columbia River (CORIE) Observation Network For Data Asignation Using Genetic Algorithms. Lecture Notes In *Computer Science*, 4549, 253-266.

[7] Martinez, K., Ong, R., & Hart, J. (2004). Glacsweb: a sensor network for hostile environments. First Annual *IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2004. IEEE SECON 2004 (pp. 81-87). IEEE.

[8] Liu, K., Yang, Z., Li, M., Guo, Z., Guo, Y., Hong, F., et al. (2010). Oceansense: monitoring the sea with wireless sensor network. *ACM SIGMOBILE Mobile Computing and Communications Review*, 14 (2), 7-9.

[9] SEMAR. (S.F.). Reporte técnico Puerto de Manzanillo, Colima. *Secretaria de Marina*. SEMAR.

Bit-Error Rate Evaluation of Hybrid MIMO LSTBC Receivers Using Limited-Precision QR Decomposition

Jorge Robles
Instituto Tecnológico de Sonora
jorgeh_rch@gmail.com

Joaquin Cortez
Instituto Tecnológico de Sonora
joaquin.cortez@itson.edu.mx

Erica Ruiz
Instituto Tecnológico de Sonora
erica.ruiz@itson.edu.mx

Adolfo Espinoza
Instituto Tecnológico de Sonora
adolfo.espinoza@itson.edu.mx

ABSTRACT
This work proposes to analyze the behavior of Sorted-QR decomposition algorithms for Hybrid MIMO systems when implemented using fixed-point arithmetic. This behavior is studied in the context of a Hybrid MIMO receiver, and the effects of the limited precision are evaluated in terms of their impact on the receiver bit error rate. Our analysis of Hybrid MIMO receiver under these conditions pretends to shows a strong dependence on the numerical precision, for both the integer and fractional parts. We trying to identified an error floor effect, where the bit error rate reaches a limit and stops improving even if the precision is increased. In addition, we will define the minimum fixed-point word size required not to adversely impact performance for a variety of antenna array sizes.

CCS Concepts
Hardware → Digital Signal Processing

Keywords
V-BLAST; STBC Codes; Hybrid MIMO Systems; QR decomposition; fixed-point arithmetic.

1. INTRODUCTION
The demand of communication systems that effectively exploit the wireless channels huge capacity [1] has grown rapidly during the last decade. In recent years, multiple-input, multiple-output (MIMO) systems have emerged as an attractive solution to these problems. MIMO systems employ multiple antennas both at the transmitter and the receiver, and are able to create both time and space diversity. Space-time coding is a set of techniques to create and exploit this diversity. Layered space-time codes, also called BLAST (Bell-Labs Layered Space-Time), were first introduced in [2]. They offer a very high spectral efficiency, ease of code design, and comparatively simple receiver architectures. Space-time orthogonal and quasi-orthogonal block codes (STBC), such as the Alamouti scheme [3], are examples of diversity codes. Linear Dispersion Codes (LDCs) were proposed in [4] with the objective of obtaining a measure of both multiplexing and diversity gain. LDCs can be designed to achieve an optimal trade-off between spatial multiplexing and diversity transmission. These codes may use any V-BLAST decoding algorithm. An alternative approach, known as Hybrid Coding (HC) [5], consists of the selection of some of the available transmit antennas to work in STBC mode, while the remaining ones operate as V-BLAST. Their principal disadvantage is high receiver complexity. In particular, the QRM HC Sorted QR scheme [6] is an interesting example of hybrid coding; it uses one Alamouti layer and one Spatial Layer, increasing spectral efficiency over pure STBC, and its structure allows to use a simplified detector. Floating-point numbers have large dynamic range and high precision, but they are expensive in terms of cost and power consumption, and have speed limitations. When planning a hardware implementation, it is natural to abandon floating point numbers in favor of fixed-point arithmetic. Furthermore, decreasing the word size almost linearly reduces the design area, hardware complexity and power consumption [7]. The number of digits used to represent fixed-point numbers will influence the accuracy of the arithmetic operations performed [8]. This will in turn affect the bit error rate of the receiver, which will diverge from the theoretical estimates if the precision of the operations is not high enough. On the other hand, it is important to avoid using more accuracy than strictly needed. In this work, we propose to study the Sorted QR algorithm used to obtain the QR decomposition required to carry the detection process. The first variant is the modified Gram-Schmidt (MGS) algorithm; the second is to use a Givens Rotations (RG) algorithm that improves the numeric stability of the former in many common cases [9]. We are interested in comparing and analyzing their behavior compared with a floating-point implementation. We will emphasize the effect of having a limited-precision numerical representation on the bit error rate of a QRM HC Sorted QR receiver. In order to better isolate the BER caused by limited-precision from that caused by noise present in the channel, our results assume a noiseless channel. Likewise, only the Sorted QR decomposition algorithm has been implemented as fixed-point; the rest of the QRM HC Sorted QR receiver continues to use full-precision floating-point numbers. Thus, all bit errors in our simulations are a consequence of using fixed-point arithmetic in the QR decomposition algorithm. Finally, we will obtain the minimum word-length required not to adversely affect the BER performance of QRM HC Sorted QR, for several antenna array sizes and for both MGS-QRM HC Sorted QR and RG-QRM HC Sorted QR algorithms. We also pretend will be design a fast, scalable hardware architecture (suitable for FPGA implementation) that performs the sorted QR detector required by the detector [10].

2. MODEL SYSTEM
When use an uncorrelated channel for the MIMO system, it is assumed that the propagation channel between each transmit and

receive antenna is a Rayleigh narrowband stationary stochastic process, with rich scattering at both the transmitter and receiver ends. This channel can be modeled as a random matrix H of size $n_R \times n_T$, where n_R and n_T are the number of antennas at the receiver and transmitter, respectively. Each h_{ij} for $i = 1, 2, \text{K}, n_R$, $j = 1, 2, \text{K}, n_T$ is a sample of a complex Gaussian random variable with zero mean and variance 0.5 per dimension. We assume quasi-static at channels, and H remains constant for at least the transmission of one space-time code word. QRM HC Sorted QR transmits four symbols each period $T = 2$ [11]. It requires three transmit antennas, divided in one Alamouti encoder and one spatial layer. All symbols are diversity coded. Its space-time mapping is shown in Table 1. The QRM HC Sorted QR system is shown in Fig. 1. The power allocated to each symbol is $P_S = \frac{1}{4}$,W. This allocation results in all symbols being transmitted with the same energy, which is the optimal strategy in the absence of CSI at the transmitter.

Table 1. QRM HC Sorted QR Space-Time Mapping

Time	Antenna 1	Antenna 2	Antenna 3
t	s_1	s_3	s_4
$t+1$	$-s_2^*$	s_4^*	$-s_3^*$

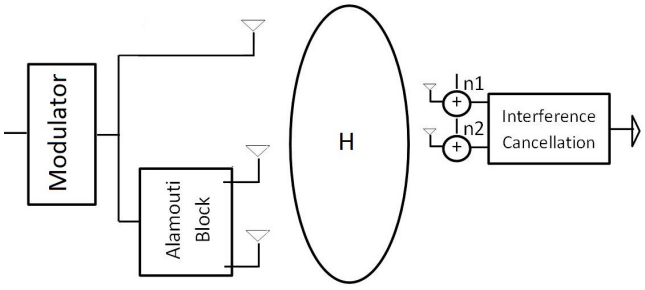


Figure 1. QRM HC Sorted QR System.

3. METHOD

This work consists of the following general phases:

- Bibliographic research. During this stage, the functioning of the QRM HC Sorted QR System will be studied deeply.
- Implementation of the algorithm in Matlab using Fixed Point arithmetic toolbox.
- Analysis of simulation results. The BER performance was obtained under several combinations of antenna array sizes and fixed-point word lengths.
- We have defined the minimum word length required to avoid or minimize the error in terms of BER of the QRM HC Sorted QR System.
- We also pretend will be design a fast, scalable hardware architecture (suitable for FPGA implementation) that performs the sorted QR detector required by the detector.
- We propose to design this architecture to carried out the QR matrix decomposition, based with a low-complexity, parallel and pipelined CORDIC [12] with fixed-point arithmetic.

4. REFERENCES

[1] I. E. Telatar. (1999) Capacity of multiple-antenna Gaussian channels. *European Transactions on Telecommunications*, Vol. 10, No. 6, p. 585-595.

[2] G. J. Foschini. (1996) Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech*. J., Vol. 1, No. 2, p. 41-59.

[3] S. M. Alamouti. (1998) Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 8, p. 1451-1458.

[4] B. Hassibi and B. M. Hochwald. (2002) High-rate codes that are linear in space and time. *IEEE Transactions on Information Theory*, Vol. 48, No. 7, p. 1804-1824.

[5] Seung-Jun Yu, Eui-Young Lee and Hyoung-Kyu Song. (2015) Combination of STBC and SM Scheme with Iterative Detection in LTE Systems. *Wireless Personal Communications*, Vol. 83, No. 2, p. 1203-1211.

[6] J. Cortez et al. (2015) A Very Low Complexity Near ML Detector Based on QRD-M Algorithm for STBC-VBLAST Architecture. *7th IEEE Latin-American Conference on Communications (LATINCOM)*.

[7] S. Kim, K.-I. Kum, and W. Sung. (1998) Fixed-point optimization utility for c and c++ based digital signal processing programs. *IEEE Transaction on Circuits Systems II*, Vol. 45, No. 11.

[8] G. H. Golub and C. F. V. Loan (1996) Matrix Computations. *Johns Hopkins University Press*.

[9] Tapan Pradhan et al. (2015), Development of numerical linear algebra algorithms in dynamic fixed-point format: a case study of Lanczos tridiagonalization, *Int. J. Circ. Theor. Appl.* , Vol. 44, pp. 1222–1262

[10] M. Karkooti, J.R. Cavallaro and C. Dick, "FPGA Implementation of Matrix Inversion Using QRD-RLS Algorithm", *Conf. Rec. 39th Asilomar Conf. on Signals, Systems and Computers*, Oct. 2005.

[11] M. Bazdresch, J. Cortez, O. Longoria and R. Parra-Michel. (2012) A Family of Hibrid Space-Time Codes for MIMO Wireless Communications. *Journal of Applied Research and Technology*, Vol. 19, No. 2, p. 122-142.

[12] J.E.Volder (1959). The CORDIC trigonometric computing technique. *IRE Transactions Electronic Computers*, Vol 8, pp. 330–334.

[13] Tong Zhou et al.(2015), Area-efficient high-throughput sorted QR decomposition-based MIMO detector on FPGA, *IEEE International Conference on Computer and Communications (ICCC)*.

Sistema de visión para conocer el campo de velocidades en un fluido mediante el trazado de partículas

Mónica Valenzuela-Delgado
Miguel E. Bravo-Zanoguera
Universidad Autónoma de Baja California. Facultad de Ingeniería, Mexicali, B.C., México
Blvd Benito Juárez s/n C.P. 21280
52 686 566 4270
monicavd74@uabc.edu.mx

Wendy Flores-Fuentes
Alejandro S. Ortiz-Pérez
Universidad Autónoma de Baja California. Facultad de Ingeniería, Mexicali, B.C., México
52 686 566 4270
flores.wendy@uabc.edu.mx

Moisés López-Rivas
Félix F. González-Navarro
Oleg Sergiyenko
Universidad Autónoma de Baja California. Instituto de Ingeniería, Mexicali, B.C., México
52 686 565 4431
mrivas@uabc.edu.mx

RESUMEN

La medición del movimiento de un fluido en un sistema magneto hidrodinámico es una fuente de información para el diseño de agitadores electromagnéticos. Por lo tanto, se requiere experimentar con los efectos de intensidad de campos eléctricos y magnéticos inducidos en un fluido, por medio de la medición de la velocidad, dirección, y aceleración del flujo. En experimentaciones realizadas, se ha recurrido al análisis visual del flujo usando tinte para revelar los vórtices y su tiempo de permanencia [1], pero se requiere de una medición y un análisis más detallado de los resultados de la experimentación, con este propósito se inicia una línea de investigación para el diseño de un sistema de visión para el análisis de los campos de velocidad presentes en un fluido a partir del trazado de partículas, por medio de procesamiento digital basado en un análisis de velocimetría por imágenes de partículas (PIV), cuya metodología de diseño se describe en este documento.

Conceptos CCS

Computing methodologies → Massively parallel and high-performance simulations.

Palabras Clave

Sistema de Visión; velocimetría por imágenes de partículas; magneto hidrodinámica.

1. INTRODUCCIÓN

La magneto-hidrodinámica (MHD) estudia la dinámica de fluidos conductores de electricidad en campos magnéticos y eléctricos. Tiene aplicaciones en micro-reactores y micro-mezcladores [2]. La medición del movimiento del fluido en un sistema MHD es una fuente de información para el diseño de micro-mezcladores. Existen estudios sobre flujos MHD en canales anulares con metales líquidos, pero pocos trabajos publicados para el caso de electrolitos. En estos estudios se menciona una inestabilidad del campo del flujo [3] y que se requiere medir la influencia de este fenómeno. Experimentos realizados por la UNAM, en colaboración con UABC, describen una inestabilidad en la superficie libre de un flujo generado por una fuerza de Lorentz azimuthal aplicada a un electrolito contenido en un agitador electromagnético abierto con geometría cilíndrica. Se ha realizado el análisis visual del flujo usando tinte para revelar los vórtices y su tiempo de permanencia. Estos vórtices son visibles claramente sobre la superficie del electrolito. Las imágenes observadas se procesaron mediante la herramienta para aplicaciones generales de Matlab, pero un análisis más detallado es requerido, a través de código personalizado a la aplicación. Una práctica común para el análisis de la dinámica de fluidos, es la velocimetría por imágenes de partículas (PIV), debido a los avances recientes en tecnologías laser y procesamiento de imágenes digitales [4]. Aunque también existen otros métodos basados en la medición de partículas tales como en [5], la de mayor interés es PIV, no solo por el avance tecnológico, sino que esta técnica identifica mayor cantidad de parámetros geométricos para la recreación de campos de velocidad en los vórtices generados en un fluido. Existen sistemas típicos PIV consistentes en una luz estroboscópica, una cámara CCD y un tanque cilíndrico en forma de torre que contiene las partículas donde estas fluyen hacia la superficie [6]. Este sistema típico no es funcional para todos los casos de análisis de la dinámica de fluidos [7] de tal modo que un sistema personalizado debe ser diseñado según la aplicación requerida. Por lo anterior, se pretende el diseño de un sistema PIV personalizado, ya que el fluido a analizar, es un electrolito en un agitador electromagnético cuyas características difieren de un sistema típico PIV. Para ello, se describen diversos sistemas PIV, se analizan los distintos algoritmos basados en correlación cruzada para medir la variación de velocidad, campos de vorticidad, espectro y otros parámetros pertinentes de turbulencia, se expone la metodología para resolver la problemática planteada y los resultados esperados.

2. METODOLOGÍA.

2.1. Montaje experimental y comportamiento del agitador electromagnético.

Al aplicar radialmente una corriente en un agitador electromagnético cilíndrico abierto, bajo un campo magnético axial aproximadamente uniforme se presenta una inestabilidad en un flujo de electrolito impulsado por una fuerza de Lorentz azimutal (ver figura 1).



Figura 1. Montaje experimental del agitador MHD. 1 {ferrita rectangular imán permanente; 2 {contenedor de plexiglás; 3 {electrodo exterior; 4 {electrodo interno; 5 {suministro de energía. Imagen tomada de Ref. [1].

La inestabilidad se caracteriza por la aparición de un número variable de vórtices anticiclónicos en un plano normal al campo magnético cerca al electrodo exterior (ver figura 2).



Figura 2. Visualización de la inestabilidad observada para un espesor de capa de 10 mm con una intensidad de 100mA y 300mA respectivamente. Imagen tomada de Ref. [1].

Con el fin de caracterizar cuantitativamente esta inestabilidad, la metodología del diseño de un sistema PIV es propuesto.

2.2 Sistema PIV

El proceso de un PIV consiste en tomar un par de imágenes consecutivas en intervalos de tiempo constante mediante una cámara CCD y una luz estroboscópica sincronizada con la cámara que ilumina las partículas sembradas en un fluido. Las imágenes se correlacionan pixel por pixel para determinar el desplazamiento y los vectores de velocidad de las partículas. Este método es conocido como correlación cruzada [8]. El sistema PIV constará de cuatro elementos básicos [9]. El primero, consiste en la superficie abierta del micro-mezclador, ópticamente transparente donde se visualiza el fluido con las partículas a evaluar. Contará con un sistema de adquisición y almacenamiento de imágenes constituido por una cámara Handy-cam Sony (HDR-XR100) con una resolución de 1080 × 720 píxeles y 30 fps, colocada encima del montaje experimental. Un sistema de iluminación inherente proveniente de tres lámparas LED colocadas en la parte superior del recipiente de plexiglás alrededor de la cavidad con el líquido. Por último, contará con un sistema de procesamiento de imágenes para la obtención de datos que permitan el cálculo del campo de velocidades presentes en el fluido mediante la implementación de algoritmos basados en correlación y seguimiento de partículas utilizando el software MatLab y la biblioteca de funciones PIV.

2.3 Algoritmos PIV

Se utilizarán los métodos de correlación debido a que reduce la complejidad del algoritmo PIV con respecto al seguimiento individual de partículas. Esta técnica proporciona una media estadística del movimiento de un conjunto de partículas dentro de pequeñas regiones conocidas como ventanas de interrogación que se obtienen al dividir los cuadros de la secuencia de imágenes en regiones de interés y para cada una de estas regiones, se asociará un vector velocidad. Este proceso permite mayor tolerancia al ruido y robustez que otros métodos conocidos. Se analizaron cuatro métodos que son descritos a continuación. El método de auto-correlación, consiste en iluminar dos veces las partículas y capturar dentro de un mismo cuadro de imagen, las dos ventanas de interrogación, y sobre ese patrón asociar una posición inicial y final a cada partícula. Esto se logra dejando el obturador de la cámara abierto durante dos pulsos laser consecutivos. Este método no puede ser utilizado en este experimento por que la iluminación será inherente. El método de correlación cruzada consiste en la captura de dos cuadros consecutivos con iluminación constante que al correlacionar de forma cruzada el primer cuadro con el segundo, se obtiene un único pico máximo en la dirección del movimiento de las partículas. En la correlación cruzada directa se obtiene un solo pico, por lo que no da lugar a ambigüedades y permite medir corrimientos nulos. Ésta posee mayor tolerancia a ruidos que el proceso de auto-correlación. La problemática, es la sensibilidad a los cambios en la intensidad entre cada cuadro adquirido, producidos por diferencia de iluminación en cada uno [9]. Por último, la correlación cruzada normalizada, resuelve la problemática de la sensibilidad a cambios de intensidad entre cuadros consecutivos, logrando un algoritmo más robusto ante situaciones donde la distribución espacial de las partículas no es homogénea, las partículas no son mono-dispersas y la intensidad de la luz entre cuadros consecutivos varía ligeramente [9]. Por lo anterior, el procesamiento digital de imágenes, estará basado en este último método. Aún con este algoritmo, existen problemáticas a resolver como los efectos de bordes, perdida del patrón a seguir y eliminación de ruido, por lo que validaciones especificas y filtros deberán ser agregados.

3. RESULTADOS ESPERADOS

Se pretende el diseño de un sistema PIV personalizado, para un electrolito contenido en un agitador electromagnético abierto con geometría cilíndrica cuyas características difieren de un sistema típico PIV. El sistema constará de un sistema de adquisición y almacenamiento de imágenes, un sistema de iluminación incoherente y un sistema de procesamiento digital de imágenes basado en seguimiento por imágenes de partículas. La finalidad es la caracterización cuantitativa de la inestabilidad presentada por el fluido electrolito ante la presencia de campos magnéticos y eléctricos radiales. También, se desea obtener el campo de vectores de velocidad en el fluido, mediante la implementación de algoritmos basados en correlación cruzada para el seguimiento por imágenes de partículas sembradas en el fluido con el objetivo de establecer el funcionamiento del agitador.

4. REFERENCIAS

[1] Pérez-Barrera, J., Pérez-Espinoza, J. E., Ortiz, A., Ramos, E., & Cuevas, S. 2015. *Instability of electrolyte flow driven by an azimuthal Lorentz force*. Magnetohydrodynamics, 51(2), 203-213.

[2] Gissinger, C., Rodriguez-Imazio, P. & Fauve, S. 2016. *Instability in electromagnetically driven flows, Part I*. arXiv:1602.02061v1 [physics.flu-dyn]. 5 Feb 2016.

[3] Rodriguez-Imazio, P. & Gissinger, C. 2016. *Instability in electromagnetically driven flows, Part II*. arXiv:1602.02061v1 [physics.flu-dyn]. 5 Feb 2016.

[4] Brossard, C. et al. 2009. *Principles and applications of particle image velocimetry*. AerospaceLab, (1), p-1.

[5] Morita, N. et al. 2015. *A micro laser Doppler velocimeter designed for a wafer-level packaging process Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*. Transducers - 2015 18th International Conference:21-25 June 2015. IEEE Xplore. Electronic ISBN: 978-1-4799-8955-3.

[6] Zhou, L., Shi, W., Li, W. et al. 2016. *Particle Image Velocimetry Measurements and Performance Experiments in a Compact Return Diffuser Under Different Rotating Speed*. Exp Tech (2016) 40: 245. doi:10.1007/s40799-016-0028-6.

[7] Johnson, M. B. 2015. *Hybrid Particle Image Velocimetry with the Combination of Cross-Correlation and Optical Flow Method*. MS Thesis submitted, Wright State University.

[8] Kong, L. F. et al. 2014. *An Improved Particle Image Velocimetry Algorithm for Velocity Measurement of Oil-Water Two-Phase Flow, Applied Mechanics and Materials*.Vols. 602-605, pp. 1654-1659, 2014.

[9] Wang, C., Xie, Z., Xu, B., Li, J., & Zhou, X. 2016. Experimental Study on EHD *Flow Transition in a Small Scale Wire-plate ESP*. Measurement Science Review, 16, (2016), No. 3, 134-1

Sincronización de Labios en Marionetas Digitales

Ma. Angélica Yáñez Castañeda
DEPI
Instituto Tecnológico de Chihuahua II
Chihuahua, Chih.
(614) 285 19 18
yacama@yahoo.com.mx

Alberto Pacheco González
DEPI
Instituto Tecnológico de Chihuahua
Chihuahua, Chih.
(614) 177 33 46
alberto.chiwas@gmail.com

Rafael Vázquez Pérez
DEPI
Instituto Tecnológico de Chihuahua II
Chihuahua, Chih.
(614) 211 28 30
rvazquez.perez@me.com

RESUMEN
Actualmente se está trabajando con marionetas digitales para apoyar al proceso cognitivo de los estudiante. Ya se tiene implementado el cuerpo, por lo que se requiere trabajar con el rostro. El objetivo de este trabajo consiste en sincronizar en tiempo real los movimientos de los labios de una marioneta digital con la señal de voz de un usuario, identificando fonemas y representarlos en los gestos de la boca de la marioneta utilizando redes neuronales. La metodología de trabajo consistió en realizar una revisión del estado del arte de diferentes métodos estadísticos y matemáticos para la sincronización de labios (lip sync), en la captura y clasificación de las señales de sonido, la correlacion con con el visema que representa a cada fonema detectado en la señal de sonido y por último la representación gráfica del gesto de la boca en el rostro de la marioneta digital, todo lo anterior en tiempo real. Se utiliza una cámara para la obtención de los puntos de referencia en el rostro y una consola de audio para la captura de voz. Como resultado de la investigación se identifica, analiza, clasifica y selecciona un método para la obtención del método que se pueda utilizar en tiempo real.

Conceptos CCS

- **Computing methodologies →Artificial intelligence → Computer vision → Image and video acquisition → Motion capture.**
- **CCS → Human-centered computing → Human computer interaction (HCI) → Interaction techniques → Gestural input.**

Palabras clave
Reconocimiento de voz; reconocimiento de imagen; sincronización de voz.

1. INTRODUCCIÓN
La marionetas digitales es el manejo interactivo de personajes virtuales en 3D o 2D, que son procesados en tiempo real y se utilizan para la enseñanza como una estrategia innovadora. En el laboratorio de m-Learning del Instituto Tecnológico de Chihuahua en el año 2012 inicia el diseño de un proceso y una herramienta de software para brindar soporte al desarrollo de animaciones, mediante la captura de movimiento a través de un sensor RGB-D. Este trabajo se desarrolla tomando en cuenta el efecto persona que surge de una narrativa interactiva con personajes animados [8] y que influye de manera positiva en el proceso congnitivo del estudiante, gracias al efecto persona se puede emerger un efecto motivacional. Una respuesta innata psicosocial del efecto persona, puede por ejemplo servir para asistir de manera afectiva a un estudiante en línea.

La herramienta desarrollada por el Tecnológico de Chihuahua consta de tres módulos:
1) Adquisición: los datos son capturados por medio del sensor RGB-D permitiendo con la librería de reconocimiento obtener las coordenadas 3D de cada una de las articulaciones.
2) Procesamiento: localización de coordenadas 2D y cálculo de los desplazamientos angulares de cada articulación.
3) Animación: localización de coordenadas de pantalla y visualización de cada miembro articulado de la marioneta, además de proyectar una imagen de fondo como parte del escenario [3].
Los avances del proyecto son una marioneta con movimiento en cuerpo, pero con rostro fijo.

Un programa en tiempo real o concurrente está formado por una colección de procesos secuenciales autónomos que se ejecutan en paralelo. Para Hanseok ko [6] en el 2008, el habla es percibida no sólo como información acústica, sino también como información visual contenida en el movimiento de los labios y las expresiones faciales existen dos factores importantes en la realización de un sistema de sincronización de labios en tiempo real, es decir, el modelado acústico dependiente del contexto en tiempo real y el reconocimiento efectivo de los fonemas de forma continua. En particular, las vocales son el factor más importante para determinar la forma del labio, que es relativamente estable durante su pronunciación y por otra parte varía rápidamente cuando se pronuncian ciertas consonantes. La animación del discurso impulsado por el habla utiliza los fonemas como unidades básicas de la oración y visemas como las unidades básicas de la animación definidas por Thavesak en 2011 [11]. Los visemas son representaciones vocales que contienen significado lingüístico y ayudan a reconocer fonemas aun cuando la persona sea débil auditiva o no cuente con la habilidad de hablar algún idioma.

Los métodos que se encuentran en la bibliografía abordan la problemática de diferentes formas. Existen métodos que utilizan colecciones de bocas (visemas) y son intercambiadas dependiendo de los fonemas registrados según Kiran [7]. Otros generan una máscara y a partir de puntos de referencia seleccionados se correlacionan con los visemas correspondientes. También se aborda la sincronización de labios utilizando redes neuronales, el modelo oculto de Markov [11] (HMM) y la combinación del modelo de cabeza-cuerpo-cola (HBT)]. Además los movimientos faciales sincronizados se pueden generar a partir del habla natural o de un motor de texto de voz.

2. METODOLOGÍA
A continuación se muestran las diversas etapas del presente proyecto de investigación de tesis que está divido en el desarrollo de los siguientes módulos:

a)Definición del problema y revisión de literatura
Partiendo del desarrollo previo que se tiene de la marioneta y que representa el cuerpo, se hace una revisión minuciosa del estado del arte, incluyendo una selección, clasificación, análisis y síntesis de los métodos que se existen relacionados con la sincronización de los labios.

***b**)Análisis, diseño e implementación del módulo de captura de voz*.

En el módulo de captura de voz, se captura la señal de audio y se selecciona según el fonema que representa. El corpus consta de 100 capturas de voces.

c)Análisis, diseño e implementación del módulo de fonemas.
El módulo de fonemas es consecutivo al módulo de captura de voz. La señal de audio se segmentará y se identificarán los fonemas que forman parte de la señal de voz.

d)Análisis, diseño e implementación del módulo de obtención de pesos.
En esta sección se extraerán las características de cada uno de los fonemas, como son la intensidad y la frecuencia, y se les otorgará un valor que posteriormente se procesará.

e)Análisis, diseño e implementación del módulo de reconocimiento de voz
A partir de la extracción de características del habla, se construirán en una serie de modelos estadísticos, los cuales se clasificarán utilizando redes neuronales.

f)*Análisis, diseño e implementación del módulo de puntos de referencia en rostro*
En este módulo se identifican los puntos de referencia en el rostro humano y se correlacionarán con los puntos de referencia colocados en el rostro de la marioneta digital para sincronizarlos con la señal de la voz.

3. RESULTADOS LOGRADOS
Se estableció el método que se evaluará para la sincronización de voz y movimiento de los gestos de la boca en una marioneta digital en tiempo real.
En la etapa b) se obtuvo una muestra de 100 voces y se seleccionaron dependiendo del fonema que representaban.
En las etapas del inciso c) se identificaron los centros de frecuencia para su posterior procesamiento.
En los incisos d), e) se ha concluido la etapa de análisis y se ha iniciado la etapa de diseño donde se utilizarán redes neuronales.
En el inciso f) se ha concluido la etapa de análisis y se ha iniciado la etapa de diseño donde se están proponiendo la colocación de puntos de referencia en la imagen para incluir un visema en la imagen svg.

4. RESULTADOS ESPERADOS
Como resultados esperados de los incisos e) y f) se tiene que los tiempos de ejecución de la señal de la voz desde el momento que se captura el sonido hasta el momento en que es identificado el fonema y seleccionado el visema correspondiente de cada uno de los sonidos dictados en la voz, coincidan con el tiempo de obtención de la imagen vectorizada, además alcanzar como máximo 1/30 de frame en la ejecución.

5. CONCLUSIONES
Se concluye que las señales de voz y video deben ser obtenidas simultáneamente y los procesos bilaterales que se ejecutan para cada una de las señales deben concluir ambas en 1/30 de frame para lograr una sincronización efectiva y una armonía en el discurso de la marioneta.
También se concluye que con la utilización de una red neuronal entrenada mejorará la percepción del sonido en tiempo real evitando ruido que ocasiona que el movimiento de los labios sea menos preciso.
En este momento está pendiente la obtención de los resultados de la red neuronal.

6. REFERENCIAS

[1] Alberto Pacheco, Inclusión Social: Impulsando la Web Móvil y su Uso Educativo. *En Memorias del XVIII Encuentro Internacional de Educación a Distancia*. Guadalajara: UdG Virtual. Doi:10.13140/2.1.3755.3284.

[2] Alberto Pacheco, Micro-Guías: Actividades de Aprendizaje en dispositivos móviles basadas en micro-contenidos. *En Memorias 3a. Conferencia* LACLO 2008. 153-161.

[3] Alberto Pacheco, Miriam Ramírez, Cynthia Guzmán, Elisabet González, Marionetas Digitales: Tecnología Emergente para Narrar Historias con Personajes Animados mediante la Captura Digital de Movimiento, *División de Estudios de Posgrado e Investigación*, Chihuahua México.1-14.

[4] Derlin Morocho Checa, Tesis Comparación de tipologías de HMMs par mejora de la segmentación fonética, *Universidad Autónoma de Madrid*, España, 2013. 3-4.

[5] Francisco Cabiedes, Ildikó Pelczer, Fernando Gamboa, Javier Bretón y Steve Rodríguez, Sincronización de labios: método sin visemas, Cd. de México, RIED, vol. 10 :1, 2007. 37-50.

[6] Hanseok Ko, Real-Time Continuos Phoneme Recognition System, Seul Corea, IEEE Transactions on multimedia, vol. 10, noviembre 2008. 1306.

[7] Kiran S. Bhat, Rony Goldenthal Yuting, Ye Ronald Mallet y Michael Koperwas, High Fidelity Facial Animation Capture and Retargeting With Contours, *Anaheim California*, SCA, julio 2013. 14.

[8] Lester, J.C., Converse. S.A. Kahler, S.E., Todd, S., Brian, B., Stone, A., & Bhogal, R.S., The persona effect:Affective impact of animated pedagogical agents, ACCM Press, 1997. 359-366.

[9] Rahim Saeidi, Paavo Alku y Tom Bäckström, Feature Extration Using Power-Law Adjusted Linear Prediction With Application to Speaker Recognition Under Severe Vocal Effort Mismatch, Finlandia, IEEE/ACM, vol. 24, enero 2016. 53.

[10] Thavesak Chuensaichol, Pizzanu Kanongchaiyos y Chai Wutiwiwatchai, Lip Synchronization from Thai Speech, Hong Kong China, ACM 978-1-4503-1060-4/11/0012, diciembre 2011. 358.

[11] Yuyu Xu, Adrew W. Feng, Stacy Marsella y Ari Shapiro, A Practical and Configurable Lip Sync Method for Games, Dublin Ireland, ACM 978-1-4503-2546-2/13/11, noviembre 2013. 118.

Sesión de Carteles



Organizadores:

Pedro César Santana Mancilla (UCol)
Juan Pablo García Vázquez (UABC)
Pablo Pancardo García (UJAT)

Estadística y Estudio Léxico-Sintáctico del Habla de la Lengua Yaqui para el Desarrollo de Herramientas de TI

Fernando Orduña-Cabrera
Instituto Tecnológico Superior de Cajeme,
SEKAI.MX
Cd. Obregón Sonora, México
644-4108650
forduna@itesca.edu.mx

Valeria Javalera-Rincón
Instituto Tecnológico Superior de Cajeme,
SEKAI.MX
Cd. Obregón Sonora, México
644-4108650
vjavalera@itesca.edu.mx

J. Manuel Barba Soto
Instituto Tecnológico Superior de Cajeme,
Cd. Obregón Sonora, México
644-4108650
jmbarva@itesca.edu.mx

RESUMEN
En México existen 68 lenguas indígenas, una de ellas es la lengua Yaqui, denominada de esta forma por ser la lengua materna de la tribu Yaqui que habita en el estado de Sonora, México y en Arizona, Estados Unidos. La lengua Yaqui se encuentra en peligro de extinción de acuerdo a los estudios del Instituto Nacional de Lenguas Indígenas. En este trabajo, se presenta un estudio de la del habla la lengua yaqui por los jóvenes, también se aborda brevemente el desarrollo del léxico del yaqui, presentando posteriormente detalles de la sintaxis y gramática ambas centradas en el estudio del uso de los pronombres en Yaqui; todo ello para el desarrollo de propuestas de apoyo a la preservación de la lengua, teniendo como base las tecnologías de información para la enseñanza del habla y escritura de la lengua en los diferentes niveles educativos.

CCS CONCEPTOS
• Artificial intelligence→Natural Languaje→Parser and grammar.

PALABRAS CLAVE
Lengua Yaqui, análisis léxico, análisis sintáctico, diagramas gramaticales.

1. INTRODUCCIÓN
Derivada de la familia Yuto-nahua, la lengua Yaqui es hablada por alrededor de 14,162 hablantes residentes, en su mayoría, en el estado de Sonora, México y Arizona, EE.UU. Esto representa poco más del 0.03% de la población conjunta de estos dos estados [1]. Bajo el cobijo de la Ley General de Derechos Lingüísticos de los Pueblos Indígenas aprobada en el año 2003, el Yaqui es considerado parte integral del patrimonio cultural y lingüístico nacional.

Con el objetivo de conocer si los jóvenes nativos y no nativos residentes en los pueblos Yaquis, hablan y escriben la lengua Yaqui, se han aplicado 707 encuestas en las secundarias, preparatorias y universidad ubicadas en la comunidad Yaqui. Para el presente estudio es de interés conocer si el joven es miembro de la tribu, si habla la lengua, si escribe la lengua y quien le enseño. Es importante conocer si estos jóvenes hablan Yaqui de manera cotidiana, dentro y fuera de casa, o bien si les da pena y cuál sería el motivo. De los jóvenes que no hablan la lengua, es de interés conocer si les gustaría aprender a hablar y escribir la lengua y porqué motivo, y finalmente conocer si les gustaría que sus hijos hablen la lengua.

2. ESTUDIOS SOBRE LA LENGUA YAQUI
Existen esfuerzos en el estudio de la lengua Yaqui, como lo es: la descripción de la lengua [7], morfología del Yaqui [8] y estudio de gramática [9]. Otros esfuerzos [3, 4, 5]. Y el diccionario Yaqui-español [6]. Según la obra de Estrada en [6] y en base a los comentarios de los nativos Yaquis, el léxico del yaqui nativo contiene los mismos caracteres léxicos del alfabeto español con excepción de los siguientes caracteres {c, d, f, h, ñ, q, v, x, z}. La simbología y números son los mismos utilizados en el Yaqui.

Anteriormente se mencionó que la lengua se ha adaptado debido a los cambios sociales y tecnológicos, lo cual ha provocado una evolución léxica y que ya no sea válido el léxico planteado en la literatura,-esto en el uso del yaqui actual-, esto debido a los préstamos lingüísticos del español e ingles al Yaqui. Al momento de analizar textos y nombres propios en Yaqui se encontró que no es posible la ejecución de un análisis léxico considerando solamente los caracteres válidos en la lengua antigua, esto debido a que el español tiene mucha influencia en el Yaqui y ha propiciado que los caracteres léxicos no validos en la lengua ahora se conviertan en válidos. Por lo tanto, en el léxico actual de la lengua Yaqui se ha transformado al grado de ser equivalente al léxico del castellano.

En [10] se examinan la estructura y funciones de las construcciones con argumento pospuesto en la lengua Yaqui, donde desarrollan un análisis de las construcciones en la que el objeto lógico realizado por una frase léxica, ocurre en posición postverbal, también presentan que la lengua es sintéticamente-aglutinante con un marco causal nominativo-acusativo y plantean que la lengua posee dos series de pronombres tradicionalmente llamados “plenos” y “reducidos”. La primera parte de la propuesta se centra en el estudio de oraciones de los pronombres nominativos: Yo(inepo), Tú(empo), El(aapo), Nosotros(itepo), Ustedes(Eme'e), Ellos(bempo). Según [10] el orden de los argumento léxicos en la lengua obedece a Sujeto Objeto Verbo (SOV). El equipo de lingüistas comparte esta aseveración. Basados en este orden sintáctico de la lengua presentamos un análisis de tiempo pasado (battu’uka), presente (iani) y futuro (yookomatchuco), así como los sufijos de uso en la lengua.

3. RESULTADOS DE ESTUDIO
A continuación presentamos en la tabla 1 los resultados obtenidos del estudio desarrollado para conocer si los jóvenes de la comunidad yaqui hablan la lengua Yaqui.

Tabla 1. Hablantes y Escritores del yaqui

	Todos		*No Nativos*		*Nativos*	
	Total	*%*	*Total*	*%*	*Total*	*%*
Total de Estudiantes que no hablan ni escriben la lengua	443	**62.66**	132	**91.67**	311	**55.239787**
Total de Estudiantes que solo hablan la lengua	264	**37.34**	12	**8.33**	252	**44.76**
Total de Estudiantes que hablan y escriben la lengua	142	**20.08**	8	**66.67**	134	**53.17**

Al analizar los tiempos presente, pasado y futuro encontramos la estructura gramatical para los pronombres la cual presentamos acontinuación:

presente: inepo->ne->chepte (yo brinco);
bempo||Aapo||Eme'e||Bempo||Itepo->chepte.

pasado:inepo->ne->cheptek(yo brinque);
Empo||Aapo||Eme'e||Bempo->chepte; Itepo ->te-> chepte.

pasado:inepo->ne->a->cheptak(yo brinque);
Empo||Aapo||Eme'e||Bempo->a->cheptak; Itepo->tea->cheptak.

futuro:inepo->ne->cheptibae(yo brincaré);
Empo||Aapo||Eme'e||Bempo->a->cheptibae; Itepo->te->cheptibae.

Con lo anterior generamos algunas reglas sintácticas para los pronombres:
Inepo -{ne, nea}-[predicado]-[verbo en {presente, pasado, futuro}]-{.}
Empo||Aapo||Eme'e||Bempo-{a}-[verbo en {Pasado, futuro}]-{.}
Empo||Aapo||Eme'e||Bempo-[verbo en {Presente, Pasado}]-{.}
Itepo -{te, tea}-[verbo en {pasado, futuro}]-{.}

Donde el guion separa los elementos de la oración, y finaliza con punto.

Con la aportación aquí presentada ya podemos evaluar si una oración básica de la lengua yaqui es correcta tanto de manera léxica y sintáctica, hemos desarrollado una primera versión del software que nos permite hacer esta evaluación y a la vez nos va indicando si lo escrito es correcto y si no lo es entonces nos indica la manera en que se debe de escribir. Otra de las funciones de esta primera versión es la traducción de estas oraciones del yaqui al español.

4. CONCLUSIONES Y TRABAJO FUTURO

La lengua Yaqui es una de las lenguas mexicanas en peligro de extinción debido a que el 20% de la población nativa es hablante de su lengua materna. Uno de los grandes problemas es que el número de hablantes va disminuyendo y solamente el 10% de los hablantes la escriben. Este fenómeno se presenta debido a la influencia de diferentes factores sociales entre ellos destaca el uso de las tecnologías y la influencia de los prestamos lingüísticos. La mayoría de los encuestados les gustaría aprender la lengua por motivos culturales, esto es una gran noticia ya que es factible aumentar el número de hablantes y escritores de la lengua, es por ello que en el presente trabajo se ha desarrollado un estudio para conocer los datos estadísticos del estatus que guarda el uso del Yaqui. También se han presentado el resultado del análisis gramatical y sintáctico del uso de los pronombres personales, logrando así un avance importante en la investigación que permitirá el desarrollo de una primera versión del software que nos permita analizar la escritura del yaqui.

5. REFERENCIAS

[1] Instituto Nacional de Lenguas Indígenas. “Catálogo de las Lenguas Indígenas Nacionales: Variantes lingüísticas de México con sus auto denominaciones y referencias geo estadísticas”. Diario Oficial de la Federación. Enero de 2008. URL: http://www.inali.gob.mx/pdf/CLIN_completo.pdf. Fecha de Consulta: 11 de agosto de 2016.

[2] A. Embriz Osorio y O. Zamora Alarcón. “México. Lenguas indígenas nacionales en riesgo de desaparición: Variantes lingüísticas por grado de riesgo. 2000”. Instituto Nacional de Lenguas Indígenas. ISBN 978-607-7538-57-8. Septiembre de 2012.
URL:http://site.inali.gob.mx/pdf/libro_lenguas_indigenas_nacionales_en_riesgo_de_desaparicion.pdf. Fecha de Consulta: 11 de agosto de 2016.

[3] E. Jelinek, F. Escalante. “Double Accusative Constructions in Yaqui. Proceedings of the Pacific Linguistics Conference”, University of Oregon, Eugene. 120-132. 1989.

[4] E. Jelinek. “Voice and transitivity as functional projections in Yaqui”. In Miriam Butt and Wilhelm Geuder (eds) “The projection of arguments: Lexical and compositional factors”. Stanford: CSLI Productions. Pp 195-225. 1998.

[5] E. Jelinek. “Quantification in Yaqui Possessive Sentences”. MIT Working Papers in Endangered and Less Familiar Languages 5. 201-214. 2003.

[6] Z. Estrada Fernández, C. Buitimea Valenzuela, A. E. Gurrola Camacho, M. E. Castillo Celaya and A. Carlón Flores. “Diccionario yaqui-español y textos: obra de preservación lingüística”. México: Plaza y Valdés. 2004.

[7] L. Guerrero. “Descripción de las cláusulas complejas en la lengua yaqui, Friends of Uto-Aztecan Working Conference”. México: Instituto Nacional de Antropología e Historia. 1997.

[8] L. Guerrero. Universidad de Búfalo, La Universidad Estatal de Nueva York, Tesis: “Foot structure and grouping harmony in Southern Tepehuan”. 2001.

[9] L. Guerrero., S. Ibáñez Cerda y A. Belloro. “Studies in Role and Reference Grammar”. México: UNAM. ISBN: 978-607-02-0659-7, 519. 2004.

[10] A. Belloro Valeria. Lilian, Guerrero. (2010). Sobre la función discursiva de la posposición de constituyentes en yaqui. Análisis lingüístico: enfoques sincrónico, diacrónico e interdisciplinario. Estudios Lingüísticos 2. Rosa María Ortíz Ciscomani (ed). Hermosillo: Universidad de Sonora, ISBN: 978-607-7782-71-1. 71-95.

Sistema de Consulta de Datos Semánticos para Dispositivos Móviles

Frida Cardenas-Morales
Facultad de Ingeniería y Ciencias
Universidad Autónoma de Tamaulipas
Victoria, México.
a2161018001@alumnos.uat.edu.mx

Ana B. Rios-Alvarado
Facultad de Ingeniería y Ciencias
Universidad Autónoma de Tamaulipas
Victoria, México.
arios@docentes.uat.edu.mx

Alan Díaz-Manríquez
Facultad de Ingeniería y Ciencias
Universidad Autónoma de Tamaulipas
Victoria, México.
amanriquez@docentes.uat.edu.mx

Edgar Tello-Lea
Facultad de Ingeniería y Ciencias
Universidad Autónoma de Tamaulipas
Victoria, México.
etello@docentes.uat.edu.mx

RESUMEN
Debido a la gran variedad de información que se encuentra en la Web, esta se ha convertido en una enciclopedia virtual de conocimiento humano, por lo que los usuarios deben ser cada vez más específicos en las consultas realizadas a través de buscadores clásicos. Esto en gran medida puede ser poco práctico, sobre todo en la época actual donde comúnmente el acceso a la información es a través de un dispositivo móvil. Una consulta recurrente por parte de usuarios de dispositivos móviles es consultar los locales comerciales relacionados a un servicio específico. Por su parte, la *web semántica* propone añadir significado a la información existente, clasificando y estructurando los recursos para que sean procesados por las máquinas y aplicaciones de cómputo. Este documento aborda una propuesta que considera el diseño y desarrollo de un sistema que incluye una aplicación web y una aplicación móvil para la consulta de locales comerciales, incorporando tecnologías de la web semántica.

Conceptos CSS
Information Systems → Information Systems Applications

Palabras clave
Web semántica, aplicación web, aplicación móvil.

1. INTRODUCCIÓN
La información en la Web ha ido creciendo incluyendo cada vez más y más datos, lo que ha generado que los resultados a las consultas en los buscadores usualmente carezcan de sentido, como consecuencia los resultados de una búsqueda no son los deseados. Además, cuando se habla de Tecnologías de la Información, uno de sus aspectos actuales para la difusión, consulta, recuperación y búsqueda de información eficiente es incorporar metadatos para llevar a cabo estas tareas. Al agregar información del dominio a los datos a consultar se pretende obtener mejores resultados en las búsquedas. Esto se puede lograr representando la información a través de un modelo semántico.

Por otro lado, debido a la necesidad de interacción entre aplicaciones, surge en 2001 el concepto de *web semántica* propuesto por Tim Bernes Lee. Él define a la web semántica como una Web dotada de significado, ya que se le adiciona sentido a la información para que pueda ser entendible por los seres humanos y también por los agentes de software, permitiendo a su vez búsquedas "más inteligentes", denominadas búsquedas semánticas [1]. Por lo tanto, la representación de la información a través de un esquema semántico permitirá usar las nuevas tecnologías de la web semántica para el modelado y el almacenamiento de la información que hace posible desarrollar aplicaciones "más inteligentes".

Esta investigación propone la especificación de una aplicación web, para la captura de información de locales comerciales con servicio de alimentos, que almacena la información en archivos RDFS y OWL. También se realizará una aplicación móvil para realizar consultas sobre los locales registrados y los resultados serán mostrados en un dispositivo móvil. Como objetivo general se propone crear un modelo semántico de información comercial, así mismo desarrollar una aplicación de búsqueda semántica desde un dispositivo móvil y una aplicación web para la captura y edición de la información de locales comerciales.

2. REVISIÓN DE LA LITERATURA
Una ontología se define como una jerarquía de conceptos con atributos y relaciones, la cual describe un dominio [2]. Según su caso de estudio, la construcción de una ontología puede variar, existe un método el cual es denominado METHONTOLOGY [3] que propone realizar una ontología en 3 pasos: especificación, conceptualización y mantenimiento. La especificación consiste en determinar por qué se construye la ontología, cuál será su uso, y quiénes serán sus usuarios finales. La conceptualización determina el modelo computable resultado de la formalización e implementación. Finalmente, la actividad de mantenimiento se encarga de la actualización y/o corrección de la ontología, en caso necesario. Además, el autor relata que no existe una metodología base para el desarrollo de una ontología pero si se puede decidir entre las existentes y las que más se acoplen a lo que se necesita.

Hoy en día existen varias ontologías disponibles en la Web. Uno de los propósitos principales en el desarrollo o implementación de la web semántica es la reutilización de ontologías existentes. Una de las ontologías más utilizadas en el ámbito del comercio electrónico es la de *GoodRelations* que describe un vocabulario estandarizado para los productos en venta que incluye atributos como precio, datos de la empresa,

dimensiones, entre otras. Usando descripciones semánticas en páginas web se aumenta la visibilidad de los productos y provee servicios de última generación a los motores de búsqueda, sistemas de recomendación y otras aplicaciones novedosas [4]. Otra ontología existente muy reutilizada es la *GeoSpatial Ontology GEO* [5] que es un vocabulario que describe puntos de referencia con base en latitud, longitud y altitud. Estas ontologías fueron seleccionadas para reutilizar los conceptos que permitan describir locales comerciales y su ubicación geografica.

3. METODOLOGÍA

A continuación se explicará el diseño de la arquitectura propuesta y posteriormente se detallará el modelado de los datos.

3.1 Arquitectura

Para el diseño de la arquitectura se definió utilizar el modelo cliente-servidor que consta de 3 niveles los cuales son: 1) Presentación: provee la interfaz al usuario y provee algunas acciones; 2) Lógica: sirve como conexión intermediaria entre la presentación y los datos; 3) Datos: Provee el acceso a la información.

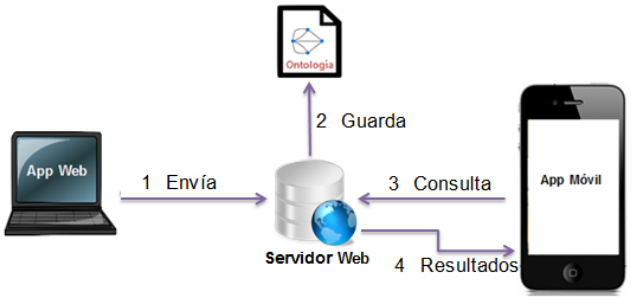


Figura 1. Diseño de la arquitectura propuesta

En la Figura 1 se presenta los procesos principales los cuales se describen a continuación: 1) la aplicación web enviará la información de los negocios al servidor, 2) la información enviada se almacena en la ontología (RDF/OWL); 3) posteriormente podrá ser consultada desde un dispositivo móvil, 4) ofreciendo resultados basados en la información semántica almacenada. El servidor web tendrá la función principal ya que será el mediador en donde se conectaran el dispositivo móvil y la aplicación web, lo cual se hará a través de peticiones de HTTP. Obteniendo como resultado la interoperabilidad entre las aplicaciones como también disminución de tiempo en las peticiones y finalmente la seguridad de los datos.

3.2 Modelo de Datos

El modelado de datos consiste en la estructuración del desarrollo de la ontología la cual es la parte esencial del proyecto, debido a que en ella se representará toda la información referente a los locales. Para la creación de una ontología se requiere de una metodología o guía de pasos, se siguio la propuesta por [3]. Se definió el área que cubrirá la ontología a desarrollar la cual es el área de locales comerciales con el servicio de alimentos. Para ello se reutilizaron etiquetas de ontologías existentes. Las etiquetas reutilizadas de *GoodRelations* son: *DayOfWeek*, *PaymentMethod*, *Location* y *OpeningHoursSpecification*. De la ontología *GEO* se reutilizaron las etiquetas de latitud y longitud para almacenamiento de la posición geográfica de la ubicación del local.

Tabla 1. Elementos de la Ontología

Elementos	Total
Conceptos	65
Relaciones	63
Atributos	43
Axiomas	423
Instancias	30

Como resultado de la etapa de conceptualización en la Tabla 1 se muestran los elementos que componen la ontología desarrollada en este proyecto. Se incluyen 30 instancias que representan los datos de comercios que han sido capturados a través de entrevistas a propietarios de locales comerciales. En la Figura 2 se muestra de lado derecho las clases con jerarquía, también se visualiza en la esquina superior derecha las propiedades de los objetos; y finalmente en la esquina inferior derecha se muestra las propiedades de los datos de la ontología creada.

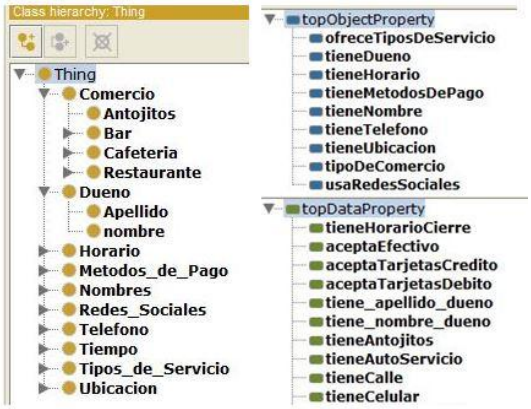


Figura 2 Clases, propiedades de objetos y propiedades de datos de la ontología creada

4. CONCLUSIONES

Se ha presentado la descripción de un sistema para consulta de información que incorpora tecnologías de la web semántica. El resultado del proyecto propuesto se prentende evaluar y dar a conocer a los propietarios de locales comerciales para su uso en un entorno real.

5. REFERENCIAS

[1] Rafael Pedraza-Jiménez, Lluís Codina y Cristófol Rovira, "Web Semántica y Ontologías en el Procesamiento de la Información Documental", El Profesional de la información, noviembre-diciembre 2014.

[2] Lluís Codina y Cristòfol Rovira, “La Web Semántica”, Departamento de Periodismo y de Comunicación Audiovisual, Instituto Universitario de Lingüística Aplicada, noviembre 2006.

[3] Oscar Corcho, Mariano Fernández-López, Asunción Gómez-Pérez, Angel López Cima, "Construcción de ontologías legales con la metodología METHONTOLOGY y la herramienta WebODE", Facultad de Informática. Universidad Politécnica de Madrid, Madrid, España, 2004.

[4] GoodRelations, "The web vocabulary for e-commerce", Disponible en: http://www.heppnetz.de/ontologies/goodrelations/v1.html, Consultado: 27 de Abril del 2016

[5] [26] GEO, "GeoNamesOntology", Disponible en: http://notes.3kbo.com/geo, Consultado: 27 Abril 2016

Predicción de la precipitación pluvial mediante algorítmos genéticos

Lorena Carmona
Universidad Autónoma de Zacatecas
Jardin Juarez 147
Centro Historico, Zacatecas
carmonalore96@gmail.com

Andrés Martinez
Universidad Autónoma de Zacatecas
Jardin Juarez 147
Centro Historico, Zacatecas
andresmtzv07@gmail.com

Daniel Cervantes
Universidad Autónoma de Zacatecas
Jardin Juarez 147
Centro Historico, Zacatecas
danielcervantesguerrero@gmail.com

Carlos E. Galván-Tejada
Universidad Autónoma de Zacatecas
Jardin Juarez 147
Centro Historico, Zacatecas
ericgalvan@uaz.edu.mx

Jorge I. Galván-Tejada
Universidad Autónoma de Zacatecas
Jardin Juarez 147
Centro Historico, Zacatecas
gatejo@uaz.edu.mx

ABSTRACT

Climate changing has affected the maximum and minimum water precipitation rates in different locations, causing alterations in the agriculture and other areas. Using 25 years of measurements from weather stations a multivariate model based on genetic algorithms is presented. The goal of this paper is to generate a multivariate predictive model for rain in future years.

1. RESUMEN

La predicción oportuna del clima es un factor importante al momento de combatir los estragos del cambio climático. A pesar de que la metereología moderna permite predecir el clima, estas predicciones siguen siendo inexactas e inclusive equivocadas. El propósito de este trabajo es el de obtener un modelo multivariado basado en algoritmos genéticos a partir de mediciones en estaciones metereológicas.

Keywords

Cambio climático; predicción de clima; precipitación

2. INTRODUCCIÓN

La predicción de clima tiene una infinidad de utilidades,tales como mejor aprovechamiento del ciclo hidrológico; puede ser utilizado como una herramienta fundamental,para determinar el tipo de cultivo idóneo en cada región de un país, así como la construcción de obras civiles e hidráulicas. El cambio climático es un problema de relevancia internacional y es la principal amenaza al medio ambiente en la edad moderna. Actualmente, incluso países con clima principalmente templado están sufriendo las concecuencias, estos cambios vienen ligados a climas extremos que rebasan los 40 grados centígrados. De acuerdo con estudios recientes los cambios climáticos han desempeñado un papel como catalizadores ante catástrofes en la biósfera tales como la extinción de las especies, desastres naturales y las crisis alimentarias [5]. Uno de los grandes problemas asociados, es la variacion de la temperatura, debido a que esto ayuda a la generación de huracanes. Esto afecta directamente a la precipitación trayendo consigo grandes desastres e inudanciones. En las ciudades pequeñas, un gran problema es la ausencia de estaciones metereológicas con tecnología capaz de hacer mediciones de precipitación, radiación ultravioleta, presión o humedad, por lo cual advertir a la población acerca de en qué momentos del día la radiación tendría repercusiones en su salud es una tarea difícil [1]. Este estudio pretende, mediante mediciones de variables sencillas, mejorar el pronóstico del clima en la ciudad donde se realicen las mediciones [3]. El uso de algoritmos genéticos para la obtención de modelos, permite el análisis de grandes cantidades de información, así como obtener asociaciones que no serían posibles a través de métodos estadísticos tradicionales. En este estudio, modelos de predicción son generados a partir de 25 años de datos en estaciones meteorológicas del estado de Zacatecas.

3. OBJETIVO

El objetivo de este trabajo es obtener un modelo de predicción de precipitación de lluvia en el estado de Zacatecas, específicamente en la zona de Calera de Víctor Rosales. Como objetivos particulares se enciuentran los siguientes:

- Identificar las variables más relacionadas con la precipitación pluvial
- Obtener un modelo de 5 variables para predicción de precipitación pluvial.
- Obtener los modelos más eficientes mediante el uso de inteligencia artificial y métodos computacionales.

4. MATERIALES Y MÉTODOS

4.1 Descripción de datos

La base de datos usada en este trabajo se encuentra disponible a través de Comisión Nacional del Agua (CONAGUA), consta de las mediciones de variables como la temperatura máxima, mínima, precipitación, humedad. Estas mediciones se realizaron a varias ciudades y comunidades del estado de Zacatecas, el modelo presentado se realizó con mediciones realizadas en Calera de Víctor Rosales.

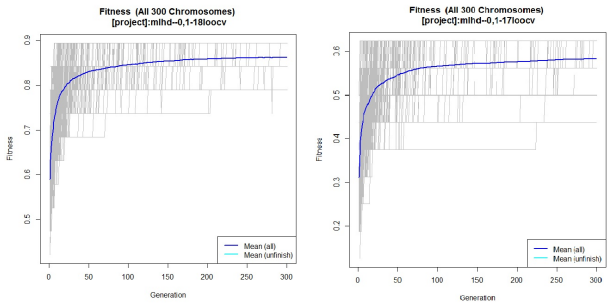


Figure 1: Progresión de eficiencia de los modelos Junio-Julio

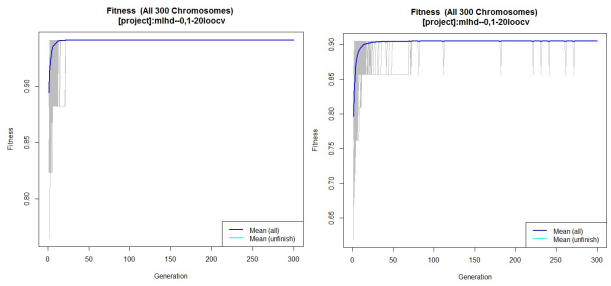


Figure 2: Progresión de eficiencia de los modelos Agosto-Septiembre.

Las bases de datos cuentan con un histórico de mediciones desde 1941 a la fecha, por lo que se tomaron 25 años de mediciones para el cálculo del modelo, de 1980 a 2005, y la variable a predecir fue la precipitación del año de 2007.

Dado que las bases de datos se encuentran agrupadas por meses, se utilizaron en este trabajo los meses de Junio, Julio, Agosto y Septiembre, siendo los meses más representativos en los periodos agrícolas.

4.2 análisis de datos

Para el análisis de datos se realizaron cuatro búsqeudas multivariadas. Cuatro modelos de predicción fueron obtenidos usando algoritmos genéticos, la función de costo o clasificación utilizada fue un discriminante lineal. El método utilizado para la validación cruzada fue ”leave one out cross validadtion” (LOOCV).El algoritmo genético se configuró para obtener 300 modelos después de recombinar 300 generaciones. Un ranking de las variables más significativas dentro de los cromosomas fue calculado en base a su aparición dentro de los modelos.

El número de variables usadas fueron 420. Cada modelo de predicción se formó de 5 variables o genes. La variable de salida fue la precipitación pluvial presentada en el año 2007 en el més correspondiente a cada modelo. Debido a la longitud propia de lso smeses, las observaciones fueron 30 o 31 dependiendo el caso del mes y su número de días.

Todo el análisis estadístico, así como los algoritmos genéticos, fueron realizados utilizando R Software version 3, así como sus paquetes de algoritmos genéticos (GALGO)[4, 2].

5. RESULTADOS

Después de el análisis realizado, se obtuvieron cuatro modelos de predicción, Tres resultaron altamente predictivos presentando una eficiencia de clasificación de más del 85 %.

En la figura 1, se presenta la evolución en el desempeo de clasificación de los modelos de los meses de junio y ulio de 2007 respectivamente. En la figura 2, se presenta el desempeño de clasificación de los modelos de agosto y septiembre. En el caso de los modelos de junio, agosto y septiembre, la eficiencia superó el 85 %, mientras qe en el més de Julio la eficiencia alcanzó un 65 % de predicción.

6. CONCLUSIONES

En base a los resultados experimentales, se concluye que es posible realizar una buena predicción de la precipitación pluvial en base a los datos históricos, ya que son robustos a los cambios en las cifras de lluvia anualizada.

Es evidente que en el més de Julio hay una alta variabilidad de lluvia en el municipio, por lo que es posible que requiera una mayor cantidad de años de estudio para una correcta predicción en base al modelo.

Es posible con ayuda de algoritmos genéticos e inteligencia artificial logra una predicción fiable del clima y sus cambios enfocados al área agro industrial.

7. TRABAJO FUTURO

Como trabajo futuro se plantea el realizar la predicción en todos los meses del año y con un mayor número de datos. También se planea realizar estrategias para refinar los modelos de predicción en base a métodos estadísticos avanzados.

El uso de otras funciones de clasificación, así como de diferentes configuraciones del algoritmo genético porían mejorar el desempeño del modelo.

8. REFERENCIAS

[1] V.˜Magaña, L.˜C. López, and G.˜Vázquez. El pronóstico de lluvias intensas para la ciudad de méxico. *TIP*, 16(1):18–25, 2013.

[2] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[3] D.˜Rogers. Predicting weather using a genetic memory: a combination of kanerva's sparse distributed memory with holland's genetic algorithms. In *Advances in neural information processing systems*, pages 455–464, 1990.

[4] V.˜Trevino and F.˜Falciani. Galgo: an r package for multivariate variable selection using genetic algorithms. *Bioinformatics*, 22(9):1154–1156, 2006.

[5] J.˜M.˜S. Viñas. Volatilidad de los mercados agrarios y crisis alimentaria. *Revista española de estudios agrosociales y pesqueros*, (229):11–35, 2011.

Análisis y Desarrollo de Algoritmo Genético para la Alineación de Cadenas de ADN

Edith A. Escarrega Payan
Universidad Autónoma de Baja California
Facultad de Ingeniería, Arquitectura y Diseño
Ensenada Baja California
Tel: 646 1135132
aescarrega@uabc.edu.mx

Eduardo Vázquez Santacruz
CONACYT-UNAM
evazquezsa@conacyt.mx

RESUMEN
En el presente trabajo se habla de las necesidades que tiene la bioinformática para ayudar a comparar cadenas de ADN con ayuda de un software especializado, capaz de detectar las diferencias y similitudes entre dos cadenas de organismos diferentes. El objetivo que se persigue es la optimización de los recursos de tiempo y espacio de los esquemas de búsqueda, para ello se utilizaron como base los algoritmos genéticos logrando de esta manera conseguir la mejor alineación de las secuencias introducidas.

Conceptos CCS
• Metodologías de computación → Inteligencia artificial.

Palabras Clave
Algoritmos genéticos; ácido desoxirribonucleico (ADN) ; genética; bioinformática; nucleótidos; optimización.

1. INTRODUCCIÓN
La comparación de secuencias biológicas (ADN, ARN (ácido ribonucleico) o proteínas) juega un papel fundamental en la investigación genética, tanto a nivel teórico (por ejemplo, para el análisis evolutivo de las especies mediante árboles filogenéticos) como para la comprensión y tratamiento de enfermedades. Cuando se analiza el genoma de una especie se obtiene un conjunto de genes cuya función no es fácil de establecer. Lo mismo ocurre cuando se determina la secuencia de aminoácidos que compone una proteína (secuencia que puede obtenerse también mediante el gen o ARN mensajero correspondiente). Sin embargo, si comparamos estas cadenas recién obtenidas con secuencias ya estudiadas y cuya función se conoce podemos deducir la función de las primeras apelando a la homología de las secuencias. Dicho de otra forma, si la composición y orden de sus componentes (nucleótidos o aminoácidos) son similares, podemos pensar que su función celular será igual o semejante [1].

En los últimos años, la Bioinformática ha atraído una gran atención de varias disciplinas, como la informática, las matemáticas y las ciencias biológicas no tradicionales. Esto se debe a la disponibilidad de enormes cantidades de datos biológicos públicos y privados, y a la necesidad imperiosa de transformar datos en información biológica útil y en conocimiento [4].

Se trata de algoritmos de búsqueda basados en la mecánica de la selección natural y de la genética. Utilizan la información histórica para encontrar nuevos puntos de búsqueda de una solución óptima del problema planteado, con esperanzas de mejorar los resultados. En este sentido, podemos decir que los algoritmos genéticos combinan la supervivencia de los más compatibles entre las estructuras de cadenas, con una estructura de información ya aleatorizada, intercambiada para construir un algoritmo de búsqueda con algunas de las capacidades de innovación de la búsqueda humana [2].

Los algoritmos genéticos permiten la exploración de opciones mucho más amplias de posibles soluciones que los programas tradicionales. La novedad que introducen los algoritmos genéticos es que explotan eficientemente la información histórica para especular sobre nuevos puntos de búsqueda, esperando un funcionamiento mejorado[3]. Los algoritmos genéticos operan según:

- Selección o reproducción: este operador escoge cromosomas entre la población para efectuar la reproducción.
- Cruce: imita la recombinación biológica entre dos organismos haploides.
- Mutación: este operador produce variaciones de modo aleatorio en un cromosoma.

2. OBJETIVOS
2.1 General
El objetivo de este proyecto es generar una herramienta automatizada que permita ayudar en la alineación de secuencias de ADN con la implementación de algoritmos genéticos.

2.2 Específicos
1. Obtener de bases de datos genomas completos de diferentes especies para ser analizados.
2. Analizar y desarrollar una herramienta basada en algoritmos genéticos para la alineación de cadenas de ADN.

3. MÉTODOS
3.1 Herramientas utilizadas
3.1.1 GenBank
Es la base de datos de secuencia genética de los Institutos Nacionales de la Salud (NIH por sus siglas en inglés), una colección comentada de todas las secuencias de ADN disponibles públicamente. Es de esta página donde se podrá localizar y descargar genomas completos de diversos organismos para su posterior análisis.

3.1.2 NetBeans
Es un entorno de desarrollo - una herramienta para que los programadores puedan escribir, compilar, depurar y ejecutar

programas. Está escrito en Java - pero puede servir para cualquier otro lenguaje de programación [6].

3.2 Solución propuesta

Mediante la utilización de algoritmos genéticos se realiza el alineamiento y la comparación de dos diferentes cadenas de ADN, con el fin de obtener el alineamiento más óptimo así como encontrar la similitud entre ambas cadenas.

Para ello se utilizo el siguiente pseudocódigo de un algoritmo genético como base:

1. Elegir la población inicial de individuos
2. Evaluar la función objetivo (fitness) de cada individuo de la población
3. Repetir en esta población hasta terminar .
4. Seleccionar los mejores individuos para la reproducción (padres).
5. Crear nuevos individuos (hijos) a través de los operadores cruce y mutación.
6. Evaluar la función objetivo en los hijos.
7. Reemplazar los padres por los hijos.

En la diagrama 1 se muestra el diagrama de flujo simplificado con las partes más importantes de un algoritmo genético clásico: selección, reproducción y sustitución [5].

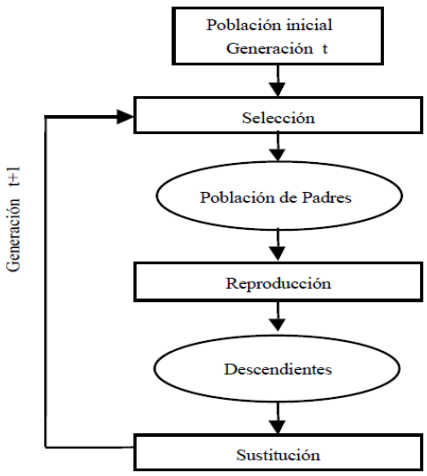


Diagrama 1: Estructura general de un algoritmo genético.

Con el fin de buscar una optimización de recursos dado que se tiene que realizar una alineación global se opto por realizar alineaciones locales de la secuencia (Ver figs. 1 y 2), para ello se han dividido las secuencias en filas las cuales se analizaran una a una, ahorrando tiempo valioso para el usuario.



Fig 1: Ejemplo de un alineamiento global (Solanilla et. al. 2005).



Fig 2: Ejemplo de un alineamiento local (Solanilla et. al. 2005).

4. RESULTADOS

Ingresando al programa y corriéndolo la pantalla principal nos mostrara la siguiente información (Ver fig. 3), en donde se le solicita al usuario ingresar el nombre de las cadenas de ADN a ser analizadas, las cuales fueron previamente descargadas de GenBank.

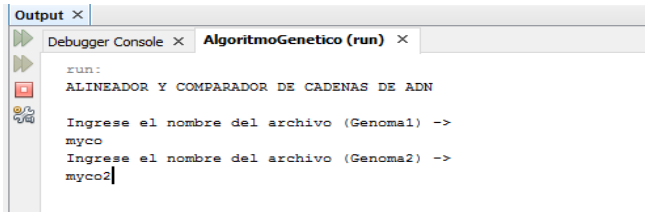


Fig. 3: Ingreso de la Mycobacterium tuberculosis H37RV y Mycobacterium tuberculosis H37RVSiena al programa para ser alineadas.

En las primeras pruebas que se realizaron del código se opto por usar fragmentos de una cadena pequeña y nos mostro los siguientes resultados en solo 5 segundos (Ver fig. 4).

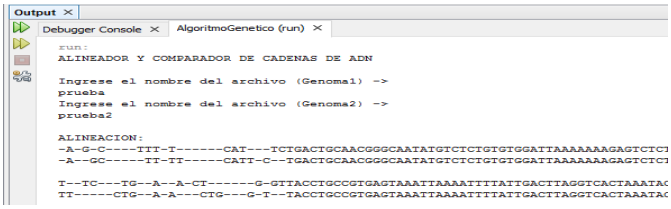


Fig. 4: Alineamiento de las secuencias de prueba.

5. CONCLUSIONES

Se cumplieron satisfactoriamente los objetivos planteados al inicio de la investigación, pues en este momento se cuenta con un código capaz de comparar secuencias de ADN de diferentes organismos utilizando algoritmos genéticos. Es necesario jugar un poco con los parámetros del algoritmo tales como el número y longitud de los gaps que se pueden insertar en las secuencias, el tamaño de la población, la probabilidad de cruce y mutación; con el fin de obtener el mejor alineamiento.

6. AGRADECIMIENTOS

Este trabajo hubiese sido imposible sin la participación de ciertas personas e instituciones que han facilitado las cosas para que llegase a un feliz término, empezando por el Programa Delfín por brindarme la oportunidad de participar en el XXI Verano de Investigación Científica, al Dr. Eduardo Vázquez por recibirme y asesorarse, así como a las facilidades brindadas por la Facultad de Medicina de la UNAM y el apoyo recibido por parte de UABC.

7. REFERENCIAS

[1] David W. Mount. Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press (2013).

[2] Goldberg, D. (1989) Genetics Algorithms in Search, Optimization and Machine Learning. Addison Wesley

[3] Piedad T. Introducción a los Algoritmos Genéticos y sus Aplicaciones. Madrid: Paseo de los Artilleros; 2008 [cited 22-05-2013]. Disponible en: http://www.uv.es/asepuma/X/J24C.pdf.

[4] Geer, Renata C. y SAYERS, Eric W. Entrez: Making use of its power. En: Briefings in Bioinformatics. vol. 4, no. 2. Junio, 2003. p. 179.

[5] Solanilla L.F., Gómez Teshima C.A., Múnera L.E. Alineamiento de múltiples secuencias de aminoácidos usando algoritmos genéticos. Sistemas & Telemática, vol. 3, núm. 5, enero-junio, 2005, pp. 53-71. Universidad ICESI. Cali, Colombia.

[6] Página Oficial de NetBeans, recuperado el 20 de Julio del 2016: https://netbeans.org/index_es.html

Implementación de mapping personalizado en Elasticsearch para determinar la importancia de los perfiles de GitHub de acuerdo a su experiencia

Joyce Selene Anaid Lozano Aguilar
Programa de Ingeniería de Software
Universidad Autónoma de Zacatecas
Ciudad Universitaria Siglo XXI
Zacatecas,México
joycseleneanaid@gmail.com

Alejandra García-Hernández
Programa de Ingeniería de Software
Universidad Autónoma de Zacatecas
Ciudad Universitaria Siglo XXI
Zacatecas,México
alegarcia@uaz.edu.mx

Alejandro Mauricio González
Programa de Ingeniería de Software
Universidad Autónoma de Zacatecas
Ciudad Universitaria Siglo XXI
Zacatecas,México
amgdark@auz.edu.mx

Resumen

Actualmente, existe una gran variedad de métodos y herramientas que permiten almacenar y buscar información, como bases de datos relacionales, no relacionales o motores de búsqueda *full-text search*. Dentro del tratamiento y análisis de la información resulta de gran interés conocer la frecuencia con la que se repite una palabra o frase dentro de una colección de datos, lo anterior nos permite determinar la relevancia que tiene un documento sobre otro, y así definir el orden en que serán mostrados los resultados de una consulta. Sin embargo, hay casos en donde las consultas requieren ser orientadas hacia un campo específico del documento, y si se mide la relevancia considerando el documento completo y no sólo ese campo, es posible que los resultados no coincidan con los resultados esperados de la consulta realizada.

Por tal motivo es necesario determinar la relevancia de los documentos en base al campo específico que abarque los datos de interés de las consultas. A través del aprovechamiento de las características del motor de búsqueda Elasticsearch, se implementó un *mapping personalizado* que permite determinar la relevancia de los documentos considerando sólo el campo que contiene los datos de interés. El mapping personalizado fue aplicado sobre un conjunto de 100 perfiles de usuarios de GitHub y las consultas fueron realizadas sobre el campo *lenguaje de programación*. Los resultados obtenidos muestran que la relevancia de los documentos coincide en mayor proporción en las consultas realizadas implementando el mapping personalizado que con el mapping por default.

CCS Conceptos

•**Information systems** → *Wrappers (data mining); Retrieval effectiveness; Retrieval efficiency;*

Palabras clave

Elasticsearch; GitHub; Big Data; Mapping; Relevancia

1. ANTECEDENTES

Hoy en día el término Big Data ha resonado en la mayoría de los sectores económicos, sociales y laborales, provocando la ardua búsqueda por manipular y aprovechar las grandes masas de datos que se encuentran circulando en la nube de internet. Los datos al provenir directamente de los usuarios, a través de redes sociales, smartphones, sitios web, en sí, cualquier medio perteneciente al mundo digital, se transforman en datos potencialmente de interés para la solución de diferentes tipos de problemas, ya sean estadísticos, predictivos, analíticos, entre otros.

Existen diversas tecnologías Big Data para realizar búsqueda de texto, una de las más recientes es Elasticsearch, un motor de almacenamiento, búsqueda y análisis open-source basado en Apache Lucene[2]. Elasticsearch es considerado como el futuro de las búsquedas de texto[4] y ha sido implementado en investigaciones de minería moderna, gracias a que cuenta con características como la esquematización *mapping*[3].

Elasticsearch muestra el orden de los resultados más relevantes de sus búsquedas de acuerdo a un peso (W), el cual es calculado considerando la frecuencia del término de la búsqueda en el documento (tf), la inversa de la frecuencia del documento (idf) y el valor de normalización (norms) sobre el campo que contiene los datos de interés. Sin embargo, cuando los datos de interés de una consulta se encuentran sólo en un campo específico del documento, no encontramos un mapping definido en Elasticsearch para estos casos.

2. OBJETIVOS

- Analizar y evaluar el algoritmo empleado para calcular la relevancia de documentos en Elasticsearch.
- Diseñar e implementar un mapping personalizado enfocado en determinar la relevancia de los documentos, contemplando los valores de tf e idf sobre un campo específico.
- Evaluar los resultados antes y después de implementar el mapping personalizado sobre el índice de Elasticsearch.

3. MATERIAL Y MÉTODOS

El diseño e implementación del *mapping personalizado* fue empleado sobre datos fuente provenientes de una muestra

de 100 perfiles de usuarios pertenecientes a la plataforma de desarrollo colaborativo GitHub[1]. El proceso que se siguió para la implementación del mapping se puede observar en la Figura 1.

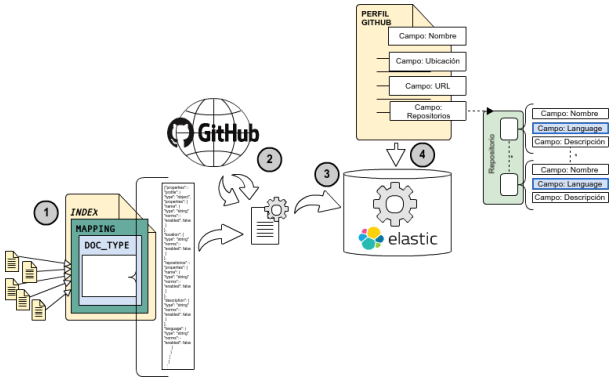


Figura 1 Proceso de implementación *mapping personalizado* en Elasticsearch: 1)Mapping, 2)Extracción de Datos, 3)Almacenamiento, 4)Consultas por campo

A continuación, se describirá cada una de las etapas del proceso.

1. ***Mapping***

Mapping es el término utilizado para referirse a los esquemas en Elasticsearch, éste al ser de "schema free" no requiere la definición de un esquema, ya que lo asigna automámicamente. Aunque en algunos casos, el emplear un mapping personalizado permite definir ciertas especificaciones de las propiedades de los campos dentro de la estructura del documento JSON[2].

Antes de definir el mapping, es importante determinar los campos que serán mapeados. En nuestro caso, fueron seleccionados los campos: nombre, ubicación, url y repositorios, éste último contiene un conjunto de repositorios, cada uno con un campo de nombre, lenguaje de programación y descripción.

Debido a que el algoritmo de Elasticsearch enfoca la determinación de la relevancia utilizando tres términos: tf, idf y norms[2], siendo éste último un valor que afecta significativamente al cálculo de la relevancia de los documentos, se optá por deshabilitarlo en cada uno de los campos definidos anteriormente, de manera que sólo serán utilizados los valores tf e idf para calcular el peso por documento.

El mapping personalizado es aplicado sobre el índice específico donde serán almacenados los documentos.

2. ***Extracción de Datos***

La extracción de los datos fue realizada a través de la automatización de un script empleando el API de GitHub. Aprovechando el flujo de datos, el script fue utilizado para mapear los datos en base al mapping que fue definido anteriormente, para después, indexarlos en Elasticsearch.

Indexar hace referencia al acto de asignar un documento JSON a un índice específico[2].

3. ***Almacenamiento***

Conforme va siendo procesado el flujo de datos, los documentos JSON generados son enviados a través de una instancia de Elasticsearch, en la cual se especifica el índice y tipo de documento donde serán almacenados.

4. ***Consultas por campo***

Finalmente, pueden realizarse consultas generales o específicas sobre el contenido de los documentos almacenados en Elasticsearch.

Puesto que el campo de repositorios contiene los datos de interés para poder determinar la experiencia práctica de cada usuario de GitHub, sobre los lenguajes de programación, éste se convierte en el campo filtro donde serán realizadas las consultas.

4. RESULTADOS

Después de realizar la implementación de mapping sobre la colección de 100 perfiles de usuarios de la plataforma colaborativa GitHub, almacenada en Elasticsearch, los resultados obtenidos mostraron que utilizando el mapping default, la relevancia de los documentos no coincide con el orden de los documentos que esperábamos, ya que el valor de tf del documento más relevante no siempre es el más alto. Por el contrario, al implementar nuestro mapping personalizado, como podemos observar en la Tabla 1, el documento con el valor de tf más alto, presenta también mayor peso (W). Lo anterior, nos permite encontrar aquellos perfiles con mayor práctica en los lenguajes de programación, como respuesta a las consultas realizadas.

Tabla 1 Comparación entre los resultados obtenidos sin y con mapping personalizado

Mapping Default		Mapping Personalizado	
TF	W	TF	W
5	1.0598	221	239.42
9	1.0156	193	209.0926
6	0.9951	135	146.2565
135	0.9834	125	135.4227

5. CONCLUSIONES Y TRABAJO FUTURO

Elasticsearch es un poderoso motor de búsqueda de texto *full-text search* que soporta la especificación de un esquema para definir el control del análisis en las búsquedas.

Implementar un *mapping personalizado* permitió especificar el control del cálculo de la relevancia de los documentos de acuerdo al campo de interés de las consultas.

Trabajo Futuro:

Como trabajos futuros, se tiene considerado implementar el mapping propuesto en diferentes redes sociales para realizar análisis de comportamiento de usuarios en redes sociales, análisis de sentimientos y de textos en sitios web considerando campos específicos.

REFERENCIAS

[1] GitHub. *GitHub.* 2016. URL: `https://github.com/about`.

[2] Gormley, C and Tong, Z. *Elasticsearch. The Definitive Guide.* O REILY., 2015.

[3] Kononenko, O., Baysal, O., Holmes, R., and Godfrey, M. *Mining Modern Repositories with Elasticsearch.* pages 1–4, June 2014.

[4] Mohan, V. *Elasticsearch Blueprints.* Packt Publishing Ltd, 2015.

Detección de Accidentes de Ciclismo Usando el Sensor de Aceleración: una Propuesta Preliminar Utilizando Redes Neuronales

Jonathan Martínez-D.Casas
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Zacatecas, Zac.
rockerjony@gmail.com

Carlos E. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería de Software
Zacatecas, Zac.
ericgalvan@uaz.edu.mx

Jorge I. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería Robótica y Mecatrónica
Zacatecas, Zac.
gatejo@uaz.edu.mx

José M. Celaya-Padilla
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería Eléctrica
Zacatecas, Zac.
jose.cpadilla@gmail.com

RESUMEN

En este trabajo se propone el uso de los datos del acelerómetro de un smartphone como fuente de reconocimiento ante la caída de un ciclista, y un posible accidente. Por las características de esta actividad física y recreativa se presentan escenarios de riesgo que pueden comprometer la vida de las personas que la realizan. Al ser frecuente la práctica en solitario se incrementa el riesgo frente a los accidentes. No obstante, el uso de un smartphone es recurrente en los ciclistas, se propone la aplicación de los sensores de este dispositivo con técnicas de "machine learning" para estimar si existe una anomalía en el comportamiento del ciclista y poder detectar un posible percance. Los resultados preliminares son positivos en la identificación y clasificación de accidentes.

Palabras Clave

acelerómetro, accidente, minería de datos

1. INTRODUCCIÓN

Los riesgos enfrentados por un ciclista cuando realiza su entrenamiento son demasiados. Conforme pasa el tiempo, se han implementado diversas medidas para prevenir los diferentes accidentes que pueden ocurrir. En México, de 2010 a finales 2014 han muerto mil 636 ciclistas en accidentes viales [1]. Esto se considera un problema de salud pública, incrementando la necesidad de proponer tecnología que incida en estas áreas de desarrollo e investigación: como puede ser el uso de sensores y sistemas de posicionamiento global para la detección de accidentes.

El uso de este sensor ha sido propuesto por Pingle et al [3] debido a que detecta la aceleración en el dispositivo,Su planteamiento presenta el uso de dispositivos como GPS (sistema de posicionamiento global), micro-controladores arduino, etc. los cuales al utilizar un algoritmo propio proponen un sistema embebido que puede ser útil en la mayoría de los accidentes.

Rajkumar, M. N. et al[4] plantea el uso del smartphone para la detección de accidentes usando el GPS y el acelerómetro ya que estos proveen suficiente información. Estos autores, proponen el manejo de una aplicación en smartphones que conjuga el uso de dispositivos a bordo como los acelerómetros y de esta forma detectar si el usuario ha caído o tirado su smartphone. Proponemos un uso similar del smartphone, de modo que nos permita explotar sus características: ser un dispositivo común, usualmente transportado por las personas y que cuenta con un acelerómetro, haciendo posible la detección de cambios repentinos o bruscos de velocidad, lo que podría significar que se produjo una caída.

2. OBJETIVOS

- Definir características extraídas de caídas con la clasificación binaria (caída 1/ no caída 0).
- Obtener modelo de predicción preciso el cual permita identificar el estado del ciclista.

3. MÉTODOLOGÍA

Para predecir si un ciclista tuvo un accidente, el modelo ha adquirido tres diferentes tareas. Usando los datos del acelerómetro como fuente de información, primero recolectamos datos de un ciclista en una condición normal y en condición de accidente; en segundo lugar aplicamos la extracción de características a estos datos sin procesar y finalmente entrenamos una red neuronal para adquirir nuestro modelo de predicción.

3.1 Recolección de Datos

Para recolectar datos se utilizó el accelerómetro de un smartphone con una frecuencia de muestreo de 100 lecturas por segundo, lo cual permite el uso de cualquier sensor, y la independencia de los sensores para conseguir una frecuencia mínima de muestreo. El sensor es colocado en el brazo derecho, como es usual en las bandas de brazo deportivas.

Los datos recolectados de los acelerómetros es para los tres ejes (x, y, z) para obtener toda la información posible del sensor.

3.2 Extracción de Características

Los datos crudos consisten en 3 vectores (x, y, z) de 100 datos por segundo por cada instancia. El número total de datos difiere en cada instancia, porque no todos los eventos de accidente o no accidente tienen el mismo tiempo. Para evitar esta restricción proponemos la extracción de características para reducir la cantidad de datos a un número fijo de características sin importar la longitud de la instancia original.

Se proponen los *cinco números*, descriptores estadísticos bien conocidos. Estas características son extraídas de los datos crudos, para obtener 15 características por cada instancia (5 por cada eje de aceleración). Características:Minimum,Lower quartile,Median,Upper quartile,Sample maximum.

3.3 Modelo de Predicción

Una vez terminada la extracción de características, adquirimos un conjunto de datos de 15 características por instancia que puede ser usado para alimentar a la red neuronal y poder obtener un modelo de predicción que nos permita identificar cuando ha ocurrido un accidente.

4. RESULTADOS

Estos son los resultados de cada una de las etapas de la metodología: Se capturaron 20 instancias. 10 en las cuales el ciclista caía de la bicicleta y 10 de un recorrido sin incidentes, el número mínimo de instancias fue determinado utilizando la ecuación de Eberhardt1 [2] para determinar el número mínimo de instancias e incrementado a 10 para tener un conjunto de datos que permitiera hacer la división para validación. Donde x es el número mínimo de instancias necesarias y N es el número de características independientes, en este caso 5 de cada eje, multiplicadas por 2 (accidente, no accidente).

$$x = log_2(N) + 1 \quad (1)$$

Se procedió a la extracción de características mostradas. Se obtuvo una matriz de 20 instancias por 15 características y 1 columna extra para describir si existía un accidente o no, fue dividida en 2 conjuntos, la validación compuesta por 6 instancias de accidente y 6 de no-accidente, seleccionadas al azar del conjunto total de datos, y el resto como conjunto de prueba, al cual le fue eliminada la columna de salida.

Se observa el total de instancias clasificada correctamente por la red neuronal. En la tabla 1 se muestra la clasificación.

5. CONCLUSIONES

Se presentó una propuesta preliminar para detectar, si un ciclista ha sufrido una caída. Los resultados obtenidos en la sección 4 nos permiten concluir lo siguiente:

- El acelerómetro genera suficientes datos para conocer el estado del ciclista

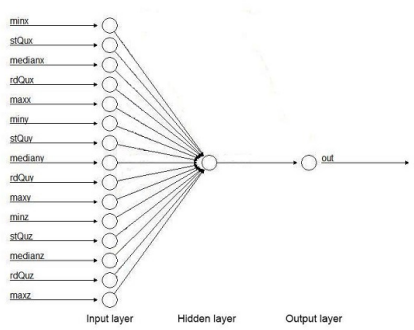


Figura 1 Configuración de red neuronal de capa simple.

Tabla 1 Prueba del modelo de predicción utilizando el conjunto de prueba.

Valor real	Valor estimado
0	0
0	0
1	1
0	0
0	0
1	1
0	0
1	1

- Una red neuronal sencilla tiene la capacidad de obtener un modelo de predicción utilizando los datos del acelerómetro a un costo computacional bajo.
- Las características estadísticas tienen información suficiente para modelar este tipo de eventos

Se propone como trabajo futuro lo siguiente:

- Incrementar el conjunto de datos para generalizar el modelo de predicción.
- Realizar un análisis de cada una de las características extraídas para determinar si relmente aportan información relevante al modelo.
- Revisar técnicas de machine learning para reducir el costo computacional.

6. REFERENCIAS

[1] Conapra: arrollan a más de mil ciclistas en 4 años. http://www.jornada.unam.mx/2015/12/17/sociedad/039n2soc, citado Abril 2016.

[2] F.˜Eberhardt. A sufficient condition for pooling data. *Synthese*, 163(3):433–442, 2008.

[3] P.˜P. Pingle, A.˜S. Marathe, and P.˜R. Ahirrao. Intelligent vehicle accident detection & notification system (vadans) using smart sensor and gps technologies. *Imperial Journal of Interdisciplinary Research*, 2(4), 2016.

[4] M.˜N. Rajkumar and P.˜Hemalatha. Automatic efficient accident detection and emergency system using smart phones:(accelerometer sensor). *Avinashilingam*, 2016.

The Aggregative K-means Clustering Algorithm

Fernando Orduña-Cabrera
Instituto Tecnológico Superior de Cajeme
Ciudad Obregón Son. México
forduna@itesca.edu.mx

Valeria Javalera-Rincón
Instituto Tecnológico Superior de Cajeme
Ciudad Obregón Son. México
vjavalera@itesca.edu.mx

Alberto Ramírez-Treviño
Instituto Tecnológico Superior de Cajeme
Ciudad Obregón Son. México
aramirez@itesca.edu.mx

ABSTRACT
In this paper, a modified version of the standard K-means algorithm to cluster a set of data is proposed. This approach adopts one policy based on the idea of adding more than one data to the same cluster in one step, driven by the strategy. Ten data sets from UCI are used to illustrate its effectiveness in reducing the computational time required without significantly affecting solution quality, and moreover improving the distribution of classes in the final clustering result.

Keywords
Clustering, k-means; Machine Learning.

1. INTRODUCTION
In the data mining field and especially in the inductive machine learning area one of the main task is to find and understand the data structure, hidden in a large amount of available data. A common procedure in dealing with these data is to classify or group them into a set of categories or clusters. The goal of clustering is to separate a finite unlabeled data set into a finite and discrete set of natural hidden data structures, rather than provide an accurate characterization of unobserved samples generated from the same probability distribution. Cluster analysis is being applied in a wide variety of disciplines such as medicine, psychology, biology, sociology, pattern recognition, and image processing. In re- search works such as [1, 2, 3, 4] can be found some overviews of clustering algorithms. Perhaps the best-known and most widely used method algorithm is the K-means algorithm. K-means is a squared error-based clustering algorithm. The algorithm is simple and can be used to solve a wide range of practical problems. The algorithm is a local search procedure, and it is well known that it suffers from the serious drawback of heavily depends its performance on the initial starting conditions [5]. In the literature a lot of extended versions of the K-means algorithm can be found. Some of these research works are [2, 3, 4, 5, 6, 7]. Several research works propose different modifications of the standard K-means algorithm. Almost all are focused on proposing a new way to compute the distance between the centroid and the data, getting a better performance. Others works are specially focused on reducing the computational load. With this later goal in mind, we propose a novel modification that adopts one policy based on the idea of adding to the nearest centroid, not only the less distance data, but all the data falling in an increased $\propto$ percentage of distance between the centroid and the data. This is summarized as, adding all data to the same cluster in one step driven by the policy. Ten data sets from UCI are used to illustrate its effectiveness in reducing the computational time required without significantly affecting solution quality. Iris dataset, and Abalone were selected to surf in the quality of this novel proposal, as described in conclusion section. In this an Aggregative K-means algorithm is proposed, sometimes referenced as policy. This algorithm reduces the computational time required without significantly affecting solution quality in comparison of standard k-means, and even more, our proposal improves the distribution of classes in the final cluster. The Aggregative K-means algorithm is presented as follows: 1. *Initialization*. We randomly choose K data points from the data set to initialize K cluster centroids, (C_1,C_2;… ;C_K). 2. *Coarse-Tuning*. Now, use the ordinary K-means algorithm with the Euclidean distance to compute the distance from data X_i = {1… n} to Centroids (C_1…C_K). 3. *Aggregative Step*. Once computed the distances, we need the minimal distance X* using the criterion$X^* = Arg\ min_{I=1\ldots N, J=1\ldots K} d(X_i, C_j)$: We find the cluster centroid C* which is nearest to X* using the minimum value criterion: $C^* = Arg\ min_{J=1\ldots K} d(X^*_i, C_j)$ If a data point X_i meets the condition $d(X_{i+1}, C^*) \leq (1+\alpha) d(X^*, C^*)$ Where $0 \geq \propto \leq 1$. Then assign the data X_i to the same cluster C* where belongs X*. 4. *Update Step* Compute the new centroids of the K clusters with: $C_K(t+1) = {}^1/_{N_K} \sum_{i \in S_K} X_i$. 5. *Continuation* if no cluster change happens then stops. Otherwise, go to the Coarse-Tuning step. The figure 1 it is a graphic view of the aggregative step proposed.

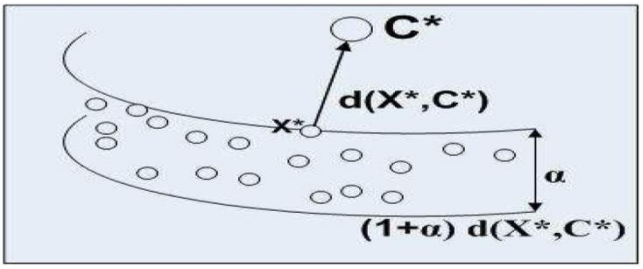


Figure 1. Aggregative function

Table 1. Datasets evaluated

Data Set	# Instances	# Class
Ecoli	337	8
Ionosphere	351	2
Vehicle	847	4
Iris	150	3
Pima	768	2
Balance-scale	625	3
Cars	1729	4
Abalone	4177	3
Waveform	5000	3

2. EVALUATION
To evaluate the performance of our proposed algorithm, we made experiments with ten datasets from the UCI machine learning repository showed in the table 1. In all data sets we conducted

experiments for the clustering problems obtained ignoring class labels. The number of instances and classes are relevant for the performance of our algorithm the results with standard k-means algorithm in terms of the total execution time and the quality of the distribution in the clusters.

The performance of our proposed aggregative K-means algorithm was evaluated in terms of the average time of the executions. In order to evaluate that, the following experiments were done: -Running ten times the standard k-means and taking note of the time of each execution. On each execution the number of k (clusters) was equal to the known number of classes the data set has. -Running ten times the Aggregative K-means with the same settings of the standard k-means, and the execution time was computed. The percentage of ∝ was set to 0.30 after some testing. Once done the executions of each algorithm the average of the ten times was computed. For instance, see the results in figures 2 and 3 and 4.

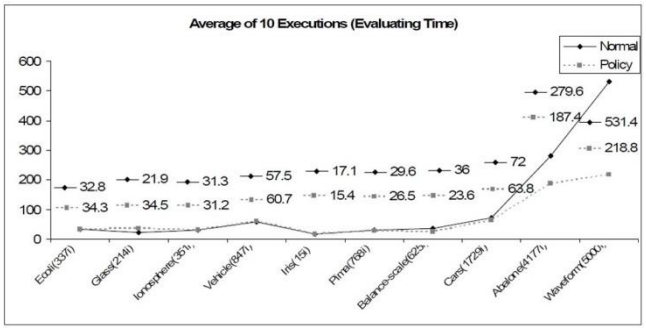


Figure 2. Results of time

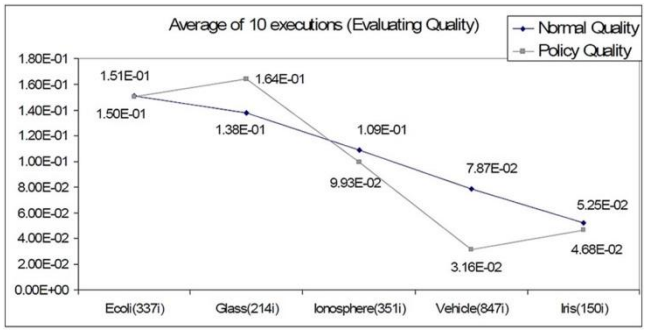


Figure 3. Quality results part 1

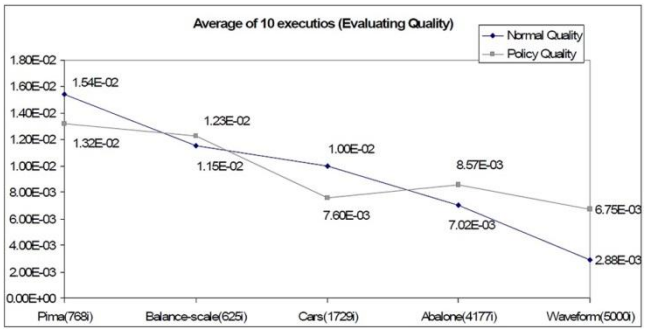


Figure 4. Quality results part 2

3. CONCLUSIONS

In this paper, a modified version of the standard K-means algorithm to cluster a set of data was presented. This approach adopts one policy based on the idea of adding more than one data to the same cluster in one step, driven by the proposed strategy. The result demonstrates how our proposal is capable to reduce the computational time in several datasets such as: Pima, Ionosphere, Abalone, Waveform, Iris, Balance-Scale and Cars, where the time required it is reduced considerably. In fact, in the data sets such as Ecoli and Vehicle, the difference of the time required in both cases is not too different. Some ideas arose in this work that consists in using intelligent agents to propose parallel intelligent clustering such [9], organizing the information in dynamical structures such in [8, 10]. The results regarding the quality of the clusters obtained shows that there is not a significant difference between the original K-Means and the proposed Aggregative K-Means method.

4. REFERENCES

[1] Rui Xu; Wunsch, D., II;, "Survey of clustering algorithms, Neural Networks", IEEE Transactions on, vol. 16, pages 645 - 678, year 2005.

[2] Member-Dimitrios Charalampidis, "A Modified K-Means Algorithm for Circular Invariant Clustering", journal. IEEE Trans. Pattern Anal. Mach. Intell., volume. 27, pages. 1856–1865, year. 2005.

[3] Mu-Chun Su and Chien-Hsing Chou, "A Modified Version of the K-Means Algo- rithm with a Distance Based on Cluster Symmetry", journalIEEE Trans. Pattern Anal. Mach. Intell., volume. 23, pages. 674–680, year. 2001

[4] Tapas Kanungo and David M. Mount and Nathan S. Netanyahu and Christine D. Piatko and Ruth Silverman and Angela Y. Wu, "An Efficient k-Means Cluster- ing Algorithm: Analysis and Implementation", journalIEEE Trans. Pattern Anal. Mach. Intell., volume.24, pages. 881–892, year. 2002.

[5] Aristidis Likas and Nikos Vlassis and Jacob J. Verbeek, "The Global K-means Clustering Algorithm", journalPattern Recognition, volume. 36, pages. 451–461, year. 2003.

[6] Ling Wang and Liefeng Bo and Licheng Jiao, "A Modified K-Means Clustering with a Density-Sensitive Distance Metric in Book Rough Sets and Knowledge Technol- ogy", pages. 544-551, year. 2006.

[7] Chinrungrueng, C. Sequin, C.H., "Optimal adaptive k-means algorithm with dynamic adjustment of learning rate", journal Neural Networks, IEEE Transactions on, volume. 6, pages. 157-169, year. 1995.

[8] A Dynamic Adaptive Framework for improving Case-Based Reasoning System Performance, Fernando Orduña-Cabrera, Ph. D Dissertation. Dept. of Comp Science, Univ. Politècnica de Catalunya, 2015.

[9] Agente Inteligente Para el Control de Clima en Invernaderos, ernando Orduña Cabrera, Valeria Javalera Rincón, Ma Guadalupe Alexandres García, Cesar E Rose Gomez.

[10] Using NIAR kd Trees to Improve the Case-Based Reasoning Retrieval Step, Fernando Orduña-Cabrera, Mexican International Conference on Artificial Intelligence, 314-325, Springer Berlin Heidelberg, 2013.

Desarrollo herramientas virtuales de terapia física y rehabilitación mediante Intel RealSense: Un enfoque preliminar

Maria L. Ortiz Pasillas
Universidad Autónoma de Zacatecas
Jardín Juárez 147
Centro histórico, Zacatecas
lulisop@hotmail.com

Cinthia B. Juárez Serrano Universidad Autónoma de Zacatecas
Jardín Juárez 147
Centro histórico, Zacatecas
cinhitajuarez@gmail.com

Kenia Miranda
Universidad Autónoma de Zacatecas
Jardín Juárez 147
Centro histórico, Zacatecas
miranda95k@gmail.com

Carlos E. Galván-Tejada
Universidad Autónoma de Zacatecas
Jardín Juárez 147
Centro histórico, Zacatecas
ericgalvan@uaz.edu.mx

Jorge I. Galván-Tejada
Universidad Autónoma de Zacatecas
Jardín Juárez 147
Centro histórico, Zacatecas
gatejo@uaz.edu.mx

ABSTRACT

Physical rehabilitation is a medical tool for retraining people affected by acquired injuries or diseases that may affect the nervous system, causing a malfunction of voluntary body control. Given the number of subjects requiring treatment, the purpose of this paper is to reduce the burden on public health facilities.

1. RESUMEN

Después de un accidente o de padecer alguna enfermedad crónica es necesario someter a un paciente a rehabilitación o terapia física, esto para manera agilizar su recuperación. En el presente artículo se pretende crear una simulación de actividades especializadas para ayudar a que el paciente pueda recibir rehabilitación virtual.

El mundo virtual es muy similar al de los juegos de video. Se requiere generar un ambiente virtual, que es el "lugar" donde se desarrollará la tarea asignada al usuario, el cual se debe sentir presente, inmerso en este ambiente simulado con el objetivo de hacer su experiencia lo más cercana y familiar a la realidad. Todo esto con el objetivo de llevar una sala de terapia física al entorno del paciente, con una aplicación directa en telemedicina y los costos de rehabilitación.

Keywords

Rehabilitación, RealSense, Software Médico, Procesamiento de imágenes

2. INTRODUCCIÓN

La rehabilitación es parte fundamental de un tratamiento, el cual incluye ejercicios físicos donde el paciente estimula la parte lesionada dado que sin atención oportuna, el daño puede ser irreparable [1]. En muchos lugares no se cuenta con centros de rehabilitación suficientes para los usuarios que lo requieren, además no se cuenta con suficientes terapeutas. Muchas de las veces el paciente no concluye el tratamiento debido a la distancia que tiene que recorrer para llegar al lugar donde se le brinda la atención médica especializada.

Nuestro proyecto ofrece rehabilitación virtual utilizando tecnología RealSense y de esta manera reduce la carga de trabajo del médico, reduce costo de infraestructura, costos de personal y al paciente le evita costos inecesarios de traslados. Las técnicas de realidad virtual, han mostrado tener un impacto en la rehabilitación de los pacientes [5, 2]. En este trabajo, y por medio de la tecnología de reconocimiento de formas de Intel RealSense se pretende desarrollar herramientas o aplicaciones para telemedicina[3, 4], basados en el reconocimiento de formas y moviemiento, es posible que por medio de entornos virtuales se puedena programar rutinas de terapia física desde un centro hospitalario a cualquier lugar del país.

En un país en desarrollo como lo es México es necesario encontrar alternativas que puedan apoyar al sector salud, ya que los costos de las terapias, y en ocaciones la falta de personal, provoca que los daños se vuelva irreversibles impactando en el costo de los servicios de salud pública.

3. OBJETIVOS

El objetivo de este trabajo es; utilizar la tecnología de reconocimiento de formas y patrones de RealSense para la creacción de aplicaciones para rehabilitación física o terapia de pacientes.

- Crear una aplicación en la cual el médico pueda programar terapia física del paciente.
- Obtener resultados concretos sobre el desempeño del paciente en su terapia.
- Implemetar los resultados obtenidos en entornos de telemedicina.

4. MATERIALES Y MÉTODOS

Para este trabajo la infraestructura y espacios necesarios serán proporcionados por el Laboratorio de Software

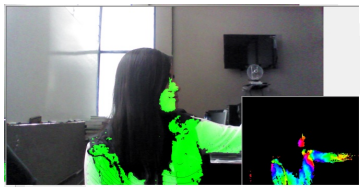


Figure 1: Captura de postura .

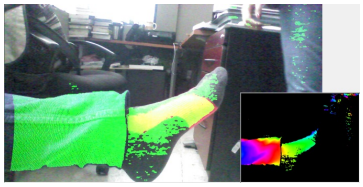


Figure 2: Extracción de formas.

libre (LabSol) del Consejo Zacatecano de Ciencia Tecnología (COZCYT) a través de los convenios existentes con la Universidad Autónoma de Zacatecas e Intel de México.

4.1 Sensores

La cámara Intel RealSense incorpora una útil tecnología en un formato pequeño. Existen tres cámaras que funcionan como si fuera una sola: una cámara HD de 1080p, una cámara infrarroja y un proyector láser infrarrojo. Estas cámaras "ven" como el ojo humano para detectar la profundidad y seguir el movimiento humano [6]. Mediante un bombardeo de luz infrarroja la cámara RealSense logra captar el movimiento del cuerpo humano

4.2 Selección de sujetos

Para el análisis, se verificarán el estado físico de los trabajadores de la industria. Por medio de un análisis de ergonomía y motricidad se determianrá cuál es la mejor terapia para cada individuo esta valoración la realizará personal calificado.

4.3 Análisis de datos

Para el análisis de imágenes, se utilizará el Software Development Kit (SDK) de intel RealSense para la captura de las imágenes. Las imágenes serán análizadas mediante métodos computacionales para obtener la información necesaria para el mapeo 3D del individuo mediante algoritmos morfológicos.

Todas las capturas de movimiento se relizarán mediante la cámara 3D Intel Creative para escritorio, modelo R200. El análisis de datos se realizará por medio de R Software y Matlab. Las aplicaciones para terapia y telemedicina serán realizadas en un entorno de Java en primera instancia.

5. RESULTADOS

Mediante el análisis preliminar se ha logrado el obtener las formas básicas del cuerpo mediante el procesamiento realizado con la cámara. En las figuras 1 y 2 se muestran ejemplos de como la cámara por medio de los algoritmos es capaz de obtener la figura humana y sus posturas separándolas del entorno.

6. CONCLUSIONES Y TRABAJO FUTURO

Con base a los resultados obtenidos en los primeros trabajos con la tecnología RealSense, es evidente que la herramienta es útil para nuestro propósito. las herramientas disponibles para la segmentación y captura permiten de manera sencilla integrar esas carácterísticas obtenidas en una aplicación programada de manera externa. Utilizando herramientas virtuales con tecnología RealSense se logró segmentar partes del cuerpo humano, en trabajos futuros se busca digitalizar el cuerpo humano en su totalidad para la simulación de la rehabilitación del paciente. Como trabajo futuro, el desarrollo de la aplicación con propósito médico, así como la aplicación a telemedicina, esto con el objetivo de llevar el proyecto a una aplicación a la vida cotidiana y solventar problemas de salud en el país.

7. REFERENCIAS

[1] V.˜Gatica˜Rojas, E.˜Elgueta˜Cancino, C.˜Vidal˜Silva, M.˜Cantin˜López, and J.˜Fuentealba˜Arcos. Impacto del entrenamiento del balance a través de realidad virtual en una población de adultos mayores. *International Journal of Morphology*, 28(1):303–308, 2010.

[2] J.˜D. Mendoza, A.˜M. Domínguez, B.˜A.˜S. Moxo, M.˜A.˜S. Acevedo, and F.˜S. Juárez. Validación de movimientos de coordinación en extremidades superiores para rehabilitación virtual usando pendientes lineales. *Revista Cubana de Informática Médica*, 16(1), 2016.

[3] F.˜Moreno, O.˜Jordan, R.˜J. Esmitt, M.˜Christiam, R.˜Omaira, R.˜Jeanlight, and Á.˜Silvio. Un framework para la rehabilitación física en miembros superiores con realidad virtual. In *Primera Conferencia Nacional de Computación, Informática y Sistemas. Universidad Central de Venezuela*, 2013.

[4] J.˜E. Munoz˜Cardona, O.˜A. Henao˜Gallo, and J.˜F. Lopez˜Herrera. Sistema de rehabilitación basado en el uso de análisis biomecánico y videojuegos mediante el sensor kinect. 2013.

[5] J.˜M. Navarrete. La realidad virtual como arma terapéutica en rehabilitación. *Rehabil. integral (Impr.)*, 5(1):40–45, 2010.

[6] J.-K. Yun, J.˜Jang, K.-R. Park, and D.-W. Han. Real-sense media representation technology using multiple devices synchronization. In *SEUS*, 2009.

Videojuego serio para promover la práctica y entendimiento de conocimientos básicos de química.

Jorge Alberto Sánchez Amaya
Universidad Autónoma de Zacatecas
Unidad de Ingeniería Eléctrica
Ingeniería de Software
lilium277@gmail.com

Luis Miguel Rada
Universidad Autónoma de Zacatecas
Unidad de Ingeniería Eléctrica
Ingeniería de Software
rada.lmc@gmail.com

Carlos E. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad de Ingeniería Eléctrica
Ingeniería de Software
ericgalvan@uaz.edu.mx

Jorge I. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad de Ingeniería Eléctrica
Ingeniería de Software
gatejo@uaz.edu.mx

J. Rubén Delgado-Contreras
Instituto Tecnológico Superior Zacatecas Sur
División de Ingenierías
ruben.dc@gmail.com

José M. Celaya-Padilla
Universidad Autónoma de Zacatecas
Unidad de Ingeniería Eléctrica
Ingeniería de Software
gatejo@uaz.edu.mx

RESUMEN

Con el paso de los años, las reformas educativas y la innovación en ciencias y tecnologías han avanzado para ir mejorando las técnicas y contenidos para la enseñanza media superior en nuestro país, sin embargo seguimos presentando severos déficits a la hora de hablar de estadísticas en cuanto a la educación media superior en México la matrícula total de este servicio es de 4.8 millones de alumnos, equivalente al 13.3 % de todo el sistema educativo escolarizado. La cobertura es del 71.5 % respecto a la población total de 15 a 17 años de edad. Un importante reto será incrementar la eficiencia terminal que en la actualidad es del 63.3 por ciento, hablamos de que solo un poco más de la mitad de los estudiantes que ingresan a las instituciones de educación media superior logran terminar satisfactoriamente la totalidad de las materias asignadas y así obtener su título pasante. El propósito de este trabajo es ayudar a incrementar la eficiencia terminal actual, apoyando al alumno de una manera en la cual sienta que se divierte mientras aprende y refuerza lo que ya ha aprendido. Se realizará a modo de material adicional e interactivo un videojuego serio[5] que sirva como un apoyo más atractivo para los estudiantes y en general para cualquier persona que quiera reforzar sus aptitudes en la materia de química a nivel medio superior, aportando un apoyo para la comprensión, retención y práctica de dicha materia por medio de este método que puede resultar atractivo para las personas interesadas.

Palabras Clave

videojuego serio; elearning; química; preparatoria; educación en México;nivel medio superior; material interactivo

1. INTRODUCCIÓN

Gran parte de la problemática de abandono o dejar inconcluso el estudio Medio Superior se debe a las dificultades para obtener las calificaciones de pasante en las diversas materias escolares siendo matemáticas, física y química las tres primeras materias en la lista de reprobación.

En evaluaciones realizadas por diversas instituciones de investigación científica a las instituciones educativas de nivel medio superior han arrojado resultados desalentadores en cuestiones científicas, en concreto se abordará el tema de la materia de Química, gracias a estas investigaciones se puede estimar que un 71.5 % de los estudiantes de nivel medio superior obtienen en sus exámenes un resultado pobre, en las encuestas realizadas estos alumnos admiten que les es difícil comprender dicha materia por diversos motivos.[1]

Gracias a las tecnologías de la informática las nuevas generaciones de estudiantes se pueden capacitar, informar y auto sustentar para un mejor desarrollo en su educación, por lo cual para este proyecto se enfocará primordialmente en el área del aprendizaje y entendimiento de la química a nivel medio superior ya que en base a mi investigación dicha materia es uno de los principales problemas para los estudiantes que cursan dicho nivel, en experiencia propia la comprensión de esta materia les parece algo complicado, tedioso o poco interesante, por lo cual hay una merma de aprobación en la materia, esto afecta ciertamente a largo plazo a la sociedad en general, es por esto que apoyándose en una de las ramas de las TIC se planea diseñar y desarrollar un videojuego serio que le permita al estudiante reforzar, comprender y familiarizarse de una manera interactiva y atractiva con dicha materia.

Es muy importante para las instituciones educativas, buscar métodos y alternativas que ayuden y promuevan la mejora en las materias que resultan complicadas para un alto porcentaje de alumnos en el nivel medio superior, ya que muchas veces los alumnos tienden a desertar definitivamente de sus estudios al no lograr aprobar dichas materias, por lo cual es importante enfocarnos en la búsqueda constante de nuevas opciones para la comprensión y aprendizaje de dichas materias, es por eso que este proyecto se enfoca en la búsqueda del desarrollo de un videojuego serio que resulte atractivo, práctico y de fácil entendimiento para las personas que se interesen o necesiten mejorar su comprensión de la materia de química.

2. OBJETIVOS

Crear un videojuego serio, orientado a estudiantes de preparatoria para apoyar y reafirmar las bases de las ciencias químicas ya aprendidas en educación secundaria, diseñado de manera que incluso ayude al aprendizaje de estas bases.

Al finalizar el proyecto, este videojuego no sólo reafirmará conocimientos ya adquiridos, además, servirá como una herramienta de apoyo para el aprendizaje de las bases de la química. Se creará para la herramienta más accesible en las aulas de clases(smartphone, pc u otros).

3. MATERIALES Y MÉTODOS

Se probaron videojuegos serios y aplicaciones que se han desarrollado anteriormente como "eterna"[3] o "foldit"[4], enfocadas en el mismo tema aun cuando no sea enfocado al mismo nivel de enseñanza que el planteado en este trabajo. Se analizaron sus mecánicas, pros y contras.

La definición de objetivos y requerimientos se realizó mediante la investigación las debilidades de la enseñanza actual sobre el tema de química.

Para determinar los campos y temas en los que se enfocará el desarrollo del videojuego serio, se investiga el plan curricular oficial de las secundarias y preparatorias, además de investigar que tan rigurosamente se sigue este plan curricular en las instituciones.

El plan de desarrollo de este videojuego serio es para dispositivos móviles por su uso tan extenso en la actualidad, pero para corroborar la viabilidad de esta plataforma se investiga el verdadero porcentaje de esta herramienta en las aulas de secundarias y preparatorias y de no ser una plataforma viable se desarrolla al mismo tiempo para una plataforma de respaldo.

Basándose en esta recopilación de información se diseña, con ayuda de técnicas aprendidas en la formación como ingenieros de software, el videojuego serio propuesto, estas técnicas son tales como, diagramas UML, Historias de usuario, Arquitecturas de software, diseños conceptuales, entre otros más[2].

Con ayuda de herramientas computacionales se desarrolla el videojuego serio, se usarán las siguientes herramientas aunque no se limitara a solo estás: Unity, Game Maker, Inkscape, Editores de texto, Lightworks, Diagramadores UML, Godot engine.

4. RESULTADOS ESPERADOS

Al término de este proyecto se presentará:

- Un videojuego serio completo y funcional.
- Un diseño de videojuego serio que pueda adaptarse para otras materias académicas con déficit de aprendizaje.
- Implementar el videojuego serio en al menos 2 instituciones locales de nivel preparatoria.
- Lanzar al mercado en modalidad gratuita como una opción de apoyo para todo estudiante que desee mejorar su desempeño.

5. CONCLUSIONES

El alumnado en la actualidad sufre de grandes problemas para lograr el objetivo de terminar sus metas académicas, uno de estos problemas es la facilidad con la que la tecnología se puede convertir en un distractor o la falta de motivación por parte del profesorado o incluso solo el desánimo del alumno.

El aprovechamiento de las herramientas tecnológicas para ayudar al alumno en su enseñanza es un paso para evitar caer en el problema de que las herramientas se vuelvan más inteligentes que sus usuarios.

Una buen videojuego serio que logre aprovechar las capacidades actuales de la tecnología para dar soporte en la enseñanza puede ser aprovechada en varios niveles de educación.

6. REFERENCIAS

[1] Arnaut, A., and Giorguli, S. Los grandes problemas de méxico. *México, DF Primera Edición. Educación* (2010).

[2] Len˜Bass, Paul˜Clements, R.˜K. *Software Architecture in Practice.Third Edition*. 2012.

[3] Sharif˜Ezzat, Johan˜Andreasson, A. K. R.˜D. eternagame.org.

[4] Zoran˜Popović, David˜Baker, F. K. S. C. J.˜M. http://fold.it.

[5] Zyda, M. From visual simulation to virtual reality to games. *Computer 38*, 9.

Reconocimiento y Ubicación de Rugas Palatinas mediante Redes Neuronales Artificiales y Visión por Computadora

Israel Velázquez Almazán
UPVM
Av. Mexiquense s/n, esq. Universidad Politécnica, Villa Esmeralda Tultitlán, 54910 Estado de México.
hades_israelsama@hotmail.com

Eduardo Vázquez Santacruz
CONACYT-UNAM
Avenida Universidad 3000, Copilco Universidad, 04510 CDMX,
evazquezsa@conacyt.mx

Ivet Gil-Chavarría
CONACYT-UNAM
Avenida Universidad 3000, Copilco Universidad, 04510 CDMX,
ivetgil@gmail.com

RESUMEN
Recientemente, el campo de la inteligencia artificial ha sido desarrollado como uno de aplicación en el que muchas cosas se hacen posibles en diferentes áreas del conocimiento. Las redes Neuronales Artificiales son por convención y capacidad de aplicación uno de los temas principales de la Inteligencia Artificial que está alcanzando una gran importancia. Una de las aplicaciones más relevantes de las Redes Neuronales Artificiales es el reconocimiento de patrones. En este artículo se presenta un modelo de una RNA para el reconocimiento de patrones que se aplicará en el desarrollo de una propuesta de solución en el área de la Odontología en el contexto forense y de esta manera se aportan herramientas a los problemas actuales que no han sido resueltos todavía.

Conceptos CCS
• Metodologías de computación → Inteligencia artificial.

Palabras Clave
Redes neuronales artificiales; reconocimiento de patrones; bioinformática; clasificación.

1. INTRODUCCIÓN
Gracias a las Ciencias Forenses se han podido establecer diversas formas de Identificación de Personas tales como las huellas dactilares, la comparación dental con todas sus variantes son las que más se han utilizado hasta nuestros días. Se entiende que la Palatoscopia [1][6] es la disciplina que se encarga del estudio en general del Paladar y su aplicación para la Identificación de Personas (Ver Figura 1).

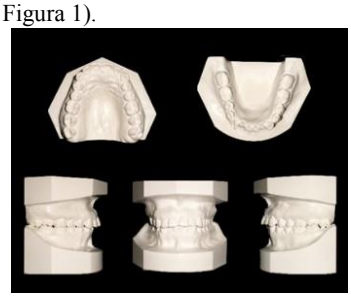


Figura 1: ***La Palatoscopia y Modelos de Paladar.***
Fuente: http://es.slideshare.net/mdaly21/tema-i-profilaxis-me-y-rx

Para poder identificar las Rugas Palatinas, la investigación y los procedimientos que se llevan a cabo en el laboratorio indican que el mejor camino es concentrarse en la forma que las caracteriza, es decir, su representación física y esto lo podemos hacer mediante una Designación Rugoscopica. Como se mencionó con anterioridad, las Rugas Palatinas son comparables y en alto grado a las huellas dactilares debido a que son:

- Inmutables: Permanecen iguales ante traumas altamente significativos.
- Perenes: Desde su formación hasta su alteración o desaparición permanecen en la misma proporción.
- Clasificables: Susceptibles a los criterios de Clasificación.

Bajo este criterio y de acuerdo con la definición de Biométrico [5], podemos concluir que las rugas palatinas en conjunto son un biométrico. Sin embargo, el proceso de identificación mediante esta técnica tiende a ser subjetivo debido a los siguientes puntos:

✓ Debido a que son susceptibles a ser clasificadas, desde principios del siglo XX, existen diversos criterios tales como C. Goria, L. Lysell que es la más usada en cuanto a investigación, Cormoy, entre otros [1].

✓ La identificación post mortem no es posible sin registros ante mortem. En circunstancias que involucran fuego, las rugas palatinas son destruidas.

La Inteligencia Artificial [2] es una ciencia que se ha desarrollado principalmente para emular la mente humana. Su estudio se compone principalmente de áreas como: Redes Neuronales Artificiales, Algoritmos Genéticos, Lógica Difusa, entre otros.

Hablar de Redes Neuronales Artificiales, es hablar de un Modelo Matemático que está diseñado para simular el comportamiento y/o funcionamiento de una Red Neuronal Biológica. Tomando en consideración lo que anteriormente se explicó, se aplicara el modelo de las Redes Neuronales Artificiales al reconocimiento de rugas palatinas. El Reconocimiento de Patrones puede definirse como la técnica empleada que tiene como propósito de extraer información que permita identificar a un individuo u objeto dentro de un entorno donde habiten sujetos de la misma clase y consta del proceso mostrado en la Figura 2:

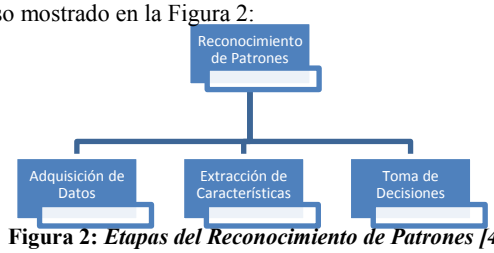


Figura 2: ***Etapas del Reconocimiento de Patrones [4].***
Fuente: Elaboración Propia.

2. OBJETIVOS
2.1 General
El objetivo general es Desarrollar una herramienta que permita el reconocimiento y la ubicación de Rugas Palatinas mediante Redes Neuronales y Visión por Computadora.

2.2 Específicos
1) Diseñar un Algoritmo eficaz que permita el aprendizaje de las formas/patrones de Rugas Palatinas y su ubicación.
2) Documentar la solución obtenida y tomarla como base para futuras investigaciones y mayor desarrollo en este campo.

3. METODOLOGÍA

3.1 Herramientas Utilizadas

Se especificó en el objetivo general que la meta de la investigación es desarrollar una herramienta que permita el Reconocimiento y Ubicación de Rugas Palatinas. Por ende y ya que se acordó que son Biométricos, se pretende implementar la Inteligencia Artificial [3] para brindar un criterio de clasificación más objetivo que el que se tiene si se realiza dicha tarea manualmente.

3.1.1 Perceptrón [2]

Ya que el punto central es la Clasificación, se optó por la implementación de un Perceptrón, definido como un clasificador supervisado que a través de una cámara permita el aprendizaje de las diversas formas que se pueden identificar en una Ruga Palatina. Los elementos del Perceptrón se aprecian en la Figura 3:

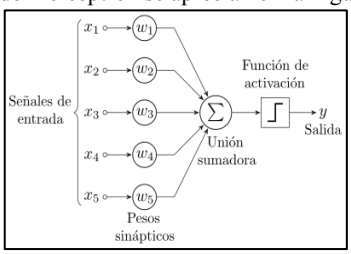


Figura 3: ***Elementos Característicos del Perceptrón.***
Fuente: https://es.wikipedia.org/wiki/Perceptr%C3%B3n

3.1.2 Comparación de Plantillas [7]

Este algoritmo es ampliamente utilizado en el procesamiento digital de imágenes, se basa en la búsqueda de ciertas características con base a una plantilla que es generada por el usuario dando como salida una imagen con la mejor aproximación a dicha plantilla.

3.1.3 MATLAB

Software creado por la compañía Mathworks Inc. cuyo acrónimo en español significa Laboratorio de Matrices. Es una herramienta realmente potente para manejar matrices y poder operar una gran cantidad de datos de manera muy eficiente. Es muy utilizado en diversos campos de la Ciencia e Ingeniería.

3.2 Solución Propuesta

La solución se obtiene mediante el entrenamiento de un Perceptrón, con 10 entradas, en este caso son las formas comunes que se aprecian en la Figura 4 y los patrones de entrenamiento son extraídos de las imágenes tomadas de los modelos de paladar mediante la Binarización de las imágenes por medio del Método de Segmentación de Otsu.

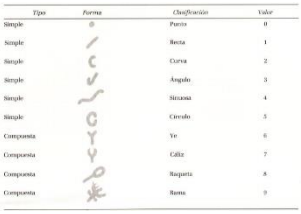


Figura 4: ***Ficha Rugoscopica General para la Clasificación de Rugas Palatinas.***
Fuente:/www.webdelprofesor.ula.ve/odontologia/isis.c/archivos/forense

El diagrama 1, muestra el proceso general para el reconocimiento de Rugas Palatinas y su Ubicación mediante Visión por computadora que se efectúa, como se mencionó, por comparación de Plantillas

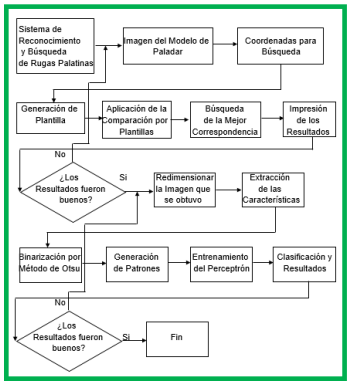


Diagrama 1: ***Diagrama de Flujo Simplificado para Reconocimiento de Patrones y Visión por Computadora***
Fuente: Elaboración Propia

4. RESULTADOS

Ya que las imágenes proporcionadas se tomaron con cámara de dispositivo móvil, es difícil apreciar un resultado realmente satisfactorio, pero la figura 7 (izquierda) muestra el resultado de la ejecución del Perceptrón y la figura 7 (derecha) la ejecución del algoritmo de comparación por plantillas:



Figura 7: ***Resultados de la Ejecución de los Programas***
Fuente: Elaboración Propia.

De con la ejecución, se observó un éxito de 93.3% en la clasificación en un tiempo considerablemente rápido que varía de 40 a 45 segundos, mientras que la herramienta de búsqueda, proporciona un nivel de asertividad bueno.

5. CONCLUSIONES

De la presente investigación se lograron establecer algunas bases objetivas y tecnológicas que se implementaran para el Reconocimiento, Búsqueda y Clasificación de Rugas Palatinas siguiendo una metodología experimental con base a la información recabada y desde luego al punto de vista de las ramas de Odontología Forense y la Ingeniería, cumpliéndose así las metas planteadas.

6. AGRADECIMIENTOS

Agradezco al Dr. Eduardo Vázquez Santacruz por su guía y orientación durante la Estancia de Investigación y la Dra. Ivet Gil Chavarría por las facilidades proporcionadas para la Realización del Proyecto.

7. REFERENCIAS

[1] "Evidence-Based Forensic Dentistry", Balwant Rai & Jasdeep Kaur, Springer.
[2] "Inteligencia Artificial con aplicaciones a la Ingeniería", Ponce Cruz, Pedro, Primera Edición, Ed. Alfaomega
[3] "Artificial Intelligence: A modern approach", Russell Stuart J., Norving Peter, Second Edition, Ed. Pearson Education.
[4] "Statistical Pattern Recognition", Webb Andrew R., Copsey Keith D., Third Edition, Ed. Wiley.
[5] "Introduction to Biometrics", Anil Jain, Arun Ross, First Edition, Ed. Springer 2011.
[6] "Odontología Forense", Correa Ramírez A. Isaac. Primera Edición, Ed Trillas.
[7] "*Template Matching Techniques in Computer Vision: Theory and Practice*", R. Brunelli, Ed. Wiley, 2009

Búsqueda de modelo univariado en características de datos descriptores de imágenes para cáncer de mama con diagnóstico asistido por computadora

Laura A. Zanella-Calzada
Universidad Autónoma de San Luis Potosí
Facultad de Ciencias
San Luis Potosí, S.L.P.
lau.zanella14@gmail.com

Carlos E. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería de Software
Zacatecas, Zac.
ericgalvan@uaz.edu.mx

Jorge I. Galván-Tejada
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería Robótica y Mecatrónica
Zacatecas, Zac.
gatejo@uaz.edu.mx

José M. Celaya-Padilla
Universidad Autónoma de Zacatecas
Unidad Académica de Ingeniería Eléctrica
Ingeniería Eléctrica
Zacatecas, Zac.
jose.cpadilla@gmail.com

ABSTRACT

Late detection of breast cancer is a worldwide health problem, decreasing the survival rate of the patients. Early detection is an important factor, recent studies have proved mammography is the most important tool for detecting early stage cancer. The purpose of this paper is to obtain an univariate model with features extracted from mammograms to detect benign tumors from malign with high accuracy.

1. RESUMEN

La detección tardía del cáncer de mama es un problema que decrementa la tasa de supervivencia del paciente, por lo tanto, la detección temprana es un factor importante en el que recientes estudios han probado que la mamografía es la herramienta actual más importante para resolver este problema. El propósito de este trabajo es el de obtener un modelo univariado a partir de características extraídas de mamografías que clasifique con alta precisión si un tumor es benigno o maligno.

Keywords

Cáncer de mama; Diagnóstico asistido por computadora; Mamografía; Modelo univariado, Clasificación de tumores

2. INTRODUCCIÓN

El cáncer de mama es una importante enfermedad mundial y es el tipo de cáncer más común en las mujeres [1]. Uno de los grandes problemas asociados a esta enfermedad es la detección tardía. Actualmente, la mamografía ha probado ser la mejor herramienta para la detección temprana de esta enfermedad, reduciendo las tasas de mortalidad en 30 - 70 % [2]. Recientes estudios han implementado algoritmos para mejorar la detección de este cáncer por medio de diagnóstico asistido por computadora (DAC), mejorando el pronóstico del estado del paciente [3]. Estos algoritmos utilizan diversas características descriptoras obtenidas de las mamografías, sin embargo, no ha sido encontrado un modelo univariado, de características, específico que clasifique con alta precisión al tumor [4]. En esta investigación se presenta la búsqueda de un modelo univariado que tenga una capacidad de predicción y clasificación con alta precisión, en donde una variable específica pueda decidir si el tumor es benigno o maligno, reduciendo el trabajo del radiólogo y proporcionando una segunda opinión.

3. OBJETIVOS

- Definir características extraídas de mamografías con la mejor precisión para la clasificación binaria de tumores (beningo/maligno). - Obtener modelo univariado que prediga de mejor o igual manera que el modelo multivariado la variable de salida. - Reducir el trabajo del médico radiólogo y obtener una segunda opinión en el diágnostico de cáncer de mama.

4. MATERIALES Y MÉTODOS

4.1 Descripción de datos

Se utilizó la base de datos Breast Cancer Digital Repository (BCDR) [4], compuesta por: Digital Mammography dataset number 1 (BCDR-D01), un repositorio de mamografías que contiene información de 64 mujeres, representando 143 segmentaciones y, Digital Mammography dataset number 2 (BCDR-D02), un repositorio basado en mamografías digitales de campo completo que contiene información de 164 mujeres, representando 455 segmentaciones. Los sujetos se

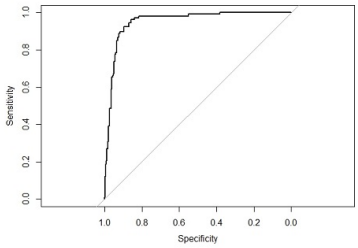


Figure 1: Curva ROC obtenida de la característica *t_corr*.

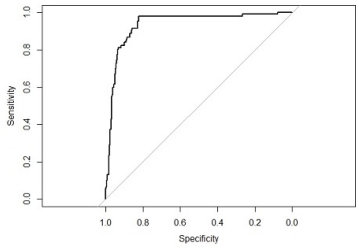


Figure 2: Curva ROC obtenida del modelo multivariado.

encuentran en un intervalo de edades de 58 +/- 11años, con una edad máxima de 89 años y una mínima de 23. La base de datos está compuesta por 79 lesiones obtenidas por biopsia, incluyendo datos clínicos y descriptores basados en imágenes. BCDR es una base de datos de clase binaria asociada a la clasificación inicial Breast Imaging Report and Database System (BI-RADS) del resultado de la biopsia (Benigno vs Maligno). De cada una de las 598 segmentaciones fueron extraídas un total de 37 características: 8 características fueron extraídas de datos clínicos y generales, 8 características fueron extraídas de descriptores de intensidad computarizados directamente de los niveles de gris de los pixeles dentro del contorno de la lesión identificada por los radiólogos, 13 características de textura fueron computarizadas de la matriz de co-ocurrencia de los niveles de gris relacionada con el cuadro delimitador del contorno de las lesiones, y finalmente 8 características relacionadas a los descriptores de la forma y localización de la lesión. El conjunto de datos total cuenta con 491 salidas benignas y 75 malignas. Todos los datos fueron normalizados con normalización Z para evitar valores atípicos.

4.2 Análisis de datos

En este trabajo se llevó a cabo la búsqueda de un modelo univariado a través del método de regresión logística (RL). La RL ha mostrado ser una de las heramientas más robustas en el caso de salidas binarias como lo es maligno benigno. La RL fue llevada a cabo con cada una de las 37 características. Se obtuvieron los valores del área bajo la curva de la característica de funcionamiento del receptor (AUC, ROC). Con el único fin de comparar, se obtuvieron el AUC y ROC del modelo multivariado. Todo el análisis y procesamiento de los datos se realizó en R (3.3.1) (http://www.r-project.org/).

5. RESULTADOS

En el modelo univariado, los resultados estadísticos obtenidos de las características analizadas fueron filtrados, basándose en el AUC con mayor valor, siendo la variable más representativa *t_corr*, obteniendo un AUC de 0.952 y la curva ROC mostrada en la figura 1. La variable *t_corr* es un descriptor de textura referido a la correlación que existe en la matriz de co-ocurrencia de los niveles de gris en relación con el cuadro delimitador del contorno de la lesión identificada por el radiólogo.

En el modelo multivariado, que se compone por las 37 características de la base de datos BCDR, se obtuvo un AUC de 0.937 y la curva ROC mostrada en la figura 2.

6. CONCLUSIONES

El resultado del modelo univariado en donde su curva ROC muestra una alta sensibilidad/especificidad como se observa en la figura 1, en conjunto con su valor de AUC confirman que *t_corr* es una característica con alta precisión para clasificar si un tumor es benigno o maligno. En la figura 2 se muestra que la curva ROC del modelo multivariado, a pesar de tener una alta sensibilidad/especificidad, no es mejor que la del modelo univariado. A pesar de la poca diferencia en los valores del AUC de ambos modelos, la propuesta del modelo univariado conlleva ventaja considerando un análisis costo-beneficio debido a que, obteniendo una sola característica de las mamografías se tiene una mejor precisión en la clasificación del tumor que si se usan las 37 características obtenidas inicialmente, además del ahorro computacional al no tener que calcular las demás características.

7. REFERENCIAS

[1] H.-D. Cheng, X.˜Cai, X.˜Chen, L.˜Hu, and X.˜Lou. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern recognition*, 36(12):2967–2991, 2003.

[2] A.-M. Dixon. Diagnostic breast imaging: Mammography, sonography, magnetic resonance imaging, and interventional procedures. *Ultrasound*, 22(3):182–183, 2014.

[3] L.˜H. Eadie, P.˜Taylor, and A.˜P. Gibson. A systematic review of computer-assisted diagnosis in diagnostic cancer imaging. *European journal of radiology*, 81(1):e70–e76, 2012.

[4] D.˜C. Moura and M.˜A.˜G. López. An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. *International journal of computer assisted radiology and surgery*, 8(4):561–574, 2013.

Taller:
Soft Computing Applications



Organizadores:
Ángel G. Andrade (UABC)
Anabel Martínez (UPP)
Guillermo Galaviz (UABC)
Oscar Montiel (CITEDI)
Roberto Sepúlveda (CITEDI)

An Optimization Based Approach for Maximizing the Information Content of Keypoints Detected on a Digital Image

Fernando Fausto
Universidad de Guadalajara
Av. Revolución 1500
Guadalajara, Jal, 44430,
México
abraham.fausto@academicos.udg.mx

Erik Cuevas
Universidad de Guadalajara
Av. Revolución 1500
Guadalajara, Jal, 44430,
México
erik.cuevas@cucei.udg.mx

Marco Pérez-Cisneros
Universidad de Guadalajara
Av. Revolución 1500
Guadalajara, Jal, 44430,
México
marco.perez@cucei.udg.mx

ABSTRACT
The task of finding point correspondences between two images of the same scene is one of the most challeging problems for a wide variety of computer vision applications. With regard to this, matching precision difficulties are related to the information content provided by the feature descriptors that are extracted from a set of interest points previously detected within a digital image. In order to overcome such difficulties, methods and techniques which aim to increase the information content (distinctivness) of such feature descriptors must be developed. In this paper, an optimization based approach for maximizing the information content provided by set of feature vectors computed by the feature description method known as Spider Local Image Feature (SLIF) is proposed. In order to test the feasibility of the proposed approach, several state-of-the-art swarm optimization algorithms, such as Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Firefly Algorithm (FA) and Social Spider Optimization (SSO) were implemented. The proposed experimental setup results show that the problem of maximizing the distinctivness of a set of feature descriptors could be efectively modeled as an optimization problem, and as such, be solved by implementing popular optimization techniques.

CCS Concepts
•**Computing methodologies → Artificial intelligence → Computer vision → Computer vision problems → Object recognition.**

Keywords
Feature detection; feature description; evolutionary algorithms; optimization; information content.

1. INTRODUCTION
Finding point correspondences between two images of the same scene or object is an important task in many computer vision applications, such as camera calibration, image registration and retrieval, object recognition, visual navigation, among others.

The first challenge of such a task is to find a distinctive and repeatable set of key points (local features). Some state-of-the-art methods, such as the Harris/Hessian based detectors or the well know Difference of Gaussian (DoG) approach allow the detection of several distinctive features within a digital image. Once a set of distinctive key points is found, a feature descriptor (or feature vector) may be constructed using pixel information contained on a neighborhood around such interest points. Some methods such as Scale Invariant Feature Transform (SIFT) [12] assign a feature descriptor to each key point detected by a scale-adapted DoG detector by computing a set of sampled magnitudes and orientations related to the image local gradients. Other approaches, such as Speeded-Up Robust Features (SURF) [3], take advantage of the properties of both the integral image and box filters to enable a time-efficient feature detection and description [16]. Finally, more recent feature description methods such as Binary Robust Scalable Key points (BRISK) [5] build feature descriptors by computing binary intensity comparisons over a set predefined locations within a particular sampling pattern. Finally, key points over different images can be matched based on the similarity of their feature vectors. Naturally, the dimension of such features vectors has an important impact on computing time, and thus, is always desirable to have low-dimensional descriptors to achieve a faster key point matching. On the other hand, lower dimensional feature vectors are usually less distinctive than their higher-dimensional counterparts, and as such are less preferred when matching precision is required.

In general, key points matching precision is related to the quality of the information provided by its respective set of feature descriptors. This problem is not trivial, since the information content (distinctiveness) of a set of feature vectors is related to both, the applied feature description method and the nature of the information contained on the image itself. In this sense, image distortions caused by image transformations (such as scale and illumination changes, image compression, noise, among others) represent a huge challenge for the task of maximizing the information content of feature vectors.

As a way to tackle such a problem, the Spider Local Image Features (SLIF) descriptor proposed in [8] features a unique orb-web sampling pattern which may be adapted to efficiently describe several specific images cases by modifying a set of user-defined parameters. However, while a give combination of SLIF parameters may enhance the descriptor's distinctively for a given image case, this

parameter selection may not yield the same favorable results for other possible image case.

In this paper, an optimization based approach for maximizing the information content of feature descriptors built by the SLIF description method is proposed. In such approach we aim to maximize the information content with regard to a wide variety of images cases (rather than a single one) by optimizing the SLIF's parameter combination which defines the structure of the orb-web sampling template. In order to solve such parameter optimization problem, well know state-of-the-art swarm intelligence algorithms such as PSO, ABC, FA, and SSO were implemented, and then compared in terms of solution quality.

This paper is organized as follows: In Section 2, we introduce some fundamental concepts related to the detection and description of key points within a digital image. In Section 3, the concept of information content provided by a set of features vectors is introduced. In Section 4, we highlight the most important traits of the SLIF description method. In Section 5, an optimization based approach for maximizing the information content provided by a set of SLIF feature vectors is proposed. In Section 6, we presents our experimental setup and results. Finally, in Section 7, conclusions are drawn.

2. FEATURES DETECTION AND DESCRIPTION

Local image features (also called interest points or key points) may be defined as locations within digital image which distinct themselves from its immediate neighborhood. Such distinctions are usually associated with a change on one or more image properties, such as pixel intensity, color, textures, etc. In some other cases, the responses to certain image filters are used to define the location of key points within an image. In the Scale Invariant Feature Transform (SIFT) [12], for example, key point locations are defined by considering the image response to a set of scale-adapted Difference of Gaussian (DoG) kernels. In general, key point may be represented by specific coordinate points $\mathbf{x} = (x, y)$ within a digital image; in this sense, the purpose of a local feature detection method is to provide a set k coordinate points $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, corresponding to each detected feature within a given image.

Once a set of interest points has been identified, a descriptor may be assigned to each of such key points in order to identify them from among the others. Typically, some kind of image information is extracted from around a region centered over each detected feature and then processed to assign a feature descriptor to each of such key points. In the case of the Scale Invariant Feature Transform (SIFT) [12] and Speeded-Up Robust Features (SURF) [3], detected key points are described by extracting sampled magnitudes and orientations from the image local gradient around the neighborhood of such interest points. Typically, the descriptor of each key point i within an image is represented by a feature vector $\mathbf{f}_i = [f_1, f_2, \dots, f_d]$ of a given length d. Such feature vectors contains an informative representation of the data extracted from each key point's neighborhood. With that being said, for a given set of k keypoints $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ detected on a given digital image, a feature description method computes a corresponding set of feature vectors $\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\}$, such that $\forall\, \mathbf{x}_n \in \mathbf{X}\ \exists\, \mathbf{f}_n \in \mathbf{F}$.

3. INFORMATION CONTENT OF A SET OF FEATURE VECTORS

The concept of information content extracted from a set of key points detected on a given image is described as the relative probability of computing a descriptor (feature vector) at any given point of an image [14].

As illustrated in Section 2, let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ denote a set of k interest points (key points) detected by a feature detector over a particular digital image, and let $\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\}$ denote its corresponding set of feature descriptors, such that $\forall\, \mathbf{x}_n \in \mathbf{X}\ \exists\, \mathbf{f}_n \in \mathbf{F}$. Also, let $\mathbf{S}$ represent the space of all possible feature descriptors. If the descriptors in $\mathbf{F}$ are all crowded within a small region of $\mathbf{S}$, then it is said that the set of key points $\mathbf{X}$ conveys a small amount of information due to their descriptors being too similar, with the converse being true in the opposite case.

Taking this into account, if we consider a partition $\mathbf{S}_j \in \mathbf{S}$, the information content of a set of feature vector $\mathbf{F}$ is given by the amount of entropy contained within such set of the descriptors, as defined as follows:

$$H(\mathbf{F}) = -\sum p_j \cdot \log_2(p_j) \qquad 1$$

where p_j represents the mass probability function of the descriptors $\mathbf{f}_n \in \mathbf{S}_j$.

Intuitively, the information content extracted from the set of keypoints $\mathbf{X}$ is intimately related with the feature description method employed to build its corresponding set of features vectors $\mathbf{F}$, and as such, selecting an appropriate keypoint descriptor is usually a crucial task for most computer vision applications.

4. SPIDER LOCAL IMAGE FEATURES (SLIF) DESCRIPTOR

Spider Local Image Feature (SLIF) is a feature description method which applies a unique description sampling template inspired in the basic structure of a spider's orb-web [8]. Such feature descriptor is applied to a set of key points detected by a scale invariant feature detector, such as the well know Fast-Hessian detector proposed in [3].

The SLIF descriptor extracts pixel intensity from a given image keypoint by employing a simplified sampling templante inspired on the highly adaptable architecture of natural spider orb-webs and its real life function for trapping small prey.

For each keypoint detected within a given digital image, a a two-dimensional orb-web sampling template is constructed by considering a set user-defined integer parameters: (1) the number radial threads N, (2) the number of spiral threads M, and (3) a magnification factor K which further amplify the web structure's sampling area. On each of such

web structures the intersections between a radial thread and a spiral thread represent a set of sampling node $\mathbf{s}_{n,m}$ which are further used in the description of the key point's local area. Each sampling nodes positions is represented by a pair of polar coordinates (θ, r) belonging such orb-web structure, as given as follows:

$$\mathbf{s}_{n,m}^{k} = \left(\frac{2\pi \cdot n}{N}, \frac{R \cdot m}{M}\right) \quad 2$$

$$\text{for } n = 1, 2 \ldots, N; \; m = 1, 2, \ldots, M$$

where $R = K\sigma_k$ represents an orb-web radius proportional to the keypoint k characteristic scale σ_k (as given by the scale-invariant feature detector applied to detect each key point) and the user defined magnification factor K.

After an orb-web structure has been assigned to each key point detected within a given image, the SLIF descriptor assigns each of such key point a corresponding feature vector, composed by a set of 8-bit binary sequences built by computing a series of one-bit weight differences over several sampling point within the orb-web sampling template [8]. This allows SLIF to build a simple and accurate rotation invariant binary feature descriptor, which may be further applied on a wide variety of computer vision applications, such as image registration, object recognition, visual navigation, 3D reconstruction, etc.

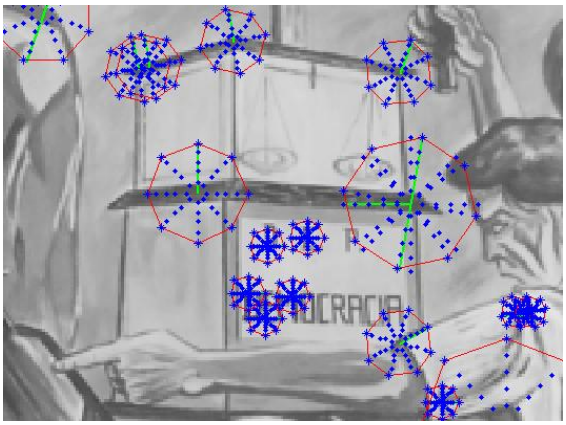


Figure 1. Visualization of the SLIF's orb-web sampling patterns used in the description of scale-invariant local image features.

5. MAXIMIZING THE SLIF DESCRIPTORS' INFORMATION CONTENT

A good feature descriptor should be able to extract a set of highly distinctive feature vectors, and as such, high amounts of information content. As stated in Section 3, the information content provided a set of features vectors is related to the key point description method employed to extract such descriptors. Spider Local Image Feature (SLIF) is a robust, low complexity key point descriptor which features a highly adaptable sampling pattern inspired in the natural structure of a spider's orb-web [8]. SLIF includes three user-definable integer parameters which represent the number of radial threads, spiral threads and magnification factor of the orb-web structure (N, M and K respectively). Such parameters modify the distribution (both numerical and spatial) of sampling points around each detected key point, and as such, they have a strong impact over the quality and quantity of the information that is extracted from such local features.

Intuitively, it is desirable to define a set SLIF parameters which could potentially allow the resulting set of feature vectors to provide the maximum possible amount of information content. Therefore, from an optimization point of view, the main objective is to maximize the amount of information contained on a set of extracted feature vectors by finding an optimal set of SLIF structural parameters N, M and K.

In order to solve such optimization problem, a cost function which considers the entropy (information content) of a given set of feature vectors is proposed. In order to illustrate such approach, let $\mathbf{F_p} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ denote a set of SLIF feature vectors extracted from a set of k keypoints by considering a given set of SLIF orb-web parameters $\mathbf{p} = [p_1 = N, p_2 = M, p_3 = K]$. We define a minimization-case similarity cost function which considers the amount of entropy (information content) contained on such set of feature vectors as follows:

$$f(\mathbf{p}) = \frac{1}{\exp(H(\mathbf{F_p}))} \quad 3$$

where $H(\mathbf{F_p})$ denotes the amount of entropy (information content) contained in the set of feature vectors $\mathbf{F_p}$ as given by equation 1

As stated above, the value of $f(\mathbf{p})$ is entirely dependent of the amount of entropy contained on the set of feature vectors $\mathbf{F_p}$, and as such, dependent of the chosen SLIF parameter combination $\mathbf{p}$.

In other words, the objective is to minimize the similarity between the feature vectors contained on the set $\mathbf{F_p}$ by choosing an appropriate SLIF parameter combination $\mathbf{p}$. In this sense, it is clear that less similar between the feature vectors on $\mathbf{F_p}$ yields to higher amounts of information content, and thus, a higher discriminative power, which is crucial to ensure a better feature's matching precision.

However, while for a certain type digital image some SLIF parameters' combination $\mathbf{p}$ may yield to high amounts of information content, this might not be true for other different images cases; in this sense, it its known that the performance of a feature description method is also related to its ability to handle a wide variety image cases, such as scale and illumination changes, image textures, distortions caused by image compression, image noise, etc.

In order to tackle this problem, an extension of the proposed optimization approach, which considers a set of different image cases, is proposed. To illustrate this, let $\mathbf{A} = \{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_N\}$ denote a set of N different digital images. Also, let $\mathbf{F}_\mathbf{p}^j$ denote a set of SLIF feature vectors extracted from a set of key points detected on a particular

image $\mathbf{I}_j$, such that $\forall\, \mathbf{I}_j \in \mathbf{A} \,\exists\, \mathbf{F}_\mathbf{p}^j$. For each set of feature vectors $\mathbf{F}_\mathbf{p}^j$ a similarity function $f(\mathbf{p})^j$ is computed as follows:

$$f(\mathbf{p})^j = \frac{1}{\exp\left(H\left(\mathbf{F}_\mathbf{p}^j\right)\right)} \quad 4$$

where $H\left(\mathbf{F}_\mathbf{p}^j\right)$ denotes the amount of entropy (information content) contained on $\mathbf{F}_\mathbf{p}^j$ as given by equation 1**.**

From what was previously illustrated, it is clear that the value of $f(\mathbf{p})^j$ is now dependent not only of the chosen SLIF parameters combination $\mathbf{p}$, but also on the given image $\mathbf{I}_j$. In this sense, we could be expect to find an optimal set of SLIF parameters $\mathbf{p}$ which allows to maximize the information content provided by the features in a particular image case $\mathbf{I}_j \in \mathbf{A}$, but it is not possible to ensure that such parameters' combination may yield to equally favorable results for the remaining image cases in **A.**

A simple approach to solve such multiple image cases problem involves the use of a weighted sum method [18], used to represent the set of the similarity cost functions $f(\mathbf{p})^j$ as a single composite cost function $F(\mathbf{p})$, such that:

$$F(\mathbf{p}) = \sum_{j=1}^{N} w_j \cdot f(\mathbf{p})^j \quad 5$$

Where $w_j \in [0,1]$, whereas N stand for the total number of different image cases, and where $\sum w_j = 1$.

In such weighted sum approach, w_j denotes a weight coefficient which represent the degree of relevance that each similarity function $f(\mathbf{p})^j$ has while computing the composite cost function $F(\mathbf{p})$. In this approach we assume that each of the N possible image case has the same degree of relevance, which yield to:

$$w_1 = w_2 = \cdots = w_N \quad 6$$

where $w_j = \frac{1}{N}$ for $j = 1,2,\ldots,N$

With that being said, we may write the composite similarity function $F(\mathbf{p})$ as follows:

$$F(\mathbf{p}) = \frac{1}{N}\sum_{n=1}^{N} f(\mathbf{p})^j \quad 7$$

where $\mathbf{p}$, as previously stated, denotes the set of SLIF parameter $[N, M, K]$ (decision variables) and where N stand for the number of considered image cases.

6. EXPERIMENTAL SETUP AND RESULTS

In order to analyze the feasibility the proposed feature's information content maximization approach, a set consisting of a wide variety of image cases was considered. Such set includes images from the well-known test images data set proposed in [13] as well as some other significant image cases. **Figure 2** shows some examples of the images considered in this optimization approach. To minimize the similarity function $F(\mathbf{p})$ defined by the equation 7 four state-of-the-art swarm optimization techniques were implemented: Particle Swarm Optimization (PSO) [10], Artificial Bee Colony (ABC) [9], Firefly Algorithm (FA) [17] and Social Spider Optimization (SSO) [7]. Such approaches are considered among the most popular swarm algorithms for many optimization applications. The parameter setting for each implemented method is as described as follows:

1. PSO: The algorithm's learning factors are set to $c_1 = 2$ and $c_2 = 2$; also, the inertia weight factor is set to decreases linearly from 0.9 to 0.2 as the process evolves [10].
2. ABC: The algorithm was implemented by considering the guidelines provided by [9], with the parameter $\text{limit} = 1000$.
3. FA: The parameters setup for the randomness factor and the light absorption coefficient are set to $\alpha = 0.2$ and $\gamma = 1.0$ respectively, as illustrated on its own reference [17].
4. SSO: The female attraction probability parameter is set to $PF = 0.7$, as proposed in [7].

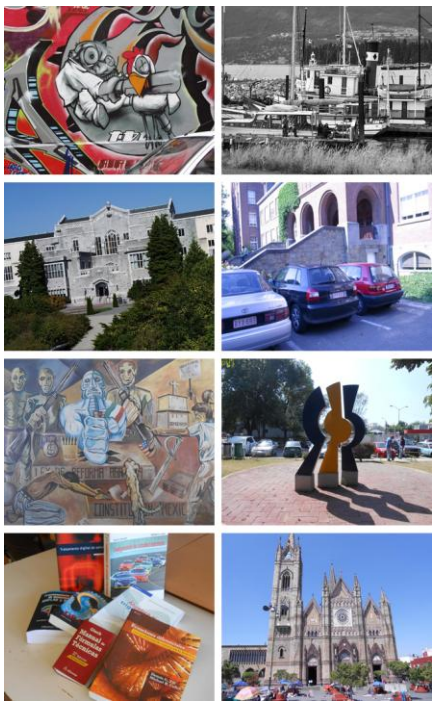


Figure 2. Examples of image case considered on the proposed optimization approach.

The experimental setup aims to compare the performance of such swarm intelligence approaches on solving the proposed optimization problem. The results for 30 individual runs are reported in **Table 1** where the best outcome are boldfaced.

Table 1. Minimization results for the proposed optimization approach, considering $n = 30$ individual runs and maximum number of iterations = 100

	AB	MB	SD
PSO	**0.0638387**	0.0644522	0.0008663
ABC	**0.0638387**	0.0645820	0.0005394
FA	**0.0638387**	0.0646908	0.0005592
SSO	**0.0638387**	**0.0638387**	**0**

The reported results consider the following performance indexes: the Average Best-so-far (AB) solution, the Median Best-so-far (MB) and the Standard Deviation (SD) of the best-so-far solution. In such comparisons, we considered a population size of 30 individual, while the maximum iteration number was set to 100. According to **Table 1**, the SSO algorithm's performance is slightly better than those of the other compared methods. From an optimization point of view such difference in performance is intuitively related to a better trade-off between exploration and exploitation. Furthermore, **Figure 3** shows the similarity function evolution curves for each of the implemented swarm intelligence approaches.

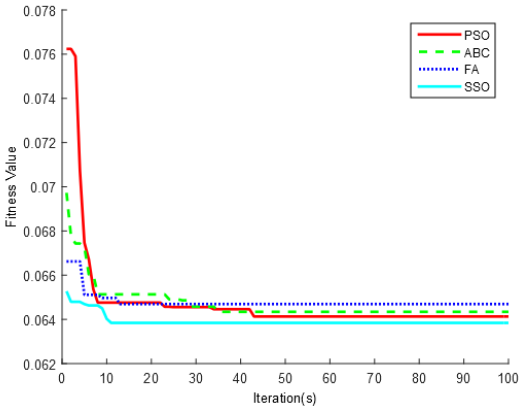


Figure 3. Evolution curves for PSO, ABC, FA and SSO for the proposed optimization approach.

7. CONCLUSIONS

In this paper an optimization based approach for maximizing the information content (entropy) of feature descriptors computed by the Spider Local Image Features (SLIF) description method was proposed.

In such approach, the problem of maximizing the entropy of a set of feature descriptors is interpreted as a minimization optimization problem in which, for a wide variety of possible image cases, the objective is to reduce the similarity of the feature vectors extracted from a set of keypoins by finding an optimal configuration on SLIF orb-web parameters N,M and K.

In order to prove the feasibility of the proposed entropy maximization approach, well known state-of-the-art swarm optimization algorithms such as Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Firefly Algorithm (FA) and Social Spider Optimization (SSO) were implemented to solve the proposed optimization problem, yielding to favorable results.

8. ACKNOWLEDGMENTS

The present research is supported under the Grand CONACYT CB 1810053. The first author acknowledge to CONACYT of México for its support under the scholarship No. 400528.

9. REFERENCES

[1] Alahi, A., Ortiz, R., & Vandergheynst, P. «A low-dimensional local descriptor incorporating TPS warping for image matching,» *Image and Vision Computing*, 28, pp. 1184–1195, 2010

[2] Banharnsakun, A., Achalakul, T. and Sirinaovakul, B., «The best-so-far selection in Artificial Bee Colony algorithm,» *Applied Soft Computing*, 11, pp. 2888-2901, 2011.

[3] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. «Speeded-Up Robust Features (SURF),». *Computer Vision and Image Understanding*, 110, (2008), pp. 346-359.

[4] Bonabeu, E., Dorigo, M. and Theraulaz, G., Swarm Intelligence: From Natural to Artificial Systems, New York, NY, USA: Oxford University Press, Inc., 1999.

[5] Calonder, M., Lepetit, V., Strecha, C., and Fua, P., «BRIEF: Binary Robust Independent Elementary Features,» *In Proceedings of the European Conference on Computer Vision (ECCV), 2010*. 1, 2, 3, 4.

[6] Deb, K., *Multi-Objective Optimization using Evolutionary Algorithms*, John Wiley & Sons, 2001.

[7] E. Cuevas, M. Cienfuegos, D. Zaldivar, M. Pérez-Cisneros, «A swarm optimization algorithm inspired in the behavior of the social-spider,» *Expert Systems with Applications*, 40, pp. 6374-6385, 2013.

[8] F. Fausto, E. Cuevas & A. Gonzalez, «A new descriptor for image matching based on bionic principles,» *Pattern Analysis and Applications (paper on revision), 2016*.

[9] Karaboga, D., «An Idea Based on Honey Bee Swarm for Numerical Optimization. Technical Report-TR06. Engine,» Engineering Faculty, Computer Engineering Department, Erciyes University, 2005.

[10] Kennedy, J. and Eberhart, R., «Particle Swarm Optimization,» *in Proceedings of the 1995 IEEE International Conference on Neural Networks*, 4, pp. 1942-1248, 1995.

[11] Leutenegger, S., Chli, M., & Siegwart, R. Y. BRISK: «Binary Invariant Scalable Keypoints,» *IEEE International Conference on Computer Vision (ICCV) 2011*, pp. 2548-2555.

[12] Lowe, D. G.. «Distinctive Image Features from Scale-Invariant Keypoints,» *International Journal of Computer Vision*, 60 (2), (2004), pp. 91-100.

[13] Mikolajczyk, K., Schmid, C., «A Performance Evaluation of Local Descriptors,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), pp. 1615-1630, 2005

[14] Olague, G., Trujillo, L., «Interest point detection through multiobjective genetic programming,» *Applied Soft Computing*, 12, pp. 2566-2582, 2012

[15] Rosten, E., & Drummond, T. «Machine learning for high-speed corner detection,» Computer Vision – ECCV 2006, *Lecture Notes in Computer Science*, 3951, (2006), pp. 430-443.

[16] Viola, P., & Jones, M. «Rapid Object Detection using Boosted Cascade of Simple Features,» *IEEE Computer Society Conference on Computer Vision and Patter Recognition (CVPR) 2001*, 1, pp. 511-518

[17] X.S. Yang, *Engineering optimization: An Introduction with Metaheuristic Applications*, John Wiley & Sons, 2010.

[18] X.S., Yang, Nature-inspired metaheuristic algorithms, Beckington: Luniver Press, 2008.

Solución del Problema PATS con el Algoritmo de Recocido Simulado

Felipe Leza Álvarez
Centro de Investigación Cientifica y de Educación Superior de Ensenada
Carretera Ensenada-Tijuana No. 3918, Zona Playitas, Ensenada
Baja California, México
fleza@cicese.edu.mx

Israel Marck Martínez Pérez
Centro de Investigación Cientifica y de Educación Superior de Ensenada
Carretera Ensenada-Tijuana No. 3918, Zona Playitas, Ensenada
Baja California, México
israelmp@cicese.mx

RESUMEN

El problema PATS consiste en diseñar un sistema de ensamblado de mosaicos de ADN con el menor número de tipos de mosaicos con los que se ensamble un patrón de entrada de dos colores y de tamaño $m \times n$. La búsqueda de soluciones mínimas para este problema se realizó con el algoritmo de recocido simulado. La solución inicial fue un conjunto de mosaicos con el mayor número de tipos de mosaicos diferentes. La función objetivo del algoritmo fue minimizar dicho conjunto. La representación del problema se realizó por medio de una matriz de mosaicos aTAM. Se diseñaron operaciones para explorar el espacio de búsqueda a partir de la solución inicial y se experimentaron con tres esquemas de enfriamiento del algoritmo: lineal, exponencial e hiperbólico. Se construyeron los sistemas de auto-ensamblado de los resultados obtenidos con el simulador ISU TAS.

Palabras clave

Ensamblado de patrones; mosaicos de ADN; aTAM; problema PATS; Recocido Simulado; RTAS; auto-ensamblado de ADN; nanoestructuras de ADN.

1. INTRODUCCIÓN

El auto-ensamblado es el proceso por el cuál pequeños componentes automáticamente se organizan entre ellos en estructuras más grandes y complejas. En la naturaleza hay ejemplos de tal proceso: los lípidos se auto-ensamblan para formar la membrana celular, el virus bacteriófago auto-ensambla una cápside para invadir a otra bacteria, etc. En general dichos sistemas implementan una serie de tareas simples de crecimiento que son controladas de manera autónoma [3].

La investigación del auto-ensamblado con ADN tiene dos enfoques importantes: por un lado, como un método para construir estructuras; por otra parte como modelo de computación [7]. En el primer enfoque, el objetivo es codificar la entrada de un problema computacional a cadenas de ADN de tal manera que las interacciones entre ellas produzcan una codificación de salida que solucione el problema [3]. En el segundo, el objetivo es diseñar secuencias de ADN que automáticamente ensamble una estructura deseada.

En este sentido, los trabajos se han centrado en la construcción de estructuras más complejas a partir de plantillas de ADN de dos dimensiones [13]. Una forma de implementar estas plantillas de auto-ensamblado es a partir de elementos llamados mosaicos de ADN. El modelo combinatorio que simplifica su comportamiento se llama *abstract Tile Assembly Model (aTAM)* y fue propuesto por Winfree en [18].

En particular, la molécula de ADN se utiliza en muchos laboratorios como bloque de construcción de nanoestructuras, debido a que es fácil de sintetizar y sus propiedades físicas son bien conocidas [13]. En la construcción de estas nanoestructuras de ADN es importante definir el conjunto de mosaicos que auto-ensamblan un patrón específico, ya que es de suma importancia obtener el menor número de tipos de mosaicos para reducir el riesgo de error en la implementación química en un laboratorio. En [9] Ma y Lombardi formularon esto como un problema de optimización combinatoria al que llaman *Pattern self-Assembly Tile set Synthesis (PATS)*. Este consiste, en dado un patrón coloreado de dos dimensiones, construir un sistema de ensamblado de mosaicos para que auto-ensamble el patrón dado, tal que se minimice el número de tipos de mosaicos diferentes necesarios.

En [2] se demostró que el problema *PATS* es *NP-Difícil* y que la búsqueda de un conjunto mínimo de mosaicos para un patrón, requiere una cantidad de tiempo exponencial en el peor de los casos. Para solucionar el problema, en el estado del arte se han propuesto dos algoritmos voraces [9], un algoritmo exhaustivo [4] y un algoritmo basado en heurísticas [7]. En este trabajo de investigación se aplicó un algoritmo de recocido simulado para resolver el problema.

2. PRELIMINARES

En esta sección se explica de manera general los conceptos y definiciones relacionados al problema PATS.

2.1 El ADN

Una molécula de ADN (ácido desoxirribonucleico) esta formada por unidades llamadas nucleótidos: Adenina (A), Citosina (C), Guanina (G) y Timina (T). La complementariedad de Watson-Crick consiste en que un nucleótido A siempre se va a unir con un T y un G a un C [14].

2.2 El modelo aTAM

El *aTAM* es un modelo formal del auto-ensamblado de moléculas de ADN [15]. Sus componentes básicos son unidades cuadradas llamadas mosaicos, con etiquetas en cada uno de sus lados como se muestra en la Figura 1(b). Cada etiqueta a, b, c, d representa un tipo de pegado molecular, que puede unir al mosaico con otros que tengan el mismo tipo de

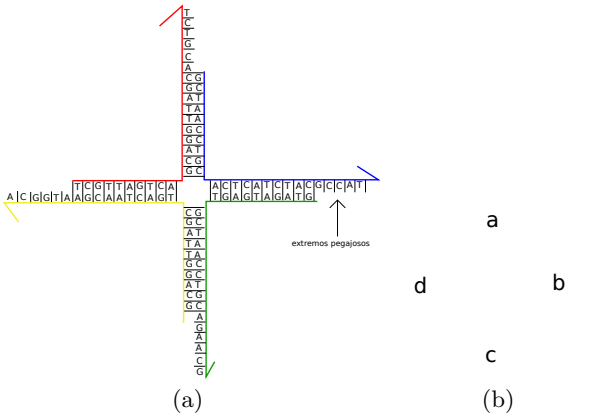


Figura 1: (a) Un mosaico de ADN, es un complejo de ADN, que tiene cuatro segmentos "pegajosos" que le permiten unirse a otros mosaicos con segmentos complementarios para formar nanoestructuras de ADN [3]. En el modelo aTAM, los mosaicos de ADN se representan por elementos cuadrados con etiquetas en cada uno de los lados (b).

pegado. Además, tienen asociada una fuerza de pegado g_i, el cual debe ser un número no negativo. A una "temperatura" τ, los mosaicos se pueden unir si la suma de las fuerzas de pegado de sus lados coincidentes son al menos τ. Los mosaicos no pueden rotarse ni reflejarse. Hay un número finito de diferentes tipos de mosaicos y uno infinito de copias de cada tipo. Los mosaicos se auto-ensamblan en un arreglo de dos dimensiones agregando un tipo de mosaico a la vez e iniciando desde una semilla [7].

2.3 Ensamblado de patrones

Formalmente un sistema de ensamblado rectilíneo de mosaicos (*RTAS, Rectilineal Tile Assembly System*) se define de la siguiente manera [5]:

Sea $\mathbb{N}$ el conjunto de enteros positivos, y $n \in \mathbb{N}$, $[n] = \{0, 1, 2, \ldots, n-2, n-1\}$. Para $k \geq 1$, un patrón k-coloreado es una función parcial que va de $\mathbb{N}^2$ al conjunto de índices $[k]$, y un patrón rectangular k-coloreado (con ancho w y alto h) es un patrón y su dominio es $[w] \times [h]$.

Sea Σ un alfabeto de pegado. Un tipo de mosaico t es una tupla $t = (g_N, g_W, g_S, g_E, c)$, donde $g_N, g_W, g_S, g_E \in \Sigma$ representa el valor de pegado del norte, oeste, sur y este, respectivamente, de t, y $c \in \mathbb{N}$ es el color de t.

Para referirse a los valores de la tupla g_N, g_W, g_S, g_E se realiza por medio de $t(N), t(W), t(S), t(E)$, respectivamente, y $t(c)$ para el color de t. Para un conjunto T de tipos de mosaicos, un ensamble α sobre T es una función parcial de $\mathbb{N}^2$ a T. Su patrón, denotado por $P(\alpha)$, cumple que $dom(P(\alpha)) = dom(\alpha)$ y $P(\alpha)(x, y) = c(\alpha(x, y))$ para todo $(x, y) \in dom(\alpha)$.

Un *RTAS* es un par $\mathcal{T} = (T, \sigma_L)$ de un conjunto de tipos de mosaicos T y una forma-L σ_L llamada semilla, la cual es un ensamblado sobre otro conjunto de tipos de mosaicos disjuntos de T tal que

$$dom(\sigma)_L = \{(-1, -1) \cup ([w] \times \{-1\}) \cup (\{-1\} \times [w])\}$$

para algún $w, h \in \mathbb{N}$. Su tamaño se define como el número de tipos de mosaicos utilizados, es decir, $|T|$. Como regla general, en todos los RTAS, el ensamblado estará delimitado por la semilla, por lo que los primeros mosaicos comenzarán a pegarse a partir de σ_L, cumpliendo con las siguientes definiciones:

Definición 1. Para un ensamble α, un mosaico $t \in T$ puede pegarse en la posición (x, y) si

1. $\alpha(x, y)$ es indefinido
2. $\alpha(x-1, y)$ y $\alpha(x, y-1)$ están definidos
3. $t(W) = \alpha(x-1, y)[E]$ y $t(S) = \alpha(x, y-1)[N]$

Definición 2. Un conjunto de mosaicos T es *dirigido* si se cumple que para cualquier tipo de mosaicos diferentes $t_1, t_2 \in T$, $t_1(W) \neq t_2(W)$ o $t_1(S) \neq t_2(S)$. Un RTAS $\mathcal{T} = (T, \sigma_L)$ es dirigido si el conjunto de mosaicos T es dirigido.

2.4 El problema PATS

El problema PATS, propuesto por Ma y Lombardi [9], consiste en minimizar un RTAS que auto-ensamble un patrón rectangular dado. El RTAS es medido por la cardinalidad del conjunto de tipos de mosaicos. Como un RTAS mínimo tiene la propiedad de ser dirigido, la definición formal es la siguiente [16]:

Definición 3. Dado un patrón rectangular P. Se requiere encontrar un RTAS mínimo que únicamente auto-ensamble P.

La entrada a este problema es un patrón o figura rectangular que contiene k colores y la salida es un conjunto mínimo de mosaicos aTAM. A cada tipo de mosaico se le asigna un color, que corresponde al color que aparece en la misma posición del patrón.

El caso general del problema es conocido como k-PATS, donde k es el número de colores únicos en el patrón de entrada. Muchas variantes del k-PATS se han demostrado que son NP-Difíciles, entre las que se encuentran valores para $k = 60, 29, 11, 3, 2$ [5]. En este trabajo de investigación se considera el caso particular para $k = 2$.

3. METODOLOGÍA

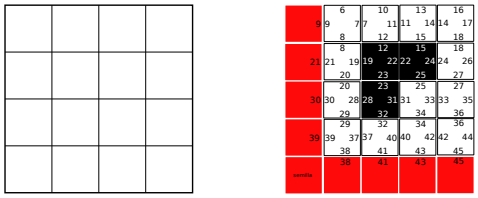
Se utilizó el algoritmo de recocido simulado [1, 6] para minimizar el conjunto de mosaicos de un RTAS. El algoritmo de recocido simulado mejora varios aspectos de las técnicas de búsqueda local, en las que una de las principales desventajas es que la búsqueda termina en un óptimo local. Recocido simulado permite seleccionar soluciones peores para que sean aceptadas y terminar la búsqueda en un óptimo global [10].

La representación del problema fue con una matriz de tamaño $m \times n$ para un patrón P. Se construyó una solución inicial y se diseñaron cinco operaciones para la generación del vecindario de soluciones.

3.1 Representación del problema

Un patrón bidimensional P, es una matriz de mosaicos aTAM de tamaño $m \times n$ tal que

$$P = \begin{pmatrix} t_{1,1} & \cdots & t_{1,n} \\ \vdots & \ddots & \vdots \\ t_{m,1} & \cdots & t_{m,n} \end{pmatrix}$$



(a) (b)

Figura 2: (a) Patrón de entrada. (b) Solución inicial generada.

donde $t_{i,j} = (n, e, s, o, c)$ son los valores de pegado para el norte, este, sur y oeste, respectivamente. Además, c es el color del mosaico $t_{i,j}$ que toma los siguientes valores:

$$c = \begin{cases} N & \text{si} \quad color = negro \\ B & \text{si} \quad color = blanco \end{cases}$$

Para cada $t_{i,j} \in P$ se cumple que

$$W(t_{i,j}) = E(t_{i,j-1}) \tag{1}$$

$$S(t_{i,j}) = N(t_{i+1,j}) \tag{2}$$

$$N(t_{i,j}) = S(t_{i-1,j}) \tag{3}$$

$$E(t_{i,j}) = W(t_{i,j+1}) \tag{4}$$

3.2 Solución inicial

La construcción de la solución inicial se hace de tal manera que, a todos los mosaicos del patrón se les asignan valores a sus fuerzas de pegado para que todos sean de diferente tipo. Se toma un valor inicial w y se comienza asignando valores a $t_{1,1}$, para que $t_{1,1} = (w, w+1, w+2, w+3)$. Para los siguientes mosaicos se revisa si comparten valores vecinos y que ya tengan valores asignados. De ser así, se les asigna el mismo valor que tienen los lados adyacentes.

En la Figura 2 se muestra la construcción de la solución inicial de un patrón de tamaño 4×4, con $w = 6$, $t_{1,1} = (6, 7, 8, 9)$, $t_{1,2} = (10, 11, 12, 7)$, etc. El patrón resultante también contiene los limites marcados con mosaicos en color rojo que corresponden a la forma-L o semilla. Estos últimos permiten comenzar y guiar el auto-ensamblado. Se puede observar que la matriz de salida contiene 36 mosaicos diferentes coloreados con negro o blanco.

3.3 Generación del vecindario de soluciones

Las operaciones que se diseñaron para generar el vecindario de soluciones, seleccionan aleatoriamente dos mosaicos del mismo color de una solución actual y se realiza una nueva asignación de los valores norte, este, sur y oeste de tales mosaicos, para convertirlos a un mismo tipo.

En la Figura 3 se muestra un ejemplo de una nueva asignación de valores a los mosaicos seleccionados para convertirlos a un mismo tipo. En (a) se tiene una solución inicial para el patrón de tamaño 4×4. Los dos mosaicos con el contorno azul son seleccionados para generar un nuevo mosaico con valores que se obtienen de una asignación que esta representada con las flechas en (b). Para que $t_1 = t_2$, se le asignan los mismo valores norte, este, sur y oeste a ambos como se muestra en (c). El siguiente paso es revisar que se cumplan las ecuaciones 1, 2, 3, y 4. Por ejemplo en (c) los rectángulos sombreados no coinciden con los valores norte, este, sur y oeste de mosaicos a su alrededor, por lo que se cambian los valores no coincidentes con los nuevos valores. Se puede observar que el número de tipos de mosaicos se reduce.

A continuación se describen a detalle las operaciones para la generación del vecindario de soluciones con el algoritmo de recocido simulado.

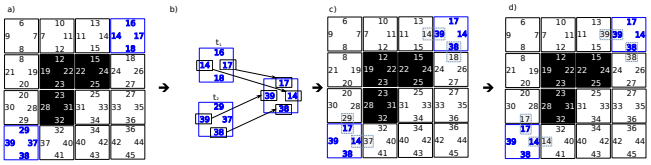


Figura 3: Unión de mosaicos para su reducción.

3.3.1 *Operación 1*

Una vez seleccionados dos mosaicos del mismo color $t_1, t_2 \in P$, se copian los valores de cada uno de los lados de t_1 a los lados correspondientes de t_2, esto es

$$t_2 = (N(t_1), E(t_1), S(t_1), W(t_1))$$

En la Figura 4 se muestra el resultado de la operación con dos mosaicos del mismo color.

3.3.2 *Operación 2*

Esta operación asigna valores mínimos a los lados de los dos mosaicos t_1 y t_2. Primero se obtienen los valores mínimos entre lados del mismo tipo, es decir

$$\begin{aligned} n &= \text{mín}\{N(t_1), N(t_2)\}, \\ e &= \text{mín}\{E(t_1), E(t_2)\}, \\ s &= \text{mín}\{S(t_1), S(t_2)\}, \\ o &= \text{mín}\{W(t_1), W(t_2)\}. \end{aligned}$$

Enseguida, se asignan los valores respectivos a los dos mosaicos para que $t_1 = (n, e, s, o)$ y $t_2 = (n, e, s, o)$. Por ejemplo en la Figura 5, $t_1 = (23, 31, 32, 28)$ y $t_2 = (12, 22, 23, 19)$. Los nuevos valores norte, este, sur y oeste son $t_1 = (12, 22, 23, 19)$ y $t_2 = (12, 22, 23, 19)$, aplicando esta operación.

3.3.3 *Operación 3*

Se selecciona aleatoriamente un mosaico t_1. La posición de t_1 en la matriz es (i, j). Primero, se toman los mosaicos que se encuentran alrededor de t_1, es decir, los mosaicos en las posiciones: $(i-1, j)$, $(i-1, j-1)$, $(i, j-1)$, $(i+1, j-1)$, $(i+1, j)$, $(i+1, j+1)$, $(i, j+1)$, $(i-1, j+1)$. En la Figura 6 se muestran las posiciones de los mosaicos con respecto a t_1. De todos los vecinos encontrados, se eliminan aquellos que no son del mismo color que t_1. Si t_1 no tiene vecinos del mismo color, como en la Figura 6(c), se elige aleatoriamente un mosaico t_k del mismo color de alguna otra parte de la matriz.

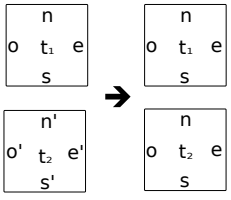


Figura 4: Operación 1.

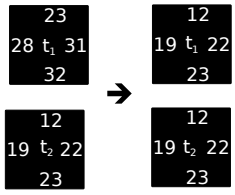
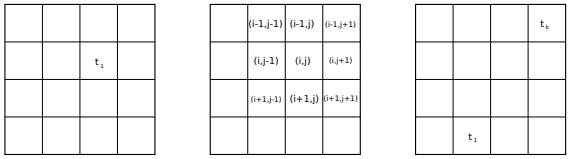


Figura 5: Operación 2.



(a) Mosaicos alrededor de t_1

(b) Posiciones de los mosaicos

(c) Mosaico sin vecinos del mismo color

Figura 6: Selección de mosaicos vecinos de t_1.

Los vecinos se almacenan en una lista L de mosaicos. Para cada vecino $t_v \in L$. Se almacenan los valores de sus lados norte, este, sur y oeste en una lista Z. También se agregan los valores de los lados de t_1. Aleatoriamente se seleccionan cuatro valores $\{a, b, c, d\} \in Z$ y se le asignan a los lados de t_1 y t_v de la siguiente manera: $t_1 = (a, b, c, d), t_v = (a, b, c, d)$.

En la Figura 7 se muestra un patrón al que se le aplica esta operación. El número de vecinos encontrados para t_1 son $v_1, v_2, v_3 . L = \{v_1, v_2, v_3\}$. Se selecciona v_1. Se agregan los valores de v_1 y t_1 a Z, esto es $Z = \{x, y, a, w, a, b, c, d\}$. Supongamos que los cuatros valores aleatorios seleccionados de Z son $\{y, w, b, c\}$. La nueva asignación de v_1 y t_1 es $v_1 = \{y, w, b, c\}$ y $t_1 = \{y, w, b, c\}$. Primero se actualizan los valores de t_1 y después los v_1. Se realizan los pasos anteriores para v_2 y v_3.

3.3.4 Operación 4

En esta operación se seleccionan dos mosaicos t_1 y t_2. Se agregan a una lista L los valores de sus lados norte, este, sur y oeste. A continuación se realiza una permutación sobre L y se seleccionan cuatro valores $\{a, b, c, d\}$ y se asignan a los lados de t_1 y t_2, para que $t_1 = (a, b, c, d)$ y $t_2 = (a, b, c, d)$. Si t_1 está una posición arriba, a la izquierda, derecha o abajo de t_2, los valores de los lados de los mosaicos se realiza de tal manera que $t_1 = t_2$. Lo mismo ocurre si t_2 esta una posición arriba, abajo, a la derecha o izquierda de t_1.

Para ilustrar esta operación se toma como ejemplo la Figura 8, donde t_1 y t_2 son los mosaicos señalados en color rojo. Este es el caso cuando t_1 se encuentra una posición arriba de t_2. Los elementos seleccionados de la permutación de Z se señalan en color verde y se le asignan a los lados correspondientes de los dos mosaicos. Primero se actualiza t_1 y después t_2, por lo que el lado sur de t_1 vuelve a cambiar, causando que $t_1 \neq t_2$. Para que sean iguales, se modifica el valor sur de t_2 de tal manera que tenga el mismo valor que el norte de t_1, y así se cumpla que $t_1 = t_2$.

3.3.5 Operación 5

Se selecciona un mosaico t_1 y se obtienen todos sus vecinos del mismo color. Si no tiene vecinos, se elige de cualquier parte de la matriz otro mosaico del mismo color. En una

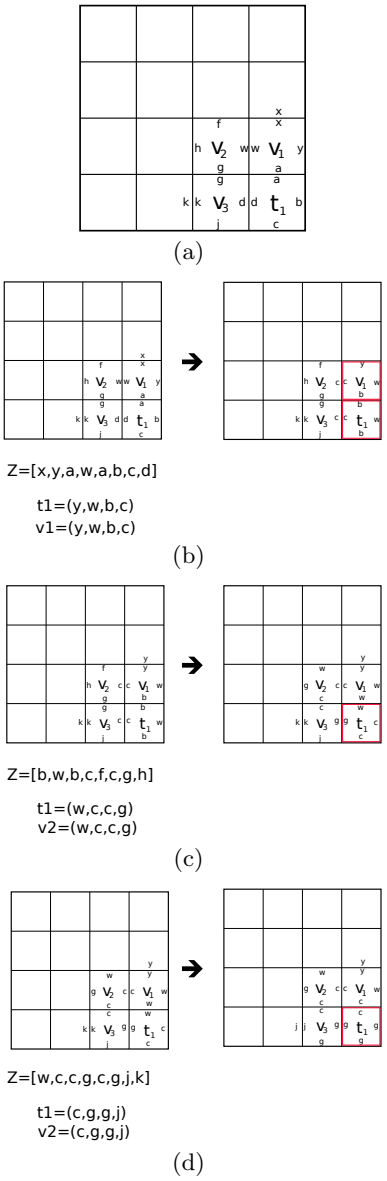


Figura 7: Ejemplo de la operación 3.

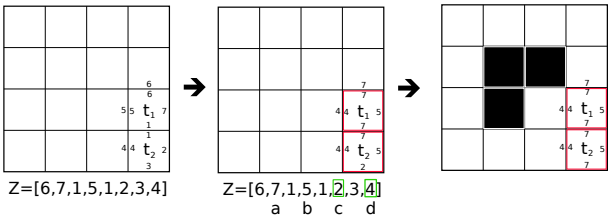
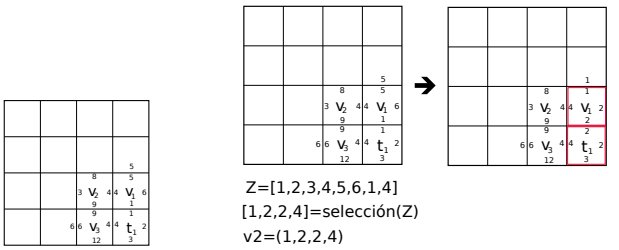
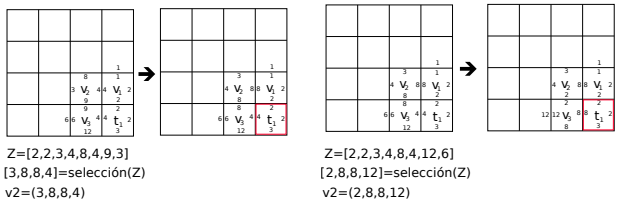


Figura 8: Ejemplo con operación 4.



(a) Buscar vecinos de t_1

(b) Selección de valores entre v_1 y t_1



(c) Selección de valores entre v_2 y t_1

(d) Selección de valores entre v_3 y t_1

Figura 9: Operación 5.

lista Z, se agregan los valores de los lados norte, este, sur y oeste de t_1 y de v_1. Posteriormente, se realiza una selección con repetición de cuatro valores de Z. Se asignan a v_1 los valores obtenidos a cada uno de sus lados, mientras que a t_1 no se le realiza ningún cambio.

En la Figura 9(a), $t_1 = (1, 2, 3, 4)$ y tiene los vecinos v_1, v_2, v_3. Para v_1 y t_1, Z y la selección de los cuatro valores es como en la Figura 9(b). Se continua la selección entre v_2 y t_1 (Figura 9(c)) y entre v_3 y t_1 (Figura 9(d)).

4. EXPERIMENTOS

Para la implementación del algoritmo de optimización, se utilizó el framework OPT4J [8]. Es de código libre y está desarrollado en el lenguaje de programación Java para computación evolutiva. Se utilizó la implementación del algoritmo de recocido simulado.

Se ejecutó el algoritmo implementado con 14 patrones de prueba (ver Tabla 1). Para cada patrón se realizaron experimentos con las cinco operaciones diseñadas. Cada operación se ejecutó con los esquemas de enfriamiento lineal, hiperbólico y exponencial [17, 11]. Para este último, con $\alpha = 0.85$. Se eligieron dos valores para la temperatura inicial de 10000 y 1000000. El valor de la temperatura final fue de uno. El número de iteraciones del algoritmo de recocido simulado fue de 1000000. Se obtuvieron 30 combinaciones de los parámetros anteriores. Cada patrón de prueba se ejecutó con las 30 configuraciones y se obtuvo el promedio de los resultados de cada patrón.

5. RESULTADOS

La configuración de parámetros con la que se obtuvieron conjuntos de mosaicos de menor cardinalidad fue utilizando la operación 5, con un enfriamiento exponencial. En la Tabla 1, se muestran la media ($\bar{x}$), la desviación estándar (σ) y la mejor solución (f), de los experimentos realizados con esta configuración. También se presentan los resultados de las mejores soluciones encontradas para cada patrón de prueba.

Tabla 1: Mejores resultados obtenidos

Patrón	Tamaño	$\bar{x}$	σ	f
Contador binario	15×15	40.70	$\pm$ 5.48	31
	32×32	189.50	$\pm$ 15.97	169
	6×6	7	$\pm$ 1.89	4
Esquina de árbol	23×23	122	$\pm$ 7.06	110
Aleatorio	12×12	46.10	$\pm$ 2.64	43
	16×16	83.30	$\pm$ 3.37	77
Triángulo de Sierpinski	12×12	42	$\pm$ 4.32	36
	17×17	74.90	$\pm$ 4.63	66
	32×32	281.70	$\pm$ 16.55	259
	64×64	1207.60	$\pm$ 28.93	1161
	6×6	10.30	$\pm$ 2.54	4
	8×8	13.20	$\pm$ 2.94	9
Tablero de ajedrez	15×15	17.80	$\pm$ 16.02	2
	6×6	4.80	$\pm$ 2.57	2

Para los patrones Triángulos de Sierpinski, Tableros de ajedrez y Contador binario se conoce que el óptimo conocido es de 4, 4 y 2 respectivamente [7]. En los experimentos realizados se obtienen las soluciones óptimas para los patrones Contador binario de 6×6, Triángulo de Sierpinski de 6×6 y Tablero de ajedrez de 6×6 y 15×15. La Tabla 2 contiene los resultados de los algoritmos propuestos en el estado del arte del problema PATS. También, se presentan los obtenidos por el algoritmo de recocido simulado (SA) implementado en este trabajo de investigación.

Tabla 2: Resultados de los algoritmos del estado del arte: Voraz [9], PS-BB [4], PS-H [7] (F^* es el valor del óptimo conocido)

Patrón	Tamaño	[9]	[4]	[7]	SA	F^*
Contador binario	15×15	121	–	–	31	4
	32×32	–	307	20	169	4
	6×6	19	–	–	4	4
Esquina de árbol	23×23	–	192	25	1103	–
Triángulo de Sierpinski	12×12	80	20	–	36	4
	17×17	172	23	–	66	4
	32×32	595	42	4	259	4
	64×64	–	95	4	1161	4
	6×6	21	–	–	4	4
	8×8	37	–	–	9	4
Tablero de ajedrez	15×15	119	–	–	2	2
	6×6	19	–	–	2	2

6. SIMULACIONES

Para verificar que un conjunto de mosaicos dado por el algoritmo de recocido simulado ensambla el patrón deseado, se realizaron simulaciones con el software ISU TAS (*Iowa State University Tile Assembly Simulator*). ISU TAS es un simulador y editor de sistemas de ensamblado de mosaicos aTAM [12]. Para la simulación de los sistemas de ensamblado de mosaicos de los casos de prueba, los datos de entrada fueron un conjunto de mosaicos y una semilla.

La Figura 10 es el resultado de la simulación para los patrones: Contador binario de (a) 32×32, (b) 15×15, (c) 6×6; (d) Esquina de árbol de 23×23; Aleatorio de (e) 12×12, (f) 16×16; Triángulo de Sierpinski de (g) 6×6, (h)

8×8, (i) 12×12, (j) 17×17, (k) 32×32, (l) 64×64; Tablero de ajedrez de (m) 6×6 y (n) 15×15.

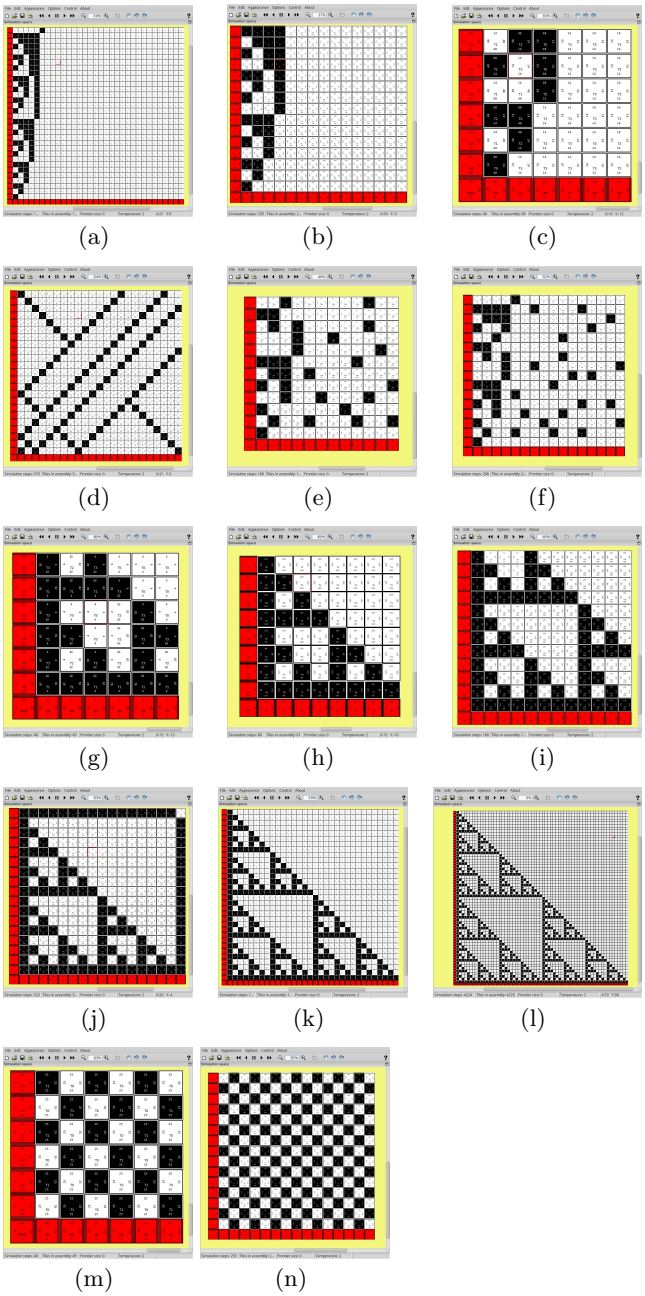


Figura 10: Resultado de las simulaciones de las soluciones obtenidas para todos los patrones de prueba.

7. CONCLUSIONES

Con los resultados obtenidos con la operación 5 para los patrones de mayor tamaño, la reducción al conjunto de mosaicos no es tan significativa como en los de menor tamaño. Una posible explicación es que las operaciones para generar soluciones vecinas realizan perturbaciones aleatorias a la matriz de mosaicos, por lo que, para una matriz pequeña esa perturbaciones se realizan sobre un espacio de búsqueda reducido de acuerdo a la solución inicial. Por otra parte, en las matrices de mayor tamaño, la perturbación hecha incrementa el espacio de búsqueda permitiendo regresar nuevamente al espacio de búsqueda de la solución inicial.

Con los resultados obtenidos se observó que los esquemas de enfriamiento más adecuados para la operación con mejores resultados fue el exponencial e hiperbólico. La selección de temperaturas iniciales se realizó a través de prueba y error, principalmente tomando en cuenta los resultados de los patrones de mayor tamaño.

El algoritmo de recocido simulado obtuvo mejores soluciones para varios patrones de menor tamaño. Para patrones de mayor tamaño se realiza una reducción en comparación con el conjunto trivial. pero no se obtiene una reducción que mejore los resultados del algoritmo propuesto en [7] para el patrón Triángulo de Sierpinski de 64×64.

8. REFERENCIAS

[1] V.˜Černý. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985.

[2] E.˜Czeizler and A.˜Popa. Synthesizing minimal tile sets for complex patterns in the framework of patterned DNA self-assembly. *Theoretical Computer Science*, 499:23–37, aug 2013.

[3] D.˜Doty. Theory of algorithmic self-assembly. *Communications of the ACM*, 55(12):78–88, 2012.

[4] M.˜Göös and P.˜Orponen. Synthesizing Minimal Tile Sets for Patterned DNA Self-Assembly. page˜12, 2009.

[5] L.˜Kari, S.˜Kopecki, P.˜Meunier, M.˜J. Patitz, and S.˜Seki. Binary pattern tile set synthesis is np-hard. *CoRR*, abs/1404.0967, 2014.

[6] S.˜Kirkpatrick, C.˜D. Gelatt, and M.˜P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):pp. 671–680, 1983.

[7] T.˜Lempiäinen. *The PATS Problem : Search Methods and Reliability Department of Computer Science*. PhD thesis, 2015.

[8] M.˜Lukasiewycz, M.˜Glaß, F.˜Reimann, and J.˜Teich. Opt4J – A Modular Framework for Meta-heuristic Optimization. pages 1723–1730, 2011.

[9] X.˜Ma and F.˜Lombardi. Synthesis of tile sets for DNA self-assembly. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 27(5):963–967, 2008.

[10] Z.˜Michalewicz and D.˜B. Fogel. *Basic Concepts*, pages 35–48. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[11] M.˜Ortner, X.˜Descombes, and J.˜Zerubia. An adaptive simulated annealing cooling schedule for object detection in images. Research Report RR-6336, INRIA, 2007.

[12] M.˜J. Patitz. Simulation of self-assembly in the abstract tile assembly model with ISU TAS. *CoRR*, abs/1101.5151, 2011.

[13] J.˜Reif and T.˜Labean. Self-assembled DNA Nanostructures and DNA Devices. *NanoFabrication Handbook*, pages 299–328, 2012.

[14] J.˜H. Reif and T.˜H. LaBean. Biomolecular devices. *Communications of the ACM*, 50(9), 2007.

[15] P.˜W.˜K. Rothemund and E.˜Winfree. The program-size complexity of self-assembled squares

(extended abstract). In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 459–468, New York, NY, USA, 2000. ACM.

[16] S.˜Seki. Encyclopedia of Algorithms. pages 8–11, 2014.

[17] D.˜Simon. Simulated Annealing. In *Evolutionary optimization algorithms*, chapter˜9, pages 227–234. John Wiley & Sons, Inc, 2013.

[18] E.˜Winfree. Algorithmic Self-Assembly of DNA Thesis by. 1998, 1998.

On-line Path Planning for Mobile Robots using Parallel Evolutionary Artificial Potential Field

Ulises Orozco-Rosas
Instituto Politécnico Nacional
Centro de Investigación y Desarrollo de Tecnología Digital (IPN-CITEDI)
Tijuana, Baja California, México
uorozco@citedi.mx

Oscar Montiel
Instituto Politécnico Nacional
Centro de Investigación y Desarrollo de Tecnología Digital (IPN-CITEDI)
Tijuana, Baja California, México
oross@ipn.mx

Roberto Sepúlveda
Instituto Politécnico Nacional
Centro de Investigación y Desarrollo de Tecnología Digital (IPN-CITEDI)
Tijuana, Baja California, México
rsepulvedac@ipn.mx

ABSTRACT

In this work, an extension of the parallel evolutionary artificial potential field (PEAPF) for on-line mobile robot path planning is presented. The PEAPF algorithm is based on the evolutionary artificial potential field (EAPF) method to derive optimal potential field functions using evolutionary computation. And, due to the parallel computation, the PEAPF algorithm is capable to find an optimal collision free path to drive a mobile robot from a start location to a goal location in a moderate runtime. Simulation results to validate the analysis and implementation of the PEAPF algorithm for on-line mobile robot path planning are provided; the experiments were specifically designed to show the effectiveness and the efficiency of the PEAPF algorithm.

Keywords

Path planning, parallel computing, evolutionary computation, artificial potential field, mobile robots

1. INTRODUCTION

At present, robotics is one of the most important technologies since it is a fundamental part in automation and manufacturing process. Mobile robotics is a young field. Its roots include many engineering and science disciplines, from mechanical, electrical and electronics engineering to computer, cognitive and social sciences [9]. Nowadays, there is a demand for mobile robots in various fields of application, such as material transport, cleaning, monitoring, driver assistance, and military applications [12]. These mobile robots must interact with their environment to accomplish their tasks. Path planning task is one of the most critical process in terms of computation time for mobile robot navigation [3, 14]. This work addresses the problem of on-line path planning for mobile robot navigation, understanding navigation as the move from a start location to a goal location through a defined path. The mobile robot will have to interact with the environment and avoid collisions with obstacles, following the path planned to achieve its established mission. All of these tasks, without the assistance of a human operator.

In this work, the parallel evolutionary artificial potential field (PEAPF) algorithm presented in [5] has been extended for on-line mobile robot path planning. The PEAPF algorithm is employed to find an optimal collision free path, taking advantage of novel processor architectures. The main objective of this work is to present an extension of the PEAPF algorithm, and a comparative performance analysis of the PEAPF algorithm for on-line path planning with unknown static and dynamic obstacles.

The organization of this work is as follows: in Section 2, the PEAPF algorithm implementation for on-line mobile robot path planning is described in detail by two flowcharts, one for the PEAPF algorithm and a second one for the complete on-line path planning algorithm. In Section 3, the experimental framework consisting of two sceneries that demonstrate the relevance of the PEAPF algorithm for on-line path planning is exposed; additionally, the performance evaluation methodology used in this work is described. Finally, in Section 4 the conclusions of this work are given.

2. ON-LINE PATH PLANNING USING PEAPF

There are three tasks that must be carried out by an autonomous robot to enable the execution of the robot navigation [7, 8]. These activities are: world perception consisting in mapping and modeling the environment; path planning for obtaining an ordered sequence of objective points and convert this sequence into a path; the path tracking in controlling the robot to follow the path [5, 9]. The selection of an appropriate algorithm in every stage of the path planning is very important to ensure that the navigation process will run smoothly [7]. Here, we use the PEAPF algorithm [5], described in Fig. 1 as a core part of the on-line mobile robot path planning algorithm described in Fig. 2, to obtain a feasible, safe and efficient path in a moderate run time to drive the mobile robot in an autonomous form to its goal. The PEAPF algorithm is the integration of the artificial potential field method [2] with evolutionary algorithms [11] and parallel computation for taking advantages of novel processor architectures [6], to obtain a flexible path planner capable of performing the on-line mobile robot path planning.

In Fig. 1, the flowchart of the PEAPF algorithm is shown. The flowchart is organized in two parts, the evolutionary algorithm to optimize the proportional gains k_r and k_a and the artificial potential field method (fitness function), that is executed in parallel form (dashed rectangle). The PEAPF algorithm starts with the creation of an initial random population $P(t)$ of chromosomes, where the population size is N. Each chromosome codifies a pair of proportional gains, so each pair is composed by one repulsion gain k_r and one

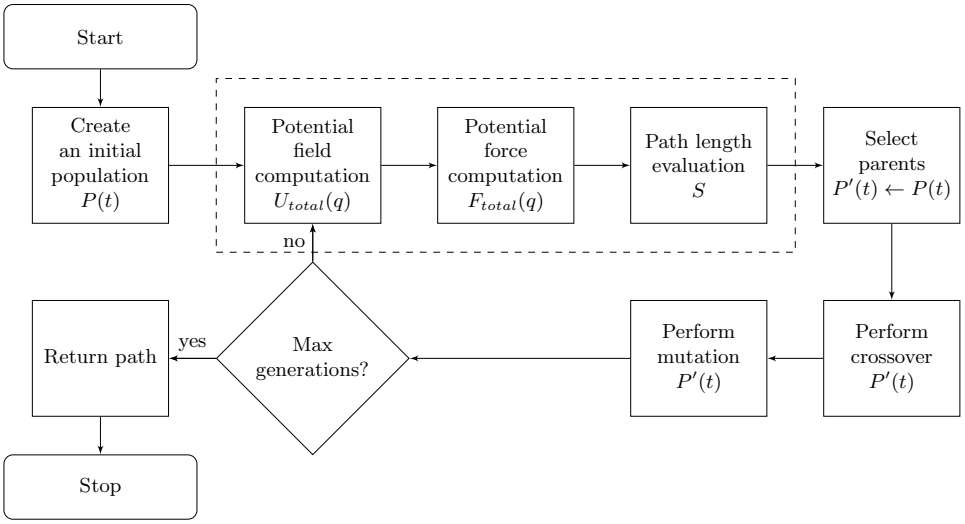


Figure 1: Parallel evolutionary artificial potential field (PEAPF)

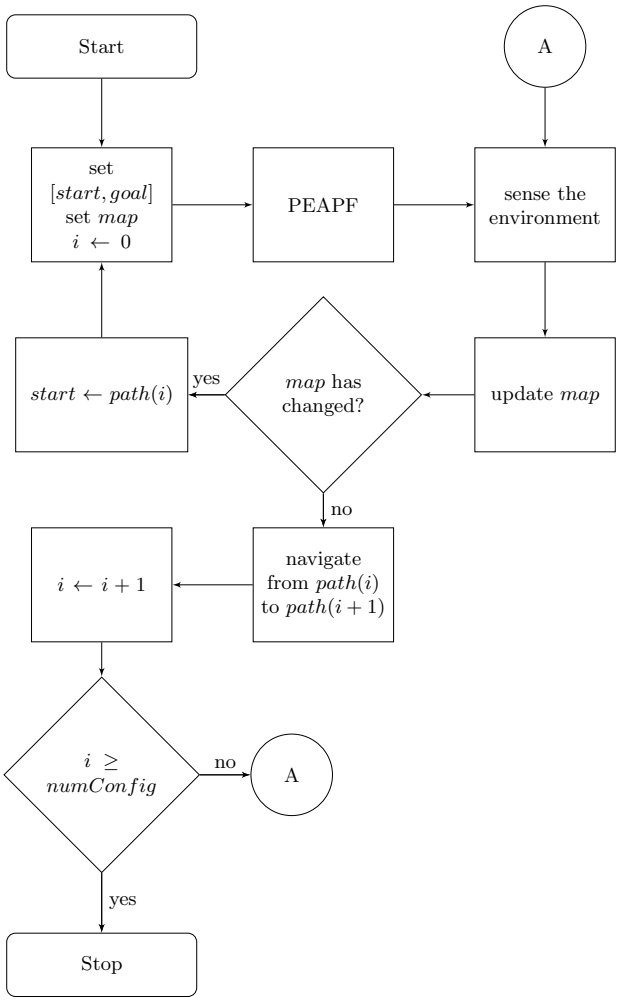


Figure 2: On-line mobile robot path planning using PEAPF

attraction gain k_a. The path generated with the solution given by the PEAPF algorithm will be composed by the best pair found.

In the fitness function the first process is to calculate the potential field $U_{total}(q)$ for each chromosome,

$$U_{total}(q) = \frac{1}{2}k_a(q - q_f)^2 + \frac{1}{2}k_r\left(\frac{1}{\rho} - \frac{1}{\rho_0}\right)^2 \qquad (1)$$

where, q is the robot position and q_f is the goal position, k_r and k_a are scalar variables that represent the attractive and repulsive proportional gains respectively; ρ_0 is the limit distance of influence of the potential field, and ρ is the shortest distance to the obstacle.

The second process of the fitness function is to calculate the potential force $F_{total}(q)$ for each chromosome,

$$F_{total}(q) = -\nabla U_{total}(q). \qquad (2)$$

The last step of the fitness function is the evaluation of the length of the path given by each chromosome. The evaluation is performed using the function S which is the cumulative Euclidean distance from point to point in the generated path, where, each path consists of n points,

$$S = \sum_{i=0}^{n-1} \|q_{i+1} - q_i\| \ . \qquad (3)$$

When the fitness function evaluation ends, the process continues with the selection where the chromosomes are chosen according to their fitness value. The selection process drives the evolutionary algorithm to improve the population fitness over the successive generations [10]. Next, the crossover is performed. The crossover operation roughly mimics biological recombination between two single chromosomes (haploid) organisms [4]. Then, the mutation process is performed, where, random mutations alter a certain percentage of the bits in the list of chromosomes. The mutation process tends to distract the evolution from converging on a popular solution [1].

The PEAPF algorithm shown in the flowchart of Fig. 1 iterates until the maximum number of generation N_{gen} is

reached. The PEAPF algorithm evolves the parameters k_r and k_a to obtain the corresponding best values. If the goal was achieved, then the algorithm returns the path (best path generated through the parameters k_r and k_a) and the number of configurations $numConfig$ to reach the goal.

In Fig. 2, the flowchart for the on-line path planning algorithm using PEAPF is shown. The algorithm requires three input parameters to work: the initial position of the mobile robot in x-y coordinates *start*, the goal point that the mobile robot must attain in x-y coordinates *goal*, and the environment configuration in a data structure *map*. The process starts with the path planning using the PEAPF algorithm, after that, the environment is sensed and the environment configuration *map* is updated. If the environment configuration has changed, we start from the beginning of the algorithm, but in this case the initial position of the robot will be the last point visited for the robot. Otherwise, if the configuration of the environment has not changed, the robot will navigate one step, and it will repeat the process from the sense of the environment, until the number of configurations $numConfig$ is achieved. When the number of configurations have been achieved, the mobile robot has reached the goal location.

3. EXPERIMENTAL RESULTS

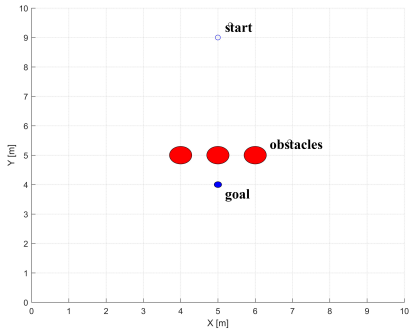
In this section, the experimental results are presented. To achieve the comparisons, all the experiments were achieved on a quad-core Intel i7-4770 CPU (3.40 GHz) with 8 GB of RAM, running Ubuntu Trusty distribution of Linux and the release R2015a of Matlab. For consistency among the experiments, we used the same PEAPF parameter values in all the experiments. The next parameters yielded the best results:

- A population size of $N = 16$ chromosomes (individuals). Each chromosome consists of two genes, k_r and k_a.
- The proportional gains are constrained, so $\{k_r, k_a \mid 0 < k_r, k_a < 10\}$.
- Single point crossover is used; a random point in two chromosomes (parents) is chosen. Offspring are formed by joining the genetic material to the right of the crossover point of one parent, and the genetic material to the left of the crossover point of the other parent.
- Mutation rate = 0.4 (40%).
- Elitist selection. Selection rate = 0.5 (50%).
- The maximum number of generations is $N_{gen} = 10$.

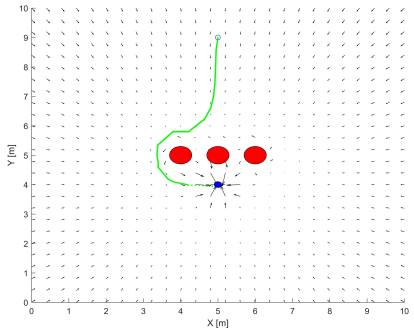
3.1 On-line path planning with unknown static obstacles

In path planning an advantage is presented when the environment is known, referred as *off-line path planning* [5]; however, real-world applications frequently face the mobile robot to environments partially known or totally unknown, where the already planned path needs to be adjusted to reach the goal. This case is referred as *on-line path planning* [13].

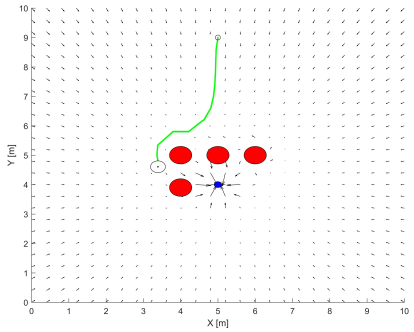
In this experiment, we are going to test the on-line path planning algorithm (described in Fig. 2) by placing an obstacle at random position, in such a way that it blocks the



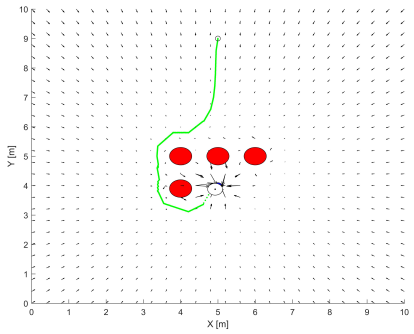
(a) Initial environment information



(b) Path planning (known environment)



(c) Obstacle detection



(d) Path planning (unknown environment)

Figure 3: On-line path planning on environments with static unknown obstacles

Table 1: Statistical results for on-line path planning on environments with static unknown obstacles

Metric	on-line path planning known environment	on-line path planning unknown environment
Best	7.0013 m	3.1904 m
Average	7.1199 m	3.2713 m
Worst	7.2663 m	3.5310 m
Std. dev.	0.0841 m	0.0905 m

path of the mobile robot, interfering with the mobile robot trajectory during the navigation. In this case, the random obstacle must remain static once it was put.

The initial information for the on-line path planning algorithm is the start of the mobile robot at (5.00, 9.00), goal at (5.00, 4.00) and the initial environment configuration, contained in *map*. At the beginning, we know the position and size of certain obstacles in the environment. For this experiment, there are three obstacles forming a barrier, see Fig. 3(a). The center position coordinates for the known obstacles are (5.00, 5.00), (4.00, 5.00), and (6.00, 5.00), all the obstacles with a radius of 0.30 m.

First, the path planning is achieved using the information of the environment, contained in *map*, see Fig. 3(b). Then, the mobile robot starts to navigate. The mobile robot follows the already planned path to reach the goal. During the navigation an obstacle is added at position (4.00, 3.90) to make a change in the environment. Next, the mobile robot at position (3.39, 4.60) detects the new obstacle, and it calculates the obstacle position to update the obstacles configuration (*map*), see Fig. 3(c). Later, the on-line path planning algorithm recalculates the new path using the new environment configuration. Finally, the mobile robot follows the updated path to reach the goal, see Fig. 3(d).

Table 1 shows the statistical results for the on-line path planning experiment on the environment with a static unknown obstacle. The results are product of thirty independent test in each case, labeled as known environment and unknown environment. On Table 1, it can be observed that the shortest path length (best) to reach the goal has a length of 7.0013 m for the known environment in Fig. 3(b). The length of the path is calculated from the start position at (5.00, 9.00) to the goal (5.00, 4.00). The proportional gains calculated for the PEAPF algorithm (described in Fig. 1) to achieve a path length of 7.0013 are $k_r = 0.9140$ and $k_a = 0.3180$. During the navigation when the mobile robot detects the new obstacle, see Fig. 3(c). The path is recalculated to reach the goal, now the shortest path length (best) from the current position of the robot to the goal has a length of 3.1904 m for the unknown environment in Fig. 3(d). The length of the path is calculated from the new start position at (3.3915, 4.6060) to the goal. The proportional gains calculated for the PEAPF algorithm to achieve a path length of 3.1904 m are $k_r = 3.4827$ and $k_a = 1.8260$.

Table 2 shows the statistical results of execution time for on-line path planning on environments with static unknown obstacles. The Table shows a summary of the obtained results. The results for the sequential implementation in Matlab, for the sequential implementation in C/C++, and for the parallel implementation in C/C++ using the application programming interface OpenMP (Open Multi-Processing). Through the results, it can be observed that the best speedup is achieved using the implementation in parallel. For the case of known environment, we have 9.8067 of speedup, and for the case of unknown environment we have 7.2589 of speedup.

3.2 On-line path planning with unknown dynamic obstacles

The interaction with the real world requires the ability to respond and taking decisions over the changes in the environment [5]. Differently to the last experiment, where the random obstacle remained motionless after it appeared in the environment; here, once the path was planned, a new obstacle appears in the environment blocking the robot trajectory, but the obstacle will not remain static, it will be moving following its own trajectory. This problem is more complicated because it faces the PEAPF algorithm to make their optimization work over a noisy surface; hence, at every iteration, the searching zone might be different.

In this experiment, we are going to test the on-line path planning algorithm by placing a dynamic obstacle at random position following its own trajectory, in such a way that in certain time it is going to block the mobile robot's path, interfering with the mobile robot trajectory during navigation.

The initial information for the on-line path planning algorithm is the start of the mobile robot at (5.00, 9.00), goal at (7.50, 4.50) and the initial environment configuration *map*. At the beginning, we know the position and size of certain obstacles in the environment. For this experiment, there are three obstacles, see Fig. 4. The center position coordinates for the known obstacles are (5.00, 7.10), (6.20, 7.50), and (6.00, 8.50), all the obstacles with a radius of 0.30 m.

First, the path planning is achieved using the information of the environment *map*, see Fig. 5(a). Then, the mobile robot starts to navigate. The mobile robot follows the already planned path to reach the goal. During the navigation a dynamic obstacle is added at position (4.00, 6.00) to make a change in the environment. Next, the mobile robot at position (5.79, 7.03) detects that the new obstacle is at (6.20, 6.42) and it is interfering with the mobile robot trajectory, at this time the on-line path planning algorithm calculates the obstacle position to update the obstacles configuration (*map*), see Fig. 5(b). Now, the on-line path planning algorithm recalculates the new path using the new environment configuration and the mobile robot will continue with its navigation. After some time, when the mobile robot is at position (5.76, 4.35) detects a second dynamic obstacle that is interfering with its trajectory, at this time the on-line path planning algorithm calculates the obstacle position to update the obstacles configuration (*map*), see Fig. 5(c). So, the on-line path planning algorithm recalculates the new path using the new environment configuration. Finally, the mobile robot follows the updated path to reach the goal, see Fig. 5(d).

Table 3 shows the statistical results for the on-line path planning experiment on environments with dynamic unknown obstacles. The results are product of thirty independent test in each case, labeled as known environment and unknown environment for the 1st and the 2nd obstacle detected. On Table 3, it can be observed that the shortest path length (best) to reach the goal has a length of 5.0477 m for the known environment in Fig. 5(a). The length of the path

Table 2: Execution time for on-line path planning on environments with static unknown obstacles

	Metric	Sequential Matlab	Sequential C/C++	Parallel C/C++/OMP
known environment	Speedup	—	3.5343 x	9.8067 x
	Average	1.7829 s	0.5045 s	0.1818 s
	Std. dev.	0.1104 s	0.0669 s	0.0326 s
unknown environment	Speedup	—	3.1625 x	7.2589 x
	Average	1.3454 s	0.4254 s	0.1853 s
	Std. dev.	0.1443 s	0.0800 s	0.0418 s

Table 3: Statistical results for on-line path planning on environments with dynamic unknown obstacles

Metric	on-line path planning known environment	on-line path planning unknown environment 1st obst. detected	on-line path planning unknown environment 2nd obst. detected
Best	5.0477 m	4.6122 m	1.9657 m
Average	5.1195 m	4.7729 m	1.9776 m
Worst	5.3210 m	5.1501 m	1.9909 m
Std. dev.	0.1103 m	0.1562 m	0.0059 m

Table 4: Execution time for on-line path planning on environments with dynamic unknown obstacles

	Metric	Sequential Matlab	Sequential C/C++	Parallel C/C++/OMP
known environment	Speedup	—	3.7313 x	8.2202 x
	Average	1.2232 s	0.3278 s	0.1488 s
	Std. dev.	0.5415 s	0.1630 s	0.0741 s
unknown environment 1st obst.	Speedup	—	3.8439 x	8.8884 x
	Average	0.9128 s	0.2375 s	0.1027 s
	Std. dev.	0.2148 s	0.0647 s	0.0311 s
unknown environment 2nd obst.	Speedup	—	3.8516 x	8.1878 x
	Average	0.7544 s	0.1959 s	0.0921 s
	Std. dev.	0.0625 s	0.0221 s	0.0137 s

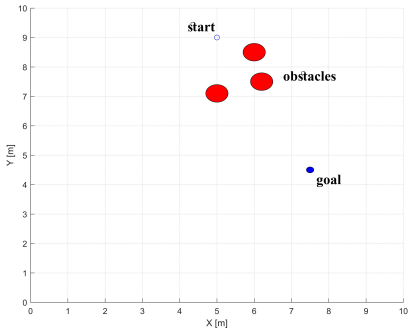


Figure 4: Initial environment information

is calculated from the start position at (5.00, 9.00) to the goal (7.50, 4.50). The proportional gains calculated for the PEAPF algorithm to achieve a path length of 5.0477 are $k_r = 4.3692$ and $k_a = 8.0989$.

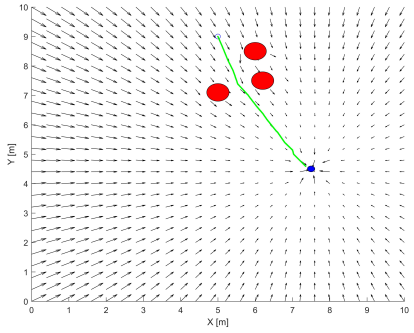
During the navigation when the mobile robot detects the first dynamic obstacle, see Fig. 5(b). The path is recalculated to reach the goal, the shortest path length (best) now to reach the goal has a length of 4.6122 m for the partially unknown environment. The length of the path is calculated from the new start position at (5.7961, 7.0337) to the goal. The proportional gains calculated for the PEAPF algorithm to achieve a path length of 4.6122 m are $k_r = 7.3963$ and $k_a = 1.7438$.

A second dynamic obstacle is detected after a while during the navigation, see Fig. 5(c). The path is recalculated to reach the goal, now the shortest path length to reach the goal has a length of 1.9657 m for the partially unknown environment. The length of the path is calculated from the new start position at (5.7643, 4.3579) to the goal. The proportional gains calculated for the PEAPF algorithm to achieve a path length of 1.9657 m are $k_r = 8.6080$ and $k_a = 8.9653$.

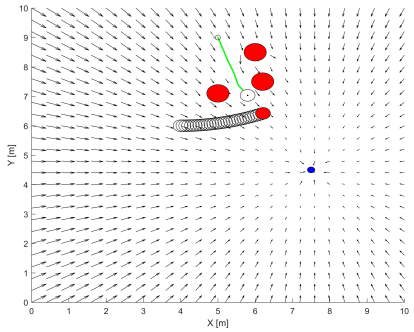
Table 4 shows the statistical results of execution time for on-line path planning on environments with dynamic unknown obstacles. Through the results, it can be observed that the best speedup is achieved using the implementation in parallel. For the case of known environment we have 8.2202 of speedup, and for the case of unknown environment we have 8.8884 of speedup when the first obstacle is detected and 8.1878 of speedup in the path computation when the second obstacle is detected.

4. CONCLUSIONS

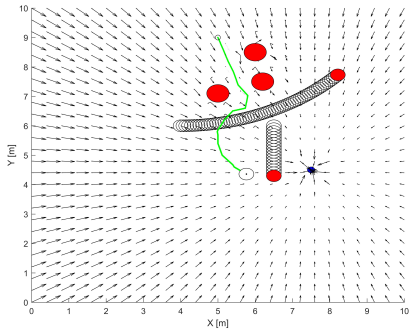
In this work, we have seen how the PEAPF algorithm solves the on-line path planning problem in all the different situations exposed. The results demonstrated that the parallel implementation is very powerful to accelerate the evaluation of the solutions, for the experiments evaluated in this work, we have reached in average an 8.47 of speedup. Due to the experimental results, we can conclude that the PEAPF algorithm can be capable of facing more complex and bigger on-line planning problems, making the implementation suitable for real-world applications in complex environments. For future work, the PEAPF algorithm can be extended for path planning in collaborative mobile robots, challenging the PEAPF algorithm in more complex environments, e.g., in robotic swarm's environments.



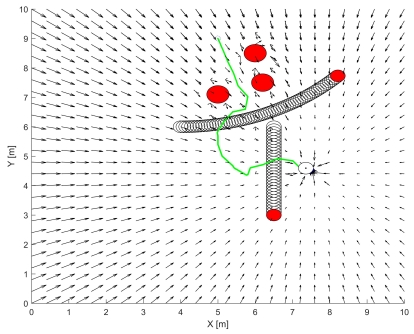
(a) Path planning (known environment)



(b) First unknown obstacle detected



(c) Second unknown obstacle detected



(d) Path planning (unknown environment)

Figure 5: On-line path planning on environments with dynamic unknown obstacles

5. ACKNOWLEDGMENTS

This work was supported in part by the Commission of Operation and Promotion of Academic Activities of the Instituto Politécnico Nacional (COFAA-IPN), and in part by the Mexican National Council of Science and Technology (CONACYT).

6. REFERENCES

[1] R. L. Haupt and S. E. Haupt. *Practical Genetic Algorithms*. Wiley, Hoboken, New Jersey, USA, 2004.

[2] O. Khatib. Real-time obstacle avoidance for manipulators and mobile robots. In *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*, volume 2, pages 500–505, Mar 1985.

[3] M. Liu, B. Xu, and X. Peng. Cooperative path planning for multi-auv in time-varying ocean flows. *Journal of Systems Engineering and Electronics*, 27(3):612–618, June 2016.

[4] M. Mitchell. *An introduction to Genetic Algorithms*. Bradford, Cambridge, Massachusetts, USA, 2001.

[5] O. Montiel, R. Sepúlveda, and U. Orozco-Rosas. Optimal path planning generation for mobile robots using parallel evolutionary artificial potential field. *Journal of Intelligent & Robotic Systems*, 79(2):237–257, 2014.

[6] U. Orozco-Rosas, O. Montiel, and R. Sepúlveda. Parallel evolutionary artificial potential field for path planning - an implementation on gpu. In P. Melin, O. Castillo, and J. Kacprzyk, editors, *Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization*, volume 601 of *Studies in Computational Intelligence*, pages 319–332. Springer International Publishing, 2015.

[7] N. Sariff and N. Buniyamin. An overview of autonomous mobile robot path planning algorithms. In *Research and Development, 2006. SCOReD 2006. 4th Student Conference on*, pages 183–188, June 2006.

[8] D. H. Shin and S. Singh. Path generation for robot vehicle using composite clothoid segments. Technical Report CMU-RI-TR-90-31, The Robotic Institute. Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, December 1990.

[9] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza. *Introduction to Autonomous Mobile Robots. Second Edition*. The MIT Press, Cambridge, Mass., 2011.

[10] S. N. Sivanandam and S. N. Deepa. *Introduction to Genetic Algorithms*. Springer, Heidelberg, 2008.

[11] P. Vadakkepat, K. C. Tan, and W. Ming-Liang. Evolutionary artificial potential fields and their application in real time robot path planning. In *Evolutionary Computation, 2000. Proceedings of the 2000 Congress on*, volume 1, pages 256–263 vol.1, 2000.

[12] P. Vyas, L. Vachhani, K. Sridharan, and V. Pudi. Cordic-based azimuth calculation and obstacle tracing via optimal sensor placement on a mobile robot. *IEEE/ASME Transactions on Mechatronics*, 21(5):2317–2329, Oct 2016.

[13] L. Yang, J. Qi, D. Song, J. Xiao, J. Han, and Y. Xia. Survey of robot 3d path planning algorithms. *Journal of Control Science and Engineering*, pages 1–22, 2016.

[14] J. Yu and S. M. LaValle. Optimal multirobot path planning on graphs: Complete algorithms and effective heuristics. *IEEE Transactions on Robotics*, pages 1–15, 2016.

Application of Genetic Algorithm for solving Fairness based Resource Allocation Problem in Mobile Systems

Adolfo Reyna-Orta
Facultad de Ingeniería, Universidad Autónoma de Baja California
Mexicali, Baja California, México
adolfo.reyna@uabc.edu.mx

Ángel G. Andrade
Facultad de Ingeniería, Universidad Autónoma de Baja California
Mexicali, Baja California, México
aandrade@uabc.edu.mx

Guillermo Galaviz
Facultad de Ingeniería, Universidad Autónoma de Baja California
Mexicali, Baja California, México
ggalaviz@uabc.edu.mx

ABSTRACT
Fairness plays and important role for resource distribution in a cellular system which seeks to maximize utility (date rate) under the scheme of dynamic resource allocation. In this sense, an unfair resources allocation among different individuals may lead to several undesired effects, such as resource starvation, wastage or even redundant allocation. In this work two scenarios with homogenous and heterogeneous users' conditions are considered. Within those scenarios Max-Min, Max-Fair and Proportional Fairness optimization problems are solved with the Genetic Algorithm and the results are compared with the optimal unfair solution and with the allocation generated with heuristics. Results show that close fairness is easily achieved by homogenous conditions, but it is virtually impossible when users' conditions are heterogeneous.

CCS Concepts
• Networks→ Network types→ Mobile networks • Computing methodologies→ Artificial intelligence→ Search methodologies→ Discrete space search.

Keywords
OFDMA; Fairness; Proportional Fairness; Genetic Algorithm; Resource Allocation;

1. INTRODUCTION
Wireless and mobile technologies are growing rapidly, leading to much more complex spectrum usage, where the supply and demand for radio spectrum vary dynamically. The conventional fixed spectrum allocation approach has been shown to be unsuitable for flexible and efficient use of the scarce radio spectrum. Orthogonal Frequency-Division Multiplexing (OFDM) is therefore designed to provide flexibility and efficiency in spectrum usage by dividing the entire bandwidth, assigned to a cellular communication system, into multiple subcarriers' signals to carry mobile data on several parallel data streams or channels. The primary advantage of OFDM over single-carrier schemes is its ability to cope with severe channel conditions (for example, attenuation of high frequencies, narrowband interference, and frequency-selective fading due to multipath). Orthogonal Frequency-Division Multiple Access (OFDMA) supports differentiated quality of service by assigning different number of subcarriers to different users. Each of these subchannels can be finely adjusted in power and/or modulation; this leads to the idea of the optimum use configuration for each transmission.

Users in an OFDM system measure the strength of each subcarrier signal and send this information to the base station as Channel State Information (CSI). The value of the CSI contains the fading phenomena and path loss of the propagated signal and it represents the gain or loss of the subchannel. Path loss and shadowing depends of the user physical conditions, i.e. the distance between the user and the base station and the objects that break the line of sight. When these user conditions are homogeneous, they are affected in the same proportion; nevertheless multipath fading still affects them differently due its random nature, but with the same probability density function.

In a dynamic resource allocation scheme, OFDMA system assigns resources to users whose CSI values are good. However, this strategy tends to be unfair since some users may be denied the communication service due to their specific channel conditions. For example, a user located at the cell edge who could be attended, might not, because the base station will prefer to serve a closer user with higher CSI value.

An unfair resource allocation among different users may lead to resource starvation, wastage or redundant allocation [1]. To avoid this situation, fairness must be considered in such scenarios. The goal of fair resource allocation schemes is to guarantee (near) equal distribution of utility, though this comes at the expense of throughput. This assumes a trade-off between the service provider's profits and user satisfaction [2].

Multiple works have addressed the problem of dynamic resource allocation for OFDMA. In 2000, Rhee [3], solves the Rate Adaptation (RA) problem by assigning each resource dimension (time or frequency) to only one user, avoiding interference and error propagation problems. The authors claim that large path loss under heterogeneous user conditions causes discrepancy between users' data rates, so the system becomes unfair for the users with poor channel gains. The authors propose an algorithm for channel allocation with a Max-Min Fairness scheme that assigns a channel to the user with the lowest current data rate, iteratively. In [4], a channel allocation scheme in OFDMA systems applying a greedy strategy is proposed. They prove that assigning each channel to only the user with the higher CSI value is optimum for RA, however, this solution leads some users to starvation, even with homogenous user conditions. In [5] and [6], the authors add Proportional Fairness to the algorithm reported in [3]. A resource allocation scheme that linearizes the power allocation problem while achieving approximate rate proportionality is developed in [5]. In a similar way, in [6] the power allocation is linearized and solved by an iterative Newton-Raphson method. After subchannel allocation, only coarse Proportional Fairness is achieved, the goal of maximizing the sum capacity while maintaining proportional fairness is achieved by the power allocation, even in heterogeneous conditions.

Various works apply Genetic Algorithms (GAs) as an alternative to solve the optimization problem related to dynamic resource allocation in OFDMA systems. For example, Wang, in [7], applies a GA for the Margin Adaptation (MA) scheme. MA aims to minimize the overall power required to achieve users' data rate requirements; unlike RA who aims to maximize overall throughput while no restrictions over individual data rates are

given. The GA evolves the subchannel assignation, searching for the lowest power consumption possible. For each subchannel allocation solution, the water-filling method is used to allocate each user's bits to the subchannel and calculate the overall transmit power as the fitness. In [8], a similar scheme is proposed, which claims to improve the GA convergence by adding high fitness chromosomes at the end of each generation. In [7] and [8], each individual chromosome is coded as an array of elements that represents the subchannels. This analogy is extended in [9] to simultaneously allocate subchannels and bits on each chromosome. With this, the RA optimization problem with Max-Min Fairness scheme is solved, but instead of a power budget, a power transmission objective for the second objective for the NSGA-II is considered. In [10], to achieve Proportional Fairness, the authors propose the power allocation with a GA after allocating subchannels. In [11], a GA that minimizes the total transmission power is reported. It satisfies the bit rate constraint for each user.

These previous works that solve resource allocation with GA consider homogeneous conditions in their evaluation scenarios; only multipath fading for channel gains is considered. That means all users have the same probability of reaching the best CSI for one or more channels. In this work we solve the resource allocation problem in an OFDMA system in both scenarios: homogeneous and heterogeneous conditions. We apply GA to evaluate the assignment of subchannels to users taking into account three fairness schemes: Max-Min, Proportional Fair and Max-Fair. The results are compared with the optimal greedy solution, as upper bound, and the algorithm proposed by Rhee.

2. SYSTEM MODEL

Our scenario consists of a single cell, with a base station in the center of it. Let an OFDM system have a total bandwidth of B, with maximum allowable total power $P_{\max}$, K users and N subchannels of bandwidth B/N.

The user data rate of each sub-channel can be expressed by the Shannon's capacity formula that express the amount of data transmitted without error:

$$R_{k,n} = \frac{B}{N}\log_2\left(1+\frac{p_{k,n}\mathrm{h}_{k,n}^2}{N_0}\right) \quad (1)$$

where $p_{k,n}$ is the power transmission for each subchannel, $h_{k,n}^2$ is the channel gain percieved for the k user to the n subchannel, and N_0 is the power density of noise.

Following the idea used in [3], [4] that if sub-channels are assigned to users that perceive high channel gains, then power allocation made a slight difference in the efficiency. Every subchannel transmits with flat PSD, i.e. each channel has a power of $p_{k,n} = p = P_{\max}/N$.

Then the system assigns a subset of these channels to each user. This is represented by the assign variable $\rho_{k,n}$ that works as follows:

$$\rho_{k,n} = \begin{cases} 1 & if\ the\ subchannel\ n\ is\ assign\ to\ user\ k \\ 0 & otherwise \end{cases} \quad (2)$$

Each user is allowed to have more than one sub-channel to transmit its data in a parallel fashion, so the user data rate R_K is the sum of the rates obtained by each sub-channel n assigned to it. Equation (2) calculates the data rate of the user k in all the sub-channels assigned to it.

$$R_k = \sum_{n=1}^{N}\frac{B}{N}\,\rho_{k,n}\,\log_2\left(1+\frac{p_{k,n}\mathrm{h}_{k,n}^2}{N_0}\right) \quad (3)$$

The main objective is then to maximize the data rate of the system, so the optimization problem is defined by:

$$\max_{\rho_{k,n}} B\sum_{k=1}^{K}\sum_{n=1}^{N}\rho_{k,n}\,\log_2\left(1+\frac{p_{k,n}\mathrm{h}_{k,n}^2}{N_0}\right) \quad (4)$$

Subject to:

$$\sum_{k=1}^{K}\sum_{n=1}^{N} p_{k,n} \le P_{Total} \quad (5)$$

$$p_{k,n} \ge 0\ \forall\ k,n \quad (6)$$

$$\rho_{k,n} = \{0,1\}\ \forall\ k,n \quad (7)$$

$$\sum_{k=1}^{K}\rho_{k,n} = 1\ \forall\ n \quad (8)$$

Equations (5) and (6) are the restrictions for the power budget and not negative power allocation, which are achieved by flat PSD. Restriction in (7) considers only two possible values for the variable of channel assignation and (8) prohibits subchannel sharing, so only one user can be assigned to each subchannel.

Maximizing the data rate of the system (as shown in (4)) by allocating each channel to the user with the highest channel gain leads to unfair resource distributions, and user resource starvation [4]. To solve this problem, it is necessary to consider fairness in the optimization problem. In this work, three well known approaches to fairness are considered: Max-Min, Proportional Fairness and Max-Fair. For each one a modified objective function is presented in such a way that a balance between utility and fairness is maintained.

1.- Max-Min approach seeks to maximize the data rate of the user who has the smallest data rate on each allocation.

$$\max_{\rho_{k,n}} \min_{\mathrm{k}} R_k \quad (9)$$

2.- Proportional Fairness searches for a balance between utility and fairness, in such a way that users with high data rates share their resources with others only if the result of this sharing ends in an equal or larger proportional gain for them.

$$\max_{\rho_{k,n}} \prod_{k=1}^{K} R_k \quad (10)$$

3.-Max-Fair approach looks to find the fairest resource allocation distribution possible. This objective uses a quantitative measure of fairness for the users' data rate.

$$\max_{\rho_{k,n}} J(R_k) \quad (11)$$

The fairness measurement used in this work was proposed by Jain, equation (12) indicates the level of fairness in a given distribution. When fairness is complete it returns a value of 1, with decreasing values as the level of unfairness increases.

$$J(R) = \frac{\left(\sum_{k=1}^{K} R_k\right)^2}{n\cdot\sum_{k=1}^{K} {R_k}^2} \quad (12)$$

For these objectives fairness is proposed to be a soft guarantee, so nearly fair distributions are accepted to be feasible solutions, similar to Wong's proposal [5] for proportional fairness.

In OFDMA systems, every user measures the CSI of each sub-channel and sends this information to the central base station. This is considered for the distribution of the frequency resources.

We considered two scenarios for the evaluation of fairness and utility trade-off. In a homogeneous scenario, all the users are located at the same distance, d, from the base station. The subchannel gain for each user is calculated by:

$$h_{k,n}^2 = d^{-\alpha} \cdot rand([0-1]) \quad (13)$$

where $\propto$ is the path loss factor, and the uniform random value represents the multipath fading gain. In this scenario all users have the same probability of receiving good channels gains.

In a heterogeneous scenario, the users are located in the cell in a random manner. In this case, distance between the user and the base station affects the channel gain differently for each user. This heterogeneity in user conditions allows the phenomena of unfairness to be more perceptible, and raises the complexity of achieving fairness in the system. Every user, k, is located at a random distance d_k in range of $[20\ m - 250\ m]$, therefore channel gain is affected as follows:

$$h_{k,n}^2 = d_k^{-\alpha} \cdot rand([0-1]) \quad (14)$$

Unlike the homogeneous case, this represents a higher probability of receiving high channel gains for users who are closest to the base station.

3. GA FOR OFDMA RESOURCE ALLOCATION

The GA follows the analogy of the evolution of species presented by Darwin. Each solution represents an individual entity that has a certain degree of adaptation to its environment. In Darwin's theory individuals that are better adapted have a higher possibility of surviving and producing offspring with their own good genes. The recombination of two strong individuals has a good chance of producing offspring with an even better gene combination. Also, new individuals have a chance of mutation over one or more genes, which can give them an advantage over others.

In the same way, the GA starts with an initial population of random solutions. The solutions are the independent variables in the optimization problem. These are coded in a gene-like prototype array that simulates a chromosome. The population is evaluated with the fitness function that assigns a real value to each one, which represents the level of adaptation. In each iteration, the algorithm chooses a set of parents in pairs from the population. Each pair produces two offspring with the recombination of their genetic information; this process is known as crossover. Each descendent has a low probability of mutation over its genes. This mutation affects the genetic information, and chances are that the process results in a better adaptation of the "good" genes that the parents had already provided them.

In our work, each solution is a vector of length N that represents the sub-channels available to be assigned, with values in the range of $[1-K]$ that represent the user assigned to each sub-channel, as the one shown in table 1. This representation fulfills the restrictions given by (7) and (8). The solution provides the information of assignation variable $\rho_{k,n}$ then every user's data rate can be calculated. The fitness functions considered in this implementation are given by (9), (10) and (11) for Max-Min, Proportional and Max-Fair fairness schemes respectively.

Table 1. Example of chromosome solution.

Channel 1	Channel 2	…	Channel N
3	1		K

4. RESULTS

In this section, the simulation results of our proposal are presented. All the experiments were implemented in Matlab®. A simple GA with a population of 50 solutions and 300 iterations is considered. The crossover probability was 0.9 and the mutation rate 0.1. The results obtained by the two scenarios were compared with the greedy (optimal) solution and the one given by the algorithm proposed by Rhee. The first scenario considers that all the users are located at 50 m from the base station, the second one places users randomly between $[20\ m - 250\ m]$. All the distances and the CSI matrix are generated at the beginning of the experiment and remain constant for all approaches and all the number of users considered. The OFDMA system considers 20 channels and users in the range of 2 to 20. The system considers a total bandwidth of 1 MHz, an overall power of 20 watts, a Noise Density of $1e^{-6}$, and a path loss factor of 3.

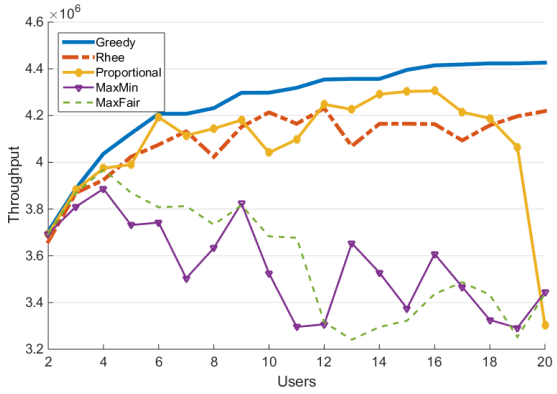


Figure 1. Evaluation of Throughput for homogeneous scenario

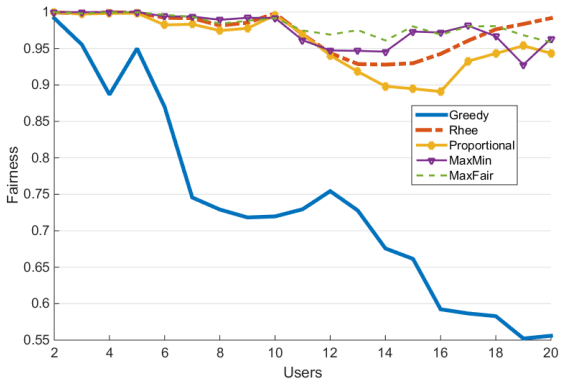


Figure 2. Evaluation of Fairness for homogenous scenario.

In Figure 1 and Figure 2, only the multipath fading is considered, so all the users have homogeneous possibilities of receiving the best CSI for one or more channels. As expected the greedy approach is optimal but is considered only as upper bound. Rhee and Proportional Fair with GA have good performance in throughput and fairness, and the other two approaches stay lower. In the case when $K = N$ it is well known that GA perform poorly

due random initialization [9], but usually $K \ll N$. It is clear that both objectives are contrary, so the approach that represents the best behavior in throughput has the worst in fairness. Rhee's algorithm uses a Max-Min strategy so it maximizes fairness, but as a heuristic, it represents suboptimal solutions. That's why GA Max-Min and GA Max-Fair tend to have better performance in fairness, and therefore less in throughput.

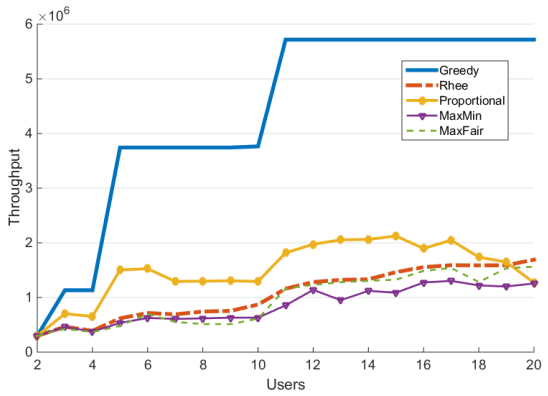


Figure 3. Evaluation of throughput in a heterogeneous scenario

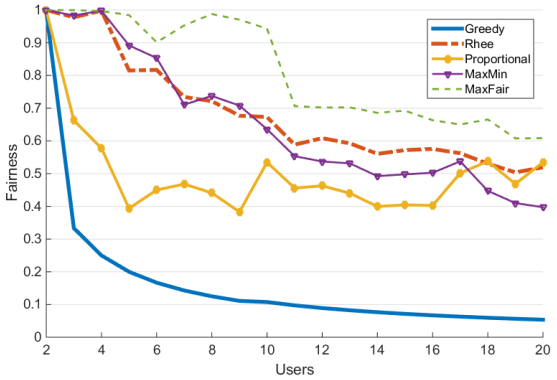


Figure 4. Evaluation of fairness in a heterogeneous scenario.

Figures 3 and 4 show the results for the heterogeneous scenario. The greedy approach increments its throughput in a fixed fashion. This occurs when a user closer to the base station is active, so most of the channels are assigned to it. In this scenario the difference between the results given for each approach becomes clear. GA Proportional Fairness presents the greatest throughput in most of the cases and the least fairness of the three, but maintains a certain level while the number of users grows, eradicating starvation. Max-Min approaches (Rhee and GA) are close in Fairness and throughput. GA Max-Fair represents the best solution in fairness in all the cases and it's close to Rhee in throughput. Nevertheless, it is visible that in this scenario the fairness decreases quickly with the number of users due to path loss differences and the considerations of flat PSD. We can see that even when the GA Max-Fair seeks the distribution with the maximum fairness possible, this has not been achieved in the heterogeneous scenario in most user cases. This means that under these conditions there is no distribution that equalizes the data rate between users. On the other hand, in homogeneous scenario GA Max-Fair always achieves fairness up to 95%, and the other approaches almost always stay above 90%.

5. CONCLUSION

In this work, a GA to solve the dynamic resources allocation problem in OFDM systems is applied. It is reported in literature that to assign resources to users with high channel gains causes resource starvation to those users with poor channel conditions. To guarantee equal resource distribution between users, the fairness principle is considered in our proposal. Simulation results show that under heterogeneous channel conditions achieving fairness in the system is limited compared to homogeneous cases. In a future work, we will consider a dynamic power allocation strategy to equalize these heterogeneous channel conditions to achieve fairness.

6. REFERENCES

[1] Prasad, R.V., Onur, E. and Niemegeers, I.G.M.M. 2014. Fairness in Wireless Networks:Issues, Measures and Challenges. *IEEE Communications Surveys & Tutorials*. 16, 1 (2014), 5–24. DOI= http://doi.org/10.1109/SURV.2013.050113.00015

[2] Lee, K., Member, S. and Yum, T.P. 2010. On Pareto-Efficiency Between Profit and Utility in OFDM Resource Allocation. *IEEE Transaction on Communications*. 58, 11 (2010), 3277–3285.

[3] Rhee, W. and Cioffi, J.M. 2000. Increase in Capacity of Multiuser OFDM System Using Dynamic Subchannel Allocation. *Vtc 2000*. (2000), 1085–1089. DIO= http://doi.org/10.1109/VETECS.2000.851292

[4] Jang, J., Member, S. and Lee, K.B. 2003. Transmit Power Adaptation for Multiuser OFDM Systems. *IEEE Journal on Selected Areas in Communications*. 21, 2 (2003), 171–178. DOI= http://doi.org/10.1109/JSAC.2002.807348

[5] Wong, I.C., Evans, B.L. and Andrews, J.G. 2004. A Low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems. *IEEE Workshop on Signal Processing Systems, 2004. SIPS 2004*. (2004), 1–6. DOI= http://doi.org/10.1109/SIPS.2004.1363015

[6] Shen, Z., Andrews, J.G. and Evans, B.L. 2005. Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints. *IEEE Transactions on Wireless Communications*. 4, 6 (2005), 2726–2736. DOI= http://doi.org/10.1109/TWC.2005.858010

[7] Ahmadi, H. and Chew, Y.H. 2009. Adaptive subcarrier-and-bit allocation in multiclass multiuser OFDM systems using genetic algorithm. *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*. (2009), 1883–1887. DOI= http://doi.org/10.1109/PIMRC.2009.5450195

[8] Reddy, Y.B., Gajendar, N., Taylor, P. and Madden, D. 2007. Computationally Efficient Resource Allocation in OFDM Systems: Genetic Algorithm Approach. *Technology*. (2007).

[9] Sharma, N., Rao, A., Dewan, A. and Safdari, M. 2008. Rate adaptive resource allocation for multiuser OFDM using NSGA - II. *Proceedings of the 4th International Conference on Wireless Communication and Sensor Networks, WCSN 2008*. (2008), 161–166. DOI= http://doi.org/10.1109/WCSN.2008.4772703

[10] Shen, Z., Andrews, J.G. and Evans, B.L. 2005. Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints. *IEEE Transactions on Wireless*

Communications. 4, 6 (2005), 2726–2736. DOI= http://doi.org/10.1109/TWC.2005.858010

[11] Ahmadi, H. and Chew, Y.H. 2009. Adaptive subcarrier-and-bit allocation in multiclass multiuser OFDM systems using genetic algorithm. *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*. (2009), 1883–1887. DOI= http://doi.org/10.1109/PIMRC.2009.5450195

Optimización de Ubicaciones de Vehículos Aéreos no Tripulados para Proveer Comunicación en Situaciones de Desastre

Gabriela L. Rodríguez Cortes
Universidad Politécnica de Pachuca
Carretera Pachuca, Cd. Sahagún km 20 Ex-Hacienda de Santa Bárbara
Zempoala, Hidalgo.
lizeth.rodriguez.cortes@gmail.com

Anabel Martínez Vargas
Universidad Politécnica de Pachuca
Carretera Pachuca, Cd. Sahagún km 20 Ex-Hacienda de Santa Bárbara
Zempoala, Hidalgo.
anabel.martinez@upp.edu.mx

Oscar Montiel Ross
Instituto Politécnico Nacional, CITEDI
Av. Instituto Politécnico Nacional 1310
Tijuana, B.C.
oross@ipn.mx

RESUMEN

Durante los últimos años las comunicaciones han jugado un papel muy importante en nuestra vida, y mantenerlas activas durante las catástrofes naturales es un reto al que se enfrentan los proveedores de comunicaciones, los servicios de emergencia e incluso el propio gobierno. Evitar que estos servicios sean interrumpidos durante una contingencia, resulta un punto crítico. Este artículo aborda el problema de encontrar la ubicación óptima de vehículos aéreos no tripulados para proveer comunicación en situaciones de desastre. Para resolverlo, se aplica un algoritmo genético con modelo de población estacionario que maximiza el número de estaciones móviles a las que se les puede proveer comunicación. Para asociar una estación móvil con un vehículo aéreo no tripulado, se utiliza la pérdida por trayectoria.

Palabras clave – Vehículos aéreos no tripulados; desastres naturales; algoritmo genético.

1. INTRODUCIÓN

Durante los últimos años las comunicaciones han jugado un papel muy importante en nuestra vida, y mantenerlas activas durante las catástrofes naturales es un reto al que se enfrentan los proveedores de comunicaciones, los servicios de emergencia e incluso el propio gobierno. Evitar que estos servicios sean interrumpidos durante una contingencia, resulta un punto crítico. Por ejemplo en la llegada del huracán Odile que afectó a cinco municipios de Baja California Sur en el 2015, el sistema de comunicaciones sufrió graves daños en postes y cableado, dejando a la región incomunicada. De acuerdo al informe de la Secretaria de Comunicaciones y Trasportes (SCT), el servicio de telefonía y todos los servicios de telecomunicaciones quedaron restablecidos 23 días después de que Odile tocara tierra en Baja California Sur [8]. Otra situación de desastre natural que tuvo gran impacto sobre el sector de las telecomunicaciones fue el huracán Katrina en el 2005 en los Estados Unidos de América (EUA), donde más de 2000 radio bases quedaron fuera de servicio. También, el servicio de emergencia 911 fue severamente dañado y tardó en restablecerse varias semanas, impidiendo que las personas pudieran comunicarse para obtener ayuda [5]. Los ejemplos antes citados, nos dan una dimensión de la importancia de desarrollar sistemas alternos de comunicaciones en casos donde la infraestructura ha sido afectada por desastres naturales.

En EUA, las comunicaciones de seguridad pública usan tecnologías de banda angosta (solo pueden transmitir voz). Esto limita su interoperabilidad, cobertura y servicio, por lo que la Comisión Federal de Comunicaciones busca implementar el uso de tecnologías de banda ancha para ofrecer servicios no solo de voz, sino también de video y datos [6]. Una de sus propuestas para enfrentar requerimientos específicos de banda ancha, es implementar Vehículos Aéreos no Tripulados (UAV's por sus siglas en inglés Unmanned Aerial Vehicles), con la finalidad de proveer comunicaciones en lugares donde la infraestructura terrestre de comunicaciones ha sido destruida [4].

Los UAV's se deben ubicar en puntos estratégicos para proveer de comunicación al mayor número de personas afectadas en un área determinada. Las ubicaciones óptimas que puede tomar un UAV dentro de un área afectada, es un problema de optimización NP-duro por el gran número de combinaciones de ubicaciones posibles. Esto lleva a pensar en la implementación de algoritmos bio-inspirados que son utilizados para tratar con este tipo de problemas de optimización. Los algoritmos bio-inspirados, exploran varias regiones en el espacio de búsqueda al mismo tiempo y buscan alcanzar la mejor solución global, es decir, no garantizan una solución óptima pero si una solución cercana a esta [9]. Los algoritmos genéticos son una rama de los algoritmos bio-inspirados, útiles para resolver problemas de combinatoria compleja. Están basados en la idea de que a lo largo de las generaciones, las poblaciones evolucionan de acuerdo con los principios de la selección natural y la supervivencia de los más fuertes.

En [1], [2], y [7] utilizan un Algoritmo Genético Simple (AGS), para ubicar estaciones base en el contexto de redes celulares bajo distintos objetivos. Por ejemplo, en [1] y [7] se maximiza la cobertura y minimiza el número de estaciones base. Por otro lado, en [2] además de tratar los problemas de maximizar la cobertura y minimizar el número de estaciones base, se incluye el problema de minimizar la potencia de la estación base. Sin embargo, las anteriores propuestas al utilizar un AGS, tienen que procesar toda la población con los operadores de cruzamiento y mutación además del método de selección. También, plantean escenarios de redes celulares convencionales, que en una situación de desastre natural resultarían dañados. En el caso de [4], utiliza un algoritmo de fuerza bruta para maximizar la cobertura de una red heterogénea donde las estaciones base son UAV´s. Este tipo de algoritmo es muy costoso computacionalmente dado que requiere hacer todas las combinaciones de ubicaciones de los UAV´s. El presente trabajo sugiere utilizar un algoritmo genético estacionario (AGE) con una representación binaria para reducir el tiempo de procesamiento para encontrar una solución satisfactoria al problema.

El trabajo se organiza como sigue: La Sección 2 describe el modelo del sistema (función objetivo y restricciones) y se caracteriza el escenario. En la Sección 3 se explica la aplicación del AGE al problema. La Sección 4 muestra los resultados

experimentales obtenidos y el análisis de ellos. Finalmente, la Sección 5 describe las conclusiones y trabajo a futuro.

2. PLANTEAMIENTO DEL PROBLEMA

Para poder proveer comunicación después de que la infraestructura de la red celular ha sido destruida por una catástrofe, los UAV's se desplegaran en la zona de desastre para proveer comunicación. Sin embargo, los UAV's al ser recursos limitados no pueden tomar cualquier ubicación, por lo que se requiere encontrar las ubicaciones óptimas de los UAV's para que puedan proporcionar el mayor número de servicios a personas afectadas. Como se muestra en la Figura 1 se define un área [-X_{max}, X_{max}] y [-Y_{max}, Y_{max}] con origen en (0,0), esta área representa la zona donde la infraestructura ha sido destruida. Los UAV's y estaciones móviles (MS por sus siglas en inglés Mobile Station), se ubicaran dentro de esta área aleatoriamente y serán representados por coordenadas cartesianas (x_k,y_k), y (x_i,y_i) respectivamente.

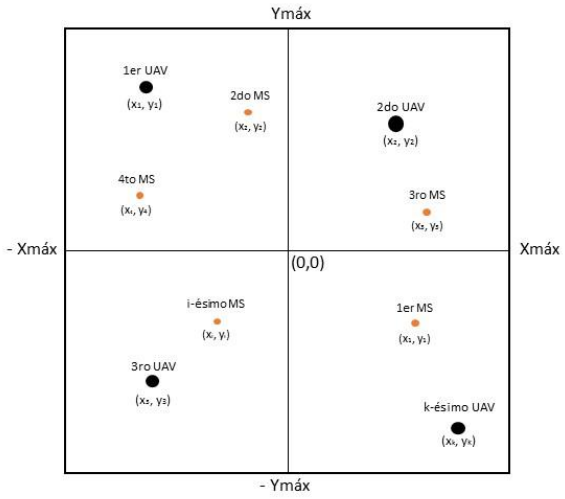


Figura 1. Representación gráfica del escenario de simulación

El individuo esta representado por una cadena binaria donde el valor del bit=1 significa que la *k*-ésima UAV fue seleccionada y un bit=0 significa que no lo es. La Figura 2 ilustra como estará representado el individuo *g*, donde el tamaño de la cadena binaria es el número máximo de UAV's.

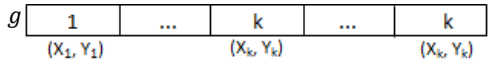


Figura 2. Representación del individuo

donde:

$g = C_1, \dots, C_k$ es el conjunto de UAV's
$u = U_1, \dots, U_i$ es el conjunto de MS´s
k es el número máximo de UAV's
i es el número máximo de MS´s
$\{(X_k, Y_k), (X_i, Y_i)\} \in R$.

Cada individuo tiene un escenario con posiciones de UAV's distintas. Las posiciones de los MS's permanecen fijas para todos los individuos. Los UAV's conocerán la posición de cada MS a través de sistema de posicionamiento global (GPS por sus siglas en inglés Global Positioning System). El sistema operara bajo la tecnología LTE (por sus siglas en inglés Long Term Evolution) lo que permitirá brindar servicios de voz, datos y video [4] .

Para decidir si un MS se asocia a una UAV, hasta el momento se ha implementado el análisis de pérdidas por trayectoria. Las pérdidas por trayectoria son las pérdidas a las que es sometida una onda electromagnética cuando es trasmitida por un medio de propagación, es decir, es la pérdida de potencia que tiene una señal al ser trasmitida en el espacio libre [10]. De acuerdo a [7] se propone utilizar un modelo de propagación Hata para zonas urbanas.

$$L_p(dB) = A + BLog_{10}(d) \quad (1)$$

donde:

d es la distancia de la UAV a la MS

$A = 69.55 + 26.16\,Log_{10}(f_c) - 13.82Log_{10}(h_b) - a(h_m)$

$B = 44.9 - 6.55Log_{10}(h_b)$

donde:

h_b es la altura de la antena del UAV

h_m es la altura de la antena de la estación móvil

$a(h_m)$ dependerá de la frecuencia de la portadora (fc).

La función objetivo es usada para evaluar al individuo y dará la calidad de la solución (aptitud). La función objetivo maximiza la cobertura de los UAV's y se utiliza la ecuación de maximización de [7]:

$$f(g) = \frac{Total\ de\ nodos\ cubiertos}{Total\ de\ nodos} \quad (2)$$

Donde el *Total de nodos cubiertos*, es la suma de las MS vinculadas a cada una de las UAV. Por otro lado, el *Total de nodos*, es el total de MS's existentes en la zona afectada.

Para que una MS sea vinculada a una UAV necesita cumplir con la siguiente condición [1]: la pérdida por trayectoria debe ser menor a 100 dB.

3. ALGORITMO PARA UBICACIÓN DE UAV'S

Un algoritmo genético localiza el espacio de búsqueda que contenga posiblemente la solución óptima al problema, aplicando técnicas de cruzamiento y mutación para obtener la descendencia que remplazara a la población. En este trabajo, se propone aplicar un AGE [3], este modelo no requiere del procesamiento de toda la población ya que reemplaza parte de ella, característica que lo hace menos costoso computacionalmente. El pseudocódigo del AGE se describe a continuación:

1: **COMENZAR** AGE
2: Generar una población inicial.
3: **HACER**
4: Seleccionar a los padres por método de torneo
5: Cruzar cuando probabilidad de cruce se cumpla
6: Mutar cuando probabilidad de mutación se cumpla
7: Evaluar a padres e hijos en la función objetivo (2)
8: Seleccionar a las mejores aptitudes entre padres e hijos
9: Reemplazar al Padre1 por mejor1
10: Reemplazar al Padre2 por el mejor2
11: **MIENTRAS** condición_paro **==**Verdadera
12: La solución es el mejor individuo de todas las iteraciones
13: **FIN** AGE

En la línea 2 se inicializa con una población aleatoria, esta población está formada por un conjunto de individuos (soluciones candidatas). Los individuos son cadenas de números binarios. En esta fase de inicialización, el algoritmo identifica la zona donde posiblemente se encuentre la solución óptima. Una vez localizada la zona, en línea 4 se seleccionan a los individuos que se aparearan (padres). La selección de padres utiliza un operador de selección por torneo [3], en la que se eligen aleatoriamente a dos individuos de la población y se evalúan para tomar al que tenga mejor aptitud. Este individuo será el padre1, para la elección del padre2 se realiza el mismo proceso. Concluida la selección por torneo, se tiene a una pareja de padres y se ejecuta el cruzamiento en línea 5. Para ello se alinean dos cadenas de individuos y de acuerdo a una probabilidad de cruzamiento (P_c), se indica si las cadenas se cruzan o no. Se aplica el operador de cruce por un punto para representación binaria [3]. Después en la línea 6 se aplica el operador de mutación, para que este proceso suceda existe un probabilidad de mutación (P_m) que debe ser cumplida, se utiliza el operador de mutación bit a bit. [3]. Cuando estos procesos terminan se crean nuevos individuos llamados hijos. La línea 7 evalúa a los padres e hijos en la función objetivo de la ecuación 2. En líneas 8-10, se obtienen a los dos individuos que contengan la mejor aptitud los cuales se insertan en la población inicial. Es por esta razón que en un AGE, no existe el concepto de generación. La condición de paro implementada es el tiempo (número de iteraciones). La solución al problema será el mejor individuo de todas las iteraciones.

4. RESULTADOS

Para demostrar y evaluar la presente propuesta, se considera un área de 14 km x 14 km [7], donde se propone ubicar a 100 MS's, distribuidos aleatoriamente con una distribución uniforme. En este caso se proponen siete UAV's, que serán ubicados aleatoriamente en el área definida con distribución uniforme. Se propone una población de 100 individuos y se determina como condición de paro 1000 iteraciones. El operador de mutación se realiza con una probabilidad de 0.1 [7] y para el operador de cruzamiento se utiliza una probabilidad de 0.7 [3]. Para vincular la MS al UAV se calcula la pérdida por trayectoria definida en la ecuación (1), si la pérdida por trayectoria cumple con la condición L_p<100 dB [2] entonces la MS se asocia al UAV. Para el cálculo de la pérdida por trayectoria se considera una altura de los UAV's (h_b) de 70 m, una altura de los móviles (h_m) de 2 m, una frecuencia de operación (f_c) de 700 MHz.

La Figura 3 muestra resultados de la convergencia del AGE de una sola ejecución. Se observa que el valor promedio de las aptitudes mejora. La peor aptitud es cuando existe un individuo con aptitud 0, es decir que ninguna UAV fue seleccionada y por lo tanto no se asocia ninguna MS.

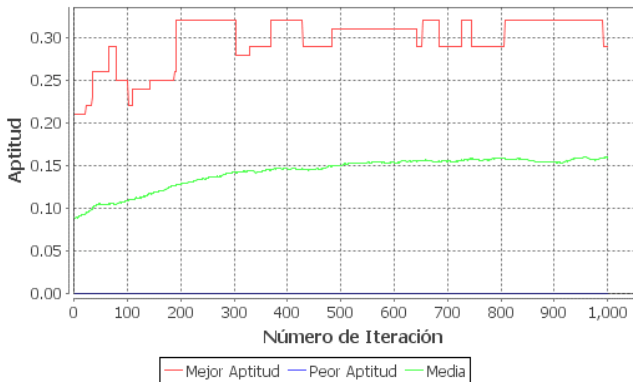


Figura 3. Aptitud vs. Número de iteraciones

La distribución óptima de los UAV's encontrada después de la simulación se muestra en la Figura 4, también se aprecia a los MS que se les brinda servicio.

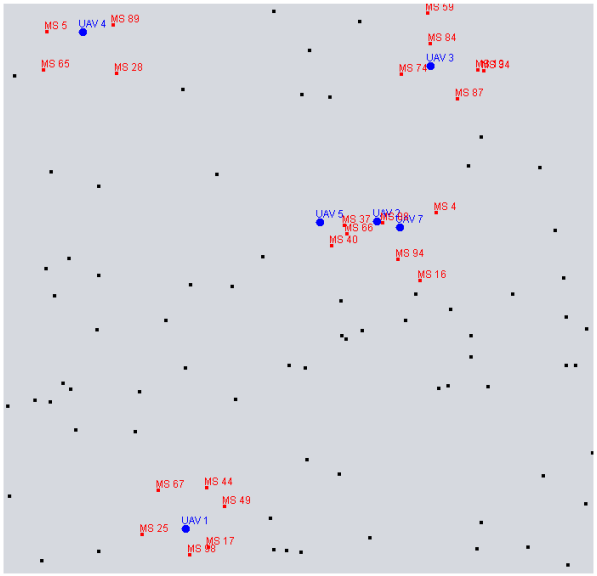


Figura 4. Distribución óptima de los UAV's.

Se realizan diez ejecuciones del algoritmo propuesto, en la Tabla 1 se observa los individuos con mejor aptitud obtenidos en cada ejecución. El porcentaje máximo de cobertura se obtuvo en la ejecución 3 en el individuo 95, ofreciendo la mayor cobertura con el 34%.

Tabla 1. Mejores encontrados

Ejecución	Mejor Aptitud	Individuo	Cobertura
1	0.32	59	32.0
2	0.31	73	31.0
3	0.34	95	34.0
4	0.33	90	33.0
5	0.31	88	31.0
6	0.33	7	33.0
7	0.33	96	33.0
8	0.34	21	34.0
9	0.32	49	32.0
10	0.32	23	32.0

5. CONCLUSIÓN

En este trabajo se utiliza un algoritmo genético con un modelo de población estacionario. Esta variante es menos costosa computacionalmente que la versión del algoritmo genético simple. De los resultados obtenidos se observa que al considerar solo las pérdidas por trayectoria, un MS puede quedar asociado a más de una UAV.

Por lo anterior, además de la pérdida por trayectoria, como trabajo futuro se pretende implementar otras restricciones como el radio de cobertura y la relación señal a interferencia. La relación señal a interferencia garantizará al MS asociado a una UAV calidad de servicio. Por otro lado, el radio de cobertura permitirá resolver el conflicto de asociación de UAV's. Con la integración de estas nuevas restricciones, se espera una mejora en la calidad de las soluciones.

6. REFERENCIAS

[1] Brizuela, C.A. and Gutiérrez, E. 2003. An experimental comparison of two different encoding schemes for the location of base stations in cellular networks. *Workshops on Applications of Evolutionary Computation* (2003), 176–186.

[2] Choi, Y., Kim, K. and Kim, N. 2008. The Displacement of Base Station in Mobile Communication with Genetic Approach. *EURASIP Journal on Wireless Communications and Networking*. 2008, 1 (2008), 580761.

[3] Eiben-Smith *Introduction to evolutionary computing*. Springer.

[4] Merwaday, A. and Guvenc, I. 2015. UAV assisted heterogeneous networks for public safety communications. *Wireless Communications and Networking Conference Workshops (WCNCW), 2015 IEEE* (2015), 329–334.

[5] Miller, R. 2006. *Hurricane Katrina: Communications & Infrastructure Impacts*. DTIC Document.

[6] National Broadband Plan - Chapter 16: Public Safety: *http://www.broadband.gov/plan/16-public-safety/*. Accessed: 2016-08-25.

[7] Park, B.-S., Yook, J.-G. and Park, H.-K. 2002. The determination of base station placement and transmit power in an inhomogeneous traffic distribution for radio network planning. *Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall. 2002 IEEE 56th* (2002), 2051–2055.

[8] .:Portal SCT:.: A un año de Odile, Baja California sur al 100 por ciento de sus capacidades: *http://www.sct.gob.mx/despliega-noticias/article/a-un-ano-de-odile-baja-california-sur-al-100-por-ciento-de-sus-capacidades/*. Accessed: 2016-08-25.

[9] Siddique, N. and Adeli, H. 2015. Nature Inspired Computing: An Overview and Some Future Directions. *Cognitive Computation*. 7, 6 (Dec. 2015), 706–714.

[10] Simon Haykin and Michael Moher *Modern Wireless Communications*. Prentice Hall.

Analysis of Taxi Destination Prediction Using a Neural Network

Normando Alí Zubia Hernández
Facultad de Ingeniería, UACh
Circuito Universitario,Chihuahua, Mx
31125
nzubiahdz@gmail.com

Luis Carlos González Gurrola
Facultad de Ingeniería, UACh
Circuito Universitario,Chihuahua, Mx
31125
lcgonzalez@uach.mx

ABSTRACT

The enormous amount of data generated nowadays opens new and challenging opportunities to improve common and daily services. Particularly, for taxi services there is an increasing interest to predict the final destination of any given user. This information could be exploited to better attend the demand of taxis, since knowing in advance where a trip may end, the dispatcher could arrange a new user service around. To tackle this problem, in this paper we used a database containing GPS coordinates of thousands of taxi trips in the city of Porto, Portugal. From this information, we are interested in detecting patterns in the very first GPS coordinates, and from that point to predict where the trip may end. An Artificial Neural Network (ANN) was implemented to perform this task, yielding an mean error of about 1.8 Km between the predicted final destination and the real one. Furthermore, an analysis of possible causes of error offers new insights about what could be an interesting path to follow, this being the combination of the former ANN and knowledge of taxi trips from individual users.

Keywords

ANN; trajectory mining; destination prediction; next location prediction

1. INTRODUCTION

One of the most common, and larger, trending initiatives in very diverse areas such as medicine, engineering and sociology, among others, is the use and exploitation of the largest datasets ever created. This tremendous amount of data has been known by the catchy name of *Big Data*. Big Data, in turn, has been generated by the participation of thousands, or even millions, of agents, that given some data collection properties, have been able to gather, acquire, measure, sense, calculate and share bits of data, altogether making possible to inferring valuable knowledge.

Specifically, in the transportation field, there is a great opportunity to exploit information that has been shared via Global Positioning System (GPS) coordinates. Taxi services, particularly, has been one of the Killer Apps nowadays, take for instance, applications like Uber or Cabify, just to mention a couple. One of the main problems lies in the city of Porto, Portugal, where the VHF-radio dispatch systems were replaced by an electronic dispatch system. With an electronic dispatcher system is easy to see where the taxi has been but not necessarily where it is going. Added to this the dispatcher's work it became harder because the broadcast-based (one to many) radio messages for service dispatching were replaced by unicast-based(one to one) messages. With unicast-messages, the dispatcher needs correctly identify which taxi they should dispatch but since taxis uses electronic dispatch is extremely difficult for dispatcher to know which taxi contact. If GPS coordinates of taxi trips are available, it would be very valuable to infer, based on the first GPS coordinates where the trip would end. With this information in hand, the taxi dispatcher would have the possibility of better manage and arrange immediate taxi services calls.

In this work, we approach the problem of analyzing the use of an Artificial Neural Network (ANN) to identify patterns in a database of taxi trips. To this end, we use a dataset of 1.5 million trips, in the city of Porto, Portugal. We test an ANN to predict the final destination of a trip, based on the very first GPS coordinates of the trajectory been taken. This problem, and the dataset, were proposed in a Kaggle[1] competition, where hundreds of teams participated worldwide. However, as shown in this competition, the most promising ways of tackling this problem relied on specific computational infrastructure, requiring the use of GPUs (Graphical Processing Unit), many of the times not available for the taxi dispatcher. So, the route that we follow in this work is to test and analyze a simple ANN, running on a regular computer, to identify final destinations from a bunch of taxi trips. Our results suggest that, although the problem is challenging, a regular ANN could very well tackle the problem, offering a mean error of about 2 Km between the prediction and the real destination. Furthermore, we also investigate what are the regions within the city that present the greatest challenge for accurate predictions, using this information, could represent a potential reduction in the error rate. Finally, we tackle a simpler problem, where we focus on detecting transportation patterns of individual users, which could be used to enhance the former model.

2. RELATED WORK

Reported in literature, there are some projects that present approaches to predict short vehicle trajectories, in this way, Wiest et al. [11] used a Kalman Filter to predict possible trajectories of vehicles when they are near to a crossroad in order to prevent a car crashing. Other works that have been proposed try to predict the complete route of a vehicle [6, 10]. In Karbassi and Barth [6], route prediction for smart

[1]Website that gathers Machine Learning enthusiast to tackle problems and challenges motivated in real world application. For further reference go to https://www.kaggle.com/



Figure 1: Results after removing spurious coordinates in a given trajectory. Left panel: A trajectory before depuration. Right panel: A trajectory after depuration.

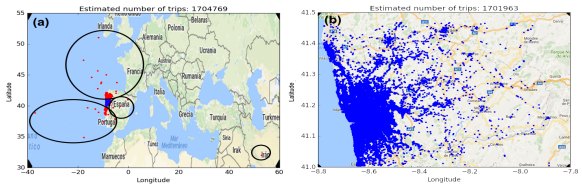


Figure 2: Left panel: location of some spurious reads. Right panel: GPS coordinates in the Porto city area.

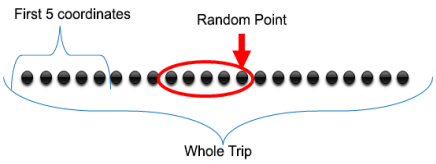


Figure 3: Election strategy for coordinates into trajectory.

vehicle was implemented. In the work of Torkkola *et al.* [10] they train a classifier to learn destinations and routes from GPS data.

Some other works employ Hidden Markov Models (HMM) to perform trajectory mining [1, 3]. In Alvarez *et al.* [1] an HMM was used to analyze the user activities from his past GPS logs to predict his next location in a certain time. In Chen *et al.* [3] three models were implemented: an individual, a general and a regional model, afterward, they were used in combination to improve the accuracy of destination prediction.

In [7], Lam *et al.* implemented an approach based on ensemble learning to predict the taxi destination and trip time from partial trajectories. Also, Brebisson *et al.* [5] implemented several kinds of ANN to predict the final destination of a trip from GPS trajectories and user metadata. This previous work won the 2015 ECML/PKDD discovery challenge on taxi destination prediction organized by Kaggle. The approach and the analysis made in this document is motivated from this work.

3. EXPERIMENTAL SETTINGS

The dataset that we employ contains approximately 1.7 million trips made by 442 taxis in the city of Porto, Portugal between July 2013 up to June 2014. The dataset is kept in csv format and it is composed by the following attributes (besides the GPS coordinates): TAXI_ID: Unique identifier for the taxi driver; ORIGIN_CALL:Unique identifier for the client; ORIGIN_STAND:Unique identifier for the taxi stand (dispatcher); TIMESTAMP:Time of Trips' start; POLYLINE.- Trip coordinates (GPS readings).

3.1 Preprocessing

After the initial dataset analysis, some discrepancies were detected on the trajectories. For example, there were some trips with illogical (and of course, spurious) GPS coordinates, that if not taken away would have added noise, resulting in false patterns or larger errors in the algorithm. To correct this issue an strategy was implemented to remove all the coordinates that were more than 5 km off the trajectory (see Figure 1). Following this strategy, we concentrate on coordinates that showed a logical order in time, so all the coordinates outside the following area were also removed: Maximum Latitude: 41.470679; Minimum Latitude: 41.00394; Maximum Longitude: -7.882110; Minimum Longitude: -8.777495.

This area represents the city of Porto, and all the trips outside this area were considered spurious readings. Figure 2 (left panel) shows that most of the eliminated trips were on other countries or even inside the ocean.

3.2 Artificial Neural Network (ANN)

An Artificial Neural Network models the relationship between inputs and outputs finding patterns inside the data, it consists on multiple layers: an input layer that receives a fixed-size feature vector, one or multiple hidden layers and an output layer. Each one of these layers receives and processes the information to return the result of the computation of the input information [8]. In this work, the Artificial Neural Network is composed by 3 layers: an input layer, a hidden layer with 500 neurons and an output layer. This model was selected based on a preliminary empirical evaluation.

Since the ANN takes as an input a feature vector with a fixed length, and most of the current trajectories have a different number of GPS coordinates, there is a need to extract the features that would make the input vector. In [5], two fixed segments of the trip were concatenated to get the feature vector, the segments represented the very first 5 coordinates and the last five coordinates taken from the middle point in the trajectory, therefore using a feature vector of dimension 10. In this work, we followed the same strategy (but selecting 20 coordinates), but instead of using the middle point as a reference to extract the last 10 coordinates, a random selected point is used (see Figure 3).

The trajectory coordinates plus the information of the taxi and client, results on an input layer of dimension 26. Table 1 presents the characteristics of the meta data that is employed. Once the feature vector is built, a process to standardize the data was followed, based on the next equation:

$$value = \frac{value - Mean}{MaxValueInColumn}$$

To define the dimension of the output layer, we run the mean-shift algorithm [4] to detect meaningful clusters of final destinations, resulting in 3310 clusters. Figure 4 shows the distributions of these clusters over the city of Porto. The

Table 1: Values amount for each feature

Feature	Amount
Taxi_id	448
Origin_call	57,106
Origin_stand	64
Day of the week	7
Week of the year	52
Quarter hour of the day	96

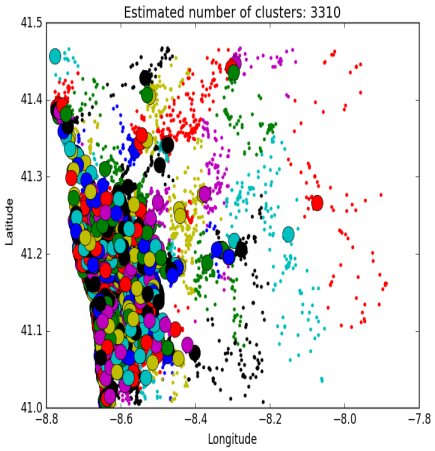


Figure 4: Clusters obtained by mean-shift algorithm. The clusters were represented with circles and the destinations were represented by different color dots depending on the group to which it belongs.

output of the ANN is a vector with the probabilities of each cluster being the final destination for a given input. With this output, a dot product was calculated with the vector of clusters coordinates, thus, the final result is the most probable coordinate of the trip destination.

The training algorithm was the stochastic gradient descent which is one of the most widely used methods for iterative minimization of a differentiable cost function, its objective is minimize the error between the real output and the predicted one through several iterations, in [9] can be seen a deeper explanation about the algorithm. The function cost was the distance between the real destinations and predicted ones. The model uses a momentum of 0.9, a learning rate of 0.01 and batch size of 200.

4. RESULTS

The model was executed 10 times in order to get an average performance. We split the original dataset in 70% and 30%, for testing and training purposes, respectively. The evaluation of the model, as mentioned above, is the measure in Km between the real destination point and the predicted point by the ANN. The model got a mean error of 1.80 Km with a standard deviation of 2.30, the maximum and minimum error was 75 Km and 0 Km respectively.

4.1 Difficulties on Trajectory Prediction

To gather some insight at possible phenomena that could be impacting the performance of ANN, we made two types of

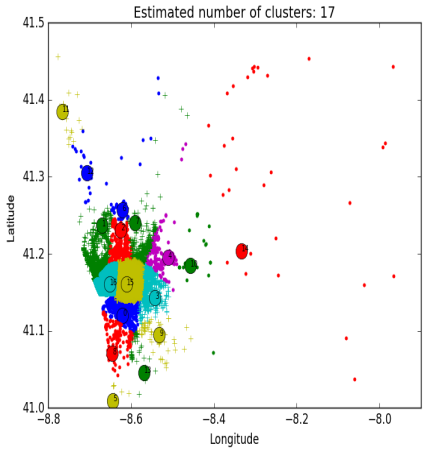


Figure 5: City of Porto Portugal divided in 17 clusters

Table 2: First 5 pairs of origin-destination clusters with the largest error.

Beginning	End	Mean Error	Influence
15	15	1.030 Km	28%
15	1	5.28 Km	15%
15	16	2.19 Km	11%
16	15	1.82 Km	9%
15	3	2.46 Km	4%

analysis. We were interested to investigate what were some of the circumstances that possibly caused a large error, but also, what could be viable ways to improve the accuracy of the model.

For the first analysis, we focused on finding what pair of origin-destination clusters resulted hard to predict, thus causing a large error. We divide Porto in 17 sections (Figure 5), the amount of section was selected after several cluster analysis and result in the minimum number of sections needed to identify the causes of the large error and facilitate the analysis of patterns inside the trips. After the city division, we generate pairs of sections (beginning and end of a trip), the mean error of all the trips contained in each pair is calculated. To represent the percent of error contribution of each section to the total error (calculated by the ANN) we used the term *influence*. This parameter could be used to know which pair of sections are more susceptible to improve, therefore, reducing the total error of the ANN. The influence of each section consists in the multiplication of the mean error by the amount of trips in that section, to get the percent of contribution we divide each value by the sum of that column:

$$Influence = MeanError * \#Trips$$

$$\%ofinfluence = \frac{influence}{TotalInfluence}$$

The trips with more influence in the general error can be seen in Table 2.

As it can be appreciated, trips where the beginning sec-

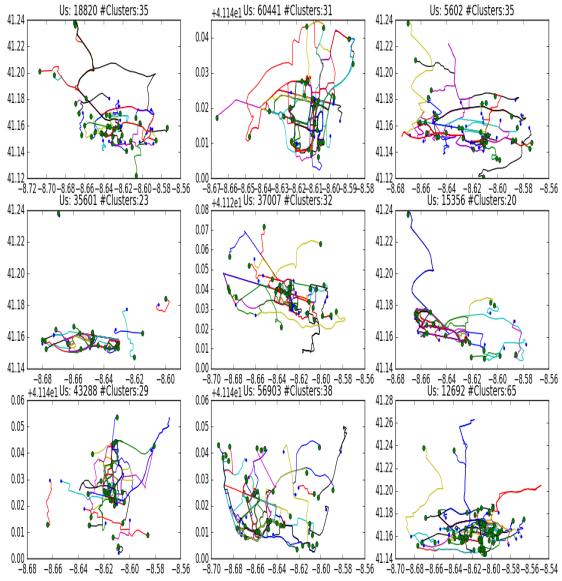


Figure 6: Trajectories of the chosen users for individual analysis.

tion is 15 and their corresponding ending section is 1, are the pairs with more potential of improvement. Analyzing a bit deeper these trips, we found that a highway is the cause of that error, most trips from section 15 to section 1 needs to use the highway to get to their destination, resulting in several similar trajectories. For the second analysis, we made a new model that receives trips of only one user as input (remember we have Origin Call attribute), then the goal is to know if with this information we could predict more accurately the final destination. Note that some works [3, 1] present an hybrid approach using several models to recognize different patterns showing a better accuracy than using each model separately. We analyze several users, and selected 9 of them to run this experiment (see Figure 6). The user's selection was made using the *inertia* parameter of the k-means algorithm [2], that parameter indicates how close the destinations inside the cluster are. This parameter was used to automatize the process of user selection. The algorithm choose 9 users with lower inertia (Figure 6)

An ANN was created for each user, these ANNs are only different by the number of clusters in the output layer. Each neural network was trained and evaluated 10 times. To put in perspective this analysis, Table 3 presents the mean error between the general model (former) and the individual model (latest).

As it can be seen, the use of individual patterns can be very useful to add valuable information to the model and to improve the accuracy, results are promising and it could be a good approach to try.

5. CONCLUSIONS

Trajectory mining is a problem with many challenges and opportunities. In this work an ANN was applied to an extensive dataset to predict final destinations. Results suggest,

Table 3: Comparison between general model and individual model.

User	General Model	Individual Model
12692	1.50 Km	1.77 Km
56903	2.42 Km	1.55 Km
43288	0.96 Km	0.92 Km
15356	1.07 Km	0.79 Km
37007	1.75 Km	0.65 Km
35601	2.36 Km	0.73 Km
5602	0.81 Km	1.64 Km
60441	0.85 Km	1.07 Km
18820	2.18 Km	1.66 Km

that this model could be potentially used, but several improvements need be done. One problem that was found in the results analysis is the similarity of trajectories caused by transportation infrastructure (e.g., highways), it is evident that some strategies need be implemented to deal with issues of this kind. One hypothesis that results from this analysis is the need to use classifiers that model individual patterns, which could improve the accuracy of a more general model.

6. REFERENCES

[1] J. Alvarez-Lozano, J. A. García-Macías, and E. Chávez. Crowd location forecasting at points of interest. *Int. J. Ad Hoc Ubiquitous Comput.*, 18(4):191–204, Apr. 2015.

[2] P. Berkhin. Survey of clustering data mining techniques. Technical report, 2002.

[3] M. Chen, X. Yu, and Y. Liu. Mining moving patterns for predicting next location. *Information Systems*, 54:156–168, 2015.

[4] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):603–619, May 2002.

[5] A. de Brébisson, É. Simon, A. Auvolat, P. Vincent, and Y. Bengio. Artificial Neural Networks Applied to Taxi Destination Prediction. *arXiv Preprint*, pages 1–12, 2015.

[6] A. Karbassi and M. Barth. Vehicle route prediction and time of arrival estimation techniques for improved transportation system management. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pages 511–516, June 2003.

[7] H. T. Lam, E. Diaz-Aviles, A. Pascale, Y. Gkoufas, and B. Chen. (Blue) Taxi Destination and Trip Time Prediction from Partial Trajectories. 2015.

[8] S. Marsland. *Machine Learning: An Algorithmic Perspective*. Chapman & Hall/CRC, 1st edition, 2009.

[9] S. Theodoridis. *Machine Learning: A Bayesian and Optimization Perspective*. .Net Developers. Academic Press, 1 edition, 2015.

[10] K. Torkkola, K. Zhang, H. Li, H. Zhang, C. Schreiner, and M. Gardner. Traffic advisories based on route prediction. In *Proceedings of Workshop on Mobile Interaction with the Real World*, number September, pages 33–36, 2007.

[11] J. Wiest, M. HÃűffken, U. KreÃ§el, and K. Dietmayer. Probabilistic trajectory prediction with gaussian mixture models. In *Intelligent Vehicles Symposium (IV), 2012 IEEE*, pages 141–146, June 2012.



Taller:
Computación Clínica e Informática Médica

Organizadores:
Ana I. Martínez García (CICESE)
Marcela D. Rodríguez (UABC)
Fernando Arámbula Cosío (CCADET, UNAM)
Alicia Martínez Rebollar (CENIDET)
Leo Joskowicz (Universidad Hebrea de Jerusalén)

LaborCheck: Sistema de Monitoreo Automático de Variables Usuario-Computadora y de Interacción Social

Misael Camaños, Wendy Sánchez, Alicia Martínez, Manuel Mejía-Lavalle
Centro Nacional de Investigación y Desarrollo Tecnológico
Interior Internado Palmira S/N, 62490 Cuernavaca, Morelos
{misael.camanos, wendy.sanchez11c, amartinez, mlavalle}@cenidet.edu.mx

RESUMEN
El estrés laboral se presenta cuando un trabajador percibe que las exigencias laborales superan sus capacidades para hacerles frente o mantenerlas bajo control. Las personas expuestas continuamente al estrés, están propensas a padecer enfermedades graves tanto físicas como psicológicas. Nuestra investigación se enfoca en el reconocimiento de episodios de estrés mediante el análisis del comportamiento del trabajador en su ambiente laboral. En este artículo se presenta LaborCheck, un sistema de monitoreo capaz de recolectar y almacenar automáticamente la información de variables relacionadas con las interacciones que se generan por medio del uso del mouse, teclado, interacciones con las ventanas de Windows, así como también con la detección del uso de tecnologías de comunicación por medio de los mensajes y llamadas. El objetivo de esta investigación es almacenar automáticamente información de variables usuario-computadora y de interacción social que puedan ser utilizadas para la detección de estrés laboral.

Palabras clave
Monitoreo automático, teclado, mouse, ventanas de Windows, mensajes, llamadas, interacciones usuario-computadora, interacción social.

1. INTRODUCCIÓN
En muchos países, el número de personas que sufren de estrés laboral se ha incrementado como consecuencia del ritmo de vida acelerado [1]. Es tanto el impacto de esta enfermedad psicológica que la OMS la reconoce como una de las grandes epidemias de la vida moderna [1].

El estrés laboral se presenta cuando un trabajador percibe que las exigencias laborales superan sus capacidades para hacerles frente o mantenerlas bajo control [2], [3].

El estrés laboral está asociado a diversos problemas en la salud física y mental. Las personas que están continuamente expuestas al estrés están propensas a padecer enfermedades cardiovasculares, respiratorias, inmunológicas, gastrointestinales, trastornos musculoesqueléticos, depresión, cáncer, entre otros [4], [5], [6]. Además, propicia hábitos negativos como fumar, beber, mala alimentación o falta de sueño [7]. El estrés crónico provoca ansiedad, tensión, baja autoestima, aislamiento social, irritabilidad y depresión [8], [9].

La detección oportuna del estrés puede ayudar a reducir la frustración y a prevenir los consecuentes problemas de salud en los trabajadores. Nuestra investigación se enfoca en el reconocimiento de episodios de estrés mediante el análisis del comportamiento del trabajador en su ambiente laboral.

El objetivo de este artículo es presentar LaborCheck, un sistema de monitoreo capaz de recolectar y almacenar automáticamente la información de variables usuario-computadora y variables de interacción social.

El artículo está ordenado de la siguiente manera: la sección 2 presenta los trabajos relacionados de nuestra investigación; la sección 3 presenta el sistema de monitoreo LaborCheck; la sección 4 presenta las pruebas y resultados; y finalmente, la sección 5 presenta las conclusiones.

2. TRABAJOS RELACIONADOS
En el área de la computación, muchas investigaciones se han enfocado en el reconocimiento del estrés mediante la medición de factores fisiológicos del ser humano, como la frecuencia cardiaca, la temperatura corporal, la conductancia galvánica de la piel, la respiración [10], [11], [4], [12], [13], [14], [15] y [16], la actividad ambulatoria y el cortisol subcutáneo [15]. Estas mediciones se realizan mediante sensores corporales que hacen contacto con la piel de la persona.
Otras investigaciones han utilizado los Smartphones para el reconocimiento del estrés, registrando las llamadas, mensajes y correos electrónicos [9] y [17], el tiempo de uso de las aplicaciones, la ubicación, las veces que enciende o que apaga la pantalla [13] y las interacciones sociales mediante el Bluetooth embebido [9].

Algunas otras investigaciones se han enfocado en la medición de las interacciones humano-computadora para detectar el estrés. Estas interacciones se refieren al cambio entre ventanas, las páginas web visitadas y el tiempo invertido en estas actividades, así como las acciones en el teclado y del mouse [18]–[20].

También se ha detectado el estrés a través de la medición de la fuerza de presión de ciertos objetos como el teclado [18], el mouse [12], la silla de oficina [21] y el volante de un automóvil [16]. Aunado a estas variables de medición, algunos de estas investigaciones han añadido mediciones de la postura de la persona, de la temperatura ambiental [11] y del clima en la ciudad [9].

Nuestra propuesta se enfoca en la detección del estrés que se presenta en el ámbito laboral, específicamente, a trabajadores que realicen su labor frente a una computadora. Por lo cual, la información que se recolecta es acerca de las interacciones del trabajador con su computadora y con su celular. LaborCheck recolecta y almacena automáticamente la información de las variables para su posterior análisis.

3. SISTEMA LABORCHECK
En esta sección se incluye el contexto general del sistema LaborCheck. La Figura 1 muestra la estructura genera del sistema, el cual permite monitorear de forma automática las interacciones que generan los usuarios con el uso de una computadora con el

sistema operativo Windows y un Smartphone con sistema operativo Android. La información recolectada por este sistema puede ser consultada por medio de una aplicación web a través de un navegador web como Google Chrome, Internet Explorer o Mozilla Firefox.

Este sistema permite recabar información que será utilizada para el reconocimiento de episodios de estrés laboral.

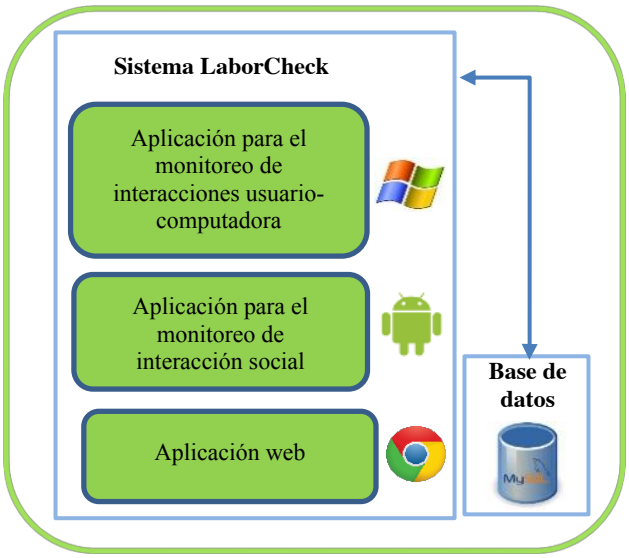


Figura 1. Contexto general del Sistema LaborCheck.

El sistema LaborCheck es un sistema distribuido que permite la interoperabilidad entre las aplicaciones desarrolladas, basado en el modelo cliente-servidor. El uso de Hardware (mouse y teclado) y Software (interacciones entre ventanas de Windows) se ha incluido en la aplicación de interacciones usuario-computadora y el uso de las tecnologías de comunicación (mensajes y llamadas) en la aplicación de interacción social. Este sistema almacena la información recolectada por las aplicaciones de monitoreo en una base de datos gestionada por medio del sistema manejador de base de datos MySQL.

3.1 Aplicación para el monitoreo de interacciones usuario-computadora

En esta sección se describe la aplicación que recolecta y almacena las interacciones usuario-computadora. La interfaz gráfica que está disponible para el usuario se oculta en la barra de notificaciones de la barra de tareas, con el objetivo de que no sea intrusiva para el usuario, como se muestra en la figura 2.

Esta aplicación permite monitorear el comportamiento que presentan los trabajadores al realizar sus tareas diarias. Este sistema permite detectar las variables mostradas en la Tabla 1. Por medio del uso generado sobre estas tecnologías se podrá determinar si los empleados que presentan estrés laboral aumentan o disminuyen la cantidad de interacciones que se están monitoreando.

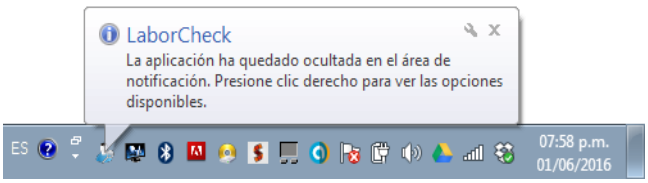


Figura 2. Aplicación de interacciones usuario-computadora.

Esta aplicación se desarrolló para ser instalada en el sistema operativo Windows y es compatible con las versiones Windows 7, Windows 8 y Windows 10.

Esta aplicación almacena localmente el registro de usuario para que la próxima vez que reinicie el sistema operativo Windows, la recolección de información inicie automáticamente sin solicitar los datos de autenticación.

Detección de interacciones del mouse. Este módulo recupera las interacciones generadas por el mouse. La clase GlobalMouseHook se programó para consumir los métodos de la librería Hook de Windows con el lenguaje de programación C#. La información monitoreada en este módulo se almacena temporalmente en una base de datos local y posteriormente se envía a la base de datos remota.

Detección de interacciones del teclado. Este módulo recupera las interacciones del teclado. La clase GlobalKeyBoardHook se programó para consumir los métodos relacionados con el teclado incluidos en la Liberia Hook de Windows con el lenguaje de programación C#. La información recuperada en este módulo se almacena temporalmente en una base de datos local y posteriormente es enviada a la base de datos remota.

Detección de interacciones de ventanas de Windows. Este módulo recupera las interacciones entre ventanas de Windows. La clase GlobalWindowsHook se programó para consumir los métodos de la librería Hook que detectan los eventos de las ventanas de Windows con el lenguaje de programación C#. La información recuperada en este módulo se almacena temporalmente en una base de datos local y posteriormente es enviadas a la base de datos remota.

Las interacciones recuperadas se almacenan localmente para no perder información cuando existen problemas en las conexiones de internet, al enviar la información hacia la base de datos remota. La base de datos local utiliza el sistema gestor de base de datos SQLite. En esta base de datos se almacenan las interacciones usuario-computadora cada cinco minutos. La base de datos local se ubica en la carpeta del disco duro del sistema operativo Windows creada para almacenar los datos de aplicación del usuario.

Tabla 1. Variables de interacción usuario-computadora

Tecnología	Variables
Mouse	Cantidad total de clics
	Cantidad de clics izquierdos
	Cantidad de clics derechos
	Cantidad de desplazamientos generados hacia arriba con la rueda del mouse
	Cantidad de desplazamientos generados hacia abajo con la rueda del mouse
	Cantidad de pixeles recorridos
Teclado	Cantidad total de teclas presionadas
	Cantidad de tecleas tipo retroceso
	Cantidad de teclas tipo eliminar
	Cantidad de teclas tipo enter
	Cantidad de tecleas tipo retroceso
Ventanas de Windows	Cantidad de ventanas abiertas
	Cantidad de cambios entre ventanas

3.2 Aplicación de monitoreo para interacción social

En esta sección se describe la aplicación que detecta la interacción generada por los usuarios a través de las tecnologías de comunicación social: mensajes y llamadas. La aplicación que se instala en el dispositivo móvil se muestra en la Figura 3.

Esta aplicación permite monitorear el comportamiento social de los empleados. Esta aplicación monitorea las variables mostradas en la Tabla 2. Por medio del uso de esta aplicación será posible encontrar comportamientos asociados al contabilizar los mensajes y llamadas de los empleados que presentan estrés laboral.



Figura 3. Aplicación de interacción social.

Esta aplicación recupera y envía automáticamente a un base de datos remota la información de las variables recuperadas. La aplicación de monitoreo social se ha desarrollado para ser instalada en el sistema operativo Android. Esta aplicación ha sido desarrollada tomando en cuenta el modelo cliente servidor.

Esta aplicación almacena localmente el registro de usuario para que la próxima vez que reinicie su dispositivo móvil la aplicación inicie automáticamente sin solicitar datos de autenticación.

Módulo de contactos. Este módulo categoriza los contactos como: contactos laborales o contactos personales. Permite agregar, eliminar o editar información de los contactos. En este módulo se integra una opción para que el usuario pueda agregar un nuevo registro utilizando los contactos de la agenda de su Smartphone o capturando su nombre y número de teléfono sobre una caja de texto. Los cambios generados en este módulo son reflejados en una base de datos local que se almacena en el dispositivo móvil.

Detección de mensajes. Este módulo detecta las variables de los mensajes que se incluyen en la tabla 2. La instancia SMSObserver, que está incluida en Android Studio, se utiliza parar detectar los cambios generados en la carpeta de mensajes del dispositivo móvil.

Detección de llamadas. Este módulo detecta las variables de llamadas que se incluyen en la tabla 2. La clase CALLSObserver, que está incluida en Android Studio, se utiliza para detectar los cambios generados en la carpeta de registro de llamadas del dispositivo móvil.

La base de datos que se encuentra almacenada de manera local en el dispositivo móvil utiliza el sistema gestor de base de datos SQLite. La información de interacción social que se recupera en los módulos de monitoreo se guarda cada dos horas. Esta información se almacena de manera local para evitar pérdida de información cuando no puede ser enviada al servidor por problemas de conexión a internet. Sin embargo, una vez que se tiene conexión, la información se envía a la base de datos remota.

Tabla 2. Variables de interacción social

Tecnología	Variables
Mensajes	Cantidad de mensajes laborales recibidos
	Cantidad de mensajes personales recibidos
	Cantidad de mensajes laborales enviados
	Cantidad de mensajes personales enviados
Llamadas	Cantidad de llamadas laborales recibidas
	Cantidad de llamadas personales recibidas
	Cantidad de llamadas laborales realizadas
	Cantidad de llamadas personales realizadas
	Cantidad de llamadas laborales perdidas
	Cantidad de llamadas personales perdidas

3.3 Aplicación web

En esta sección se describe la aplicación web que permite que los empleados consulten la información recolectada por la aplicación de interacciones usuario-computadora y la aplicación de interacción social. La pantalla de inicio de la aplicación web se muestra en la figura 4. La información se muestra por medio de gráficas de series que incluyen la cantidad de información recolectada cada cierto periodo de tiempo.

Esta aplicación permite analizar de manera gráfica el uso de las tecnologías que están siendo monitoreadas en tiempo real. Se puede observar la interacción de los empleados que presentan estrés laboral y compararla con empleados mentalmente saludables por medio de los periodos en donde registran menor o mayor actividad computacional.

Módulo de consulta: Este módulo se desarrolló para recuperar la información de las interacciones usuario-computadora y de interacción social almacenada en la base de datos remota. Este módulo incluye un campo fecha que permite filtrar la información de fechas pasadas y actuales. El Framework de programación de PHP Yii2 se utilizó para programar los métodos que recuperan la información de las variables recolectadas.

La información que se recupera en este módulo se muestra en gráficas de series que son intuitivas para el usuario. Estas gráficas se desarrollaron usando la librería interactiva Highcharts.



Figura 4. Aplicación

3.4 Base de datos

En esta sección se describen las características de la base de datos remota que se diseñó para almacenar las interacciones usuario-computadora y de interacción social.

La base de datos incluye una tabla para almacenar cada una de las tecnologías que son monitoreadas y una tabla para identificar al usuario que se está monitoreando.

Por medio de esta base de datos se puede filtrar la información necesaria que permita detectar el comportamiento de los empleados que presentan episodios de estrés laborales como puede ser: cantidad de errores que comenten los usuarios durante el uso del teclado por medio de las teclas suprimir y retroceso, la cantidad de espacios o pulsaciones excesivas con la tecla enter, el tiempo exacto en que se recibe un mayor número de llamadas o mensajes, etc.

La base de datos sólo almacena información relacionada con el usuario para fines de investigación, no almacena información privada como correos, teléfonos etc. La identidad del usuario se mantiene segura debido a que sólo él conoce su nombre de usuario y contraseña. La base de datos se gestiona por medio del sistema gestor de base de datos MYSQL.

4. PRUEBAS Y RESULTADOS

En esta sección se presentan las pruebas de las aplicaciones que monitorean las interacciones usuario-computadora, las variables de interacciones social y la aplicación web que presenta la información recolectada por las aplicaciones de monitoreo. En estas pruebas participaron seis trabajadores con edades entre 25-35 años, tres mujeres y tres hombres.

Las pruebas funcionales se llevaron a cabo siguiendo la estrategia de pruebas del estándar IEEE 829-2008, en la cual se especificó la descripción de los casos de prueba, el enfoque general de las pruebas, los requisitos ambientales de Hardware y Software y los criterios de suspensión y reanudación.

4.1 Pruebas de la aplicación de interacciones usuario-computadora

En esta sección se muestran las pruebas de la aplicación de monitoreo de interacciones usuario computadora que se realizaron después de haber instalado la aplicación en la computadora del usuario con sistema operativo compatible con las versiones Windows 7, Windows 8 o Windows 10.

Las interacciones generadas con el mouse fueron probadas solicitando a los participantes que presionaran el botón izquierdo y derecho del mouse varias veces, que desplazaran hacia arriba y hacia abajo la rueda del mouse y que movieran el cursor del mouse de un lado a otro dentro de la pantalla de su computadora. La aplicación fue capaz de detectar correctamente las interacciones generadas con el mouse como se muestra en la figura 5.

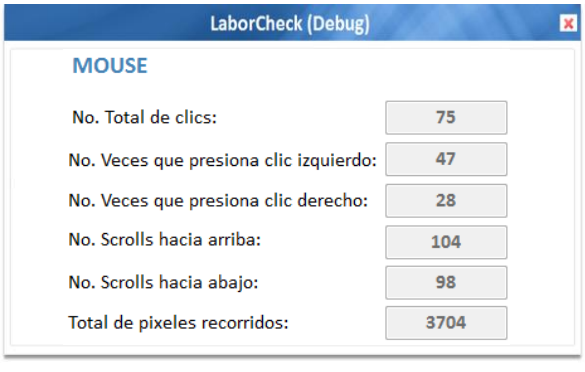


Figura 5. Pruebas con las interacciones del mouse.

Las interacciones generadas con el teclado fueron probadas solicitando a los participantes ejecutar cualquier programa editor de textos como Word, Bloc de Notas, etc. Los participantes presionaron las teclas retroceso, eliminar, enter y barra espaciadora. La aplicación detectó correctamente estas interacciones como se muestra en la figura 6.

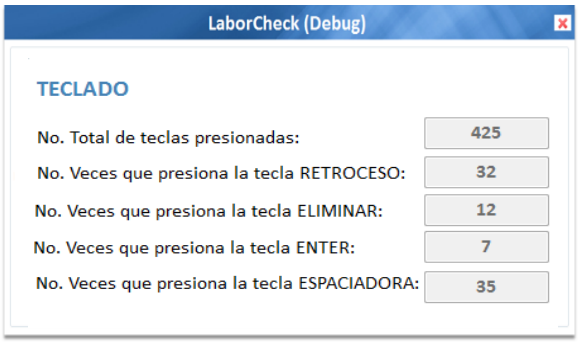


Figura 6. Pruebas con las interacciones del teclado.

Las interacciones generadas con las ventanas de Windows fueron evaluadas solicitando a los participantes abrir y cerrar diferentes ventanas del sistema operativo Windows. Los participantes ejecutaron diferentes programas como Word, Excel, PowerPoint, Calculador, Bloc de notas o carpetas del sistema. La aplicación detectó correctamente las acciones generadas por el usuario como los cambios entre ventanas (programas) y el número de ventanas abiertas, como se muestra en la figura 7.

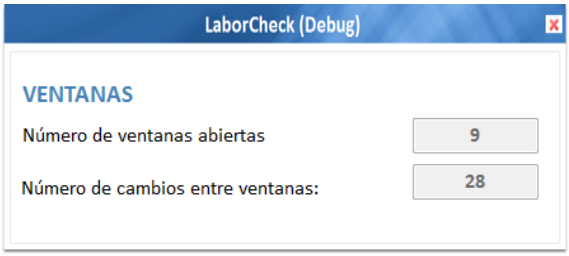


Figura 7. Pruebas con las interacciones entre ventanas de Windows.

4.2 Pruebas de la aplicación de interacción social

En esta sección se muestran las pruebas de la aplicación de monitoreo de interacción social que se realizaron después de haber instalado la aplicación en un dispositivo Android del usuario, con sistema operativo compatible con las versiones: Android 4.1.2, Android 4.3, Android 5.0.2 y Android 6.0.1.

Las pruebas funcionales del módulo de contactos se realizaron solicitando a los participantes que, después de instalar la aplicación, agregaran nuevos contactos dentro de la aplicación de interacción social, con el objetivo de clasificar sus contactos como laborales o personales como se muestra en la figura 8. Los contactos se agregaron correctamente después de capturar los datos o de seleccionar algún contacto directamente de la agenda.

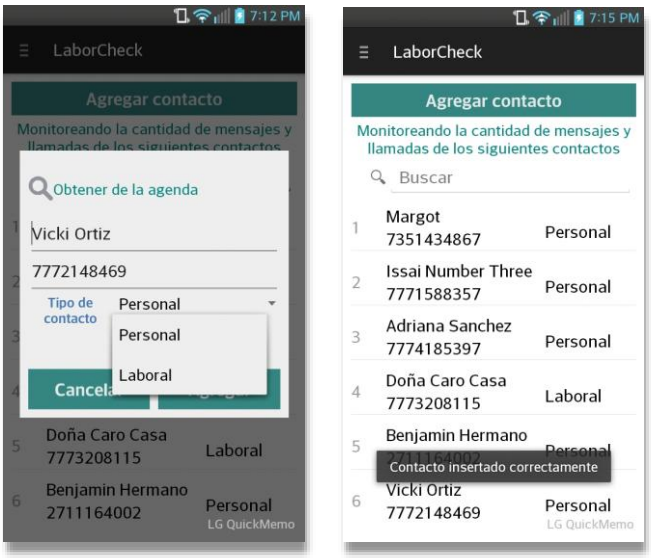


Figura 8. Pruebas con el módulo de contactos.

La interacción social se comprobó enviando y recibiendo mensajes y llamadas de prueba como se muestra en la figura 9. Estos mensajes y llamadas fueron contabilizados correctamente por la aplicación de monitoreo de interacción social. Los mensajes y llamadas se categorizaron de acuerdo al tipo de contacto que se asignó al momento de registrarlo en el Smartphone, el tipo de contacto pudo ser Personal o Laboral.

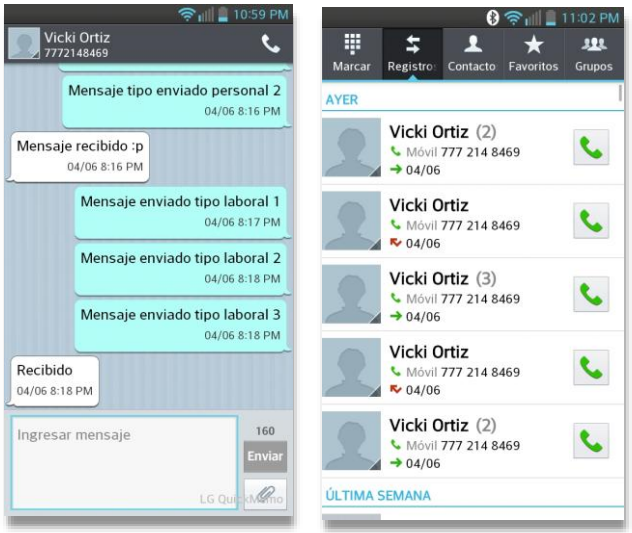


Figura 9. Pruebas con la interacción de mensajes y llamadas.

4.3 Pruebas con la aplicación web

Esta sección se muestran las pruebas realizadas a la aplicación web después de que los participantes ingresaron la dirección web en la barra de búsqueda. Los participantes utilizaron diferentes navegadores como: Google Chrome, Firefox, Internet Explorer y Android Browser.

La información de las interacciones usuario-computadora se recuperaron correctamente, como se puede observar en las gráficas de series presentadas en la figura 10 donde se muestran los valores recuperados de las interacciones generadas con el mouse y teclado, así como en la figura 11 donde se observan los valores recuperados de la base de datos remota con las interacciones generadas de las ventanas de Windows. Los usuarios pudieron observar la actividad computacional que generaron y los lapsos de tiempo en los que no registran ningún movimiento con el uso de las tecnologías monitoreadas: mouse, teclado y ventanas de Windows. La cantidad de interacciones generadas se mostró correctamente en intervalos de 5 minutos.

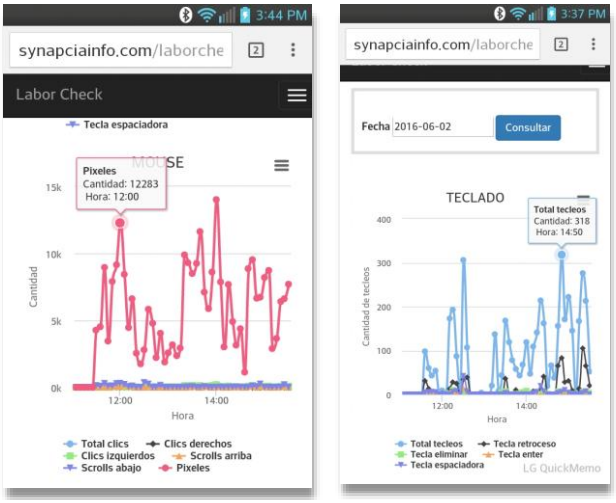


Figura 10. Gráficas que muestran la información de las interacciones usuario-computadora del mouse y teclado.

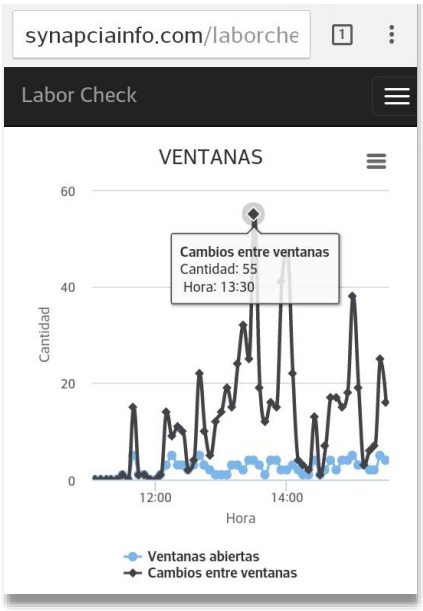


Figura 11. Gráfica que muestran la información de las interacciones usuario-computadora de ventanas de Windows.

La información de la interacción social que genera el usuario se recuperó correctamente, como se puede observar en las gráficas

de series presentadas en la figura 12. Los usuarios lograron observar la cantidad de mensajes y llamadas detectados con la página web, las cuales mostraron la cantidad correcta de mensajes y llamadas generados por los usuarios. Las variables de mensajes y llamadas se clasificaron de acuerdo al tipo contacto laboral o personal; al tipo de mensajes: recibidos y enviados; y al tipo de llamadas: recibidas, perdidas y realizadas.

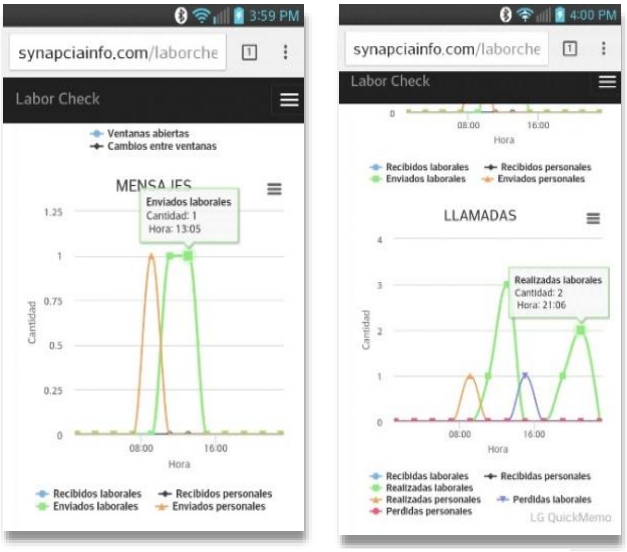


Figura 12. Gráficas que muestran la información almacenada de la interacción social.

5. CONCLUSIONES

En esta investigación se almacenaron un total de 23 variables por medio de dos aplicaciones no intrusivas. La primera aplicación recupera las interacciones usuario-computadora y se ejecuta en el sistema operativo Windows. La segunda aplicación recupera la información de la interacción social y se ejecuta en el sistema operativo Android. La información se mostró por medio de una aplicación web que puede ser consultada en cualquier navegador web. Nosotros aportamos una solución automática de recolección que detecta correctamente la cantidad de variables relacionadas con estrés laboral.

En estas aplicaciones se utilizaron las tecnologías: teclado, mouse, ventanas de Windows, mensajes y llamadas para detectar el comportamiento de los usuarios que presentan episodios de estrés laboral. Sin embargo, es posible contabilizar información de otras tecnologías incluidas como: cantidad de eventos en la agenda, cantidad de lugares que se visitan con el uso de GPS, cantidad de pestañas que se abren en los navegadores, etc.

Este sistema de monitoreo distribuido permite almacenar información de varias tecnologías en una sola base de datos. En la cual, se puede filtrar información por periodos de tiempo de forma accesible y así detectar el uso excesivo o carente con estas tecnologías que pueden revelar comportamientos en tiempo real de una gran cantidad de usuarios.

La información almacenada por medio de este sistema de monitoreo puede ser utilizada no solo para propósitos de investigación, sino que también permite que las personas comparen la actividad generada entre las tecnologías que monitorea el sistema LaborCheck y de esta forma crear hábitos saludables que reduzcan problemas de salud ocasionados por el uso excesivo de estas tecnologías.

REFERENCIAS

[1] Leka S., Griffiths A., and Cox T. 2004. La organización del trabajo y el estrés: estrategias sistemáticas de solución de problemas para empleadores, personal directivo y representantes sindicales. *Paris, Francia: Organización Mundal de la Salud*.

[2] Blaug R., Kenyon A., and Lekhi R.. 2007. *Stress at Work: A report prepared for The Work Foundation's Principal Partners*. London.

[3] Secretaría de Salud de México. 2010. *Guía sobre el manejo y prevención del estrés laboral*. México, DF.

[4] Plarre K., Raij A., Monowar S., et al. 2011. Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment. In *Information Processing in Sensor Networks (IPSN)*, 97–108.

[5] Organización Internacional del Trabajo. 2013. *La organización del trabajo y los riesgos psicosociales: una mirada de género*.

[6] Bauer ME. 2005. Stress, glucocorticoids and ageing of the immune system. *Stress* 8, 1: 69–83. http://doi.org/10.1080/10253890500100240

[7] Glanz K. and Schwartz M. 2008. Stress, coping and health behavior. In *Health behavior and health education: Theory, research, and practice* (4th ed.). Jossey-Bass, San Francisco, California, 211–236.

[8] Sysoev M., Sedlar U., Kos A., and Pogačnik M. 2014. Stress-sensors classification and stress-analysis algorithms. *Stress* 81, 5: 263–271. Retrieved from http://ev.fe.uni-lj.si/5-2014/Sysoev.pdf

[9] Bogomolov A., Lepri B., Ferron M., Pianesi F., and Pentland A. 2014. Pervasive Stress Recognition for Sustainable Living. 345–350. http://doi.org/10.1109/PerComW.2014.6815230

[10] Bakker J., Pechenizkiy M., and Sidorova N. 2011. What's Your Current Stress Level? Detection of Stress Patterns from GSR Sensor Data. *2011 IEEE 11th International Conference on Data Mining Workshops*, 1: 573–580. http://doi.org/10.1109/ICDMW.2011.178

[11] Raij A., Blitz P., Ahsan A., et al. 2010. mStress : Supporting Continuous Collection of Objective and Subjective Measures of Psychosocial Stress on Mobile Devices. In *ACM Wireless Health*.

[12] Hernandez J., Paredes P., Roseway A., and Czerwinski M. 2014. Under Pressure: Sensing Stress of Computer Users. In *ACM SIGCHI Conference on Human Factors in Computing Systems*, 51–60. http://doi.org/10.1145/2556288.2557165

[13] Sano A. and Picard R. 2013. Stress Recognition Using Wearable Sensors and Mobile Phones. In *Humaine Association Conference on Affective Computing and Intelligent Interaction*, 671–676. http://doi.org/10.1109/ACII.2013.117

[14] Hernandez J., Morris R., and Picard R. 2011. Call Center Stress Recognition with Person-Specific Models. In *Affective*

Computing and Intelligent Interaction, 125–134. http://doi.org/10.1007/978-3-642-24600-5_16

[15] Majoe D., Bonhof P., Kaegi-trachsel T., Gutknecht J., and Widmer L. 2010. Stress and Sleep Quality Estimation from a Smart Wearable Sensor. In *Pervasive Computing and Applications (ICPCA)*, 14–19. http://doi.org/10.1109/ICPCA.2010.5704068

[16] Hernandez J., Benavides X., Maes P., McDuff D., Amores J., and Picard R. 2014. AutoEmotive: Bringing Empathy to the Driving Experience to Manage Stress. In *Designing interactive systems*, 53–56. http://doi.org/10.1145/2598784.2602780

[17] Harwood J., Dooley J., Scott A., and Joiner R. 2014. Constantly connected – The effects of smart-devices on mental health. *Computers in Human Behavior* 34: 267–272. http://doi.org/10.1016/j.chb.2014.02.006

[18] Karunaratne I., Atukorale A., and Perera H. 2011. Surveillance of Human-Computer Interactions : A Way Forward to Detection of Users' Psychological Distress. In *Humanities, Science and Engineering (CHUSER)*, 491–496. http://doi.org/10.1109/CHUSER.2011.6163779

[19] Karunaratne I., Atukorale A., and Perera H. 2010. Reflection of stress in working with computers: A case study. *2010 Fifth International Conference on Information and Automation for Sustainability*: 484–488. http://doi.org/10.1109/ICIAFS.2010.5715710

[20] Karunaratne I., Atukorale A., and Perera H. 2015. Trust Models in Cloud: A Survey on Pros and Cons. 471–478. http://doi.org/10.1007/978-3-319-06764-3

[21] Arnrich B., Setz C., La Marca R., Troster G., and Ehlert U. 2010. What Does Your Chair Know About Your Stress Level? *Information Technology in Biomedicine, IEEE Transactions on* 14, 2: 207–214. http://doi.org/10.1109/TITB.2009.2035498

Interacción entre Adultos Mayores con Demencia y Robots de Compañia: Un Estudio Cuantitativo

Dagoberto Cruz-Sandoval, Luis González & Jesús Favela
Departamento de Ciencias de la Computación
CICESE, México
{dagoberto, lgasso}@cicese.edu.mx, favela@cicese.mx

RESUMEN

Los robots de asistencia social (RAS) son utilizados para proporcionar una estimulación social, afectiva y cognitiva a los adultos mayores con demencia (AcD). Uno de los aspectos más importantes del éxito de este tipo de robots es su aceptación y adopción por el adulto mayor. Este trabajo presenta los resultados de un estudio cuantitativo de la interacción de los AcD con un RAS de compañía. El estudio constó de observar y analizar la interacción de diez AcD y un RAS. El estudio tuvo como objetivo analizar las formas de interacción entre el AcD y el robot, para determinar la viabilidad de utilizar este tipo de tecnología como apoyo en el cuidado de los AcD.

Palabras Clave

Interacción Humano-Robot, adultos mayores con demencia, robots de asistencia social.

1. INTRODUCCIÓN

Los robots han comenzado a ser utilizados como dispositivos de asistencia en nuestra vida diaria, siendo la robótica un método para expandir las capacidades físicas y cognitivas del ser humano. Hay avances en nuevos campos de la robótica, donde la meta no es proporcionar solo asistencia física sino brindar estimulación a los humanos a través de interacción con un robot [6].

Los robots para los cuidados de la salud pueden tener diferentes funciones y pueden ser categorizados como robots de rehabilitación y robots de asistencia social (RAS). En contraste con los robots de rehabilitación, los RAS deben proporcionar interfaces de comunicación más fáciles y amigables para que los humanos puedan interactuar con este tipo de robots e inclusive servir como compañía o distracción.

Entre los diferentes servicios de salud que deben de proporcionarse a los adultos mayores con demencia (AcD); la terapia de ejercicios físicos, la interacción social y la compañía se pueden abordar por medio del uso de RAS [2]. En materia de compañía, los RAS buscan cubrir algunas de las percepciones y necesidades del AcD, incluyendo la atención y compañía por largos periodos, estimulación emocional y social positiva; y al mismo tiempo proporcionar tiempo libre a los cuidadores para realizar otras actividades [5]. Sin embargo, el éxito de este tipo de terapias se basa en la aceptación e interacción por parte de los adultos mayores.

Este trabajo presenta un estudio cuantitativo para evaluar la interacción y adopción de un RAS por parte de AcD. El estudio consistió en la comparación de un mascota robot comercial y un juguete de peluche inanimado, ambos utilizados con el objetivo de proporcionar compañía a los AcD. El estudio se realizó bajo la hipótesis de que los adultos mayores con demencia en casas geriátricas tienen mayor aceptación de una mascota robot que con un juguete de peluche de similar apariencia.

2. DISEÑO EXPERIMENTAL

El experimento tuvo como objetivo analizar cuantitativamente la interacción de una mascota robot por parte de los AcD. El estudio cuantitativo se enfoca en la comparación de los niveles de adopción e interacción que tienen los AcD con un RAS y un juguete de peluche. La naturaleza del estudio fue intra-sujetos, es decir, un grupo de sujetos elegidos al azar interactuaron con ambos objetos. La interacción con cada objeto fue grabada en video para su análisis posterior. Además, la interacción con cada objeto fue controlada en cuestión de tiempo y condiciones -el tiempo y lugar de la prueba fueron los mismos para todos los participantes. En las siguientes secciones se definen a detalle cada uno de los materiales y métodos utilizados durante el estudio.

2.1 Objetos de Interacción

La variable independiente del estudio fue el objeto de interacción. Para esto se definieron dos opciones. La primera opción es un RAS con forma de gato. El robot es fabricado por la empresa americana "Hasbro" y distribuido comercialmente bajo el nombre de *Joy for All: Companion Pets* [4]. La mascota robot responde a caricias, abrazos y movimientos. Dentro de las características más importantes mencionadas por el fabricante son: 1) Los sensores integrados (cara, cuello, espalda y pecho) responden al movimiento y al tacto, 2) La piel realista se ve y se siente como un gato real, 3) Movimientos y sonidos como un gato, 4) El sistema "VibraPurr" suena y se siente como un ronroneo real. El segundo objeto, es un juguete de peluche tradicional. La apariencia del peluche es de una foca recién nacida, del cual su principal característica es la forma y material agradable al tacto.

Se estableció una diferencia de apariencia entre el robot y el juguete de peluche para resaltar ambos objetos, y así evitar que los participantes tuvieran la percepción de que se trataba del mismo objeto. Sin embargo, se consideraron similitudes en los aspectos de tamaño, peso y material de los objetos.

2.2 Hipótesis del Experimento

A partir de los objetos de interacción, se plantearon las hipótesis que guiaron el experimento.

H1. El adulto mayor con demencia disfruta más la interacción con la mascota robot.
H2. La mascota robot motiva más la interacción verbal en el adulto mayor con demencia.
H3. El adulto mayor con demencia pierde de forma más evidente el interés de interactuar con el juguete de peluche.
H4. La mascota robot motiva al adulto mayor con demencia a proporcionar mayor afecto físico.
H5. El adulto mayor con demencia se mantiene más enfocado en la mascota robot.
H6. A menor nivel de demencia, mayor es la interacción entre el adulto mayor y la mascota robot.

2.3 Definición de Comportamientos

Las variables dependientes del experimento son los comportamientos observados en los AcD durante la interacción con la mascota robot y el juguete de peluche. Cada uno de los comportamientos fue definido en base a la relación que existe con cada una de las hipótesis definidas. A continuación se presenta la descripción de los comportamientos medidos durante el experimento, además de mostrar la hipótesis con la que se relacionan.

- **Expresión positiva (Sonrisa).** Se presenta un cambio de expresión facial positiva en el AcD. Esto incluyó sonrisas, expresiones de sorpresa y risas. La expresión debe denotar un cambio positivo y no simplemente una mueca. (H1)
- **Habla.** El AcD habla o intenta hablar, ya sea al objeto, al acompañante o con él mismo. Cuando hable con él mismo o al acompañante, tiene que hacerlo en referencia a un tema desprendido de la interacción con el objeto en turno. (H2, H6)
- **Mirada fija.** El AcD enfoca su mirada en el objeto de interacción. (H5)
- **Rechazo.** El AcD expresa de manera explícita que no quiere interactuar más con alguno de los objetos. El rechazo puede presentarse de manera verbal o por medio de lenguaje corporal (mover la cabeza, tratar de quitar o entregar el objeto al acompañante). (H3)
- **Afecto físico (Caricia).** El AcD demuestra afecto físico al objeto como abrazos, besos, palmadas o caricias. Esta reacción debe ser un afecto físico claro y no solo un contacto físico para moverlo o sostenerlo. (H4, H6)
- **Palabras positivas.** El AcD expresa claramente una palabra de afecto o cariño al objeto de interacción. (H1)

2.4 Participantes

Los participantes del estudio fueron adultos mayores diagnosticados con demencia en una residencia geriátrica. Los criterios de inclusión se basaron en la edad (entre 70 y 95 años) y contar con aptitues físicas para manejar los objetos de interacción. Referente a la exlusión, se tomaron en cuenta el rechazo a los animales -debido a la apariencia de los objetos- y padecimiento de alergias o algún tipo de impedimento asociado a los objetos del estudio. En base a expedientes clínicos y la información obtenida a través de la convivencia y cuidados, los cuidadores proporcionaron la información requerida para definir la inclusión o exlusión de participantes.

Sin embargo, para el criterio de exclusión relacionado con el rechazo hacia los animales y en particular para conocer si existía un rechazo hacia los objetos de interacción, se realizó un estudio exploratorio. El estudio consistió en presentar ambos objetos (robot y peluche) a los residentes, y evaluar quienes mostraban un rechazo instantáneo hacia los objetos. De un total de 20 residentes, sólo el 15% (n=3) mostraron un rechazo al o los objetos, o expresaron su temor a los animales en general.

De los participantes restantes (n=17) se eligieron 10 participantes al azar. Dentro del estudio participaron hombres (n=3) y mujeres (n=7). Con una media (M) de edad de 81.77 años y una desviación estándar (DE) de 6.51. Los participantes elegidos han sido diagnosticados con algún tipo de demencia, dentro de las que se encuentran deterioro cognitivo, demencia mixta y enfermedad del Alzheimer.

Además, se realizó una prueba para obtener una medida estandarizada acerca del estado cognitivo de los participantes. El método utilizado fue una prueba basada en la examinación Mini-Mental (MMSE) [3]. La prueba utilizada se basa en la evaluación de 1) orientación temporal, 2) orientación espacial, 3) fijación de recuerdo, 4) atención y cálculo, 5) recuerdo diferido, y 6) lenguaje (denominación, repetición, órdenes, lectura, escritura y copia). Siendo 30 el puntaje máximo y un puntaje menor a 24 representa un estado de deterioro cognitivo. La Figura 1 muestra los resultados del MMSE. Los 10 participantes obtuvieron puntuaciones relacionadas con deterioro cognitivo y demencia (M=12, DE=4.47).

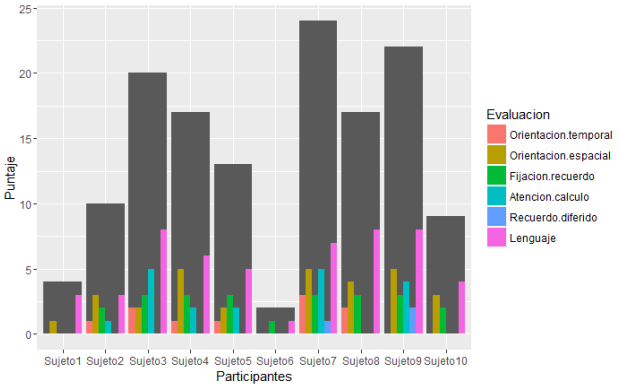


Figura 1. Resultados de la prueba MMSE de los 10 participantes.

2.5 Adquisición de Datos

Cada uno de los 10 participantes interactuó con ambos objetos (mascota robot/juguete de peluche) en un cuarto privado de la residencia geriátrica. Un miembro del equipo de trabajo (siempre la misma persona) entregó el robot y el peluche, y estuvo con cada uno de los participantes por alrededor de 10 minutos. La interacción de los participantes con cada uno de los objetos fue grabada en video para su análisis posterior. Las pruebas de interacción se hicieron a puerta cerrada evitando, en lo posible, elementos adicionales dentro de la habitación, ambas medidas para evitar la distracción del AcD en la interacción con los objetos. En algunos casos, el área de interacción fue modificada; debido a que algunos de los participantes utilizan silla de ruedas para su desplazamiento.

La interacción efectiva con cada uno de los objetos fue de 3 minutos, es decir, se toma en cuenta sólo el lapso en el cual el AcD interactúa con el objeto, despreciando los lapsos en donde se hace la presentación y cambio del objeto, o cualquier otra actividad que no estuviera relacionada directamente con la interacción. El protocolo para guiar la interacción con los objetos fue el mismo para todos los participantes. El orden de presentación de los objetos fue definido para el primer participante al azar. A partir de esa elección, se combinó el orden de presentación de los objetos. Se presentó primero el juguete de peluche y después la mascota robot (n=5), siendo en orden inverso para el resto de los participantes (n=5).

2.6 Codificación

Se grabaron dos videos por participante, uno para cada objeto de interacción y con una duración de 3 minutos. Cada video fue codificado en base a los comportamientos previamente definidos: Expresión positiva, Habla, Mirada fija, Rechazo, Afecto físico y Palabras positivas. A excepción del comportamiento de Mirada fija que fue medido por tiempo, los demás fueron contabilizados por evento.

Dos personas realizaron la codificación de los videos, en base a un mismo esquema de codificación cargado previamente en la herramienta BORIS. En primera instancia se analizaron videos de manera conjunta (n=4) para refinar los criterios del esquema de codificación. Posteriormente, se tomaron 2 videos al azar y se codificaron de manera individual, en donde la tasa de concordancia de la evaluación de los comportamientos entre los dos evaluadores fue del 93.33%. Esta tasa de concordancia, mostró la viabilidad de codificar los videos restantes (n=4) de manera individual.

2.7 Método de Análisis

Con el objetivo de comprobar las hipótesis planteadas, se recabaron los datos asociados a los comportamientos definidos. Cada uno de esos comportamientos son las variables dependientes en razón de una variable independiente (mascota robot/juguete de peluche). Se obtuvieron las frecuencias o tiempo de cada comportamiento para cada uno de los participantes. Posteriormente, se calcularon la media y desviación estándar por comportamiento y objeto de interacción. De esta forma se tiene una primera comparación de la interacción con ambos objetos.

Debido a la naturaleza de los datos obtenidos - en donde no se presentó una distribución normal - se decidió utilizar un método de prueba no paramétrico, por lo cual se eligió la prueba de rangos con signo de Wilcoxon para las hipótesis 1 a 5. Para la hipótesis 6 fue utilizada la prueba de correlación de Spearman – se eligió esta prueba debido a que es una prueba no paramétrica - para comprobar si existe una correlación entre el nivel de interacción con la mascota robot y el estado cognitivo – en base a los resultados del MMSE - de los participantes. Para ambos métodos se estableció un valor de significancia de α=0.05. El procesamiento estadístico se realizó utilizando la plataforma estadística R.

3. RESULTADOS

Los eventos y tiempos relacionados con las variables dependientes fueron recabadas por medio de la herramienta BORIS. Posteriormente fueron tratados por medio de la plataforma R. La Figura 2 muestra las frecuencias medias observadas para cada una de las variables dependientes.

La Figura 3 muestra la dispersión de las variables dependientes para el juguete de peluche. También se puede observar que la distribución de los datos no es simétrica.

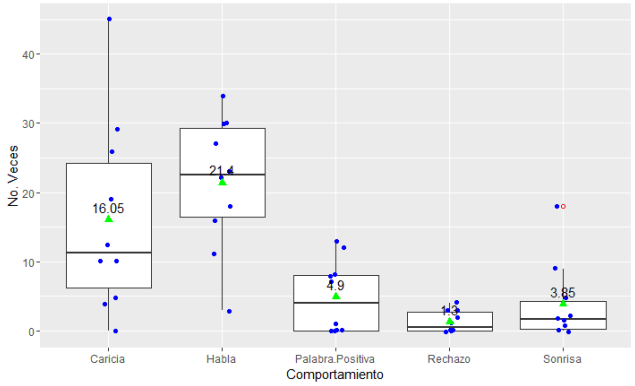


Figura 3. Diagrama de caja para las variables dependientes de la interacción con el juguete de peluche.

En el caso de la interacción con la mascota robot, se encontraron un mayor número de valores atípicos en las variables dependientes, en concreto en las variables relacionadas con los comportamientos relacionadas con afecto físico y sonrisas (véase Figura 4). Lo anterior, muestra que la distribución de los datos no es normal, por lo cual para el análisis de prueba de hipótesis se utilizaron métodos no paramétricos.

3.1 Prueba de Hipótesis

Utilizando la prueba de rangos con signo de Wilcoxon con un valor de significancia de 0.05, se procedió a probar las hipótesis. Las pruebas se realizaron mediante la plataforma R.

La Tabla 1 muestra los resultados para las hipótesis 1 a 5. Solamente la hipótesis 3 muestra una diferencia significativa, lo cual se interpreta como que los AcD pierden más el interés de interactuar con el juguete de peluche. En el caso de las hipótesis restantes, no existe una diferencia significativa que lleve a aceptar los postulados.

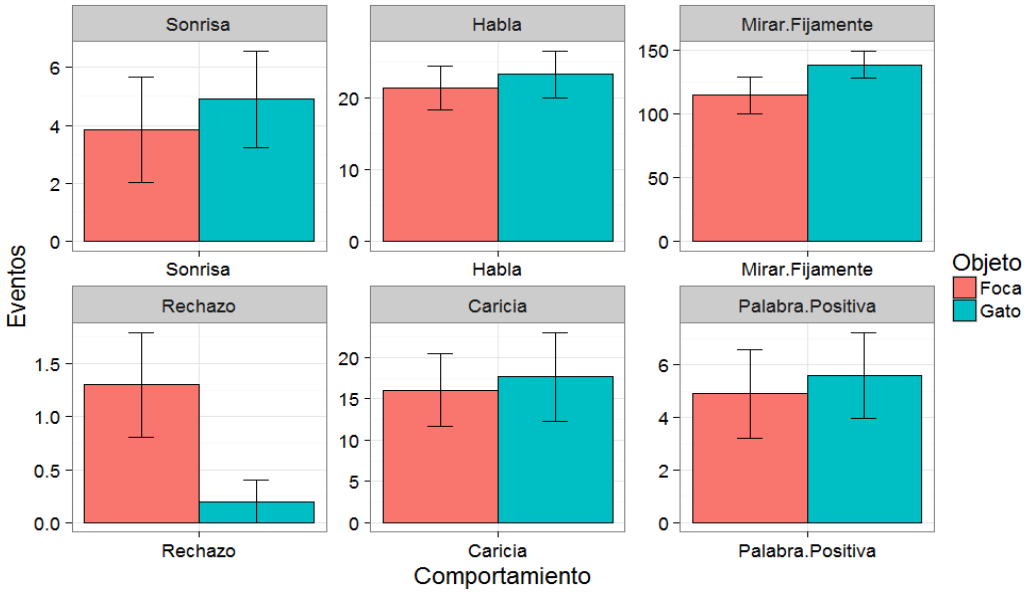


Figura 2. Frecuencias medias de los comportamientos observados durante la interacción a) Expresión positiva (sonrisa, b) Habla, c) Mirada fija, d) Rechazo, e) Afecto físico (caricia) y f) Palabra positiva.

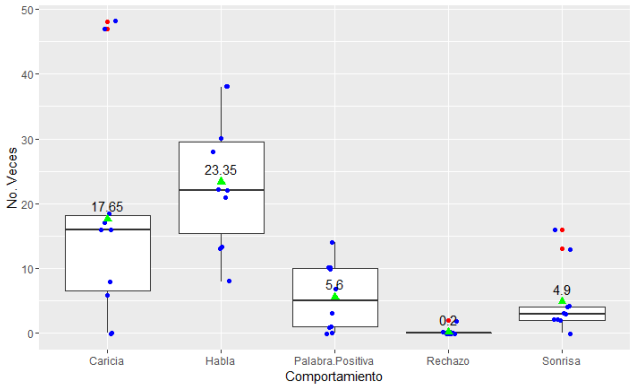


Figura 4. Diagrama de caja para las variables dependientes de la interacción con la mascota robot.

La Figura 5 muestra la relación que existe entre las puntuaciones obtenidas en el MMSE y la interacción – definida por las variables dependientes de expresión verbal y afecto físico. Para probar la sexta hipótesis, se verificó si existía una correlación positiva entre los conjuntos de datos. Calculando el coeficiente de correlación de Spearman se obtiene que $\rho = 0.3547$, lo que significa que existe una correlación positiva débil entre ambos conjuntos de datos. Además, se obtuvo que $p - valor = 0.3145$. Dado que $p - valor > 0.05$, nos permite aceptar el valor de ρ calculado, siendo estadísticamente significativo.

Tabla 1. Resultados del análisis estadístico (prueba de rangos con signo de Wilcoxon) para las hipótesis planteadas.

	Hipótesis	Resultado
(1) Hay diferencia significativa	H3	$p - valor = 0.0421$
(2) No hay diferencia significativa	H1	$p - valor = 0.0809$
	H2	$p - valor = 0.1824$
	H4	$p - valor = 0.3738$
	H5	$p - valor = 0.0744$

3.2 Observación cualitativa

Ante la pregunta explícita de qué objeto les había gustado más a los AcD, 5 respondieron que la mascota robot, 2 que ambos y 3 no respondieron. Al preguntar el por qué preferían la mascota robot, los AcD expresaron que debido a sus movimientos y sonidos.

"Me gustó más este [gato], porque cierra los ojitos. Este [gato] se mueve más y el otro [foca] no." [S3]

"Este [gato] es el que más me gusta, porque se mueve. Como quisiera tener uno. Me gusta para tenerlo en mi cama encima de mi almohada." [S4]

"Me gustó más este [gato] porque llora y platica." [S5]

Algunos AcD percibieron a los objetos como animales reales. Durante la interacción con la mascota robot, algunos AcD detectaron que se trataba de un robot, y esto en lugar de causar rechazo, generaba asombro y admiración.

"Que inteligentes los que hacen esto [gato]. Como los hacen inteligentes ahora, antes ni pa' cuando." [S7]

"Esto [gato] ha de costar 4 mil pesos. Abre los ojos de vez en cuando, ¿verdad?, pero es por su máquina que trae adentro." [S10].

Otro aspecto observado, fue que, en aquellos participantes que primero interactuaron con la mascota robot, al interactuar con el juguete de peluche, lo acariciaban o movían buscando tener el mismo tipo de respuesta (movimientos, sonidos) que el robot.

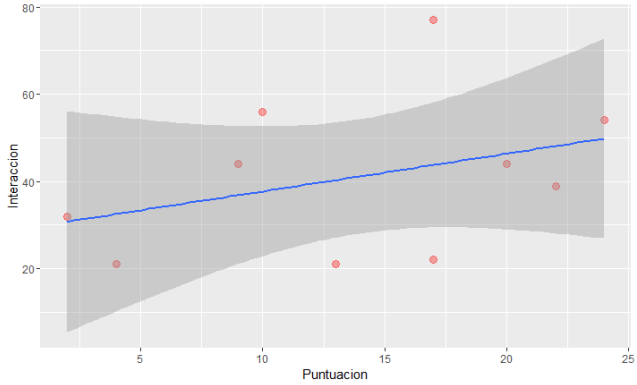


Figura 5. Gráfica de Correlación entre puntuación de MMSE e interacción con la mascota robot.

4. CONCLUSIONES

Durante el análisis, solamente dos de las hipótesis planteadas (H3 y H6) son aceptadas por tener significancia estadística.

Aunque varios de los participantes manifestaron de manera explícita su preferencia por la mascota robot, esto no se ve reflejado en los resultados del estudio, ya que los datos obtenidos para cada variable dependiente son muy cercanos. Por lo tanto, podemos decir que no parece haber diferencia en la interacción con ambos objetos.

Sin embargo, es importante considerar es que no existe un rechazo hacia el RAS por parte de los AcD demostrando que puede existir una interacción entre ambos, lo cual es un resultado favorable para superar el conocido "*uncanny valley*" [1]. Esto proporciona un buen antecedente para el uso de RAS más sofisticados que proporcionen interfaces de interacción más naturales como la voz y expresión de sentimientos, en vías de poder generar interacción más efectiva, y así proporcionar estimulación social, afectiva y cognitiva al AcD.

5. REFERENCIAS

[1] DiSalvo, C.F. et al. 2002. All robots are not created equal. *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '02* (New York, New York, USA, 2002), 321.

[2] Fasola, J. and Mataric, M.J. 2012. Using Socially Assistive Human–Robot Interaction to Motivate Physical Exercise for Older Adults. *Proceedings of the IEEE*. 100, 8 (Aug. 2012), 2512–2526.

[3] Folstein, M.F. et al. 1975. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*. 12, 3 (Nov. 1975), 189–98.

[4] Hasbro's Joy For All: Companion Pets: 2016. *http://joyforall.hasbro.com/en-us*. Accessed: 2016-07-30.

[5] Shibata, T. 2012. Therapeutic seal robot as biofeedback

medical device: Qualitative and quantitative evaluations of robot therapy in dementia care. *Proceedings of the IEEE*.

[6] Wada, K.. and Shibata, T.. 2007. Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House. *IEEE Transactions on Robotics*. 23, 5 (Oct. 2007), 972–980.

Propuesta de Adquisición de Datos Para el Problema Inverso Electrocardiográfico

Esteban Herrera Hernandez
Beneméreita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación
Av. San Claudio y 14 Sur, Cd Universitaria, 72592
Puebla, Pue.
artanis_ii@hotmail.com

Rafael Lemuz López
Beneméreita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación
Av. San Claudio y 14 Sur, Cd Universitaria, 72592
Puebla, Pue.
rlemuz@cs.buap.mx

Carlos Guillén Galván
Beneméreita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación
Av. San Claudio y 14 Sur, Cd Universitaria, 72592
Puebla, Pue.
cguillen@fcfm.buap.mx

Ángel Ramos del Olmo
Universidad Complutense de Madrid
Facultad de Ciencias Mátematicas
Av. Séneca, 2, 28040 Madrid, España
España, Madrid
angel@mat.ucm.es

Resumen

En este trabajo se presenta una propuesta para la adquisición de datos del problema inverso en electrocardiografía. Obtener una imagen electrocardiográfica resolviendo el problema inverso para estudiar la actividad eléctrica en la superficie del corazón, es importante en el estudio y diagnostico de enfermades cardiacas. En el trabajo se propone un método de captura de las señales electrocardiográficas utilizando un sistema de adquisición de bajo costo inalámbrico. Ademas se describe la solución del problema inverso mediante el metodo de regularización de tikhonov.

Palabras Clave

Adquisición de datos, Ecuaciones diferenciales, Electrocardiograma (ECG), EPOC Headset, Problema Inverso.

1. INTRODUCCIÓN

Las enfermedades cardiovasculares (ECV) son la principal causa de muerte en todo el mundo. Cada año mueren más personas por ECV que por cualquier otra causa.
En México es el segundo lugar como causa de muerte, durante el año 2014 se registraron 82,334 decesos por enfermedades isquémicas del corazón lo cual es preocupante [4],

ya que por nuestro ritmo de vida esto se incrementa año con año y muchos de los estudios que existen para diagnosticar este tipo de enfermedades son excesivamente caros al igual que sus tratamientos y resolviendo el problema inverso de la electrocardiografía podría tener un impacto muy fuerte tanto social como económico a favor de todos.

El problema inverso de electrocardiografía consiste en obtener potenciales miocárdicos a partir de los registros de superficie de manera no invasiva. Los registros superficiales se obtienen con la técnica de BSPM (Body Suface Potential Mapping) con alrededor de 32 a 300 electrodos alrededor del torso. Con ello se pretende encontrar regiones del corazón donde están ocurriendo patologías cardiacas, como son el bloqueo de rama, insuficiencia cardiaca, infarto del miocardio, el flutter auricular, potenciales tardíos ventriculares o el síndrome de Brugada, todas ellas causantes de la denominada muerte súbita. En la literatura se dice que el problema inverso de electrocardiografía está mal planteado resultando en múltiples soluciones posibles e inestabilidad. Por tanto, es necesario el desarrollo de técnicas y algoritmos nuevos que simplifiquen y permitan aplicar estos resultados a una mayor cantidad de población.

2. METODOS DE MEDICIÓN

Existen dos tipos diferentes para poder medir el potencial eléctrico, unipolar y bipolar. Para medir un potencial eléctrico, es necesario utilizar al menos dos electrodos. En la medición unipolar, un electrodo se coloca cerca del corazón para medir la señal y el otro electrodo se coloca lejos del corazón como referencia. En la medición bipolar, dos electrodos tienen aproximadamente la misma distancia desde el corazón. Existen varios sistemas de conducción para medir el potencial eléctrico. Algunos incluyen cables bipolares, algunos incluyen sólo las derivaciones unipolares, y algunos contienen ambos tipos.
El estándar de 12 derivaciones de ECG, es un sistema que incluye tanto conductores bipolares como unipolares y es estándar en la práctica clínica actual. En este caso diez electrodos se colocan en el cuerpo y 2 son referencias. En

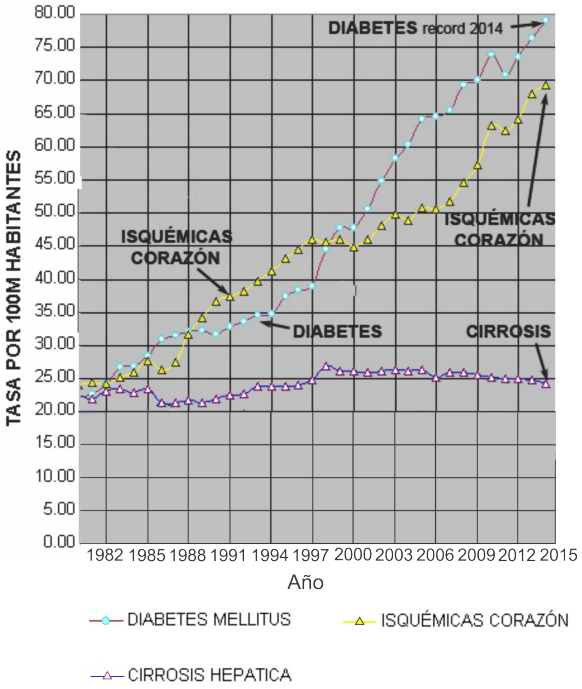


Figura 1: Principales 3 causas de muerte en México

ellas se registran doce señales eléctricas diferentes en aproximadamente el mismo tiempo [2]. Mientras más electrodos superficiales se utilicen mejores resultados se obtendran. Los electrodos se colocan directamente sobre la superficie del cuerpo y son capaces de tomar registros sobre la actividad eléctrica. Los materiales de los que se constituyen la mayoría de los electrodos son muy variables, entre ellos están la plata , el oro, el acero inoxidable, platino, entre otros [3].

Existen 3 tipos de señales muy importantes que son necesarias para poder monitorear el corazón:

- La onda P registra la actividad eléctrica de las dos cámaras superiores, las aurículas.
- El complejo QRS registra la actividad eléctrica de las dos cámaras inferiores, los ventrículos.
- La onda T registra el regreso del corazón al estado de reposo.

Estas señales, dependiendo de su forma y tamaño o el tiempo que existe entre ellas dan una gran cantidad de información sobre el corazón y su ritmo. Los cambios en estas ondas nos indican si el corazón sufre algún tipo de enfermedad.

3. ALGUNAS ENFERMEDADES CARDIOVASCULARES DE INTERES

Las arritmias son problemas con la frecuencia cardíaca (pulso) o el ritmo cardíaco. Esto ocurre cuando el sistema eléctrico del corazón no funciona correctamente. El corazón puede palpitar demasiado rápido, demasiado lento o en forma irregular. Algunos problemas del corazón, como un ataque cardíaco o una insuficiencia cardíaca, pueden causar problemas con el sistema eléctrico del corazón. Algunas personas nacen con una arritmia [6].

Las enfermedades de las válvulas cardíacas ocurren cuando una de las cuatro válvulas en el corazón no funciona correctamente. La sangre puede escaparse a través de la válvula en la dirección equivocada (llamado regurgitación), o es posible que una válvula no se abra lo suficiente y bloquee el flujo sanguíneo (llamado estenosis). Un latido cardíaco inusual, llamado soplo cardíaco, es el síntoma más común. Algunos problemas del corazón, como un ataque cardíaco, una cardiopatía o una infección, pueden causar enfermedades de las válvulas del corazón. Algunas personas nacen con problemas de válvulas cardíacas [6].

4. PROBLEMA INVERSO

Para poder obtener datos y resolver el problema inverso en electrocardiografía, tenemos que interpretar los potenciales registrados en la superficie del cuerpo debido a la actividad eléctrica del corazón. El objetivo general de los problemas "inversos" llamados "hacia adelante" de electrocardiografía es una mejor comprensión cualitativa y cuantitativa de la actividad eléctrica del corazón [1].
Para poder medir la actividad eléctrica en la superficie del corazón utilizaremos un sistema de adquisición de bajo costo el cual es el EPOC Headset, ya que este sistema cuenta con sensores de humedad a base de solución salina (sin geles pegajosos) y 14 canales de EEG, más 2 canales de referencia y nos ofrece un posicionamiento óptimo para la adquisición de señales, es inalámbrico y recargable.

Este sistema nos ayudara a obtener la información necesaria sobre las señales P,T y el complejo QRS que mencionamos anteriormente, atraves de electrodos superficiales puestos en 4 partes del torso, de los cuales 2, estarán cerca del brazo derecho y del brazo izquierdo, otro electrodo irá junto a la pierna izquierda y por último el electrodo que va a estar junto a la pierna derecha sera nuestra referencia, ver figura 3. Con esto se podremos obtener los datos necesarios para poder resolver el problema inverso mediante ecuaciones parciales y la descopomposición de valores singulares regularizados y así poder determinar si existe algún tipo de enfermedad en el corazón.

Sabemos que cuando existe alguna alteración en las señales P,T y el complejo QRS, el paciente puede sufrir de alguna enfermedad del corazón. En la figure 3 se muestra los signos vitales de un corazón sano, uno con arritmia isquemica y otro cuando tiene un infarto. Se puede observar que, cuando ocurre una arritmia hay un cambio en la onda T y cuando ocurre un infarto hay un cambio en el intervalo ST.

A continuación se muestra otra grafica donde se muestran las diferentes enfermedades que el corazón puede sufrir y



Figura 2: 12 Derivaciones de ECG y EPOC Headset

esto lo podemos observar ya que las señales P, Q, R, S y T muestran cambios mas radicales en su forma, tamaño y tiempo.



Figura 3: Enfermedades del corazón

5. METODOS MATEMATICOS

Para comprender el enfoque matemático del problema inverso de electrocardiografía se requiere un conocimiento básico. Aquí explicaremos algunas definiciones y teoremas, que se necesitan saber para manejar un problema mal planteado.

Definición
Un modelo matemático es un mapeo

$$A : X \to Y \tag{1}$$

Del conjunto de causas (parámetros) X para el conjunto de la dirección (datos) Y. En el problema directo el efecto se calcula en función de la causa, es decir, para $x \in X$ calculamos $Ax \in Y$.A la inversa, en el problema inverso tenemos que observar $y \in Y$ y tratar de encontrar una $x \in X$ cuando $Ax = y$.

La siguiente definición debido a Hadamard es una clasificación de los problemas bien planteados y mal planteados.

Definición
Sea $A : X \to Y$ un mapeo entre espacios topológicos X y Y. El problema (A, X, Y) se llama bien planteado, si las siguientes condiciones son cumplidas.

a) Para toda $y \in Y$ la ecuación $Ax = y$ es solución de $x \in X$.

b) La solución es única.

c) La aplicación inversa $A^{-1} : Y \to X$ es continua, es decir la solución de x depende de forma continua de los datos en y (pequeños errores en y harán pequeños errores en x).

5.1 Inversa generalizada

Las dificultades en la resolución de un problema mal planteado son obvias. No podemos asumir que una solución es única, o que siquiera existe. Por lo tanto, tenemos que encontrar una manera de satisfacer las condiciones a) y c).
En lo que sigue, suponemos que X y Y sean espacios de Hilbert y A para ser un operador lineal, $A \in L(X, Y)$. Analizamos la solución de la ecuación [10].

$$Ax = y \tag{2}$$

Con un determinado arbitrario $y \in Y$. Si el rango de $AR(A)$ no es el espacio completo de la imagen, no es solucionable para $y \in Y$. Por otro lado, si A tiene un espacio nulo trivial $N(A)$, la solución no es única. Así que si hay varias soluciones, es razonable elegir el que tenga la norma mínima entre ellos. Este pensamiento lleva a la siguiente definición:
Un elemento $x \in X$ es llamado:

a) Mínimos cuadrados solución de la ecuación $Ax = y$ si

$$||Ax - y||_Y = inf\{||Az - y||_Y | z \in X\} \tag{3}$$

b) Solución de norma mínima de la ecuación $Ax = y$ si

$$||x||_X = inf\{||z||_X | z \text{ es una solución de mínimos cuadrados de } Ax = y\} \tag{4}$$

Todavía no podemos asegurar que existe una solución de mínimos cuadrados para una y arbitraria, ya que $R(A)$ no necesita ser cerrado. Pero por ahora, esto se puede suponer. En este caso, la solución de norma mínima es única, debido a la convexidad de la minimización relativa funcional. A continuación, se define la inversa generalizada Moore-Penrose, con la que podemos calcular la solución de la norma mínima [2].

Definición
Sea A un continuo funcional lineal entre espacios de Hilbert X y Y, es decir, $a \in L(X, Y)$. Sea $A : N(A)^{\perp} \to R(A)$ sea su restricción. A continuación, la inversa generalizada de Moore-Penrose $A^{\dagger}$ se define como la extensión lineal única de A^{-1} a su dominio

$$D(A^{\dagger}) := R(A) \oplus R(A)^{\perp} \tag{5}$$

Con $N(A^{\dagger}) = R(A)^{\perp}$. Así $(A^{\dagger})$ son mapas de $D(A^{\dagger})$ para $N(A)^{\perp}$. (6)

Como se ha mencionado antes, la inversa generalizada se puede utilizar para calcular la solución mínima. Esto lo podemos ver en el siguiente teorema.

Teorema
La ecuación $Ax = y$ tiene una solución única por la norma-mínima dada por

$$x^{\dagger} := A^{\dagger} y \qquad \forall y \in D(A^{\dagger}) \tag{7}$$

El conjunto de todas las soluciones de mínimos cuadrados está dada por $\{x^\dagger\} + N(A)$.
Con este Teorema, podemos encontrar una solución única para la ecuación $Ax = y$ si $y \in D(A)$. Por lo tanto, las condiciones a) y b) de la definición anterior son cumplidas. Si $R(A)$ no está cerrado y por lo tanto no existe $y \in Y \backslash D(A^\dagger)$, porque se viola la tercera condición. Hay algunas técnicas de regularización, que ayudan a manejar este caso [2].

6. TÉCNICAS DE REGULARIZACIÓN

Como ya se ha mencionado, el problema inverso no posee una solución única y como hemos visto el problema inverso está mal planteado, es decir, el operador R^{-1} no tiene límites, por lo tanto, los pequeños errores en los datos medidos d pueden ser críticos e introducir grandes errores en g. Por esa razón, tenemos que encontrar una manera de manejar esta dificultad. Hay algunas técnicas de regularización, que se aproximan al problema mal planteado mediante el uso de ecuaciones bien planteadas. A continuación, se presenta la Regularización de Tikhonov [10].

6.1 Regularización de Tikhonov

Este método de regularización se inspira en la formulación discreta de mínimos cuadrados [10]. Por esta razón, se introduce la función

$$J(g) = ||R(g) - d||^2_{L^2(\sigma_1)} \qquad \text{para } g \in H^1(\sigma_3)$$

en la que usamos la misma denotación como antes.
Ahora volvemos a escribir la ecuación $R(g) = d$ como el problema de minimización

$$\min_{g \in H^1(\sigma_3)} J(g) \tag{8}$$

Por supuesto, mediante la formulación del problema inverso de esta forma, el "mal planteado" aún está presente. Con el siguiente paso, nos aproximamos a (8) por un problema de buen comportamiento.
Se define una nueva función $J_\in(g)$ mediante la adición de un término de regularización a $J(g)$:

$$J_\in(g) = J(g)+\epsilon||g||^2_{L^2(\sigma_3)} = ||R(g)-d||^2_{L^2(\sigma_1)}+\epsilon||g||^2_{L^2(\sigma_3)} \quad para \quad g \in H^1(\sigma_3) \tag{9}$$

con $\epsilon > 0$ como el parámetro de regularización. Con (9) se formula una aproximación de (8):

$$\min_{g \in H^1(\sigma_3)} J(g) \tag{10}$$

Esto es un compromiso entre minimizar la norma de la residual y mantener g pequeña.
Por último, se demuestra que la adición del término de regularización a nuestra función original hace cumplir la estabilidad. Recordamos que $g \in H^1(\sigma_3)$ se puede escribir en la forma

$$g(x) = \sum_{K=o}^{\infty} pk \cos(k\pi x) \qquad para\ la\ adecuada\ p0, p1, ... \in \mathbb{R}$$

Por otra parte, con

$$J_\epsilon = J_\epsilon(p0, p1, ...)$$

y desde el primer orden la condición necesaria

$$\frac{\partial J_\epsilon}{\partial p_i} = 0 \qquad para\ \ i = 1, 2, ...$$

durante un mínimo, se deriva la fórmula

$$pk = \frac{d_K \cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)}$$

para los coeficientes de Fourier de la solución g de (10), véase [9]. Por lo tanto, podemos concluir, que

$$g(x) = \sum_{k=0}^{\infty} \frac{d_K \cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)} \cos(k\pi x)$$

resuelve (10) con

$$d_k = 2 \int_\epsilon d(x) \cos(k\pi x) dx, \quad para\ k = 1, 2, ...,$$
$$d_0 = \int_\epsilon d(x) dx$$

siendo los coeficientes de Fourier los datos de medición d. Ahora obtenemos la siguiente aproximación R_ϵ^{-1} del operador inverso R^{-1} del mapeo de R:

$$R_\epsilon^{-1}(d) = R_\epsilon^{-1}\left(\sum_{k=0}^{\infty} d_k \cos(k\pi x)\right) = \sum_{k=0}^{\infty} dk \frac{\cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)} \cos(k\pi x)$$

y con

$$\lambda_k = \frac{1}{\cosh(k\pi)} \qquad para\ k = 0, 1, ...$$

siendo los valores propios de R, obtenemos:

$$R_\epsilon^{-1} = \sum_{k=0}^{\infty} d_k \frac{\lambda_k}{\lambda_k^2 + \epsilon} \cos(k\pi x) \tag{11}$$

Si comparamos R^{-1} y el R_ϵ^{-1} aproximado, vemos que, por una parte, para los componentes de baja frecuencia de d, la acción de R^{-1} y R^{-1_ϵ} son casi idénticos. Por otra parte, los componentes de alta frecuencia de d se amortiguan de manera eficiente por R_ϵ^{-1}. Además de eso, podemos concluir, que los pares de valores propios de la función propia del R_ϵ^{-1} son

$$\lambda_{\epsilon,k}^{-1} = \frac{\cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)}, \quad \epsilon_{\epsilon,k} = \cos(k\pi x) \ \ para\ \ k = 0, 1, ... \tag{12}$$

Mediante el uso de este, se muestra que R_ϵ^{-1} es un operador lineal acotado para cada $\epsilon > 0$ es fijo. Vemos que para $k = 0, 1, ...$

$$\lambda_{\epsilon,k}^{-1} = \frac{\cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)} \leq \frac{\epsilon^{-1}(1 + \epsilon \cosh^2(k\pi))}{(1 + \epsilon \cosh^2(k\pi)} = \frac{1}{\epsilon}$$
$$\lambda_{\epsilon,k}^{-1} > 0$$

y de eso, podemos concluir

$$\lambda_{\epsilon,k} \geq \epsilon > 0$$

En consecuencia, el cero no es un punto adherente de $\lambda_{\epsilon,k}$ valores propios de R_ϵ. El resultante R_ϵ^{-1} es entonces un operador lineal acotado y por lo tanto continua, a condición de que $\epsilon > 0$. De esta manera, el problema (10) esta bien planteado para cada $\epsilon > 0$.

Todavía hay un reto en encontrar el equilibrio adecuado entre el tamaño de ϵ y el orden de los errores presentes en d. Nosotros no nos extenderemos sobre este tema. Un ejemplo referente a esta materia se pueden encontrar en [9].

7. CONCLUSIONES

En este trabajo se presenta una introducción al problema inverso electrocardiografico, se muestra un diseño del sistema de adquisición de las señales electrocardiográficas a través de un sistema de adquisición de bajo costo para obtener las 12 derivaciones que en un trabajo futuro permitirán estudiar problemas asociados a los segmentos P,T y el complejo QRS las cuales pueden indicar si un corazón sufre de alguna enfermedad cardiovascular, por medio de su tamaño y forma dentro de una gráfica. Ademas se describe como se puede resolver el problema inverso mediante el método de regularización de Tikhonov. El principal resultado del trabajo es la propuesta de diseño del sistema de captura de bajo costo de señales ECG. En un trabajo futuro se utilizará el sistema propuesto para resolver numéricamente el problema inverso electrocardiográfico que podrá servir para concientizar a la población de algunas de las principales enfermedades cardiacas y así poder ayudar en el entrenamiento médico.

8. AGRADECIMIENTOS

Los autores agradecen el apoyo financiero de la beca de tesis asociada al proyecto de investigación VIEP-BUAP y a la red PRODEP de modelación matemática y computacional en medicina y dinámica poblacional.

9. REFERENCIAS

[1] Gulrajani, R. M. (1998). The forward and inverse problems of electrocardiography. IEEE Engineering in Medicine and Biology Magazine, 17(5), 84-101.

[2] Kochinke, J. (2012). Einfluss von Normierung auf die Analyse von EKG-BSPM Daten.

[3] Correa-Figueroa, J. L., Morales-Sánchez, E., Huerta-Ruelas, J. A., González-Barbosa, J. J., Cárdenas-Pérez, C. R. (2016). Sistema de Adquisición de Señales SEMG para la Detección de Fatiga Muscular. Revista Mexicana de Ingeniería Biomédica, 37(1), 17-27.

[4] Recopilación: Ing. Manuel Aguirre Botello, con datos de INEGI, OMS y SINAIS, MÉXICO, Principales Causas de Mortalidad 1938 - 2014 Recuperado de: http://www.mexicomaxico.org/Voto/MortalidadCausas.htm

[5] Brooks DH, Ahmad G, MacLeod RS: Multiply constrained inverse electrocardiography: combining temporal, multiple spatial, and interative regularization. In: Proc 16th Ann IntConf IEEE Eng Med Biol Soc, EEE Press, New York, pp 137-138, 1994.

[6] Goldman L. Approach to the patient with possible cardiovascular disease. In: Goldman L, Schafer AI, eds. Goldman-Cecil Medicine. 25th ed. Philadelphia, PA: Elsevier Saunders; 2016:chap 51.

[7] Zemzemi, N. (2014, June). An iterative method for solving the inverse problem in electrocardiography in normal and fibrillation conditions: A simulation Study. In International Congress on Electrocardiology (ICE) (pp. 53-56).

[8] Berrier, K. L., Sorensen, D. C., Khoury, D. S. (2004). Solving the inverse problem of electrocardiography using a Duncan and Horn formulation of the Kalman filter. IEEE Transactions on biomedical engineering, 51(3), 507-515.

[9] Lines, G. T., Cai, X., Nielsen, B. F., Mardal, K. A., Tveito, A. (2006). Computing the Electrical Activity in the Heart. Monographs in Computational Science and Engineering. Springer.

[10] Kochinke, J. (2013). Einfluss von Normierung auf die Analyse von EKG-BSPM Daten.

Implementación de un modelo de detección de aislamiento social en adultos mayores

Virginia Ortiz
CENIDET
Int. Internado Palmira S/N, 62490 Cuernavaca. Morelos
vicky.ortiz@cenidet.edu.mx

Wilfrido Campos
Facultad de Ingeniería, UAGro
Instituto Tecnológico de Chilpancingo, TECNM
wilfrido@uagro.mx

Alicia Martínez
CENIDET
Int. Internado Palmira S/N, 62490 Cuernavaca. Morelos
amartinez@cenidet.edu.mx

Hugo Estrada
INFOTEC
Ciudad de México DF, México
hugo.estrada@infotec.mx

RESUMEN
El aislamiento social es un problema que se acentúa en la etapa de la vejez, poniendo en riesgo la integridad física y mental de los adultos mayores. En este artículo se presenta la implementación de un modelo de detección de aislamiento social para adultos mayores a través del monitoreo de las actividades de comunicación y movilidad que realiza un adulto mayor, además de considerar aspectos demográficos. Como resultado de la implementación se generaron dos aplicaciones: una que permite monitorear las actividades del adulto mayor y otra que envía notificaciones sobre el nivel de aislamiento social que presenta el adulto mayor.

Palabras clave
Adulto mayor, aislamiento social, modelo predictivo, monitoreo de actividades.

1. INTRODUCCIÓN
En la actualidad, casi 700 millones de personas en el mundo son mayores de 60 años. Para 2050, las personas de 60 años o más serán 2,000 millones, esto es, más del 20% de la población mundial. El crecimiento de este segmento de la sociedad ha ocasionado que se preste mayor atención a las necesidades particulares de las personas de la tercera edad y los problemas a los que se enfrentan muchas de ellas[1].

El envejecimiento de la sociedad ha despertado el interés de instituciones de salud y áreas afines por mejorar el bienestar de los adultos mayores, por lo que, encontrar formas efectivas de proveer cuidado a los adultos mayores ha llegado a ser uno de los mayores retos para la comunidad científica[2].

En este contexto, el aislamiento Social es uno de los problemas que más se acentúa en la etapa de la vejez, donde los problemas de socialización son los más recurrentes y los que más ponen en peligro la vida independiente del adulto mayor. El aislamiento social se define como la falta de contacto e interacción con otras personas. Los riesgos asociados a este padecimiento pueden ser comparados con los riesgos asociados al tabaquismo, alcoholismo u obesidad, así mismo una persona aislada socialmente puede morir prematuramente [3].

El objetivo de este artículo es mostrar la implementación de un modelo predictivo que permite detectar de manera automática el nivel de aislamiento social que presenta un adulto mayor [4]. Para lograr este objetivo, se desarrolló una aplicación que permite llevar a cabo el monitoreo de las actividades cotidianas de un adulto mayor definidas en el modelo predictivo. La aplicación permite informar, a familiares, amigos o cuidadores del adulto mayor si el adulto mayor presenta aislamiento social.

Este artículo se encuentra organizado como sigue: la sección 2 presenta los trabajos relacionados; la sección 3 presenta una vista general de la metodología de solución; la sección 4 presenta la implementación del modelo de detección de aislamiento social; la sección 5 presenta la evaluación y finalmente, la sección 6 presenta las conclusiones.

2. TRABAJOS RELACIONADOS
En esta sección se describen las propuestas más relevantes para hacer frente al problema del aislamiento social en los adultos mayores.

De manera particular a esta investigación presentada, existen dos trabajos previos. En [2] se llevó a cabo un análisis estadístico para identificar actividades de interacción social que tienen una correlación, estadísticamente significativa, con el grado de aislamiento social en los adultos mayores. En [3] se describe un modelo predictivo desarrollado para servir como línea base para la determinación de los niveles de aislamiento social, dicho modelo hace uso de las variables relevantes obtenidas en [2]. La principal diferencia que existe entre los trabajos mencionados con el trabajo presentado en este artículo es que en el trabajo presentado se aborda la implementación y evaluación del modelo predictivo propuesto en un sistema computacional, mediante la recolección automática de los atributos relevantes.

En [5] y [6] se desarrollaron frameworks que analizan los estados emocionales del adulto mayor y de acuerdo a la emoción detectada, se realiza una acción para poder cambiar el estado emocional del adulto mayor.

En [7] se presenta un sistema basado en tecnología móvil que se centra en el apoyo a la relación entre los adultos mayores que viven solos en México y sus familiares que viven en el extranjero.

En [8], [9] y [10] se monitorean las actividades diarias del adulto mayor y se informa a sus cuidadores el estado en el que se encuentra el adulto mayor.

En todos estos trabajos, los autores recalcan la importancia de hacer frente a la gestión de las enfermedades crónicas, al mantenimiento de la salud física y mental que afectan a los adultos mayores. La diferencia de nuestro enfoque a los ya mencionados, es el tipo de variables que se utilizan para el monitoreo de la actividad del adulto mayor.

3. VISTA GENERAL DE LA METODOLOGÍA DE SOLUCIÓN
En esta sección se presenta, de forma sintética, la metodología de solución utilizada para la implementación de un modelo predictivo para la detección automática de aislamiento social de un adulto mayor. La figura 1 presenta la metodología de solución

propuesta, la cual está formada de dos fases: (1)Fase de implementación del modelo de aislamiento social; (2)Fase de evaluación.

El modelo predictivo que fue implementado analizó un conjunto de variables relacionadas con las actividades cotidianas realizadas por un adulto mayor. El resultado de este análisis fue un conjunto de variables que resultaron relevantes para detectar el aislamiento social, y que se agruparon en variables de comunicación, movilidad y demográficas (Tabla 1). El primer grupo de variables analiza el uso del teléfono móvil por parte del adulto mayor obteniendo la siguiente información: número de llamadas entrantes de la familia, duración promedio de las llamadas entrantes de la familia (Min.), número de llamadas entrantes de los amigos, duración promedio de las llamadas salientes a la familia (Min.), número de mensajes enviados a los amigos, cantidad de mensajes recibidos de la familia. El segundo grupo de variables está orientado a la movilidad del adulto mayor dentro y fuera de su hogar, analizando el número de lugares que visita, analizando además el tiempo que pasa el adulto mayor en cada una de las áreas del hogar. El tercer grupo está formado por la variable que resulto ser relevante en el caso de aspectos demográficos: género.

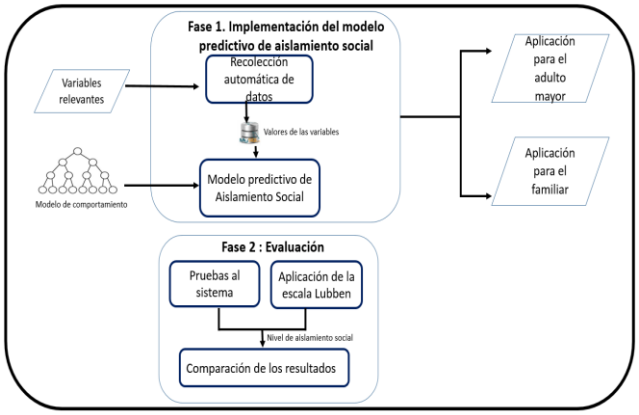


Figura 1. Metodología de solución

A continuación se describen cada una de las fases desarrolladas para la implementación del modelo predictivo de aislamiento social.

4. IMPLEMENTACIÓN DEL MODELO PREDICTIVO DE AISLAMIENTO SOCIAL

En esta sección se detalla la implementación del modelo predictivo de aislamiento social de un adulto mayor.

En la primera fase de esta metodología de solución se lleva a cabo la recolección automática de las variables relevantes del modelo predictivo, por lo que se desarrollaron aplicaciones de software que permiten obtener la información de monitoreo.

Tabla 1. Relación de variables relevantes

Tipo de interacción	Variable relevante
Variables de Comunicación	Cantidad de llamadas entrantes de la familia
	Duración promedio de las llamadas entrantes de la familia (Min.)
	Cantidad de llamadas entrantes de los amigos
	Duración promedio de las llamadas salientes a la familia (Min.)
	Cantidad de mensajes enviados a los amigos
	Cantidad de mensajes recibidos de la familia
Variables de Movilidad	Tiempo total de estancia en recámara (Hrs.)
	Tiempo total de estancia en sala (Hrs.)
	Tiempo total de estancia en comedor (Hrs.)
	Tiempo total de estancia en jardín (Hrs.)
	Tiempo total de estancia en otra área del hogar (Hrs.)
	Cantidad de lugares que visita
Variable demográfica	Género

Además se realizó la implementación del modelo predictivo propuesto, que permite detectar el aislamiento social con base en los datos recolectados por las aplicaciones.

4.1 Recolección automática de variables relevantes

La implementación del modelo predictivo implicó la recolección automática de las variables relevantes del modelo (Tabla 1). La información de las variables relevantes se extrae mediante el uso de un Smartphone y de sensores empotrados en el hogar del adulto mayor.

Para lograr esto fue necesario el desarrollo de una aplicación para dispositivos móviles que permite recolectar los valores de las variables. Se desarrolló además otra aplicación móvil que servirá al cuidador para monitorear el nivel de aislamiento social. Ambos sistemas fueron desarrollados utilizando la arquitectura cliente servidor.

La figura 2 muestra la arquitectura de estas dos aplicaciones de monitoreo y recolección de datos.

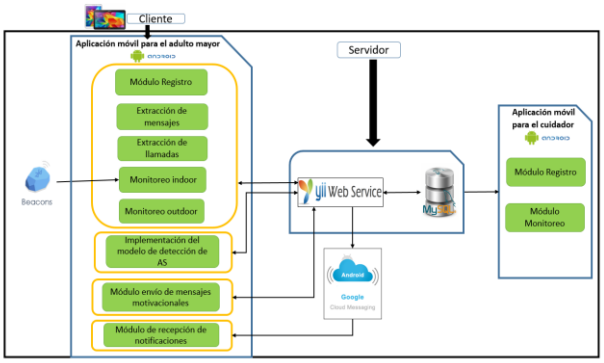


Figura 2. Arquitectura del sistema.

Los objetivos principales de la aplicación desarrollada para monitorear la actividad del adulto mayor son los siguientes:

1) Extraer la información del teléfono móvil relacionada a las llamadas y mensajes de texto.
2) Realizar el vínculo entre hardware externo (Beacons) para extraer información de monitoreo.

3) Realizar el vínculo entre la aplicación móvil y el servidor de base de datos mediante la utilización de servicios web.
4) Mostrar frases motivacionales al adulto mayor, que dependerán del grado de aislamiento social detectado.

La aplicación diseñada para el cuidador, visualiza el nivel de aislamiento social que presenta el adulto mayor y cumple con las siguientes funciones:

1) Implementa los servicios web que permiten extraer la información de la base de datos.
2) Presentar la información al usuario.

En el servidor web se encuentran los servicios y el servidor de base de datos. Las principales acciones que se ejecutan desde el servidor son las siguientes:

1) Implementar los servicios web que permitan extraer la información de las variables relevantes del teléfono móvil para almacenarla en la base de datos y que posteriormente permitan extraer la información que será mostrada al cuidador.
2) Proporcionar información, a la aplicación móvil del adulto mayor, para que se ejecute el sistema que implementa el modelo de detección de aislamiento social.

A continuación se describe brevemente los módulos de la aplicación que se ejecuta en el teléfono del adulto mayor:

a) Módulo de registro: este módulo obtiene la información demográfica del adulto mayor, como es la edad, género y nombre. También debe añadir a sus familiares y amigos con los que se comunica con mayor frecuencia.
b) Módulo de extracción de mensajes: este módulo recupera la información de los mensajes del teléfono del adulto mayor. Las variables relevantes que se extraen en este módulo son: cantidad de mensajes enviados a los amigos y cantidad de mensajes recibidos de la familia.
c) Módulo de extracción de llamadas: Este módulo obtiene información de las llamadas del teléfono del adulto mayor. Las variables relevantes que se obtienen en este módulo son: cantidad de llamadas entrantes de la familia, duración promedio de las llamadas entrantes de la familia, cantidad de llamadas entrantes de los amigos, duración promedio de las llamadas salientes a la familia.
d) Módulo de monitoreo dentro del hogar: en este módulo se captura la información de la cantidad de tiempo, medido en horas que el adulto mayor pasa en cada una de las habitaciones de su casa. Las variables relevantes que se obtienen de este módulo son: tiempo total de estancia en recamara, tiempo total de estancia en sala, tiempo total de estancia en comedor, tiempo total de estancia en jardín, Tiempo total de estancia en otra área del hogar. Para obtener esta información se utilizaron dispositivos *bluetooth* llamados *Beacons* que se comunican con la aplicación desarrollada.
e) Módulo de monitoreo fuera del hogar: en este módulo se hace uso del GPS embebido en el Smartphone y monitorea las actividades que el adulto mayor realiza fuera de su hogar. En este módulo se obtiene la variable: Cantidad de lugares que visita.
f) Módulo de envío de mensajes motivacionales: este módulo envía mensajes emotivos al adulto mayor acordes al nivel de aislamiento social que se detecta, pueden existir 3 niveles de aislamiento social: Nulo, Bajo o Alto.
g) Módulo de notificaciones: este módulo permite generar una notificación emergente en la aplicación móvil del adulto mayor de que existe una persona que estará monitoreando sus actividades. Esto se realizó con el fin de que los cuidadores puedan intervenir en cambiar el estado del adulto mayor.

La aplicación para el cuidador del adulto mayor tiene como objetivo monitorear el nivel de aislamiento social que presenta el adulto y poder tomar acciones preventivas en caso de requerirlo.

A continuación se describe brevemente los módulos de la aplicación móvil del cuidador:

a) Módulo de registro de usuario: este módulo obtiene los datos personales del cuidador, su nombre completo, edad y género. Estos datos son utilizados por el sistema para posteriormente notificar al adulto mayor los datos de la persona que estará monitoreando sus actividades.
El acceso a los datos del adulto mayor, que ya se encuentran almacenados en el servidor remoto, se realiza cuando el cuidador conoce el código de ingreso que se le proporcionó al adulto mayor cuando instaló la aplicación en su teléfono.

b) Módulo Monitoreo: este módulo contiene la interfaz para que el usuario pueda ingresar a los datos del adulto mayor, las opciones que tiene disponibles el cuidador son las siguientes:

- Evaluar el nivel de aislamiento social: al ingresar a esta opción se le mostrará al cuidador una pantalla donde puede visualizar el nivel de aislamiento social detectado, se mostrará una imagen que tenga relación al nivel de aislamiento social encontrado.
- Consultar historial: A partir de que se ha instalado la aplicación en el teléfono del adulto mayor cada determinado tiempo se analizarán los datos para detectar el nivel de aislamiento social, cuando el familiar quiera consultarlo podrá hacerlo en la opción “Consultar historial”.

4.2 Modelo predictivo de aislamiento social

El modelo predictivo de aislamiento social está basado en un árbol binario. La determinación del grado de riesgo de aislamientos social se realiza recorriendo el árbol nodo por nodo, hasta llegar a la hoja que indica el nivel aislamiento social: alto, bajo o nulo. Cada nodo del árbol representa un atributo relevante. El recorrido se inicia en el primer nodo localizado en la parte superior del árbol. Se decide avanzar hacia un lado u otro del nodo, mediante las condiciones que posee cada rama. Esta condición se evalúa de acuerdo al valor del atributo. Se continúa recorriendo el árbol de la misma manera, nodo por nodo, evaluando las condiciones de las ramas, hasta llegar a una hoja. El grado de aislamiento social será el que indique la hoja final.

La implementación del modelo de detección de aislamiento social se realizó en uno de los módulos que conforman la aplicación del adulto mayor. Una vez que se ha realizado la recolección automática de los datos, la implementación del modelo será la

encargada de determinar el nivel de aislamiento social de acuerdo al valor recolectado de las variables relevantes.

5. EVALUACIÓN DEL SISTEMA DE SOFTWARE

En esta sección se detalla la segunda fase de la metodología propuesta. Se realizaron dos tipos de pruebas del enfoque propuesto: a) pruebas al sistema, que consiste en realizar pruebas funcionales a las dos aplicaciones desarrolladas y posteriormente a un grupo de control. B) aplicación de la escala de Lubben para contrastar los resultados obtenidos por la aplicación.

5.1 Pruebas al sistema

Para esta tipo de pruebas se realizaron pruebas funcionales a ambas aplicaciones. La tabla 2 muestra las características de los equipos adonde se instalaron las aplicaciones.

Tabla 2. Datos de los equipos de prueba.

Aplicación probada	Modelo	Versión del sistema operativo Android.
Aplicación para el adulto mayor	Samsung Galaxy note 2	4.4.2
	ZTE Blade v6	5.0.2
Aplicación para el cuidador	Motorola X play	6.0.1
	ZTE Blade V6	5.0.2
	Galaxy note 2	4.4.2
	Galaxy note 3	5.0.2

Las pruebas funcionales se llevaron a cabo siguiendo la estrategia de pruebas del estándar IEEE 829-2008.

La Figura 3 muestra la pantalla de registro de la aplicación del adulto mayor.

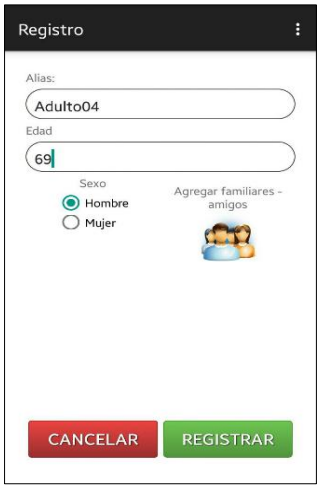


Figura 3. Pantalla de registro de la aplicación del adulto mayor

La pantalla de Registro de la aplicación móvil del cuidador se muestra en la Figura 4 y las opciones que tiene disponibles el cuidador se muestran en la Figura 5.

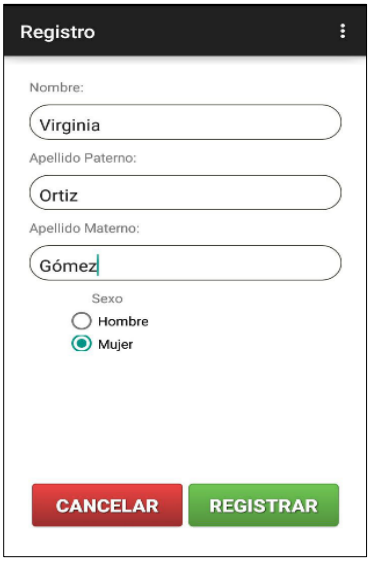


Figura 4. Pantalla de Registro de la aplicación del cuidador

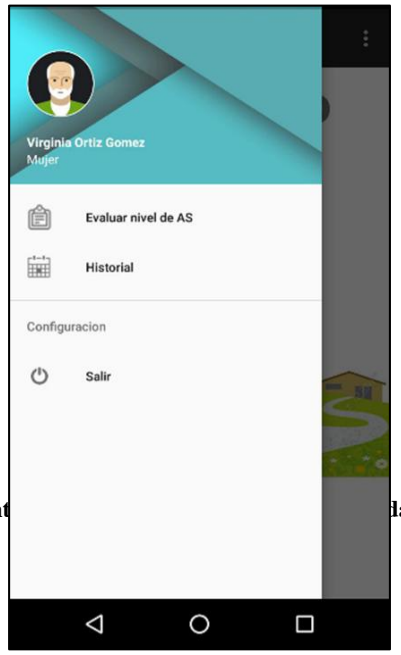


Figura 5. Pant lador.

5.2 Pruebas realizadas al grupo de control

El grupo experimental estuvo formado por siete adultos mayores, quienes cubrían el perfil requerido:

- Adulto mayor (60 años o más).
- Tener plenas capacidades cognitivas y motrices.
- Poseer un smartphone.

Tabla 3. Datos recolectados

Género	Cantidad de llamadas entrantes de la familia	Duración promedio de las llamadas entrantes de la familia(min)	Cantidad de llamadas entrantes de los amigos	Duración promedio de las llamadas salientes a la familia (min)	Cantidad de mensajes enviados a los amigos	Cantidad de mensajes recibidos de la familia	Tiempo total de estancia en recamara (hrs)	Tiempo total de estancia en sala (hrs)	Tiempo total de estancia en comedor (hrs)	Tiempo total de estancia en jardín (hrs)	Tiempo total de estancia en otra área del hogar (hrs)	Cantidad de lugares que visita	Nivel de Aislamiento Social
Femenino	128	256	64	40	0	0	182	111.87	150.40	36.80	149.80	84	Bajo
Masculino	4	120	24	24	0	0	410.8	60.20	177.27	44.13	1.47	156	Bajo
Femenino	12	60	64	24	0	0	364.93	68.20	21.47	1.93	0.00	100	Bajo
Femenino	20	260	44	0	68	48	48.25	0.72	7.20	0.65	3.92	164	Bajo
Femenino	8	12	4	20	0	0	601.04	24.80	36.80	12.56	129.80	4	Alto
Masculino	56	560	4	0	0	0	275.07	79.33	36.27	63.33	242.27	8	Alto
Femenino	24	270	16	48	12	32	214.4	53.93	8.27	109.27	57.40	64	Bajo

El procedimiento de la prueba consistió en que el adulto mayor debía traer consigo un Smartphone con la aplicación instalada durante el tiempo que durara la prueba. No se necesitó de ninguna interacción por parte de adulto mayor con la aplicación. El adulto mayor únicamente participó en el registro de sus datos y el de sus contactos, el monitoreo de las actividades se realizó de manera automática.

Para llevar a cabo las pruebas se instaló la aplicación móvil en el teléfono del adulto mayor y además se colocaron en las habitaciones de su hogar los sensores de presencia denominados Beacons para contabilizar la cantidad de tiempo que el adulto mayor pasa en cada una de las áreas de su hogar. Se le indicó al adulto mayor que estos sensores no obtienen información de audio o video.

Estas pruebas se realizarán en dos fases. En la primera fase, se recolectaron los datos de las variables relevantes. Para que las pruebas fueran realizadas correctamente se tomaron en cuenta algunas recomendaciones

El encargado de aplicar el experimento realizó las siguientes actividades:

a) Colocó los Beacons en las siguientes habitaciones: recámara, sala, comedor, jardín y otra área del hogar (estudio, garaje, cocina, bodega, etc.). Para la facilidad de la colocación de los sensores, estos se encontraban etiquetados con el nombre de la habitación en la que se deben colocar. P
b) Para evitar la interferencia de la señal de los Beacons se colocaron al menos con dos metros de distancia uno del otro.

Al adulto mayor se le proporcionaron las siguientes recomendaciones: llevar con él todo el tiempo el teléfono que se le proporcionó, al finalizar el día debía dejar conectado el teléfono a la corriente eléctrica para que al día siguiente pudiera seguir recolectando información.

En la segunda fase después del monitoreo se aplicó la escala LSNS-R al grupo experimental y se obtuvo el nivel de aislamiento social evaluando esta escala.

5.2 Aplicación de la escala de Lubben

La escala de Lubben Social Network Scale [11] fue el instrumento psicológico utilizado como escala de referencia para obtener el nivel de aislamiento social en el adulto mayor. Esta escala posee las siguientes características:

La escala hace diferencia entre familiares y amigos, además se cuenta con una versión en español

Esta escala consta de 6 preguntas tipo likert con un rango de puntuación de 0 a 30, donde 30 puntos indican la ausencia de aislamiento social, 12 puntos o más, indican un bajo riesgo de aislamiento social y menos de 12 puntos un alto riesgo de aislamiento social[12]. Además esta escala cuenta con una versión es español, adecuada al perfil del grupo de estudio.

5.3 Comparación de resultados

En la tabla 3 se muestra la tabulación de los registros realizados al grupo de control en la cual se puede observar el total de cada una de las variables y nivel de aislamiento social detectado por la aplicación móvil.

La Figura 6 muestra la relación de actividades para cada uno de los usuarios del grupo de control.

La Figura 7 muestra los niveles de aislamiento social del grupo experimental obtenidos con la escala de valoración LSNS-R y con la aplicación móvil desarrollada en este trabajo de investigación. El modelo de detección de aislamiento social implementado en la aplicación móvil clasificó correctamente siete de los siete participantes pertenecientes al grupo de control.

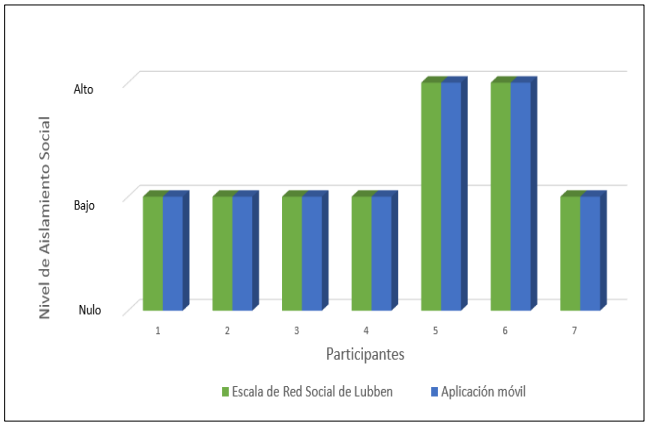


Figura 7. Comparación de resultados entre la escala de red social de Lubben y la Aplicación Móvil.

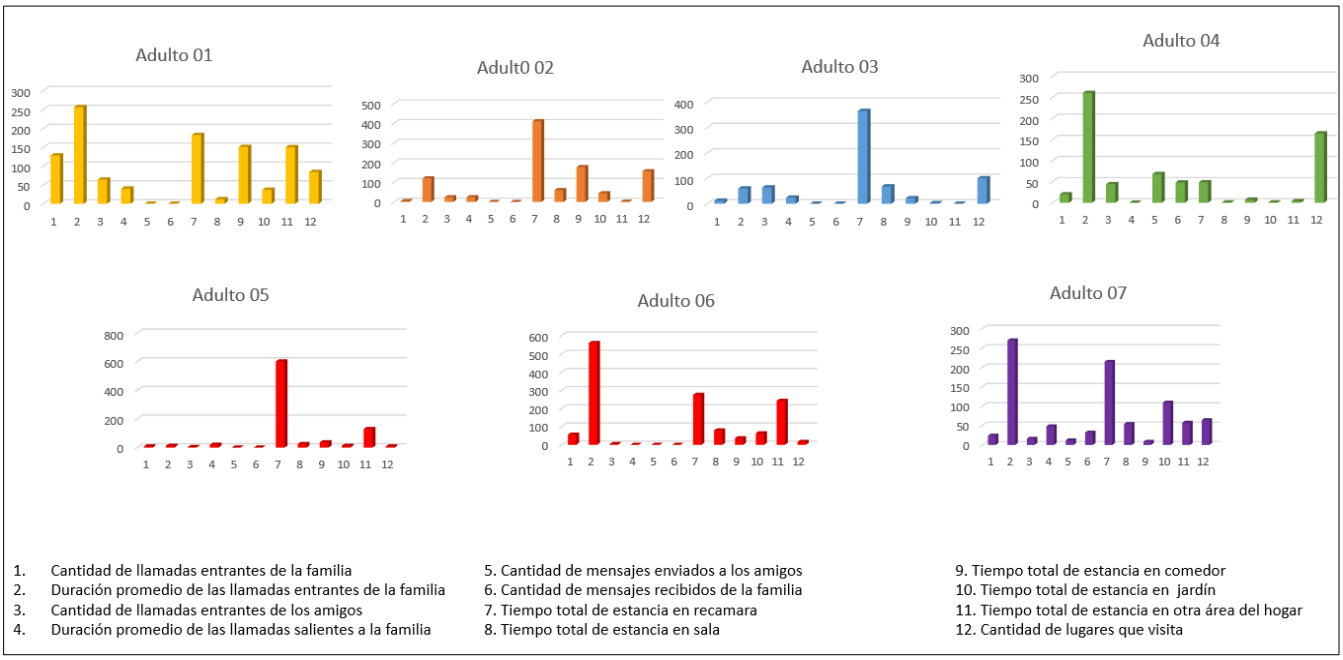


Figura 6. Relación de actividades de los participantes.

El modelo implementado fue probado en un trabajo anterior [3], en el cual la recolección de los datos fue de forma manual, los datos recolectados en ese trabajo se muestran en la tabla 4.

Tanto los datos recolectados manualmente como los datos recolectados de forma automática fueron introducidos al modelo predictivo con el fin de incrementar el número de los sujetos considerados en el experimento. Los resultados obtenidos de la comparación de los resultados obtenidos por la escala de Lubben y el modelo propuesto en el trabajo de investigación se muestran en la figura 8, donde se puede notar que el modelo clasifica correctamente 14 de las 15 instancias proporcionadas, con lo que se obtiene una precisión del 93.33%.

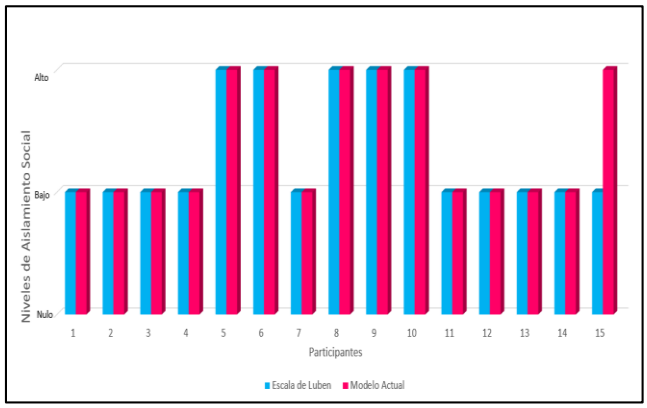


Figura 8. Comparación de resultados entre la escala de red socia de Lubben y el modelo propuesto.

6. CONCLUSIONES

En este trabajo de investigación se implementó un modelo de detección de aislamiento social desarrollando una aplicación móvil que de manera automática, detecta el nivel de aislamiento social de un adulto mayor.

Se desarrolló una aplicación móvil para que el cuidador del adulto mayor pueda estar informado del grado de aislamiento social que este presenta. La evaluación inicial del enfoque propuesto permitió determinar que, con el sistema que implementa el modelo de detección de aislamiento se obtuvo una precisión del 100%, aunque debe realizarse pruebas con más casos para determinar con precisión la efectividad del enfoque. Por otra parte, con los datos recolectados en un trabajo anterior, la precisión fue del 93.3%.

Tabla 4. Datos recolectados en el trabajo antecedente.

Género	Cantidad de llamadas entrantes de la familia	Duración promedio de las llamadas entrantes de la familia(min)	Cantidad de llamadas entrantes de los amigos	Duración promedio de las llamadas salientes a la familia (min)	Cantidad de mensajes enviados a los amigos	Cantidad de mensajes recibidos de la familia	Tiempo total de estancia en recamara (hrs)	Tiempo total de estancia en sala (hrs)	Tiempo total de estancia en comedor (hrs)	Tiempo total de estancia en jardín (hrs)	Tiempo total de estancia en otra área del hogar (hrs)	Cantidad de lugares que visita	Nivel de Aislamiento Social
Femenino	2	90	8	450	9	0	0	90	14	28	1	90	Alto
Masculino	2	2	2	10	60	0	0	45	91	43	153	0	Alto
Masculino	2	4	0	120	300	0	0	30	42	15	16	0	Alto
Femenino	1	10	0	150	0	0	0	30	10	29	45	0	Bajo
Masculino	3	0	0	0	180	0	0	0	178	2	0	0	Bajo
Masculino	2	0	60	0	120	0	0	0	60	13	0	0	Bajo
Masculino	4	16	16	240	240	0	0	0	122	45	0	15	Bajo
Femenino	2	4	4	8	8	0	0	0	5	60	135	0	Bajo

7. REFERENCIAS

[1] «Día Internacional de las Personas de Edad». [En línea]. Disponible en: http://www.un.org/es/events/olderpersonsday/. [Accedido: 06-abr-2015].

[2] A. Martinez y W. Campos, «Correlación entre Actividades de Interacción Social Registradas con Nuevas Tecnologías y el grado de aislamiento Social en los Adultos Mayores», *Revista Mexicana de Ingeniería Biomédica*, pp. 181-190, 2015.

[3] W. Campos, A. Martinez, W. Sanchez, H. Estrada, J. Favela, y J. Perez, «Inferring Social Isolation in Older Adults through Ambient Intelligence and Social Networking Sites», *Comput. y Sist.*, vol. 20, pp. 143–152, 2016.

[4] W. Campos, «Detección de aislamiento social en adultos mayores a través de inteligencia ambiental y sitios de redes sociales en linea», Centro Nacional de Investigación y Desarrollo Tecnológico, 2015.

[5] D. Tang, B. Yusuf, J. Botzheim, N. Kubota, y C. S. Chan, «A novel multimodal communication framework using robot partner for aging population», *Expert Syst. Appl.*, n.º January, 2015.

[6] J. C. Castillo, A. Fernández-Caballero, Á. Castro-González, M. A. Salichs, y M. T. López, «A framework for recognizing and regulating emotions in elderly», pp. 320-327, 2014.

[7] M. Rodríguez, V. Gonzalez, J. Favela, y P. Santana, «Home-based communication system for older adults and their remote family», *Comput. Human Behav.*, vol. 25, n.º 3, pp. 609-618, 2009.

[8] D. Dasgupta, K. Feldman, D. Waghray, P. Willaert, D. A. Raybold, y N. V Chawla, «An Integrated and Digitized Care Framework for Successful Aging», pp. 440-443, 2014.

[9] Y.-F. Tan-Hsu Tan-IEEE Senior Member, Munkhjargal Gochoo, Ke-Hao Chen, Fu-Rong Jean y and C. F. H. Chen, Fu-Jin Shih, «Indoor Activity Monitoring System for Elderly Using RFID and FitBit Flex Wristband».

[10] S. S. and J. L. Bruno M. C. Silva, Joel J. P. C. Rodrigues, Tiago M. C. Simoes, «An Ambient Assisted Living Framework for Mobile Environments», 2014.

[11] J. L. and M. Gironda, «Measuring Social Networks and Assessing their Benefits», *Soc. Networks Soc. Exclusion Sociol. Policy Perspect. ed Burlingt.*, pp. 20-34, 2004.

[12] and A. E. S. J. Lubben, E. Blozik, G. Gillmann, S. Iliffe, W. v. R. Kruse, J. C. Beck, «Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations», *Gerontologist*, vol. 46, p. 11, 2006.

Uso de Red Neuronal para Estimación del Estrés en Personas Jóvenes en Actividades Sedentarias

Jorge H. Robles
Instituto Tecnológico de Sonora
Ciudad Obregón, México
jorgeh.rch@gmail.com

Cesar R. Encinas
Instituto Tecnológico de Sonora
Ciudad Obregón, México
cesar_rer@hotmail.com

Jesús G. Zarate
Instituto Tecnológico de Sonora
Ciudad Obregón, México
jesuszarate175@gmail.com

Joaquín Cortez
Instituto Tecnológico de Sonora
Ciudad Obregón, México
joaquin.cortez@itson.edu.mx

Jessica Beltrán-Márquez
Instituto Tecnológico de Sonora
Ciudad Obregón, México
jessica.beltran@itson.edu.mx

Luis A. Castro
Instituto Tecnológico de Sonora
Ciudad Obregón, México
luis.castro@acm.org

RESUMEN

En el presente trabajo se presenta la implementación de un sistema para la detección de estrés basado en sensores para la medición del pulso cardiaco y flujo respiratorio. Para las mediciones de las señales físicas mencionadas, se utilizan sensores que envían información a una computadora donde se realiza el procesamiento a través de una red neuronal previamente entrenada para clasificar si una persona está experimentando estrés o no. Aun cuando no se encontró en la literatura trabajos que implementen el método presentado, existen trabajos con similitudes los cuales se discuten en el presente trabajo. Los resultados de este trabajo son alentadores (96.7% precisión) para detectar cuando personas jóvenes estén experimentando episodios de estrés en sus vidas.

CCS Concepts

Applied computing→ Life and medical sciences→ Health informatics

Palabras clave

Estimación de Estrés; Redes neuronales; Sensores.

1. INTRODUCCIÓN

El estrés es una consecuencia del ritmo de vida de en la sociedad actual [1]. En la actualidad las personas se ven sometidas a más exigencias y presiones que son difíciles de afrontar, las cuales son generadoras de este problema. Las causas del estrés son muy variadas, entre ellas podemos mencionar el trabajo, las finanzas personales, las relaciones, el entorno, distintos cambios y pérdidas personales [1]. En el trabajo presentado en [2], se menciona que las evaluaciones y el tiempo limitado, son las principales causas de estrés en estudiantes universitarios.

El estrés trae consigo algunas consecuencias que tienen impacto en la vida de las personas. Entre los principales problemas psicológicos y fisiológicos se encuentran: episodios depresivos, brotes esquizofrénicos, problemas sexuales, problemas de sueño, hipertensión, alteraciones del sistema inmunológico, deterioro del rendimiento laboral, trastornos cardiovasculares, problemas respiratorios, y problemas gastrointestinales [3]. Además, también puede ser causa de sobrepeso en algunos grupos de población, si es experimentado por tiempo prolongado [4].

De igual forma, el estrés es una respuesta del organismo que pone al individuo en disposición de afrontar las situaciones interpretadas como estresantes que pueden ser de tres tipos: pérdida, amenaza y desafío [5]. Para afrontar dichas situaciones, el cuerpo debe proveerse de manera rápida de suficiente energía. Es por ello que las células comienzan a catabolizar la energía acumulada en el organismo [6]. Las células, al trabajar tan rápido, requieren nutrientes de forma continua. La sangre, que contiene esos nutrientes, debe de circular más rápido, lo que ocasiona que el ritmo cardiaco se acelere. Se ha reportado que existe una fuerte correlación entre el ritmo cardiaco y el nivel de estrés [7].

Otra consecuencia del estrés es los cambios en la velocidad de respiración. De esta forma el ritmo cardiaco y la respiración se sintonizan para poder llevar el máximo de recursos disponibles a las células. En [8] se reporta una relación entre el ritmo cardiaco, la respiración y el estrés.

En otras áreas de estudio se ha reportado que existen diferentes formas para combatir este problema. Algunas de ellas son: la relajación, desarrollar un estilo de vida sano y activo o incluso la medicación. Una persona debe poder identificar si está padeciendo de estrés para poder tomar las medidas necesarias para combatirlo. Existen varias tecnologías que se han desarrollado para encontrar los niveles de estrés, algunos de los cuales se presentan en la sección 2 del presente trabajo.

El presente trabajo muestra el desarrollo de un sistema para determinar si una persona está experimentando estrés o no, basándose en señales del flujo de aire (respiración) y la frecuencia cardiaca. Los datos obtenidos al medir estas señales se introducen a un algoritmo diseñado para clasificar si una persona está experimentando estrés, el cual consiste principalmente en una red neuronal artificial (RNA) [10]. Además, se presentan los resultados del experimento realizado para verificar el funcionamiento del sistema. Para finalizar se muestran las conclusiones y el trabajo futuro propuesto.

2. TRABAJOS RELACIONADOS

El uso de dispositivos tecnológicos para tratar de estimar estrés en individuos ha sido abordado con anterioridad. Por ejemplo, en [9], se presenta una investigación de diferentes trabajos para medir estrés, los cuales utilizan distintas señales como la presión sanguínea, ritmo cardiaco, variabilidad del ritmo cardiaco, conductancia de la piel y el diámetro de la pupila. Generalmente para determinar el estrés, se han utilizado distintas pruebas que se basan en instrumentos psicométricos [5].

En un trabajo propuesto en [11], se realizó un sistema el cual permite medir el estrés en un ser humano a partir de los

parámetros: respiración, postura, tensión muscular, conductancia de la piel, y los niveles de cortisol y prolactina. El séptimo parámetro corresponde al estado psicológico en el que se encuentra la persona, valorado a través de un cuestionario. Este sistema proporciona como salida el nivel de estrés con el que la persona se encuentra. Una desventaja de este trabajo, a diferencia de nuestra propuesta, se debe a que es necesario un análisis de la muestra de sangre, que es analizada separada del sistema, lo cual ocasiona que la respuesta del sistema no sea en tiempo real.

En otro trabajo, se considera el efecto del parpadeo involuntario como indicio de que una persona puede estar estresada [12]. De esta manera, es posible detectar de una manera no intrusiva indicios de estrés en una persona. En este trabajo, se realizó un detector de parpadeos a través de visión por computadora y técnicas de aprendizaje máquina. A diferencia del proyecto que aquí se presenta, este sistema no detecta si la persona padece o no de estrés, sino que sólo muestra las conexiones entre éste y el parpadeo, dejando pendiente investigar más a fondo esta relación.

Hay otros trabajos que utilizan los teléfonos celulares para detección de estrés. En [13], utilizan los teléfonos celulares como detectores, esto mediante información derivada del usuario, tales como: llamadas, SMS, e-mails, además de condiciones del clima y rasgos de personalidad. Al igual que en el proyecto presentado en este artículo, se realiza una clasificación de 2 clases, si se está o no experimentando un episodio de estrés en la persona. En [12] y [13], utilizan el algoritmo Random Forest.

3. IMPLEMENTACIÓN DEL SISTEMA

Los sensores utilizados se obtuvieron del paquete biométrico *e-Healt Sensor Platform Complete Kit V2.0 for Arduino*, *Rasperry Pi and Intel Galileo [14]*, el cual cuenta con 8 sensores biomédicos, sensor de pulso y oxígeno en la sangre (SPO2), sensor de flujo de aire, sensor de temperatura corporal, sensor de electrocardiograma (ECG), sensor de presión de sangre, sensor de posición (acelerómetro), sensor de electromiografía (EMG), y sensor de respuesta galvánica de la piel. Este paquete es utilizado en múltiples sistemas para el monitoreo de pacientes en tiempo real, como se presenta en [15], que ofrece una respuesta propia, a base de un sistema experto, o como se presenta en [16] que realiza una recolección de datos para ser interpretados por médicos. En la Figura 1 se presenta información del sistema implementado en este trabajo.

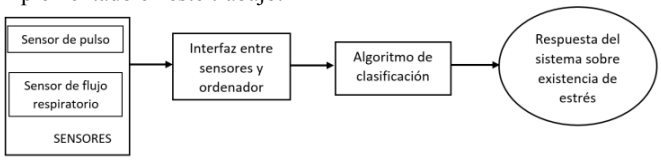


Figura 1. Diagrama a bloques del sistema



Figura 2. Sensor de respiración (izquierda); Sensor de pulso cardiaco (derecha). Fuente: [14].

El sensor de flujo respiratorio (Figura 1) se utiliza para medir la proporción de respiración en un paciente. El sensor consiste de un hilo flexible, el cual se coloca detrás de las orejas y dos pequeñas púas que se colocan en las fosas nasales, las cuales miden la respiración [17]. En la Figura 2 (izquierda) se puede observar el sensor y su colocación.

El sensor de pulso cardiaco es un sensor no invasivo que mide pulsaciones por minuto y la saturación de oxígeno en la sangre [17]. En el presente trabajo solamente se utiliza el pulso. De manera que éste coloca su dedo índice dentro del sensor, éste último tiene un tiempo para estabilizarse y encontrar la frecuencia cardiaca. El sensor y su colocación se pueden observar en la Figura 2 (derecha). Como interfaz entre los sensores y el ordenador (Figura 1), se utilizó el shield para Arduino Uno (Figura 3).



Figura 3. Conexión de sensores e interfaz

Para la obtención de las señales de los sensores se implementó un programa en Arduino. El algoritmo consiste en un ciclo que realiza una medición del flujo de aire cada 50ms, esto durante 4 segundos, resultando en 80 mediciones. Las mediciones son leídas como cadenas de bits, con valores entre 0 y 1023, el valor de las 80 mediciones es promediado y se entrega como salida para la entrada del programa que ejecuta la red neuronal.

El sensor de pulso realiza 50 mediciones para reducir la latencia y estabilizar el sensor, el cual, al dar la lectura en Arduino, entrega el valor correspondiente a los latidos por minuto como un valor entero.

Enseguida, se inicializa la comunicación serial en Arduino a 9600 baudios y en un ciclo infinito se colocan las rutinas de medición de los datos, las cuales imprimen el valor por el puerto serie de cada sensor. Debido a que el sensor de respiración requiere de 4 segundos para la toma de mediciones, se envían datos por el puerto serie cada 4 segundos.

Estos datos que se envían al puerto serie son enviados a una computadora, el cual lee los datos mediante un programa en MATLAB (Figura 1, Algoritmo de clasificación).

Estos datos se procesan mediante una RNA. Una RNA es una forma de computación inspirada en modelos biológicos, que permiten tener aprendizaje mediante entrenamiento o experiencia inicial. Cuando una RNA aprende, ésta puede diferenciar patrones para la clasificación de elementos [10]. La RNA es realizada en MATLAB utilizando la Neural Network Toolbox, misma que provee de algoritmos, funciones y aplicaciones para crear, entrenar, visualizar y simular redes neuronales. En la Figura 4 se puede observar un esquema de la RNA creada en MATLAB, misma que cuenta con 2 neuronas de entrada, 4 neuronas en la capa oculta y 1 neurona de salida. Se optó por utilizar 4 neuronas en la capa oculta, debido a que para este problema en específico proporciona un buen desempeño (96.7% de precisión) sin la necesidad de utilizar un gran número de neuronas.

Entonces, el sistema funciona introduciendo como entradas de la RNA los parámetros promediados del sensor de respiración y el

sensor de pulso cardiaco, produciendo como salida la respuesta si la persona está experimentando un episodio de estrés o no.

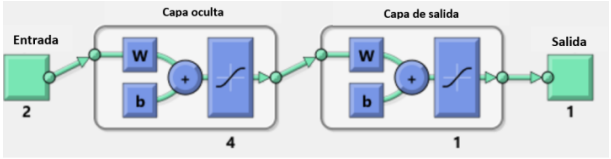


Figura 4. Esquema de la RNA entrenada

4. EXPERIMENTO

Para la realización del experimento se utilizó una muestra de 10 personas (10=Masculino) de entre 18 y 25 años, con una edad promedio de 22.4 años. A cada persona se le realizaron 10 diferentes mediciones en 2 condiciones. La primera cuando se encontraba en una situación catalogada no estresante y la segunda cuando se encontraba en situación catalogada como estresante. En total, cada participante proveyó 20 mediciones.

En el caso de la situación catalogada como no estresante, a los participantes se les colocó en un asiento cómodo, con música de relajación y jugando videojuegos con un nivel de dificultad baja. De esta forma los participantes podían estar relajados y al mismo tiempo con un nivel de concentración bajo. Se optó por someteros a esta concentración baja para evitar factores que pudieran afectar el experimento, tales como problemas personales por los que la persona estuviera pasando. Se tomaron 10 mediciones a cada participante, realizando cada medición con una diferencia aproximada de 3 minutos.

En el caso de las situaciones catalogadas estresantes, se colocó a los participantes en un asiento incómodo, con niveles elevados de ruido y presión psicológica (en forma de distracciones) por parte de los investigadores mientras los participantes jugaban los mismos videojuegos pero con un alto nivel de dificultad.

5. RESULTADOS

Al final de la recolección de datos, se obtuvieron 200 tuplas. Cada tupla contiene mediciones promedio del flujo de respiración y las del pulso cardiaco de los 10 participantes, en ambas condiciones. Es decir, 100 tuplas pertenecían a estados catalogados estresados y las restantes de los catalogados no estresados.

En la Figura 5 se observan los datos obtenidos en las mediciones. Los círculos de color azul representan mediciones cuando el sujeto se encontraba en condiciones catalogadas como sin estrés, mientras que las equis rojas son las mediciones cuando el sujeto estaba en condiciones consideradas como de estrés.

De la Figura 5, se puede observar que los participantes en condiciones catalogadas como estresantes, tienden a tener mayor pulso cardiaco y menor índice del flujo de exhalación. Tener un menor índice de flujo de respiración significa que el participante está inhalando y exhalando con mayor frecuencia. Por otro lado, los sujetos con condiciones catalogadas como no estresantes tienen menor pulso cardiaco y mayor intensidad de flujo de exhalación. Además, se puede ver el traslape entre los dos conjuntos donde en pulso cardiaco se tiene entre 74 y 76 y en respiración se tiene de 5 a 20, en el índice de respiración.

Estos valores se introdujeron a la red neuronal junto con la salida deseada en cada caso, ya sea esperando que la persona estuviera estresada o no, lo anterior con el objetivo de entrenar a la red. Las 200 tuplas recolectadas se distribuyeron aleatoriamente de la siguiente manera: 75% fueron utilizados para el entrenamiento de la red, 15% para pruebas y 15% para validación.

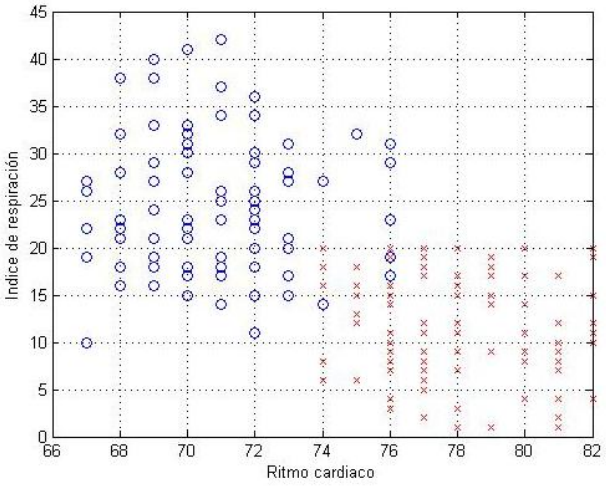


Figura 5. Mediciones correspondientes y clasificación en estresado (X rojo) y no estresado (O azul).

Una vez realizado el entrenamiento de la red, se verificó. Para esto se utilizó la matriz de confusión, que es una herramienta que permite la visualización del desempeño de un algoritmo de aprendizaje. Los resultados del análisis se muestran en la Tabla 1. Como se puede observar, se tiene un porcentaje de error bastante bajo, sólo del 3.3%, lo que demuestra que el sistema es bastante efectivo al momento de realizar la clasificación.

Tabla 1. Matriz de confusión (Clase 1= no estresado, Clase 2= estresado)

		Clase esperada		
		Clase 1	Clase 2	Precisión/Error
Clase obtenida	Clase 1	**14** **46.7%**	1 3.3%	93.3% 6.7%
	Clase 2	0 0.0%	**15** **50%**	100% 0.0%
	Precisión Error	100% 0.0%	93.8% 6.2%	**96.7%** **3.3%**

Algo importante a notar es que, de acuerdo a los resultados del experimento, sólo ocurrieron errores cuando la persona se encontraba en condiciones que fueron catalogadas como estresantes. Por el contrario, cuando la persona estaba en situaciones catalogadas como no estresantes se obtuvo un 100% de precisión en esta clase. Sin embargo, existieron errores en la segunda clase que fueron confundidos con la clase 1 (no estresado), lo que significa en realidad la predicción de una persona considerada sin estrés puede estar errónea en algunas ocasiones, y realmente tener estar en una situación de estrés.

6. DISCUSION Y TRABAJO FUTURO

El sistema que se implementó es el primer trabajo para medir estrés que utiliza la medición del pulso cardiaco y el flujo de aire junto a una red neuronal. A pesar de las limitaciones, este trabajo sienta precedentes para trabajos futuros.

En particular, el sistema es capaz de clasificar si una persona se encuentra en estado de estrés con una buena precisión. Sin embargo, el sistema no define el nivel de estrés con el que la persona se encuentra, solamente refiere si la persona se encuentra

o no en situaciones catalogadas como de estrés. Además, el algoritmo de clasificación utilizado es una red neuronal. Es recomendable comparar los resultados utilizando diferentes métodos como Maquina de Soporte Vectorial (SVM, por sus siglas en inglés) [18], que también puede ser adecuado para este problema en particular.

Las mediciones del rimo cardiaco y de respiración, dependen de factores como la edad y el sexo, al igual que del peso, o incluso difieren en personas con problemas circulatorios. Estos puntos se pueden tomar como parámetros de entrada en trabajos futuros. Además, para hacer las mediciones más precisas podrían tomarse otros parámetros como es el de la conductancia de la piel y la tensión muscular. En cuanto al experimento para la toma de muestras, éste se hizo en condiciones donde los participantes no realizaban actividad física considerable, por lo que las mediciones de ritmo cardiaco y respiración podrían variar de manera considerable en otro tipo de ambientes que requieran que del individuo ejecutar considerable esfuerzo físico.

En cuanto a la generalización de resultados, el sistema, se probó con 10 personas de 18 a 25 años, por lo que es necesario contar con una muestra más grande, así como poder estudiar cómo se diferencia el estrés en distintos grupos de población y en condiciones distintas (Ej., altura de la ciudad donde viven).

7. CONCLUSIONES

Para la muestra de participantes del estudio, se obtuvo un sistema para la verificación del estrés con una precisión del 96.7%.

Las redes neuronales son algoritmos que operan eficientemente para el reconocimiento de patrones, ya que si estas son entrenadas de forma correcta, éstas permitirán obtener una salida muy cercana a la deseada en la gran mayoría de las ocasiones. Sin embargo, también se pueden utilizar otros tipos de algoritmos como SVM.

Hay otros parámetros que se pueden incluir como la edad, sexo, estatura, o conductancia de la piel. Se espera que se le pueda dar seguimiento a este tipo de trabajos para la medición del estrés, ya que en la actualidad es un problema que está ligado a la causa de distintas enfermedades, y es importante que la persona pueda detectarlo, reconocerlo, y poder tomar medidas para combatirlo.

8. AGRADECIMIENTOS

Este trabajo fue financiado parcialmente por el Consejo Nacional de Ciencia y Tecnología (CONACYT) en México y el Instituto Tecnológico de Sonora (ITSON).

9. REFERENCIAS

[1] Lazarus R., Folkman. S., “Estrés y proceso cognitivos”, Roca, México, 1991

[2] Mendoza L., Cabrera M., Gonzáles D., Martínez R., “Factores que ocasionan estrés en estudiantes universitarios”, Revista de enfermería 4(3):35-45, dic 2010

[3] Labrador F., “Estrés: trastornos psicofisiológicos”, Eudema, Madrid, 1993

[4] Korkeila M, Kaprio J., Rissanen A., “Predictors of major weight gain in adult Finns: stress, life satisfaction and personality traits”, International Journal of obesity (1998) 22: 949-957, 1998

[5] González F., “Instrumentos de evaluación psicológica”, Ciencias médicas, La Habana, 2007

[6] Joseph-Bravo P., De Gortari P., “El estrés y sus efectos en el metabolismo y el aprendizaje”, Biotecnología, 14(3), 65-76, 2007

[7] Emil J., O’Donnell A., Raskovic D., Frank A., “Stress Monitoring Using a Distributed Wireless Intelligent Sensor System”, IEEE Engineering in Medicine and Biology Magazine, Volume: 22, Issue: 3, May-June 2003

[8] Sundblad P., Linnarsson D., “Relationship between breath-synchronous arterial pressure and heart rate variations during orthostatic stress”, Clin Physiol Funct Imaging., 23(2):103-9, Marzo 2003.

[9] Soniya Lakudzode, Prof.S.M.Rajbhoj, “Review on human stress monitoring system using wearable sensors”, International Research Journal of Engineering and Technology Volume: 03 Issue: 04, Apr-2016.

[10] Haykin Simon, “Neural networks”, Pearson education, Hamilton, 1999

[11] Acevedo M., Luca A. “Sistema de medición de estrés”, Revista Mexicana de Ingeniería Biomédica 22(1):20-25, 2001

[12] Marcos-Ramiro, A., Pizarro-Perez, D., Marron-Romera, M., Pizarro-Perez, D., & Gatica-Perez, D., “Automatic Blinking Detection towards Stress Discovery" In Proceedings of the 16th International Conference on Multimodal Interaction (pp. 307-310), ACM, 2014

[13] Bogomolov, A., Lepri, B., Ferron, M., Pianesi, F., & Pentland, A. S., “Daily stress recognition from mobile phone data, weather conditions and individual traits”, Proceedings of the 22nd ACM international conference on Multimedia (pp. 477-486). ACM, 2014

[14] e-Health Sensor Platform Complete Kit V2.0, disponible: https://www.cooking-hacks.com/ehealth-sensors-complete-kit-biometric-medical-arduino-raspberry-pi

[15] Makisimovic M., Vujovic V., Persic B., “A custom internet of things healthcare system”, Information systems and technologies (CISTI), 2015

[16] Absullah A., Ismael A, Rashid A., Abou-ElNour A., Rarique M., “Real time wireless health monitoring application using mobile devices”, International Journal of computer Networks & Communication (IJCNC), May 2015

[17] Fernández M., Sanchez D., “Desarrollo de materiales docents para asignaturas de Bioingeniería”, Universidad de Cádiz, Cádiz, 2014

[18] Cristianini N. Shawe-Taylor J. “An introduction to support vector machines”, Cambridge University Press, Cambridge, 2000

Una Propuesta de Ontología de Desórdenes de Conducta en Infantes

Andrea Tamborrell Hernández, Hugo R. Lasserre Chávez, María J. Somodevilla García, Ivo H. Pineda Torres, Concepción Pérez de Célis.
Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación
{andietamborrell, hugoraziel, mariajsomodevilla, ivopinedatorres,mcpcelish}@gmail.com

RESUMEN

Los desórdenes de conducta son un comportamiento de patrones persistentes en niños y adolescentes en el cual los derechos de otros o reglas básicas son violadas. Considerando este problema, se propone la creación de una ontología de desórdenes de conducta en el idioma español, que permitirá a quien la consulte, el poder obtener un diagnóstico de acuerdo a los síntomas presentados, así como también una orientación a un tratamiento que pueda llevar a mejorar dicha condición.

Palabras Clave
Ontología; Desorden de Conducta; Web Semántica.

1. INTRODUCCIÓN

Los desórdenes de conducta son una condición severa caracterizada por un comportamiento hostil y, en ocasiones violencia física, sin preocuparse por los demás. Infantes con problemas de conducta exhiben crueldad, desde empujar, pegar o morder hasta aprovecharse de los demás e intimidar (*bullying*). Se puede también involucrar el lastimar animales, buscar peleas, robar o incendiar. Desde la infancia y adolescencia los problemas de conducta normalmente desarrollan en el adulto una personalidad antisocial, que necesita ser corregida con tratamiento tan pronto como sea detectada, si este problema es detectado a tiempo, mejor será el pronóstico [3]. Una ontología representa un modelo conceptual describiendo cierto dominio, tipos de objetos y conceptos que existen, así como sus propiedades y relaciones.

Considerando la gravedad de los desórdenes de conducta infantil, las tecnologías de la información podrían utilizarse como un soporte para su detección, así como un de orientación para padres e hijos. La Web se ha convertido en el principal medio de comunicación en nuestros días y son las ontologías las que proveen su base de conocimiento semántica. En este trabajo se propone desarrollar una ontología de desórdenes de conducta infantil.

2. MARCO TEÓRICO

2.1 Ontologías

De acuerdo a [9] una ontología es una especificación explícita de una conceptualización, algunos de sus posibles aplicaciones y usos pueden ser:

- Repositorios para la organización del conocimiento.
- Servir como herramientas para la adquisición de información.
- Permitir compartir conocimiento.
- Posibilitar el trabajo cooperativo al funcionar como soporte común de conocimiento entre organizaciones, comunidades científicas, etc.

Las ontologías se pueden clasificar en:

- Ontologías generales: Describen conceptos generales, espacio, tiempo, materia, objeto.
- Ontologías de dominio: Describen el vocabulario de un dominio concreto del conocimiento.
- Ontologías específicas: Son ontologías especializadas para un campo limitado del conocimiento o una aplicación concreta.

Los elementos que componen a una ontología son:

- Clases: También conocidas como conceptos, suelen organizarse en taxonomías.
- Atributos: Conforman la estructura interna de la clase.
- Relaciones: Representan la interacción entre los conceptos del dominio a través de la taxonomía de dominio, por ejemplo: *subclase de*, *parte de*, etc.
- Axiomas: Expresiones asumidas como verdaderas, suelen usarse para definir el significado de los componentes ontológicos, definir restricciones etc.
- Instancias: Representan objetos o individuos de una determinada clase.

2.2 Publicación de Ontologías

La publicación de ontologías médicas se puede realizar en sitios como *biontology.org* para que estas puedan ser de libre acceso en todo el mundo. Dicho sitio contiene una gran base de datos con ontologías del área de la medicina y permite a cualquier usuario subir su trabajo ontológico para que este pueda ser consultado por personas alrededor del mundo. Además, este sitio permite al usuario que lo visita ver a detalle los datos de la ontología tales como el nombre, el identificador único e incluso tiene un apartado para visualizar la ontología de manera gráfica en línea.

3. ESTADO DEL ARTE

Existe trabajo previo en la FCC en el desarrollo de ontologías espaciales biomédicas a través de los proyectos en [10] [11] [12] [13].

Diferentes ontologías se han creado en el ámbito médico como la *Disease Ontology*, la cual tiene como finalidad agrupar un vocabulario médico controlado desarrollado en el campo bioinformático, en colaboración con el Centro de Medicina Genética de la Universidad Northwestern, Chicago. *Disease Ontology* fue diseñada para facilitar una relación de las enfermedades y las condiciones asociadas a ellas para códigos médicos particulares como ICD9CM[1], SNOMED[2] y otros. [9]

La ontología SNOMED[3], esta ontología incluye una terminología de 364,000 conceptos de cuidado médico, con significados únicos y definiciones basadas en lógica formal organizadas en jerarquías [9].

La mayor parte de los trabajos ontológicos desarrollados en el ámbito de medicina están desarrollados en el idioma inglés y los pocos trabajos hechos de desórdenes de conducta no abordan específicamente este problema en infantes, por lo cual este trabajo representaría una gran diferencia en dicha área y en los países de habla hispana.

4. DESORDENES DE CONDUCTA

Los desórdenes de conducta son una serie de problemas comportamentales y emocionales que se presentan en niños y adolescentes. Los problemas pueden involucrar comportamiento impulsivo o desafiante, consumo de drogas o actividad delictiva [7].

De acuerdo a [4], infantes con desorden de conducta pueden presentar cuatro tipos de comportamiento negativo como se detalla en 4.1.

4.1 Tipos de comportamiento

- Agresión hacia personas o animales
  - Acoso, intimidación y amenaza
  - Iniciar peleas físicas
  - Daño físico serio con el uso de armas (ladrillos, palos, cuchillos o armas)
  - Asalto sexual
- Destrucción de propiedad ajena
  - Daño intencional a la propiedad de otros
  - Iniciar incendios de manera deliberada
- Acciones como engañar, mentir y robar
  - Allanamiento
  - Mentir para obtener cosas o favores
  - Mentir para evitar obligaciones
  - Robar sin confrontación
- No respetar las reglas
  - Faltar a clases
  - Escapar de casa
  - Estar fuera de casa por las noches sin permiso de los padres

4.2 Causas de los desórdenes de conducta

Diferentes factores contribuyen a que una persona desarrolle desórdenes de conducta, incluyendo daño cerebral, abuso o negligencia, problemas genéticos, problemas escolares y experiencias traumáticas [4].

Existen diferentes tipos de factores de riesgo como:

- Biológicos:
  - Padres con diagnósticos de esquizofrenia, desorden de conducta, déficit de atención, hiperactividad, comportamiento antisocial y alcoholismo.
  - Hermano con desorden de comportamiento.
  - Historial familiar de comportamiento negativo, desafiante o desorden de comportamiento.
- Modo de vida
  - Disciplina dura o violenta.
  - Padres solteros
  - Divorcio
  - Abuso emocional, físico o sexual
  - Negligencia
  - Uso de drogas o alcohol por parte de algún familiar.
  - Exposición a violencia
  - Falta de supervisión
  - Rechazo familiar

4.3 Tratamientos e intervenciones

4.3.1 Intervenciones psicosociales

Sin intervención es probable que los desórdenes de comportamiento negativo (*Disruptive Behavior Disorders*) progresen [6]. Hay un número de tratamientos prometedores que están disponibles y que cuando son completados pueden llevar a grandes beneficios. Una investigación en [1] demostró que las personas que completan la "Terapia de Intervención Padre-Hijo" muestran un cambio significativo al terminar dicha terapia.

En [5] se han identificado 16 tratamientos basados en evidencias para los desórdenes de comportamiento. Dos ejemplos son:

- PMT[4] "Entrenamiento para padres". Está dirigido a padres y les enseña a identificar antecedentes, comportamientos resultantes y consecuencias asociadas a sus hijos así como también a ellos mismos. Por último, el entrenamiento se enfoca en reforzar conductas deseadas.
- PCIT[5] "Terapia de interacción padre-hijo". Enfatiza mejoras en la relación entre los padres e hijos y ofrece herramientas para ayudar a manejar comportamientos que son negativos.

4.3.2 Antipsicóticos

Medicamentos antipsicóticos o neurolépticos son ampliamente utilizados en el tratamiento para el tratamiento de agresión aguda y crónica en varias poblaciones [14].

Estudios indican que antipsicóticos atípicos generalmente son más eficaces que un placebo en el tratamiento de agresiones, pero tienen diferentes efectos en cada persona.

En menores de edad con problemas de desórdenes de comportamiento, la *risperidona* es el medicamento más estudiado.

5. PROPUESTA DE SOLUCIÓN

5.1 Definición del problema

Como se explica previamente en la introducción, se realizará la creación de una ontología que trate de los desórdenes de conducta en infantes. Se optó por utilizar la metodología propuesta por [2] para la creación de dicha ontología. Las características de la metodología que se siguieron para la realización de esta ontología son:

- Diseño orientado a dominios, es decir la ontología está enfocada a un dominio en específico.
- Ontología Orientada a la reutilización, que la ontología pueda ser reutilizada por otros.
- Ontología Modular.
- Ontología Incremental, que sea posible aumentar el conocimiento almacenado en la ontología sin modificar su estructura general.
- Ontología Iterativa.

La técnica que propone [2] para la extracción de términos que poblarán la ontología es la creación de preguntas de competencia relacionadas con el problema a resolver las cuales son preguntas que se espera que la ontología responda al final de su realización.

Algunas de las preguntas de competencia para esta ontología son las siguientes:

- ¿Cuáles son los síntomas del desorden de conducta?
- ¿Es "X" un tratamiento adecuado para la actitud "Y"?
- ¿Qué causas originan un desorden de comportamiento?

[1] The International Classification of Diseases, Ninth Revision, Clinical Modification
[2] Estándar para el intercambio de información de salud clínica en Estados Unidos [8])
[3] Systematized Nomenclature of Medicine Clinical Terms
[4] Parent Management Training
[5] Parent-Child Interaction Therapy

- ¿Cuáles son los tipos de evaluación para diagnosticar desórdenes de comportamiento?
- ¿Cuáles son las consecuencias de no tratar un desorden de conducta?

Para la identificación de relaciones semánticas para el poblado de la ontología, de acuerdo a [2] es necesario:

1. Identificar las relaciones jerárquicas (taxonómicas).
2. Identificar las relaciones de tipo *data properties*. Es decir, preguntarse ¿Cuáles características son suficientes y necesarias para definir a un concepto?
3. Identificar las relaciones de tipo *object properties*. Es decir, ¿Se han cubierto todas las relaciones explicitas de las preguntas de competencia?
4. Identificar las características de las relaciones.

5.2 Evaluación de la Ontología

De igual manera, [2] propone tres métodos de evaluación para la ontología, los cuales consisten en:

1. Evaluación del conocimiento y conceptualización, es decir, el grado de cumplimiento del listado de las preguntas de competencia previamente establecidas.
2. Evaluación de la calidad, grado de cumplimiento con los principios de diseño.
3. Evaluación de rendimiento, velocidad de las respuestas que el razonador y/o inferenciador generarán.

El objetivo de la metodología que [2] propone es producir ontológicas reutilizables, usables y actualizables.

5.3 Diseño Conceptual

La ontología creada consta de diferentes clases, individuos, propiedades de individuos y propiedades de objeto (relaciones entre clases) como se muestra en la Figura 1. Las clases son:

- Causa
- Desorden de conducta
- Paciente
- Síntoma
- Tratamiento

Las relaciones entre clases se presentan en la Tabla 1

Tabla1. Relaciones entre clases

Relación	Dominio	Rango
tieneCausas	Desorden de conducta	Causa
tieneSintomas	Desorden de conducta	Síntoma
tieneTratamiento	Desorden de conducta	Tratamiento
recomiendaTratamiento	Paciente	Tratamiento
tieneDesorden	Paciente	Desorden de conducta

Dentro de cada una de las clases se tienen individuos, los cuales cuentan con propiedades que los identifican. Hasta la fecha se ha realizado un poblado manual parcial de la ontología. Los individuos más representativos por clase se muestran en la Tabla 2.

Tabla 2. Individuos por clase

Clase	Individuos
Causa	**Bilógicas** causaPadresAntisocial: causaPadresAlcoholismo causaPadresDeficitAtencion causaPadresDesordenConducta causaPadresEsquizofrenia causaPadresHiperactividad **Modo de Vida** causaAbuso causaDisciplina causaDivorcio causaFaltaDeSupervision causaNegligencia causaRechazoFamiliar causaUsoDrogas causaViolencia
Desorden de conducta	DesordenDeOposicionDesafiante desordenesDeConducta
Paciente	Contiene 100 individuos
Síntoma	sintomaAcoso sintomaActitudDesafiante sintomaAgresion sintomaAllanamiento sintomaAmenaza sintomaAsaltoSexual sintomaDestruccionPropiedad sintomaDiscutirAdultos sintomaEnojo sintomaEscaparDeCasa sintomaFaltarAClase sintomaIntimidacion sintomaMentir sintomaMolestar sintomaNoRespetarReglas sintomaPeleaFisica: sintomaProvocarIncendio sintomaRencorsintomaRobosintomaUsoDeArma sintomaRobo sintomaUsoDeArma
Tratamiento	tratamientoPCIT tratamientoPMT

Las propiedades de las clases se describen en la tabla 3.

Tabla 3. Propiedades de las clases

Clase	Propiedad
Causa	causaNombre causaTipo causaDescripcion

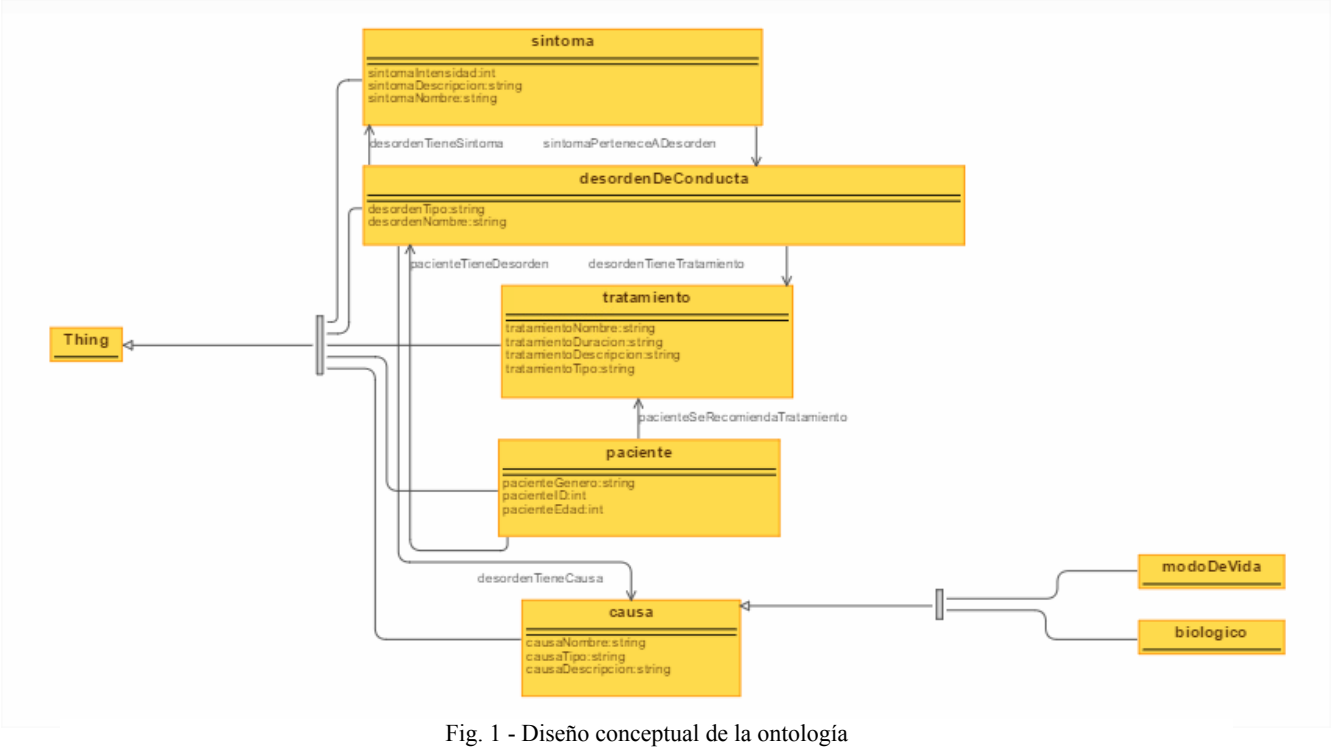


Fig. 1 - Diseño conceptual de la ontología

Desorden de conducta	desordenNombre desordenTipo
Paciente	pacienteEdad pacienteGenero pacienteID
Síntoma	sintomaNombre sintomaIntensidad sintomaDescripcion
Tratamiento	tratamientoNombre tratamientoTipo tratamientoDuración

5.3 CONSULTAS EN DL QUERY

Las reglas SWRL se definieron para responder las preguntas de competencia enlistadas en la sección 5.1. En la Tabla 4 se muestran dichas preguntas de competencia, así como la regla que le da respuesta.

Tabla 4. Definición de SWRL

Pregunta de Competencia	Sintaxis de Consulta
¿Cuáles son los síntomas del desorden de conducta?	sintoma and sintomaPerteneceA value "Desorden de Conducta"^^ xsd:string
¿Qué causas originan un desorden de oposición desafiante?	causa and causaOriginaUn value "Desorden de Oposicion Desafiante"^^ xsd:string

En la sección 6 se presentan los resultados de la clasificación reportados por Prótégé 5.0.0.

6. RESULTADOS

Al ejecutar un razonador sobre la ontología, esta no registra ningún error por lo cual se concluye que la ontología es consistente.

6.1 Ejecución de DL Query

La figura 2 y 3 muestran la ejecución de la primera y segunda preguntas de competencia descritas en la tabla 4.

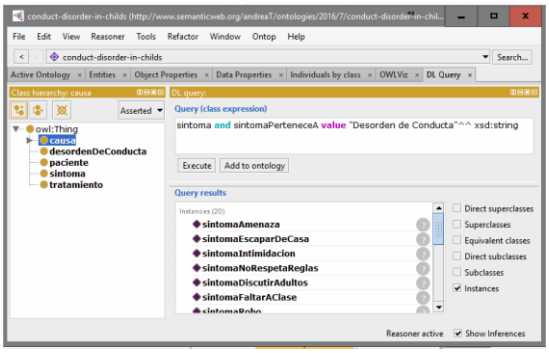


Fig. 2 - Ejecución de primera pregunta.

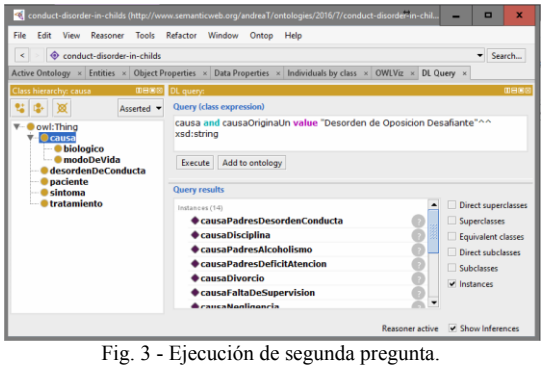


Fig. 3 - Ejecución de segunda pregunta.

7. CONCLUSIONES

Las alteraciones de la conducta en la infancia, han pasado a ser no solo el motivo de consulta más frecuente en psiquiatría infanto-juvenil, sino un motivo de alarma para la sociedad en general. La sociedad está experimentando un aumento de la presencia de comportamientos alterados en sus niños y adolescentes, siendo especialmente estremecedor el contenido de agresividad y violencia de dichos comportamientos. Se propuso una ontología de desórdenes de conducta infantil que considera las clases causa, desorden de conducta, paciente, síntoma y tratamiento. Se implementan además las relaciones entre clases que posteriormente se usaron en la definición de reglas SWRL para dar respuesta a las preguntas de competencia. La ontología propuesta es consistente de acuerdo con la validación realizada por Protégé 5.0.0.

La presente propuesta tendrá un impacto social puesto que se plantea incorporarla como base de conocimiento en un sistema para la toma de decisiones relacionado con desórdenes de conducta infantil en México.

Como trabajo a futuro se planea realizar una metodología para el poblado automático de la ontología generada en esta propuesta. Esta metodología comprende desde la búsqueda de casos clínicos, síntomas y tratamientos en diferentes sitios Web de la especialidad hasta la inserción de individuos de manera automática. La población automática permitirá el rápido crecimiento de individuos en la base de conocimientos para lograr una mejor precisión y orientación en los casos de desorden de conducta.

AGRADECIMIENTOS

Agradecemos a la Vicerrectoría de Investigación y Estudios de Posgrado de la BUAP por el soporte ofrecido para este proyecto.

REFERENCIAS

[1] Boggs SR, Eyberg SM, Edwards DL, Rayfield A, Jacobs J, Bagner D, Hood KK. Outcomes of parent-child interaction therapy: A comparison of treatment completers and study dropouts one to three years later. Child & Family Behavior Therapy. 2004.

[2] Bravo M, Perez J, Velazquez J, Sosa V, Montes A & Lopez M. Design of a shared ontology used for translating negotiation primitives. International Journal of Web and Grid Services 2, no. 3 (2006): 237-259.

[3] Child Mind Institute. Guide to conduct Disorder. What is it?

[4] Desórdenes de la Conducta. American Academy of Child Adolescent Psychiatry. 2004.

[5] Eyberg SM, Nelson MM, Boggs SR. Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. 2008.

[6] Gathright M & Tyler L. Disruptive Behaviors in Children and Adolescents. 2012.

[7] Medline Plus. “Transtorno de Conducta” Biblioteca Nacional de Medicina de los EE.UU. 2015

[8] National Institutes of Health. SNOMED CT. U.S. National Library of Medicine. 2016

[9] Sarrión A, Corrales Y. Un acercamiento a las ontologías médicas y su importancia. IV Jornada Científica de la SOCECS.

[10] Somodevilla M. “Desarrollo de Sistemas de Ontologías para descubrir patrones de estilo de vida asociados con enfermedades crónicas no transmisibles” SOGM-ING15-I VIEP 2015.

[11] Somodevilla M. “Desarrollo de técnicas para el procesamiento de consultas con datos espaciales biomédicos” SOGM-ING14-I VIEP 2014.

[12] Somodevilla M. “Desarrollo de Técnicas para Poblado Automático de Ontologías Espaciales Biomédicas” SOGM-ING15-I VIEP 2016.

[13] Somodevilla M. “UbiTIC Salud Ubicua Cómputo Ubiquo” CONACYT 2013

[14] Werry J.S. & Aman M,g. Methylphenidate and haloperidol in children: Effects on attention, memory and activity. 1975

Designing visualization tools to support older adults care process

Amanda Tapia
Sonora Institute of Technology
Ciudad Obregón
Sonora, México
amandatapiaz@hotmail.com

Jessica Beltrán
Sonora Institute of Technology
Ciudad Obregón
Sonora, México
jessica.beltran@itson.edu.mx

Valeria Soto-Mendoza
Center for Scientific Research and Higher Education at Ensenada
Baja California, México
vsoto@cicese.edu.mx

Karina Caro
Center for Scientific Research and Higher Education at Ensenada
Baja California, México
karicaro@cicese.edu.mx

ABSTRACT

Visualization enables analytic reasoning supported by an interactive visual graphical interface designed to easily interpret and assess large amounts of data. In the health care domain, visualization might help in the decision-making process and could provide clearer understanding of patients data. In this paper, two visualization tools are proposed to be used by the caregivers from a geriatric residence. Those tools are designed to show information related to the activities of the residents and identify possible abnormal behaviors. This paper describes the design of the two visualization tools and ends with plans for future work.

CCS Concepts

•Human-centered computing → Visualization design and evaluation methods;

Keywords

Visualization tools; older adults; geriatric residence, activity visualization, behavior automatic detection.

1. INTRODUCTION

Nowadays, data is generated in a large scale from multiple sources. In the health field, data can come from the readings of medical devices sensing vital signs [7], from manual annotations about the daily activities from patients [13], from laboratory tests [14], among others sources. While the vast volumes of data being stored are helpful to access to specific records, the main purpose for collecting health data is the extraction of information to build knowledge. For example, in the last decades health data has been used to automatically detect different behaviors of interest through machine learning techniques (e.g., abnormal behaviors) [11]. Another alternative for obtaining knowledge from health data is through visualization, which enables analytic reasoning supported by an interactive visual graphical interface designed to easily interpret and assess large amounts of data [9]. Visualization has gained importance in the health care domain because is helpful for decision-making and clearer understanding of patients data [4].

Visualization tools can be customized according to the needs from specific users. For example, by depicting information relevant to medical doctors [2], patients [10], nurses [6], family members [3] or even a whole community [1]. In this paper, two visualization tools are proposed to be used by the caregivers from a geriatric residence. Usually, the number of older adults living in geriatric residences surpasses the number of available caregivers. As a result, one caregiver can be in charge of several residents at the same time. In addition different shifts during the day might cause that one resident receives assistance from different caregivers. The caregivers constantly record older adults' activities generating large data sets of historical data [12]. While these data are valuable, they are usually not followed up because the shifts or just forgotten and never consulted again. We propose to display the data recorded from the residents in visualization tools, so the caregivers can easily have awareness about the overall well-being from each resident. Additionally, we propose to enhance the visualization tools by complementing the human reasoning with an automatic prediction module to know more about the state of the residents.

In the next section related work is mentioned. Then, the data collected from a geriatric residence is described followed by a description of the automatic module in section 4. In section 5, the two proposed prototypes of visualization tools are detailed. Finally, future work is presented.

2. RELATED WORK

Although raw data might seems meaningless, a proper visualization helps to identify trends and patterns turning raw data into knowledge. There are proposals that mainly show the information summarized to visualize key indicators that could be used to make decisions. However, such designs are not arranged to allow the identification of trends and patterns at a glance. An example is given by a community visualization dashboard encountered in [1]. The purpose of the dashboard is serving as a tool to understand the indicators that affect adults' quality of life based on their wealth. An example of the information depicted are statistics about adults with diabetes, by age, gender, ethnicity and the presence of cases of diabetes distributed by locations in maps

from the community. Another example of a tool designed to eyeball changes in routines or possible diseases is presented by the ORCATECH life laboratory in [8]. The data depicted are obtained through motion sensors installed in homes to detect the location of a person. The data are displayed using a two-dimensional spiral plot. However, it lacks on highly labeled semantic information, for example the activities performed by the person. Usually visualization tools are evaluated with the end-users to achieve an appropriated visualization. For example, in [10] an evaluation of a visualization tool related to personal health data is conducted. The authors designed a visualization of health data including physical, cognitive, social and spiritual information from an individual living in a smart home with the purpose of self monitoring of wellness and to provide a meaningful overview of long-term health progress. For the evaluation, the authors conducted focus groups with older adults to know the value of the visualization tool, the cognitive process and the integration of the data into the visualization tool. Another example is presented in [7]. The authors propose far-view displays to support nurses to recognize changes in patients condition, to identify alarms and to monitor infusion pumps. The tool was evaluated by asking to nurses to complete specific tasks and by analyzing the usability and the interpretation of the data, as well as the decision making speed and accuracy. Up to our knowledge, there is not related work about visualization tools designed for caregivers in geriatric residences that can be used to monitor the activities from the residents, to eyeball changes in their routines and that be complemented with the results from automatic classifying algorithms as is proposed in this paper. Existing tools support nurses with the information from patients in hospitals but not in geriatric residences nor with the activities performed by the patients. In the next part, a background about the type of data displayed is provided. Also, how those data are collected and the process to attain the automatic classification.

3. DATA FROM A GERIATRIC RESIDENCE

For about 6 months, a sensing campaign was conducted to collect data through electronic activity logs using smart phones in a geriatric residence in Rosarito, Baja California, Mexico. The activities were annotated manually by the caregivers, clicking a button from an application described in [13]. The total data obtained consist in 19,182 activity records from 15 residents with an average of 91 days annotated per resident and an average number of 8.63 activities per day. The recorded activities were: Anomaly, Check over, Toilet, Inside, Outside, Medication, Hygiene, Mood, Recreation, Social interactions, Feeding, Visits and Vitals.

4. AUTOMATIC PREDICTION MODULE

An approach to analyze the collected annotations consist in designing algorithms to automatically discover patterns that can be used to detect behaviors. In a previous work we developed an algorithm to detect abnormal days in residents daily life [11]. A time-based analysis is performed according the occurrence hour of the activities of the resident. Clustering algorithm is used to identify the hours when an activity occurred. Then every activity is evaluated and classified as normal or abnormal. Finally, depending on the number of each classified activity, the day is also classified as normal or abnormal. However, a time-based analysis to identify abnormalities is not enough. The analysis of the sequences of activities is also proposed in [11] through the sequential pattern mining method. A day is classified as normal or abnormal depending on how well the activities sequence of that day fits with the most frequent sequence (which is considered as normal). Nevertheless, this automatic classification of abnormalities remains as an open problem. We use the output of the developed algorithms to enhance the visualization tool by showing an abnormality light that serves as additional information to the caregiver.

5. VISUALIZATION TOOLS

In this section, the designs and descriptions of two visualization tools for activity records generated in a geriatric residence are presented. The proposed designs and elements from each tool were obtained from brainstorms with our research team including experienced researchers based on their familiarization with the context and the data collected, graphic designers, HCI and UBICOMP experts. We used the collected data and the output from the prediction module as the input of the visualization tools. The variation of each visualization tool relies in the time frame to display. For example, the first proposed visualization tool aimed to show the activities of the residents per hours of a day, showing the last 10 days of a specific resident. The second visualization tool aimed to show the whole sequence of activities per day, but with emphasis on observing patterns within a lager period of time, i.e. weeks. These designs are preliminary and we plan to validate each design with caregivers working in a geriatric residence.

Daily abnormalities spiral.

The Figure 1 depicts the activities performed from an older adult (i.e., resident 1) living in a geriatric residence during 10 days. The information displayed corresponds to data collected in a real scenario. The column of the left shows the symbols and colors of each activity and the 3D spiral graph in the right shows at what time the activities were annotated for the resident 1 during the last 10 days. Each circle which forms the spiral, corresponds to a specific day and the time over the day is distributed clockwise. The circle near to the center (i.e., most inner circle) corresponds to the day that was 10 days ago while the circle in the front (i.e., most outside circle) corresponds to the current day. As we can see from the distribution of the annotated activities, we can eyeball that resident 1 suffered an increase of medication and medical attention since the last 4 days. In this case, this happened because the resident passed through a surgery procedure and was indeed being medicated and receiving a follow up from the medical doctor. In the upper right is shown an abnormality light (colored yellow circle) that would be red if the current day is classified as abnormal, green if is classified as normal and yellow as neutral. Nevertheless, the caregivers or the medical doctor are the ones prepared to interpret properly what the information means. In this sense, the visualization tool permits the spotting and interpretation from persons related to the residents which enrich the decision making process comparing to the decision that an automatic recognition algorithm could take.

Colored daily maps.

The objective of this tool is to determine the similarities among days at a glance using different colors. The sequence

of activities from a day is represented with a sequence of colors. In Figure 2, activity records from one resident are presented. The general information about the resident is displayed on the top of colored bars. Every row is a different day, the dates are indicated in the first column. In the figure the period from *2013-07-05* to *2013-07-16* is shown. The colors to the right of the dates are the sequence of activities. Every color matches a specific activity and its corresponding code is exhibit in the chart. The number of activities per day is not standard, it depends on the resident care needs. However, the caregiver can observe certain differences among days. For example, *Feeding* records shown in color gray are within the first three records of the most of the days. Also, *Medication* records appear three consecutive times in all days. This means that the resident is being medicated with three different drugs. In the case that these three records missed one day, that might be abnormal. In some days *Toilet* records do not appear, which is also abnormal because it might be an indicator of a constipation of the resident or a simple forgetfulness from the caregiver. In any case, actions and decisions must be making about this fact.

6. FUTURE WORK

The above visualization tools are designed to display the same information considering different time frames. Their designs emerged from the analysis of data collected and experienced researchers in the context of older adults care. However, feedback and validation from final users of the designs are also needed. As future work, we plan to validate the proposed designs with caregivers from a geriatric residence. From caregivers, we want to obtain opinions and perceptions about what are the main useful characteristics from the prototypes? are the elements of each tool easy to understand? What elements need to be improved or changed? Is there other relevant information that is important to show? In addition, in the validation with the caregivers, we expect to answer specific questions about the design of the visualization tools, for example: Do these tools support the making decision process? why and how? Do these tools improve the quality of attention from caregivers to older adults? why and how?

Although, the validation phase with the potential users (i.e., caregivers) will bring improvements to the visualization tools, there are some methods reported in the literature that we can consider for improving the designs. For example, the daily activities contained in the collected data can be cast to a wellness model [5] to reduce the set of activities and make the visualization tools simpler. Moreover, every visualization tool can be evaluated taking into account the design goals to design visualization tools for older adults proposed in [10]. Then, these design goals from the perspective of older adults can be contrasted with the caregivers design goals extracted from an evaluation conducted with real caregivers. This could enable that the visualization tools are designed from both perspectives: older adults and caregivers. The application of these mixed design goals can be in a scenario where the older adult lives at his own home, but occasionally or certain hours per day he receives a visit of a caregiver to take care of him.

7. ACKNOWLEDGMENTS

We thank the residents, caregivers, nurses, staff and managers at Serena Center in Rosarito for their willing and collaboration in this project. We also appreciate the suggestions from Luis Castro, Edgar Chavez and Eduardo Quintana for their recommendations and ideas to improve the design of the visualization tools. This work was partially supported by National Council for Science and Technology (CONACYT) under grant number 18383.

8. REFERENCES

[1] Be Well Placer community dashboard. http://www.placerdashboard.org/index.php?module=indicators&controller=index note = Accessed: 2016-09-06.

[2] S. Amri, H. Ltifi, and M. B. Ayed. Towards an intelligent evaluation method of medical data visualizations. In *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, pages 673–678. IEEE, 2015.

[3] C. Bock, G. Demiris, Y. Choi, T. Le, H. J. Thompson, A. Samuel, and D. Huang. Engaging older adults in the visualization of sensor data facilitated by an open platform for connected devices. *Technology and Health Care*, (Preprint):1–10, 2016.

[4] M. Craft, B. Dobrenz, E. Dornbush, M. Hunter, J. Morris, M. Stone, and L. E. Barnes. An assessment of visualization tools for patient monitoring and medical decision making. In *Systems and Information Engineering Design Symposium (SIEDS), 2015*, pages 212–217. IEEE, 2015.

[5] H. L. Dunn. High-level wellness for man and society. *American journal of public health and the nations health*, 49(6):786–792, 1959.

[6] M. Görges, K. Kück, S. H. Koch, J. Agutter, and D. R. Westenskow. A far-view intensive care unit monitoring display enables faster triage. *Dimensions of Critical Care Nursing*, 30(4):206–217, 2011.

[7] M. Görges, D. R. Westenskow, and B. A. Markewitz. Evaluation of an integrated intensive care unit monitoring display by critical care fellow physicians. *Journal of clinical monitoring and computing*, 26(6):429–436, 2012.

[8] J. A. Kaye, S. A. Maxwell, N. Mattek, T. L. Hayes, H. Dodge, M. Pavel, H. B. Jimison, K. Wild, L. Boise, and T. A. Zitzelberger. Intelligent systems for assessing aging changes: home-based, unobtrusive, and continuous assessment of aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(suppl 1):i180–i190, 2011.

[9] T. Le, B. Reeder, H. Thompson, G. Demiris, et al. Health providersâĂŹ perceptions of novel approaches to visualizing integrated health information. *Methods Inf Med*, 52(3):250–258, 2013.

[10] T. Le, B. Reeder, D. Yoo, R. Aziz, H. J. Thompson, and G. Demiris. An evaluation of wellness assessment visualizations for older adults. *Telemedicine and e-Health*, 21(1):9–15, 2015.

[11] V. Soto-Mendoza, J. Beltrán, E. Chávez, J. Hernández, and J. A. García-Macías. Abnormal behavioral patterns detection from activity records of institutionalized older adults. pages 119–131, 2015.

[12] V. Soto-Mendoza and J. A. Garcia-Macias. Lessons learned from a long-running assistive system for

geriatric care. In *International Workshop on Ambient Assisted Living*, pages 83–90. Springer, 2014.

[13] V. Soto-Mendoza, J. A. García-Macías, E. Chávez, A. I. Martínez-García, J. Favela, P. Serrano-Alvarado, and M. R. Z. Rojas. Design of a predictive scheduling system to improve assisted living services for elders. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(4):53, 2015.

[14] P. C. Tang, J. S. Ash, D. W. Bates, J. M. Overhage, and D. Z. Sands. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, 13(2):121–126, 2006.

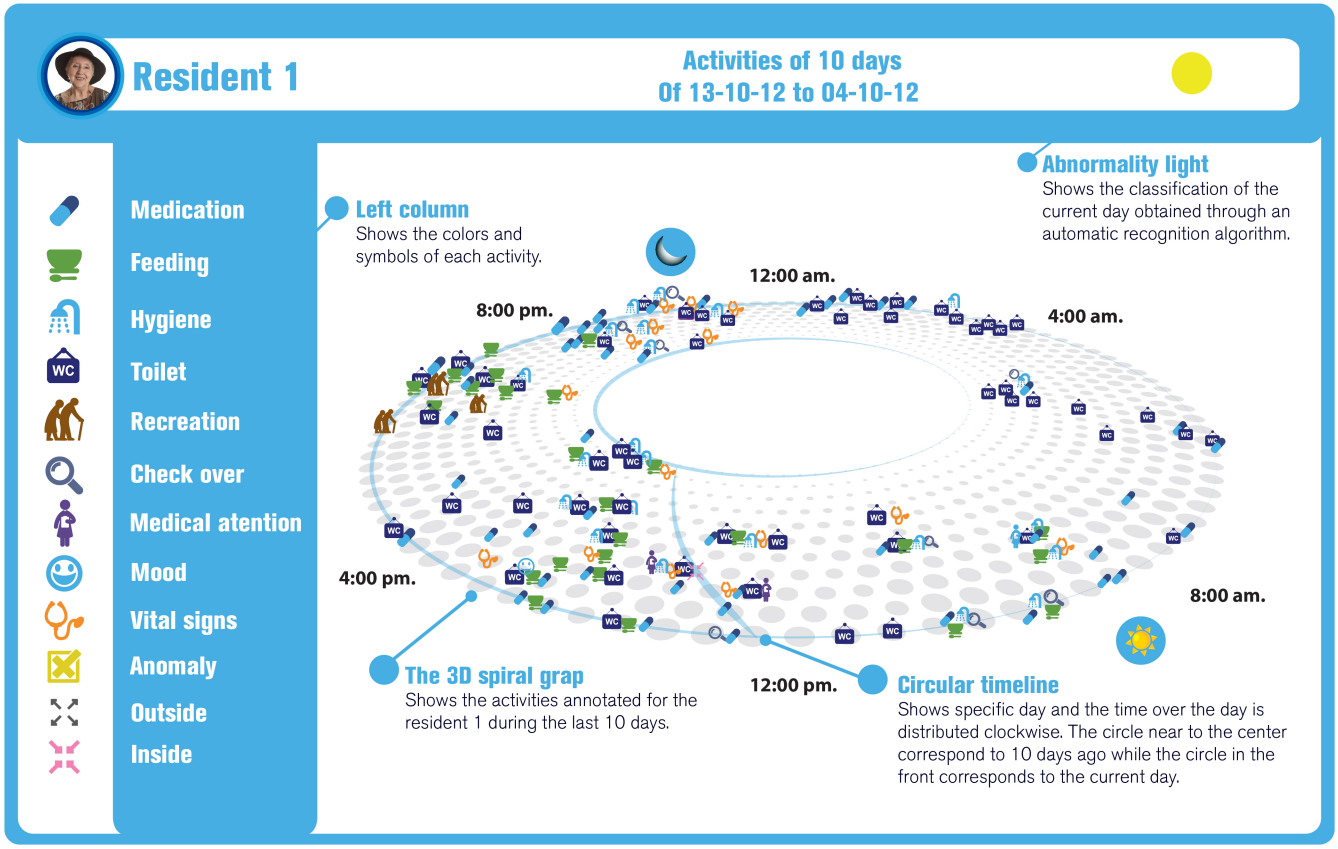


Figure 1: Daily abnormalities spiral.

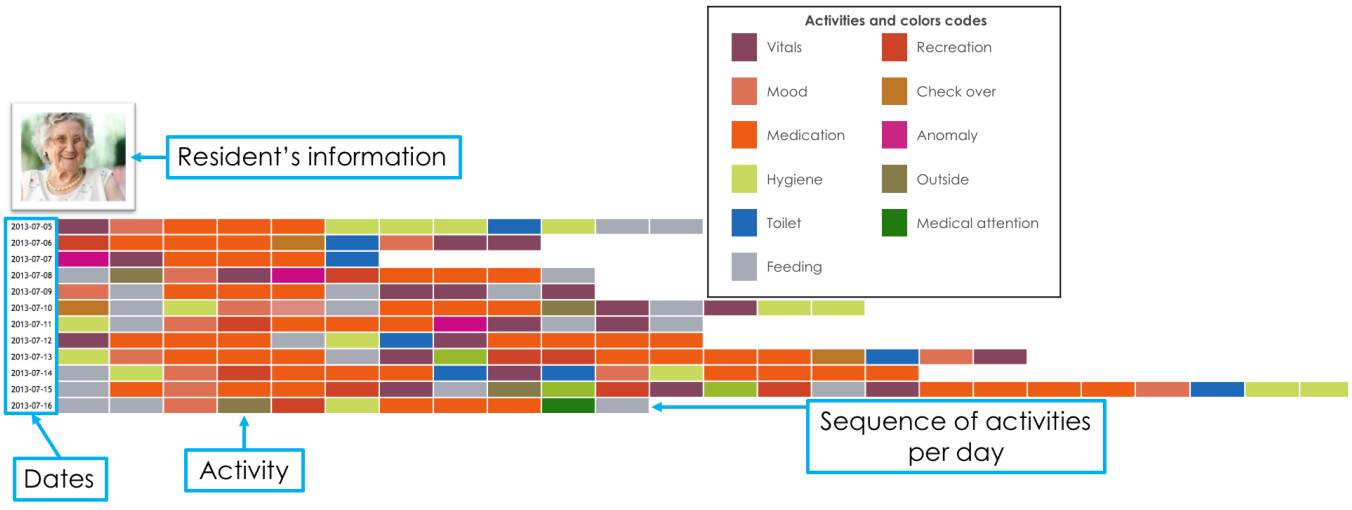


Figure 2: Colored daily maps.



Taller:
Inteligencia Artificial: Métodos y Aplicaciones

Organizadores:
Félix Fernando González Navarro (UABC)
Juan Carlos Nieves Sánchez (Umea University)
Fernando Orduña Cabrera (ITSC)
Jorge Eduardo Ibarra Esquer (UABC)

Aplicación de un Algoritmo de Clasificación para Generar un Mapa de Radio para Localización en Interiores

Isabel Lebasi Ambriz Silva
Universidad Autónoma de Baja California
Blvd. Benito Juárez S/N
Mexicali,Baja California
iambrizsilva@uabc.edu.mx

Juan Pablo García Vázquez
Universidad Autónoma de Baja California
Blvd. Benito Juárez S/N
Mexicali,Baja California
pablo.garcia@uabc.edu.mx

Ángel Gabriel Andrade Reátiga
Universidad Autónoma de Baja California
Blvd. Benito Juárez S/N
Mexicali,Baja California
aandrade@uabc.edu.mx

Luis Vizcarra Corral
Universidad Autónoma de Baja California
Blvd. Benito Juárez S/N
Mexicali,Baja California
luisvi@uabc.edu.mx

RESUMEN

La colaboración abierta distribuida es un nuevo paradigma que empodera a los usuarios a contribuir con datos detectados o generados en sus dispositivos móviles. Este paradigma está siendo utilizado en el área de localización para la creación automática de planos de un entorno de interiores, y para facilitar la colección de datos. En este artículo se presenta un experimento realizado para conocer la eficacia de la colaboración abierta distribuida para crear un mapa de radio con mediciones de intensidad de campo magnético en un entorno de interiores. Para generar el mapa de radio se experimentó con el algoritmo de clasificación llamado k vecinos más cercanos (del inglés: kNN-k Nearest Neighbors). Los resultados indican que el mapa de radio generado con kNN puede estimar la localización simbólica de un usuario con 92.3% de precisión.

Conceptos CCS

•Information systems → Data mining; Nearest-neighbor search;

Palabras Clave

mapa de radio; localización en interiores; colaboración abierta distribuida.

1. INTRODUCCIÓN

La localización física de un usuario se ha convertido en uno de los aspectos más importantes de su contexto, ya que permite generar sistemas que se adapten a su situación actual [5]. En escenarios al aire libre, el sistema de posicionamiento global (del inglés: GPS-Global Positioning System) ha tenido un tremendo impacto en nuestra vida cotidiana, ya que no sólo proporciona una guía para llegar a los lugares en automóvil o caminando, sino también como un apoyo en una amplia variedad de aplicaciones, en las que destacan las guías de turismo, sistemas para personas con capacidades diferentes (p. ej. invidentes, adultos mayores), sistemas de noticias, entre otros [17, 12, 10]. Sin embargo, en interiores los sistemas de localización basados en GPS, están limitados, debido a la complejidad de estos entornos, ya que en estos existe una mayor concentración de lugares, en comparación con los escenarios al aire libre. Además, en la literatura se reporta que sin la visibilidad de 3 o 4 satélites los sistemas de localización basados en GPS podría cometer errores mayores a 100 metros [15]. Por lo anterior sería una tarea difícil tratar de localizar a un sujeto en un entorno de interiores, ya que con un error de 50 metros o más, el individuo podría ya no estar ubicado en el mismo edificio.

Con el desarrollo de sensores y dispositivos más sofisticados, han surgido diversos enfoques tecnológicos que abordan el problema de localización en interiores, que van desde el uso de Wi-Fi, RFID, ZigBee, Bluetooth, acelerómetro, magnetómetro, sensor de luz, por mencionar algunos [7, 3, 2]. Así como diversas técnicas de localización para desarrollar este tipo de sistemas, tales como: (i) *triangulación* que utiliza las propiedades de los triángulos para obtener la posición del objeto o persona [4]; (ii) *proximidad* que consiste en determinar cuándo un objeto está cerca de una localización conocida [4]; y (iii) *Firmas de señal* que consiste de dos fases: entrenamiento y determinación de la posición [13]. En la fase de entrenamiento, se genera un mapa de las señales de radio mediante la captura de una medición de la señal de interés (p. ej., intensidad de campo eléctrico) en un entorno de interiores (ver figura 1). En la fase de determinación de la localización, una muestra de la medida de la señal es recolectada por un dispositivo que es portado por el usuario; la medida se compara con los valores del mapa de radio generado en la clase de entrenamiento utilizando algoritmos de proximidad, por ejemplo vecinos más cercanos, con el objetivo de inferir la localización actual del usuario.

Las técnicas descritas anteriormente, han demostrado ser de utilidad en la estimación de la localización. Sin embargo, presentan limitaciones, por ejemplo, los sistemas de localización que utilizan la técnica de proximidad o triangulación se caracterizan por requerir de infraestructura disponible en el entorno de interiores, es decir, dispositivos emisores y recep-

tores desplegados en el entorno de interiores, lo que limita su rango de cobertura. Mientras que, los que utilizan la técnica de firmas de señal, presentan la desventaja de una etapa de colección de datos laboriosa, debido a que se deben tomar mediciones de una señal, de manera espaciada, en una determinada posición y dirección en todo el entorno de interiores (ver figura1). Además, si el entorno cambia, la fase de recolección de datos debe realizarse nuevamente, ya que la información recolectada podría no ser útil para generar un modelo de localización.

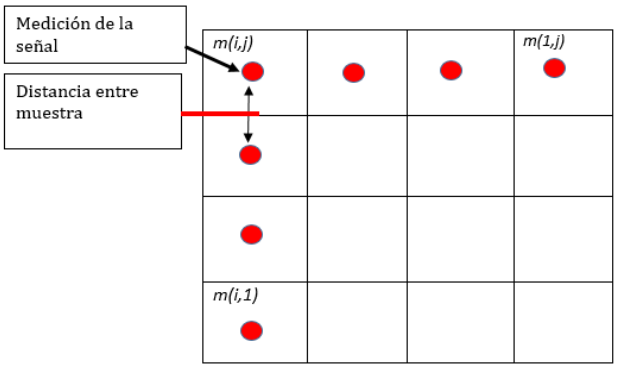


Figura 1: Mapa de radio

Para abordar las limitaciones de la técnica de firmas de señal, en este trabajo se propone utilizar la colaboración abierta distribuida, que es definida como un nuevo paradigma que empodera a los usuarios a colaborar en la realización de una tarea específica [9], en el ámbito de la localización es la captura de datos. En este sentido surge la hipótesis de que la colaboración abierta distribuida, permite generar un mapa de radio de intensidad de campo magnético para estimar la posición simbólica (p. ej. "cocina", "aula A") de un sujeto en un entorno de interiores, con un menor número de muestras de señal, las cuales son recolectadas de manera aleatoria. Por lo que se espera como resultado una precisión igual o similar a la que proporciona la técnica de firmas de señal. En este experimento se parte del supuesto de que la precisión de la técnica de firmas de señal es el 100%.

El artículo se organiza de la siguiente manera. En la siguiente sección se presentan proyectos que utilizan colaboración abierta distribuida en localización en interiores. Las tareas realizadas para alcanzar el objetivo de está investigación son presentadas en la sección 3. Enseguida se presentan los resultados de este trabajo de investigación (sección 4); y finalmente, las conclusiones y el trabajo futuro son presentadas en la sección 5.

2. TRABAJO RELACIONADO

Se han identificado dos maneras en las que la colaboración abierta distribuida es aplicada en el área de localización para interiores. La primera es para automáticamente construir planos de un entorno de interiores, por ejemplo, en [1] proponen un sistema llamado *CrowdInside* que utiliza las mediciones del acelerómetro de un dispositivo móvil para construir los trazos de movimientos generados por un usuario al desplazarse naturalmente dentro de un entorno de interiores. Estos datos son fusionados con información semántica para construir y describir los caminos recorridos por los usuarios en entorno de interiores. De manera similar, en el proyecto *Piloc* se capturan los movimientos del desplazamiento de un usuario en un entorno de interiores (distancia y dirección), así como también los nombres de los puntos de acceso y la intensidad de la señal de Wi-Fi. La información capturada es fusionada para construir un plano que indica el desplazamiento de un usuario dentro del entorno de interiores [8]. Con respecto a señales ambientales, se identificó el proyecto llamado *GROPING*, en este recolectan muestras de campo magnético provenientes de colaboración abierta distribuida, las cuales son capturadas cuando un usuario camina en un entorno de interiores y son utilizadas para generar el plano de un entorno de interiores[18]. El segundo enfoque identificado es la aplicación de la colaboración abierta distribuida para la construcción de los mapas de radio. En [16] propusieron un sistema de localización llamado *freeloc* el cual estima la posición de un usuario utilizando un mapa de radio de Wi-Fi que fue creado a partir de la colección abierta distribuida. De manera similar, en el proyecto llamado *Zee* se genera un mapa de radio con información de Wi-Fi y datos de aceleración y dirección del desplazamiento de un usuario [11].

En este trabajo de investigación de manera similar que [16, 11] se propone utilizar la colaboración abierta distribuida para generar un mapa de radio, a diferencia de estos trabajos, se propone utilizar una señal ambiental, ya que se encuentran disponibles en todos los entornos de interiores; y no requieren de un infraestructura disponible en el entorno, es decir, dispositivos emisores (p. ej. puntos de acceso) y receptores desplegados en el entorno de interiores [2].

3. DISEÑO DEL EXPERIMENTO

3.1 Descripción del entorno de interiores

El experimento se realizó en el laboratorio de sistemas computacionales de la Facultad de Ingeniería de la UABC, campus Mexicali. Este se compone de varias aulas y espacios abiertos (p. ej. pasillos). Las aulas consideradas en el experimento se muestran en la figura 2, estas fueron etiquetadas con las letras A, B, C, D, E; sus dimensiones en son: A=64 m^2, B=32 m^2, y 55.25 m^2 para las aulas C, D y E.

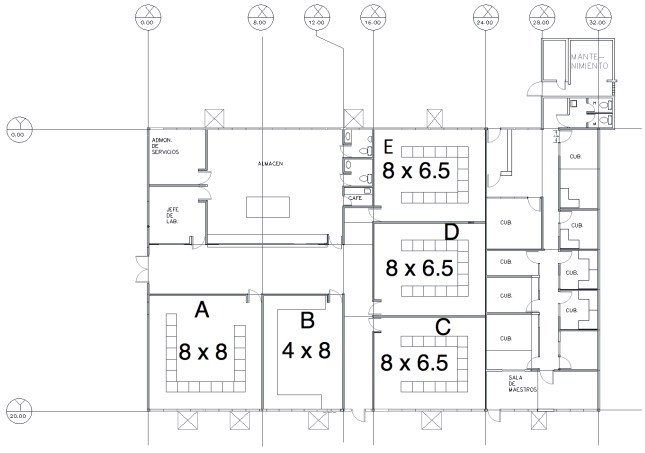


Figura 2: Plano de laboratorio de sistemas computacionales

3.2 Desarrollo de una aplicación para la recolección de datos

Para recolectar la señal de campo magnético de un entorno de interiores, desarrollamos una aplicación para dispositivos móviles Android compatibles con versiones 4.3 en adelante. La aplicación utiliza el sensor magnetómetro del dispositivo móvil configurado en modo no calibrado, a una tasa muestreo de 20 Hz.

Los datos recolectados son: intensidad del campo magnético en tres ejes (x, y, z) en μT (Teslas). Adicionalmente, se capturan datos como: la hora y fecha, el nombre del lugar donde se tomó la muestra, la versión Android utilizada, el modelo de magnetómetro y teléfono inteligente, y la longitud y latitud propocionada por el GPS con respecto a la localización actual del usuario.

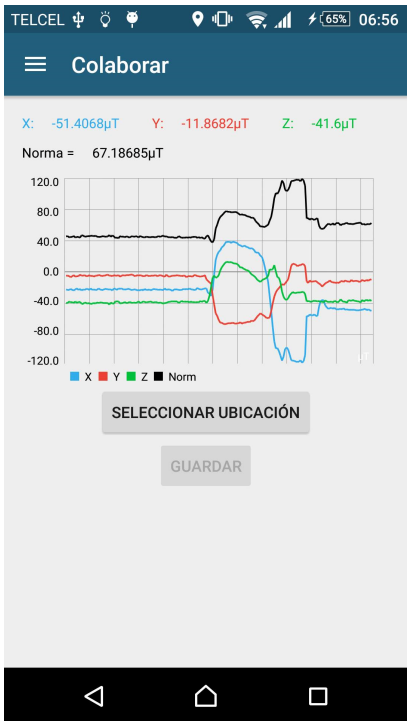


Figura 3: Interfaz de la aplicación desarrollada

Los datos recolectados son almacenados en *Firebase*[1], una base de datos en la nube del tipo NoSQL que proporciona una API (del inglés: Application Programming Interface) a los desarrolladores para implementar aplicaciones que puedan almacenar y sincronizar los datos a través de múltiples clientes en un documento JSON (del inglés: JavaScript Object Notation).

3.3 Recolección de datos

A través de la red social Facebook se invitó a los estudiantes del programa educativo de Licenciado en Sistemas Computacionales a participar en la actividad de recolección de datos. Para fomentar la participación se ofrecía una comida y dos entradas para el cine.

Para participar los estudiantes descargaban una aplicación para identificar si el dispositivo móvil tiene el sensor magnetómetro. Realizada la identificación del sensor, el estudiante instalaba la otra aplicación que desarrollamos para medir y recolectar la intensidad del campo magnético en un entorno de interiores.

Los estudiantes realizaban la captura seleccionando en la pantalla del dispositivo el aula donde se ubicaban, e iniciaban la recolección de datos al tocar el botón guardar (ver Fig. 3). El estudiante tenía que caminar en el entorno de interiores con el dispositivo móvil en la palma de la mano al nivel de la cintura, con la pantalla hacia arriba (ver Fig.4).



Figura 4: Estudiante recolectando datos

3.4 Preprocesamiento y Transformación de los datos

La calidad de los datos influye en la estimación de la localización. Para asegurar la calidad de los datos capturados y consecuentemente, obtener una mayor precisión en los resultados de la estimación de la localización, los datos almacenados en el documento tipo JSON fueron convertidos a un documento en formato abierto para representar datos en forma de tabla, en las que las columnas se separan por comas (CSV del inglés: Comma Separate Value). Para realizar esto se utilizó la herramienta *Opal convert*[2].

Los valores de intensidad de campo magnético y la magnitud de la intensidad fueron normalizados, mediante la Eq. 1 de normalización Z. Para mantener los valores de los datos en un rango de 0 a 1, se utilizó la ecuación de rango percentil (Eq.2).

$$normalizacion.z = \frac{(x - mean(x))}{sd(x)} \quad (1)$$

$$rango.percentil = \frac{trunc(rank(x))}{length(x)} \quad (2)$$

3.5 Generación del mapa de radio

[1]https://uabcsensor.firebaseio.com/?page=Hosting

[2]http://opal-convert-csv-to-qif-to-csv.en.softonic.com/

Para generar un mapa de radio válido para estimar la localización simbólica de un usuario en un entorno de interiores. Se experimento con el algoritmo llamado k Vecinos más Cercanos (del inglés: kNN- k Nearest Neighbors), un algoritmo simple que almacena todos los casos disponibles y clasifica los nuevos casos sobre la base de una medida de similitud (p. ej. funciones de distancia) [14].

Para hacer el análisis se utilizó R[3], un entorno de software libre para computación y gráficos estadísticos. En R se instaló el paquete *class*, el cual contiene diversos algoritmos de clasificación, entre los que se incluye kNN.

4. RESULTADOS

4.1 Recolección de datos

Un total de seis alumnos participaron en la actividad de recolección de mediciones de intensidad de campo magnético en las aulas del laboratorio de licenciado en sistemas computacionales. Durante un período de diez días, los alumnos recolectaron un total de 403, 874 muestras, de las cuales 84, 708 correspondían a la aula A, 63, 679 a la aula B, 62, 000 a la aula C, 98, 135 a la aula D y 95, 352 a la aula E.

Diversos dispositivos móviles fueron utilizados durante la recolección. Los dispositivos identificados se listan en la tabla 1

Tabla 1: Dispositivos Móviles utilizados en la captura

Dispositivo	Modelo de Magnétometro
Sony C6906	AK8963
Sony D6503	AK09911
Google Nexus 4	LGE
Samsung SM-G531H	YAS537
Google Nexus 7	Invensense MPL
Verizon SM-G900V	AK09911C

4.2 Generación del mapa de radio

Para generar el mapa de radio se utilizó el algoritmo de clasificación kNN. Para entrenar al algoritmo se utilizó el 70% de las muestras de intensidad de campo magnético recolectadas (282,711 registros). Estas muestras fueron seleccionadas aleatoriamente a través de la función *sample* del paquete R. El resto de las muestras se consideraron para realizar pruebas (121, 163 registros).

En el algoritmo utilizó un valor de k igual a 100. La precisión obtenida para estimar la localización simbólica con este valor fue de 92.3%; con valores de k mayores a 100 la precisión dismuinuía. En la tabla 2 se presenta la matriz de confusión obtenida. Las columnas representan las clases reales y los renglones las clases estimadas [6].

En la tabla 2 se puede identificar que con el mapa de radio actual se enfrentan problemas para localizar a los usuarios principalmente en el aula C y D ya que presentan una precisión de 53.61% y 56.82% respectivamente.

Tabla 2: Clasificación de las aulas

	A	B	C	D	E	Precisión
A	36,005	5,297	4,792	5,392	4,669	64.15%
B	1,168	30,182	1,551	292	2565	84.40%
C	6,235	2,701	22,862	7,027	1,409	56.82%
D	10,938	4,456	11,097	47,815	14,883	53.61%
E	4,862	2,140	3,252	8,014	43,147	70.18%

5. CONCLUSIONES Y TRABAJO FUTURO

En este artículo se presentó un experimento para identificar la eficacia de la colaboración abierta distribuida para generar un mapa de radio de mediciones de intensidad de campo magnético, que será utilizado para estimar la localización simbólica de un sujeto en un entorno de interiores. Los resultados muestran que es posible generar este mapa de radio sin tener que recolectar muestras de todo el entorno de interiores. La precisión de este mapa para localizar a un sujeto es de 92.3%. Como trabajo futuro, se tiene considerado realizar experimentos con otros algoritmos (p. ej. máquinas de soporte vectorial, algoritmos genéticos, etc.) para mejorar la precisión de la clasificación. Asimismo, generar un mapa de radio de señal de intensidad de campo magnético para estimar la localización de un sujeto en un entorno de interiores compuesto por varios niveles.

[3]https://www.r-project.org/

6. AGRADECIMIENTOS

A los estudiantes que participaron en la etapa de recolección de datos. Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca otorgada para la realización de mis estudios de maestría (CVU-692550).

7. REFERENCIAS

[1] M. Alzantot and M. Youssef. Crowdinside: automatic construction of indoor floorplans. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, pages 99–108. ACM, 2012.

[2] C. E. Galván-Tejada, J. P. García-Vázquez, J. I. Galván-Tejada, J. R. Delgado-Contreras, and R. F. Brena. Infrastructure-less indoor localization using the microphone, magnetometer and light sensor of a smartphone. *Sensors*, 15(8):20355–20372, 2015.

[3] Y. Gu, A. Lo, and I. Niemegeers. A survey of indoor positioning systems for wireless personal networks. *IEEE Communications surveys & tutorials*, 11(1):13–32, 2009.

[4] J. Hightower and G. Borriello. A survey and taxonomy of location systems for ubiquitous computing. *IEEE computer*, 34(8):57–66, 2001.

[5] J.-y. Hong, E.-h. Suh, and S.-J. Kim. Context-aware systems: A literature review and classification. *Expert Systems with Applications*, 36(4):8509–8522, 2009.

[6] C. Leacock, G. Towell, and E. Voorhees. Corpus-based statistical sense resolution. In *Proceedings of the workshop on Human Language Technology*, pages 260–265. Association for Computational Linguistics, 1993.

[7] H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(6):1067–1080, 2007.

[8] C. Luo, H. Hong, and M. C. Chan. Piloc: A self-calibrating participatory indoor localization system. In *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*, pages 143–153. IEEE, 2014.

[9] L. Pei, M. Zhang, D. Zou, R. Chen, and Y. Chen. A survey of crowd sensing opportunistic signals for indoor localization. *Mobile Information Systems*, 2016, 2016.

[10] H. Petrie, V. Johnson, T. Strothotte, A. Raab, S. Fritz, and R. Michel. Mobic: Designing a travel aid for blind and elderly people. *Journal of navigation*, 49(01):45–52, 1996.

[11] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen. Zee: zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 293–304. ACM, 2012.

[12] L. Ran, S. Helal, and S. Moore. Drishti: an integrated indoor/outdoor blind navigation system and service. In *Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on*, pages 23–30. IEEE, 2004.

[13] A. Taheri, A. Singh, and A. Emmanuel. Location fingerprinting on infrastructure 802.11 wireless local area networks (wlans) using locus. In *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*, pages 676–683. IEEE, 2004.

[14] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

[15] G. Xu. Applications of gps theory and algorithms. *GPS: Theory, Algorithms and Applications*, pages 219–241, 2007.

[16] S. Yang, P. Dessai, M. Verma, and M. Gerla. Freeloc: Calibration-free crowdsourced indoor localization. In *INFOCOM, 2013 Proceedings IEEE*, pages 2481–2489. IEEE, 2013.

[17] K. Yelamarthi, D. Haas, D. Nielsen, and S. Mothersell. Rfid and gps integrated navigation system for the visually impaired. In *2010 53rd IEEE International Midwest Symposium on Circuits and Systems*, pages 1149–1152. IEEE, 2010.

[18] C. Zhang, K. P. Subbu, J. Luo, and J. Wu. Groping: Geomagnetism and crowdsensing powered indoor navigation. *IEEE Transactions on Mobile Computing*, 14(2):387–400, 2015.

A Machine Learning Pipeline to Automatically Identify and Classify Roadway Surface Disruptions

M. Ezra Aragón
Universidad Autónoma de Chihuahua - Facultad de Ingeniería, Circuito Universitario,Chihuahua, Mx 31125
aragon.ezra@gmail.com

M. Ricardo Carlos
Universidad Autónoma de Chihuahua - Facultad de Ingeniería, Circuito Universitario,Chihuahua, Mx 31125
ricardo.carlos@gmail.com

Luis C. González Gurrola
Universidad Autónoma de Chihuahua - Facultad de Ingeniería, Circuito Universitario,Chihuahua, Mx 31125
lcgonzalez@uach.mx

Hugo J. Escalante
Instituto Nacional de Astrofisica, Optica y Electronica
Sta María Tonanzintla, Puebla, Mx 72840
hugojair@inaoep.mx

ABSTRACT

Applications for Intelligent Transportation Systems (ITS) based on common smartphones have become a real possibility because of their sensing and computing capabilities. Particularly, in this work we employ smartphones' accelerometers to sense the quality of roads, collecting in the accelerometers' axes the perturbations encountered by the vehicle. The ultimate goal of this line of work is to correctly identify, classify and geo-reference all obstacles so alleviating measures can be taken. Having a continuous series of accelerometer readings, the first problem is to identify the region where a perturbation was sensed (segmentation). To approach this problem, we propose using a Support Vector Machine (SVM) that accomplishes an accuracy performance of about 82%, outperforming other ad-hoc techniques such as Simple Mobile Average (SMA) and four other competitors. After having successfully segmented the candidate acceleration readings, the next problem is to classify the event in one out of four different categories. To this end we apply a Bag of Words representation and a Random Forest (RF), obtaining an accuracy of about 75%. These results were obtained by exhaustively training and testing this classifier over a newly created dataset that comprises signals for 30 different roads. Altogether, the use of a SVM followed by a RF seem to be a viable option to create a pipeline to automatically recognize and identify Roadway Surface Disruptions.

Mexican International Conference on Computer Science 2016 November 14–16, 2016, Chihuahua, Mexico

ACM ISBN 978-1-4503-2138-9.

DOI: 10.1145/1235

Keywords

Mobile Sensing, Smartphone, Accelerometer, Machine Learning, Bag of Words

1. INTRODUCTION

Intelligent Transportation Systems (ITS) are defined as technological solutions that improve operation and safety in transportation networks. One issue that is shared by thousands of cities worldwide is the presence of anomalies in roads, that can potentially cause damage to tires and suspensions in vehicles and lead to tragic accidents. In spite of the ubiquity of these anomalies, the vast majority of cities attend this problem using traditional methods, i.e. a workers crew fortuitously find and (hopefully) fix the anomaly.

Now imagine a scenario where an enormous amount of vehicles serve as sensing agents, collecting samples to measure the quality of kilometers of roads, and automatically sending the report to the office in charge, altogether with the GPS (Global Positioning System) coordinates of the anomaly. Without a doubt, this improvement would speed up maintenance and repairs, reducing costs as a side effect. This scenario is in the horizon with the omnipresence of common smartphones, since these devices can be used to sense the road while the user is driving. In particular, the accelerometer (now present in almost every smartphone) has been deemed as the *de facto* sensor to measure impacts suffered by the vehicle. When the smartphone experiences an inertial force in any of its axes, the resulting acceleration is measured in m/s^2. From this data, after applying some rotations to the axes to identify the axis that is perpendicular to the road, it is now possible to infer what caused the impact. This axis, hereinafter *Z-axis*, reflects every anomaly, protuberance, bump, or hole that was passed over by the vehicle.

The literature reports a large number of works that have been focused on identifying these anomalies. Nonetheless,

there are some shortcomings that have been detected. For instance, the proposals do not distinguish among different anomalies (pothole, speed bump, cracks, etc.), or employ simplistic techniques to extract possible anomalies from a continuous series of accelerometer readings. Furthermore, to the best of our knowledge there is not a complete (and fine-grained) pipeline that, given a series of accelerometer readings, first segments the readings of interest and then unequivocally classifies those candidates in one out of several different categories.

In this work we propose the creation of a Pattern Recognition Pipeline to segment and classify different types of anomalies present in roads. Figure 1 shows a schematic representation of this pipeline. First we need to extract those segments that correspond to anomalies in roads, from accelerometer data obtained by the drivers' smartphones. Once the segmentation is performed and candidate acceleration readings are obtained, a classification is applied to unequivocally identify the particular category of the anomaly. To create the training and testing dataset, we sensed potholes, speed bumps and asphalt bumps in real roads, avenues and streets. We compare five different classifiers for the segmentation problem, considering as well a traditional competitor. We apply a Random Forest to perform the classification on the segments suspected of containing an anomaly, since it has proved to be a robust technique hard to overfit. We present results for the segmentation and classification task, which show very competitive performance.

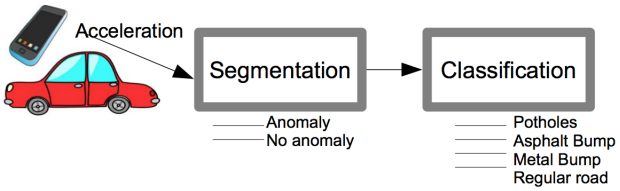


Figure 1: Interface of the web platform Pothole Lab.

The rest of this paper is divided as follows: section two presents related work, detailing different previous approaches for the segmentation of acceleration signal and the detection of events of interest. In the third section we present our experimental setup and our approach to solve the problem, divided in two stages. We present in the fourth section the results obtained with our pipeline method and, finally, the last section presents our conclusions and future work.

2. RELATED WORK

Several vehicular sensing systems for road anomaly detection have been proposed in the literature, the vast majority of them use accelerometers and GPS data. The simplest and most widely used way to detect road anomalies is the usage of thresholds over accelerometer data. Four different methods of this kind are proposed in [6], applying thresholds on samples captured at 100 Hz. The first method they proposed is Z-THRESH, the simplest reported way to detect events in acceleration data, making a detection if a sample is above or below predefined threshold values. Z-DIFF is the second method they propose, applying an upper threshold over the first derivative of the acceleration. STDEV(Z) employs a sliding window and calculates the standard deviation, if this metric is above a threshold, it is assumed an event has been found. The last method proposed is G-ZERO, and uses triaxial acceleration.

Another approach based on a threshold is presented in [2]. They capture triaxial acceleration and GPS data, installing dedicated hardware in seven taxis, to collect data while driving in the Boston area. In their work, they look for values above a threshold in highpass-filtered vertical acceleration to detect road anomalies, and use lateral acceleration to differentiate between potholes and other types of events. Their detection strategy was called Z-PEAK.

The usage of two threshold-based heuristics applied on vertical acceleration is presented in [7]. The first heuristic (Z-PEAK, presented in [2]) is meant to be used when traveling at speeds higher than 25 km/h and a new one (Z-SUS) at lower speeds. Z-SUS searches for a sustained dip in vertical acceleration: consecutive sample values below a threshold, lasting at least 20 ms.

Other way to process the data for detection is the usage of Machine Learning strategies [1, 3–5, 8, 9], this kind of approaches is still growing and the reported results are still incipient. We expect that with this contribution we set the first exhaustive testing experiments of Machine Learning classifiers for this problem.

3. EXPERIMENTAL SETUP

To evaluate the performance of the classifiers, we built a database that comprises different kind of anomalies (see Table 1). These anomalies were manually labeled and curated. For the collection process we used smartphones Moto G with Android at a 50 Hz sampling frequency.

Table 1: Anomalies included in the Pothole Lab web platform.

Category	Samples
Speed Bump	81
Pothole	56
Metal bump	50
Plane road	50
Total	137

3.1 Pothole Lab

To evaluate our proposal we created 30 virtual roads, each one with different number and nature of anomalies, so that we could test our proposal on a heterogenous dataset. Table 2, presents the main characteristics for each of the 30 roads. To create this dataset, we developed a web platform that receives as input the number and nature of anomalies, producing as an output a csv file with real acceleration samples that correspond to the requested anomalies. This platform is called *Pothole Lab* and we make it publicly available at //www.accelerometer.xyz/pothole_lab. Figure 2 presents a picture of the Pothole Lab interface.

3.2 First Stage - Segmentation Problem

To extract a segment of accelerometer readings and evaluate if it corresponds to an anomaly, we used a sliding window technique. Every segment (window) is passed to the classifier in turn, so that a category is assigned to the segment. The categories are: anomaly or non anomaly. For this this problem, we applied and compared five popular Machine

Table 2: Thirty virtual roads generated. The number of every anomaly follows the next key: p = potholes, b = bumps, mb = metal bumps.

Road	Training	Testing
1	16 p	40 p
2	20 p	50 p
3	30 b	50 b
4	20 b	50 b
5	20 mb	30 mb
6	10 mb	40 mb
7	50 p and b	50 p and b
8	20 p and b	80 p and b
9	20 b and 20 mb	50 b and mb
10	20 p, 20 and 20 mb	100 p, b and mb
11	5 b	10 b
12	5 p	10 p
13	5 mb	10 mb
14	3 b, 3 p	5 b, 5 p
15	3 b, 3 mb	5 b, 5 mb
16	3 p, 3 mb	5 p, 5 mb
17	3 b, 3 p, 3 mb	5 b, 5 p, 5 mb
18	15 b	25 b
19	15 p	25 p
20	15 mb	25 mb
21	8 b, 8 p	15 b, 15 p
22	8 b, 8 mb	15 b, 15 mb
23	8 p, 8 mb	15 p, 15 mb
24	8 b, 8 p, 8 mb	15 b, 15 p, 15 mb
25	30 b	40 b
26	20 p	25 p
27	25 mb	25 mb
28	15 b, 15 p	20 b, 20 p
29	15 b, 12 mb	20 b, 20 mb
30	15 p, 15 mb	20 p, 20 mb

Learning Techniques: Support Vector Machine (SVM), Artificial Neural Network (ANN), Nearest Centroid (NC), Decision Tree (DT) and Random Forest (RF). The segments that are deemed as anomalies are passed to the next stage for classification. Figure 3 shows an schematic representation of this stage.

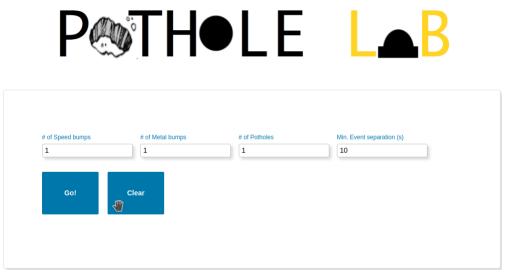


Figure 2: Interface of the web platform Pothole Lab.

3.3 Second Stage - Classification Problem

When we have the segments, the next step is to classify them in order to know what type of anomaly they represent. For this stage we have 4 different classes: class 0 corresponds to normal road, class 1 to pothole, class 2 to Asphalt bump and class 3 to Metal bumps.

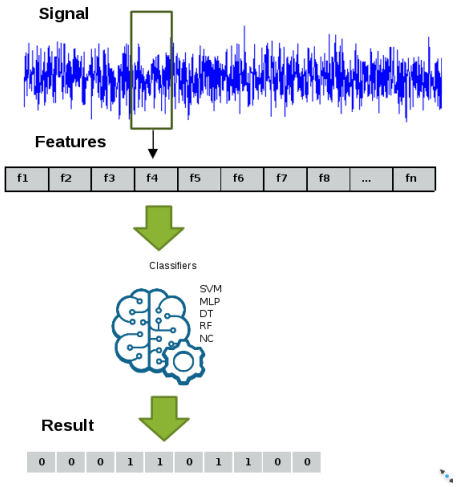


Figure 3: Visualization for the segmentation task.

In this stage we used the Bag of Words (BoW) technique to construct the feature vector. BoW is a simplified representation used in natural language for document classification where the occurrence or frequency of each word form a histogram for training a classifier. A brief explanation of this procedure is as follows. For this explanation consider that we have several candidate signals (accelerometer readings) that we want to preprocess (Fig. 4A). For each one of this signals, short segments are "cut" (Fig. 4B). With all these cuts, the k-means algorithm is applied, from this step centroids are found, they no represent "codewords" (Fig. 4C). With all the centroids, considering all the categories of interest, a codebook is generated (Fig. 4D). Finally, to represent a complete signal, this is cut, and its short segments is compared, via Euclidian distance, against all codewords, resulting in a representation histogram-like, which is the feature vector (Fig. 4E). The classifier employed in this part is the Random Forest.

4. RESULTS

For the first stage we select as the performance metrics the Accuracy and F-measure of each classifier. The F-Measure[1] (equation 1) is defined as the harmonic mean of Recall (equation 2) and Precision (equation 3). Table 3 shows the accuracy of the five classifier over the dataset. Accuracy sometimes could be misleading, this is one of those cases, since four out of the five classifiers offers very competitive results (as seen in this table). Nonetheless, when we consider the F-measure (see Table 4), difference in the classifiers are more evident. From this table, it could be seen that the best result, a better trade-off for Recall and Precision, is given by the SVM.

$$F - Measure = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (1)$$

[1]Just a reminder, TP stands for True Positives, FN and FP correspond to False Negatives and False Positives, respectively.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Table 3: Accuracy.

SVM	ANN	NC	DT	RF
0.8151	0.8051	0.7939	0.7337	0.8083

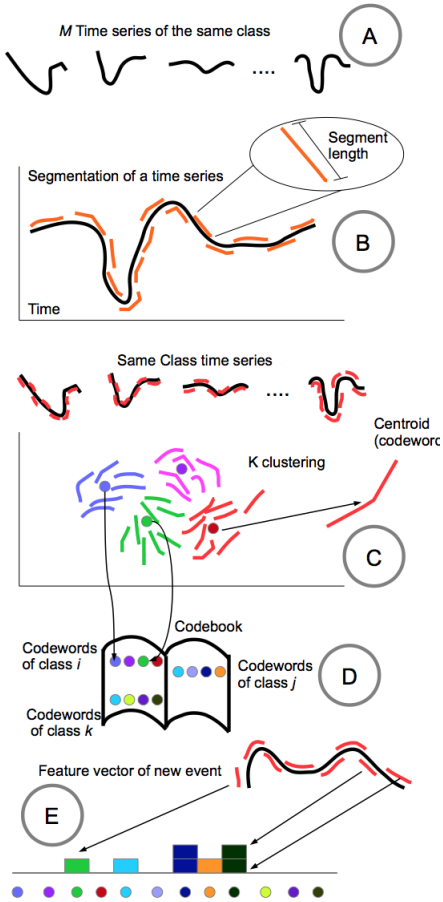


Figure 4: Schematics of Bag of Words.

Table 4: F-Measure.

SVM	ANN	NC	DT	RF
0.7857	0.7749	0.5875	0.4083	0.7360

A traditional contender for segmentation purposes is known as the Simple Moving Average (SMA), which detects perturbation on the data based on deviation from some calculated average. To contrast the results obtained by SVM Table 5 presents the metric performance for SMA on the segmentation problem. It could be seen that SMA is not able to accurately identify sections of real anomalies in the road.

We also compared and evaluated our results using classic threshold methods of the related literature (described on Section 2), and tested over our roads. This heuristics are useful for detecting a point of interest, but not really good for segmentation and classification of the event road. Table 6 shows the F-Measure results of the heuristics.

Table 5: SMA Results.

Precision	Recall	F-Measure
0.5934	0.5170	0.5240

Table 6: F-Measure Threshold Heuristics Results

Z-Thresh	Z-Diff	STDEV	GZERO	Nericell	Pothole
0.7060	0.5280	0.7525	0.6760	0.6750	0.4575

For the classification task, we apply the Random Forest classifier to candidate segments. Here, we followed two paths. In the first one, a human expert manually extracted a candidate anomaly from a series of accelerometer readings and the RF is applied over this segments (Table 6, first column). The accuracy for this case achieves a 83%. When the proposed pipeline is applied, and the segments are automatically extracted by the SVM, the classifiers obtains a 71% accuracy percent (multi-class scheme). Now, if the classifier discriminates one class at the time (one-vs-others scheme), the classification performance achieves 74%.

Table 7: Accuracy Average for Multiclass and One vs The Rest

Full Event	Segmented Event MC	Seg Event OvsTR
0.8323	0.7111	0.7425

One final calculation is focused in obtaining a complete picture of the total accuracy of the pipeline, meaning, what

is the percentage of correctly identified (from a series of readings) events and in turn, be classified in a respective category. The total error of our proposal obtained an error of 0.6584

5. CONCLUSIONS AND FUTURE WORK

Using accelerometer signals we can identified and represent different road events. In this work we can appreciate the usage of Machine Learning methods for the segmentation of the accelerometer signals and the usage of the BoW representation for the classification of the different events on the Roadway Surface. On the segmentation stage the average of the results are good and promising, resulting on the idea for new exploration methods of Machine Learning, the enhance over the features, and try to achieve a better segmentation of the signal.

When we used the BoW representation we can see a good performance over the full event and a little less performance over the segmented event. This makes a good choice for exploration methods in the future. As a future work we want to explore improvements in the representation of the BoW using n-grams, different kind of histograms, and different vocabulary for the classes.

6. REFERENCES

[1] T. Brisimi, C. Cassandras, C. Osgood, I. Paschalidis, and Y. Zhang. Sensing and Classifying Roadway Obstacles in Smart Cities: The Street Bump System. *IEEE Access*, 3536(c):1–12, 2016.

[2] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan. The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring. In *Proceeding of the 6th international conference on Mobile systems, applications, and services - MobiSys '08*, pages 29–39. ACM Press, 2008.

[3] L. C. Gónzalez, F. Martínez, and M. R. Carlos. Identifying Roadway Surface Disruptions Based on Accelerometer Patterns. *IEEE Latin America Transactions*, 12(3):455–461, 2014.

[4] M. Jain, A. Singh, S. Bali, and S. Kaul. Speed-Breaker Early Warning System. In *Proceedings of the 6th USENIX/ACM Workshop on Networked Systems for Developing Regions*, 2012.

[5] C. Kyriakou, S. E. Christodoulou, and L. Dimitriou. Roadway pavement anomaly classification utilizing smartphones and artificial intelligence. In *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, pages 1–6. IEEE, 2016.

[6] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo. Real time pothole detection using Android smartphones with accelerometers. In *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, pages 1–6. IEEE, jun 2011.

[7] P. Mohan, V. N. Padmanabhan, and R. Ramjee. Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones. In *Proceedings of the 6th ACM conference on Embedded network sensor systems - SenSys '08*, pages 323–336. ACM Press, 2008.

[8] M. Perttunnen, O. Mazhelis, F. Cong, M. Kauppila, T. Leppanen, J. Kantola, J. Colling, S. Pirttikangas, J. Haverinen, T. Ristaniemi, and J. Riekki. Distributed Road Surface Condition Monitoring Using Mobile Phones. In *Ubiquitous Intelligence and Computing*, pages 64–78. LNCS, 2011.

[9] F. Seraj, B. J. van der Zwaag, A. Dilo, T. Luarasi, and P. Havinga. RoADS : A road pavement monitoring system for anomaly detection using smart phones. In *1st International Workshop on Machine Learning for Urban Sensor Data, SenseML 2014*, 2014.

Automatic Characterization of Vertically Aligned Nanotubes by Image Analysis

Jesús Caro-Gutiérrez
Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
c17jesus@gmail.com

Félix F. González-Navarro
Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
fernando.gonzalez@uabc.edu.mx

Oscar M. Peréz-Landeros
Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
oscar.manuel.perez.landeros@uabc.edu.mx

Mario A. Curiel-Álvarez
Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
mcuriel@uabc.edu.mx

Benjamín Valdez-Salas
Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, México
benval@uabc.edu.mx

ABSTRACT

The study of vertically aligned nanotubes is a great interest for the development of new technology in different areas. Such a study involves a feature extraction of these structures that is carried out manually by a person in the computer, which causes problems as subjectivity and unrepresentative data collection. This work proposes a methodology based on image analysis for automatic characterization of vertically aligned nanotubes that includes counting, average diameter and area, gap between nanotubes and density; with the aim to face the drawbacks mentioned. The results obtained show that the methodology works well on good quality images, and moderately acceptable on low contrast and noisy images because these affect the nanotubes detection which impacts the measurement of features as a consequence. It is concluded that image analysis is a viable option to create characterization methodologies of nanotubes. Finally, some recommendations for future work are given.

Keywords

data mining; nanotubes characterization; feature extraction

1. INTRODUCTION

Nowadays, vertically aligned nanotubes (VANT) are widely used to develop new technologies in different areas such as electronics, biomedical and environmental [1]. Study of VANT is performed by microscopes like the scanning electron (SEM), which allows the image acquisition of these structures for further characterization. Feature extraction includes number of nanotubes, average diameter and area, etc; and it is carried out by a nanotechnologist in the computer manually. When the characterization of a large number of VANT (see Fig. 1) is needed, certain disadvantages arise with it. On the one hand, the manual measurement of features is a very time-consuming activity. On the other hand, the data may not be representative if the nanotechnologist considers only a part of nanotubes. Finally, the image feature extraction is subject to subjectivity because it is made by a human observer.

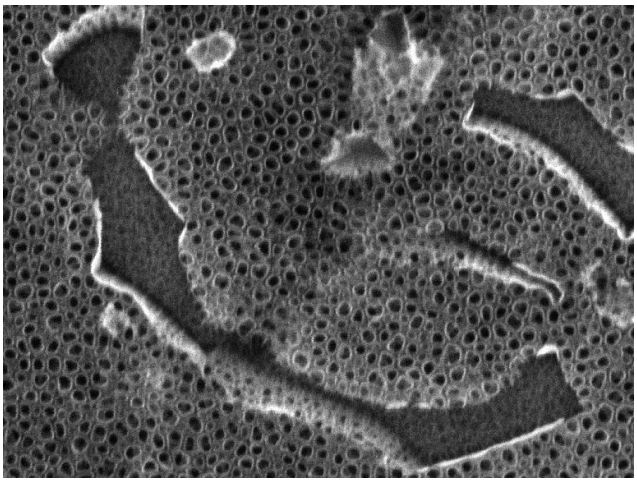


Figure 1: SEM image of vertically aligned TIO_2 nanotubes.

Problems mentioned above have generated great interest in image analysis for characterization of VANT because it can succeed in solve them. First, image analysis reduces times because it is made with powerful computers. Second, the data obtained are representative since image analysis makes possible the feature extraction of all nanotubes in the image. And thirdly, the subjectivity drawback is overcome for the reason that the human observer's intervention is not necessary.

The main objective of this paper is to present a methodology for automatic characterization of vertically aligned nanotubes by image analysis. The body of the work is organized

as follows: section 2 describes the materials and methods used in the experiments, section 3 shows the results and their discussion, and section 4 the conclusions and future work.

2. MATERIALS AND METHODS

Next, a description of the image database used in this work is presented. After that, the software and hardware utilized for experimentation is described. Finally, some stages of image analysis process are detailed.

2.1 Image Database

Database consists of 4 images of vertically aligned TIO_2 nanotubes acquired using a LYRA3 SEM with the following parameters: accelerating voltage 10 kV, magnification 100 kx and work distance 9.0 mm. All images are in JPEG format with a size of 514x314 pixels. Correct characterization (number of nanotubes, average diameter and area, etc) of these images are known and shown in the results section.

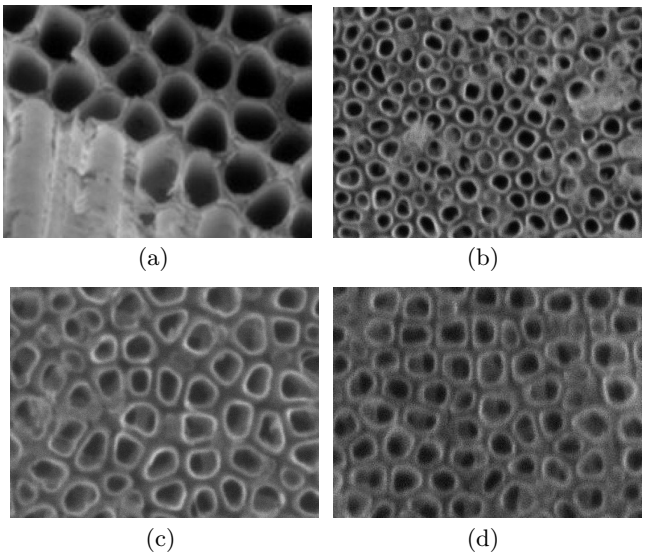


Figure 2: Database of 4 SEM images of vertically aligned TIO_2 nanotubes.

2.2 Hardware and Software

Experimental part and software development were carried out by using the software MATLAB R2013a, run-ning on Windows 7 operative system, and a CPU @ 64-bit AMD TO6-3600 APU @ 2.10 GHz and a 6GB RAM memory.

2.3 Methodology

The proposed methodology in [5] developed in this study, which corresponds with the image analysis process (see Fig. 3). The activities carried out in some of its stages are described below.

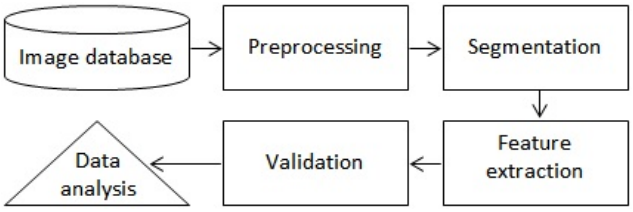


Figure 3: Stages of image analysis process.

2.3.1 Preprocessing

To reduce processing time, the image database was converted from RGB to grayscale using the formula (1). Where Gr is the grayscale image; R,G and B are the color channels red, green and blue of the original images.

$$Gr = 0.3R + 0.59G + 0.11B \quad (1)$$

This conversion formula was preferred instead of the average $(R+G+B)/3$ because it produced a better contrast in the image database, which favors the nanotubes segmentation. Later, it was applied to grayscale images an adaptive median filtering with neighborhood size of 9x9 as shown in [2] to reduce the additive noise and improve their quality. Principal sources of noise in VANT images originate during the SEM image acquisition process by the poor quality of equipment, external influences such as high temperatures and the operator's lack of experience.

2.3.2 Segmentation

Once the images were converted to grayscale and restored, the next step was to transform these to binary images with the aim of dividing their nanotubes from the background. To achieve this, were quantitatively compared the segmentation methods: thresholding, region growing, region splitting and merging and K-means clustering. This methods comparison was performed based on their percentages of correct classification estimating false positives and negatives. The method which obtained the best segmentation result was K-means clustering, for this reason it was chosen in this step; besides that it has an easy implementation, lower computational cost and a lot of flexibility. The K-means clustering algorithm is given below and can be found in [4].

1. Determine the number of desired clusters K and randomly assign to each cluster a pixel intensity from the image as its center c_k.
2. Group every pixel to the cluster whose center is closest. The closest normally mean the intensity value is similar.
3. Recompute the cluster centers. The cluster center is obtained by averaging the intensity values of the pixels belonging to the cluster.

$$c_k = \frac{1}{N_k} \sum_{p \in K} p$$

4. Repeat step 2 and 3 until no reassignment of pixels or a convergence criterion is met as certain number of iterations.

2.3.3 Feature Extraction

The features of the detected nanotubes in the previous segmentation are extracted at this stage, which includes: number of nanotubes, average diameter and area, gap between nanotubes and density. The methods used to calculate these features are presented below.

Nanotubes Counting: The number of nanotubes was obtained by applying a connected components labeling algorithm on the binary images. This algorithm groups the pixels of each nanotube and label them. At the end, the number of labels assigned corresponds to the number of nanotubes. A complete description of this algorithm is shown in [6].

Inner Area: This feature was extracted for each nanotube from the labeled images. The pixels of each nanotube are counted to obtain a measure of the area in terms of pixels A_p. This measure is converted to derived units such as square nanometers divide it by the number of pixels per square unit N, as in (2). Where N can be obtained from the bar scale provided at the bottom of a SEM image.

$$A = A_p/N \tag{2}$$

Diameter and Density: The method used to obtain an approximation of the nanotubes' diameters is based on the assumption of equal area circles. This equivalence is made for using the known formula (3), where d is the diameter of a nanotube and A its area in derived units.

$$d = \sqrt{\frac{4A}{\pi}} \tag{3}$$

For the density, defining it as the number of nanotubes per square micrometer, this was calculated by the ratio shown in (4), where C is the count of nanotubes in a region and A its area in square units. This feature is useful for characterization of nanotubes arrays.

$$D = C/A \tag{4}$$

Gap Between Nanotubes: The average distance between nanotubes in the binary images was calculated tacking into account for each nanotube the distances with its N neighbors. It is shown below the algorithm applied that can be found in [3].

1. Get the centroid coordinates (x, y) of each nanotube.
2. Fit on each nanotube a circle of radius r and center (x, y).
3. Calculate the distance between two nanotubes as: $d = \sqrt{(Y_2 - Y_1)^2 - (X_2 - X_1)^2} - (r_1 + r_2)$.
4. Store N distances of each nanotube with its neighbors.
5. Calculate the average gap between nanotubes.

3. RESULTS AND DISCUSSION

Figure 4 presents the results of applying the preprocessing and segmentation stages on the image database. The binary images obtained show qualitatively that the proposed methodology achieved to detect most of the nanotubes. Table 1 shows the feature extraction results (characterization) of the image database. The extracted features were number of nanotubes (count), average diameter and area, gap between nanotubes (GBN) and density. The results obtained by the proposed methodology are presented in the column approximation while the correct results for these features in the column ground truth.

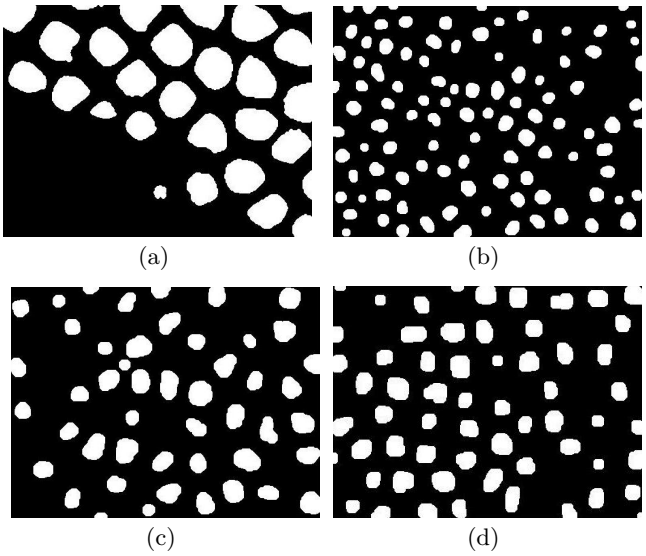


Figure 4: Preprocessing and segmentation of the database to binary images for characterization.

Table 1: Characterization results of the image database with the proposed methodology

Image in Fig. 2(a)		
Feature extracted	Approximation	Ground Truth
Count	29 NTs	28 NTs
Diameter	70.28 nm	70.08 nm
GBN	27.95 nm	28.25 nm
Area	3879.30 nm^2	3857.26 nm^2
Density	90.62 NTs/μm^2	87.50 NTs/μm^2
Image in Fig. 2(b)		
Feature extracted	Approximation	Ground Truth
Count	101 NTs	121 NTs
Diameter	102.71 nm	102.12 nm
GBN	22.31 nm	17.55 nm
Area	8285.45 nm^2	8190.53 nm^2
Density	17.41 NTs/μm^2	20.86 NTs/μm^2
Image in Fig. 2(c)		
Feature extracted	Approximation	Ground Truth
Count	48 NTs	52 NTs
Diameter	129.87 nm	123.73 nm
GBN	21.86 nm	20.17 nm
Area	13246.72 nm^2	12023.77 nm^2
Density	14.37 NTs/μm^2	15.56 NTs/μm^2
Image in Fig. 2(d)		
Feature extracted	Approximation	Ground Truth
Count	61 NTs	69 NTs
Diameter	90.70 nm	115.91 nm
GBN	21.79 nm	18.81 nm
Area	6461.08 nm^2	10551.94 nm^2
Density	32.44 NTs/μm^2	36.70 NTs/μm^2

As can be seen in Table 1, the approximation obtained for the image in Fig. 2(a) was acceptable compared against

the ground truth. This result is due to the good quality of the image. The little noise and high contrast allowed an precise image segmentation with well-defined nanotubes as shown in the image of Fig. 4(a). This accuracy of the image segmentation was crucial for the success of the posterior feature extraction. With this example it can be seen that the proposed methodology works well under ideal conditions.

In the case of the image in Fig. 2(b) the approximation was moderately acceptable. One of the main causes of this result was the nanotubes detection (101 NTs of 121 NTs). The image in question belongs to a nanotubes array with very different heights that usually results in a lack of contrast. In other words there are clearly defined nanotubes and others not so much; these last ones were confused with the background and did not allow the correct image segmentation by the proposed method (K-means). A total of 20 nanotubes were not detected in the count, which affected others derived features such as the gap between nanotubes (22.31 nm against 17.55 nm) because this depend on it.

The next feature extraction is about an image that, beside of lack of contrast also contains noise, see Fig. 2(d). In this image characterization was not achieved a good approximation of the diameter since there was a measurement error of 25.21 nanometers (90.70 nm against 115.91 nm), that impacted the area value as a consequence. An error of this magnitude is certainly unacceptable for any practical purpose. Apparently, this error is because the proposed preprocessing (adaptive median filtering) failed to remove noise which affected the size of the nanotubes in the segmentation. Finally, the lack of contrast decreased again the count.

4. CONCLUSION AND FUTURE WORK

Image analysis is a viable option to develop tools for automatic characterization of vertically aligned nanotubes because it could solve the problems associated with the manually feature extraction. This work shows that the proposed methodology works well and achieves acceptable characterization results on good quality images, but not in other situations. In the case of low contrast images, it demonstrated that the segmentation stage did not a good nanotubes detection, which affected the count and the GBN. And for the noisy image, it was observed that the preprocessing stage did not eliminate the noise, what impacted the size of nanotubes (area) and the diameter.

The results obtained in this work show the need to improve certain aspects, some recommendations are: create a larger image dataset to demonstrate better the effectiveness of the proposed methodology. Include contrast techniques (e.g. adaptive histogram equalization) in the preprocessing stage to enhance the nanotubes detection (segmentation) and image restoration methods in the frequency domain to also correct disturbances such as image blurring. Finally, it would be good to add a postprocessing stage with morphological operations after the segmentation to a better definition of the shapes of the nanotubes, and get so a feature extraction more approximate to the ground truth.

5. REFERENCES

[1] M. F. L. De-Volder, S. H. Tawfick, R. H. Baughman, and A. J. Hart. Carbon nanotubes: Present and future commercial applications. *Science*, 339:535–539, February 2013.

[2] R. Gonzalez and R. Woods. *Digital Image Processing*. Prentice-Hall, Reading, New Jersey, 2008.

[3] M. Haeri and M. Haeri. Imagej plugin for analysis of porous scaffolds used in tissue engineering. *Journal of Open Research Software*, 3:1–4, January 2015.

[4] A. M. Khan and S. Ravi. Image segmentation methods: A comparative study. *International Journal of Soft Computing and Engineering*, 3(4):84–92, September 2013.

[5] K. Madhumathi and A. S. Thanamani. Image mining: frameworks and techniques. *International Journal of Innovative Research in Computer and Communication Engineering*, 2(1):1818–1822, March 2014.

[6] B. Sugandi, H. Kim, J. K. Tan, and S. Ishikawa. A block matching technique for object tracking based on peripheral increment sign correlation image. *InTech*, pages 3–22, February 2011.

Aceleración en GPU de un algoritmo de Evolución Diferencial para el problema de Inferencia en Redes Reguladoras de Genes

Angel Omar Flores Olivas
Universidad Autónoma de Chihuahua, Facultad de Ingeniería
Circuito Número I s/n, Nuevo Campus Universitario II
Chihuahua, Chihuahua, México
afv9988@gmail.com

Luis Carlos González Gurrola
Universidad Autónoma de Chihuahua, Facultad de Ingeniería
Circuito Número I s/n, Nuevo Campus Universitario II
Chihuahua, Chihuahua, México
lcgonzalez@uach.mx

Eduardo Rodríguez Tello
CINVESTAV-Tamaulipas, Laboratorio de Tecnologías de la Información
Parque Científico y Tecnológico TECNOTAM, Km 5.5 carretera Cd. Victoria
Cd. Victoria, Tamaulipas, Mexico
ertello@tamps.cinvestav.mx

RESUMEN

En este documento se presenta un modelo de implementación paralela en GPU (utilizando CUDA), para reducir el tiempo de ejecución de un algoritmo de evolución diferencial para el problema de inferencia en redes reguladoras de genes.

Utilizando el modelo de paralelización de GPU (SIMD), es posible acelerar los procesos que componen la evolución diferencial, por ejemplo, es posible evaluar todos los individuos de una población al mismo tiempo, o realizar una comparación masiva de manera paralela.

Utilizando recursos como hilos, streams y paralelización clásica de la arquitectura (bloques y mallas) se pueden distribuir múltiples procesos completos de evolución diferencial en múltiples tarjetas gráficas, con lo cual se logran obtener resultados de evoluciones independientes con la posibilidad de combinarlos y modificarlos entre si, esto es útil al hacerlos aplicables a búsqueda local en el proceso de evolución diferencial y de múltiples soluciones.

Palabras clave

Paralelización; Evolución Diferencial; SIMD; CUDA

1. INTRODUCCIÓN

Las Redes Reguladoras de Genes (GRN) son representaciones que determinan a través de sus niveles de expresión, la influencia que tiene un gen con cada uno de los genes que comprenden una red (incluso con si mismo). Entender este tipo de relaciones es de suma importancia para el estudio de los sistemas biológicos. Un ejemplo de aplicación es poder predecir la reacción de un organismo con un medicamento, también es ampliamente utilizado en el desarrollo de sustancias y organismos sintéticos.

Estas relaciones se obtienen empleando instrumentación muy avanzada y compleja en un proceso tedioso y costoso, por lo que se han propuesto una variedad muy diversa de enfoques para describir y simular este tipo de redes [8], uno de los más prácticos e importantes es usar Evolución diferencial (DE) para crear un modelo por medio de ecuaciones diferenciales ordinarias de tipo S (S-systems) [7].

De esta manera es posible obtener un conjunto de datos mediante la experimentación para luego aplicar ingeniería inversa para generar un modelo computacional que describa el mecanismo completamente.

Desde hace algunos años la paralelización de algoritmos iterativos en GPU ha ido en auge debido a la creación de arquitecturas y herramientas que facilitan la implementación de algoritmos más complejos. Aunque existen trabajos que tratan el problema de aceleración al problema de inferencia de redes reguladoras de genes desde la perspectiva de GPU, en ellos se utilizan diferentes metodologías como la búsqueda exhaustiva [3] [4] y la evolución diferencial [6] sin las mismas aplicaciones que se integran en este trabajo como lo son la aplicación de búsqueda local, diversidad y skeletalizing. La inclusión de estas características permite probar más opciones y configuraciones, lo que permite mejorar los resultados reportados en esos trabajos.

2. EVOLUCIÓN DIFERENCIAL

Teniendo como pilar principal la teoría de la evolución natural de Darwin, la evolución diferencial basa su funcionamiento en un conjunto de procesos estocásticos que evolucionan generación a generación con base en la predominancia del más apto. Dada su formulación, es ideal para resolver problemas de optimización y clasificación [2].

La variante de DE utilizada en este trabajo es *DE/rand/2/bin*, es decir, se perturba un vector aleatorio (/rand) y se consideran 2 vectores para la mutación (/2) con un esquema de cruza binomial (/bin) con un parámetro de Cr = 0.002 . Las funciones de DE utilizadas se presentan en el Algoritmo 1.

2.1 Sistemas S (S-Systems)

Comúnmente utilizados para calcular los niveles de expresión de los genes de una GRN a través de ecuaciones diferenciales ordinarias, se representan de la siguiente manera:

$$\frac{dX_i(t)}{d(t)} = \alpha_i \left(\prod_{j=1}^{N} X_j^{g_{ij}(t)} \right) - \beta_i \left(\prod_{j=1}^{N} X_j^{h_{ij}(t)} \right) \quad (1)$$

donde: N es el número de genes de la GRN, g_{ij} y h_{ij} son los ordenes cinéticos de síntesis y degradación respectiva-

```
Resultado: Población
crea población;
evaluar;
Mientras generaciones hacer
    cruzar;
    evaluar;
    comparar original con cruza;
    seleccionar mejores;
    si diversidad entonces
        calcula parecidos;
        aplica diversidad;
        evaluar;
    fin
    si búsqueda local entonces
        seleccionar individuos;
        búsqueda local;
        evaluar;
    fin
fin
```

Algorítmo 1: Métodos de la evolución diferencial

mente y α y β son las constates de ritmo. $X_j(t)$ representa el nivel de expresión del gen j en el tiempo t.

El total de parámetros en un sistema de este tipo es *2N(N+1)*, por lo que la cantidad de genes de una red aumenta sustancialmente el número de parámetros que deben ser estimados. El objetivo es obtener los mejores valores de g_{ij}, h_{ij}, α y β para modelar la red. El valor del nivel de expresión se calcula utilizando el método numérico Runge-Kutta.

2.2 Función objetivo

La función objetivo más común (ecuación 2), consiste en la suma del cuadrado de los valores de los genes calculados en cada tiempo ($X_{i,calc,t}$) menos el valor obtenido de forma experimental ($X_{i,exp,t}$), dividido sobre el mismo valor experimental del gen para el mismo tiempo t, esto para cada uno de los genes de la red (T es el número de tiempos) [1].

$$f = \sum_{i=1}^{N} \sum_{t=1}^{T} \left\{ \left(\frac{X_{i,cal,t} - X_{i,exp,t}}{X_{i,exp,t}} \right)^2 \right\} \tag{2}$$

2.2.1 Estructura (Skeletalizing)

Con motivo de hacer la estructura de la red más dispersa, se agregó un componente a la función de fitness para dar preferencia a los valores de ordenes cinéticos más pequeños, dicho de otra manera, los valores de g y h más cercanos a cero determinan una menor afectación de un gen a otro. Partiendo del trabajo de Noman [7] se definió la función de la siguiente manera:

$$\lambda \left\{ \sum_{ij}^{N} |g_{ij}| + \sum_{ij}^{N} |h_{ij}| \right\} \tag{3}$$

Donde λ es un factor que determina la importancia del agregado, y N representa la cantidad de genes que se van a evaluar y $N \leq totaldeGenes$, este agregado además permite obtener la estructura de la red.

Otra forma de lograr más dispersión en la red, también propuesto en el trabajo de Noman [7] es aplicar un parámetro de reducción, en el cual, cualquier valor de g y h menor a este parámetro será cambiado a 0, ayudando al propósito descrito anteriormente.

3. CUDA

El gran poder de procesamiento de las tarjetas gráficas había sido aprovechado de manera burda y complicada. Para motivar su uso de manera general NVIDIA creo CUDA, poniendo a disposición el poder de cómputo de sus productos de alto rendimiento utilizando lenguaje C.

3.1 Modelo de paralelización SIMD

Según las diferentes clasificaciones propuestas por Flynn [5], el cómputo paralelo de una GPU puede trabajar con el modelo *SIMD (Single Instruction Multiple Data)*, el cual realiza una misma operación sobre un conjunto de datos diferentes, esto lo hace ideal para trabajar con matrices y vectores.

3.2 Herramientas de CUDA

La función básica de paralelización con CUDA es el *Kernel*. Se define como una función estándar en lenguaje C que permite ejecutarse de manera paralela, siempre con la misma instrucción. Cada ejecución es denominada *hilo*.

Cada *hilo* contiene los datos que se modificarán con la instrucción, y se organizan en conjuntos denominados *bloques*, los cuales a su vez se organizan en *mallas*. La cantidad de *mallas* y *bloques* están limitadas por la arquitectura de la tarjeta gráfica. Los *Kernels* pueden ejecutarse de manera asíncrona y paralela por medio de *Streams*. En la nomenclatura de CUDA, un CPU es definido como *Host* y la tarjeta gráfica como *Device*.

3.3 Paralelización de los Procesos de DE

Utilizando este modelo de programación, cada componente de cada individuo se evalúa en un *hilo*, por lo que la inicialización, cruza, evaluación, comparación, selección y sustitución se hace para todos los individuos de la población al mismo tiempo en n bloques.

Al evaluar el parecido de un individuo con la población se crea un bloque de dos dimensiones que contienen los valores de cada individuo de la población, por lo que cada individuo del bloque puede compararse con cualquier hilo del mismo y obtener el parecido en una sola comparación $n \times n$ hilos a la vez.

Haciendo uso de hilos en CPU para controlar la ejecución de Kernels de manera paralela, es posible evaluar varios procesos completos de evolución diferencial que son independientes entre si, pero que pueden combinarse, por ejemplo, para ejecutar sobre ellos búsqueda local.

4. RESULTADOS

Para contextualizar el ambiente sobre el cual se realizaron las pruebas, las características técnicas del equipo utilizado se muestran en la tabla 1.

4.1 Fitness

Se utilizó una instancia de 2 genes y 50 tiempos encontrada en el trabajo de Tominaga [1] para analizar los resultados, teniendo como objetivo replicar el valor de los niveles de expresión de los genes en todos los tiempos utilizando los mismos parámetros.

Tominaga reporta que tardo un total de 4085.407 segundos en encontrar los valores con 1000 individuos, mientras

Población	Fitness D	Fitness P	Fitness CPU	Tiempo D	Tiempo P	Tiempo CPU	Aceleración D	Aceleración P
32	5.982277	3.170987	4.39158	0.36945	0.2592	0.68	x1.84	x2.62
64	1.681027	1.718314	5.17165	0.72365	0.49715	1.36	x1.87	x2.73
128	1.692924	1.903811	5.23209	1.7329	0.9964	2.66	x1.53	x2.66
256	0.902072	0.311762	5.17518	3.959	2.0216	5.33	x1.34	x2.63
512	0.4753	0.232287	5.15556	8.66255	4.0921	10.71	x1.23	x2.61
1024	0.366	0.177123	4.45661	17.1332	7.5872	21.56	x1.25	x2.84
2048	0.174122	0.138175	4.15864	46.70625	15.7489	42.72	x0.91	x2.71

Tabla 1: Valores obtenidos al ejecutar la instancia Tominaga 2 con el equipo de pruebas (P) y con el equipo de desarrollo (D)

	Desarrollo	Pruebas
Modelo de GPU	GeForce GTX 775M	Tesla C2070
Número de GPUs	1	4
CUDA Capability	3.0	2.0
GPU Clock	797 MHz	1.15 GHz
Memoria global	2048 MB	5375 MB

Tabla 2: Características técnicas del equipo de Desarrollo y del equipo de Pruebas, este ultimo consiste en un clúster con tarjetas gráficas específicamente diseñadas para cómputo paralelo.

	Tominaga	1 GPU	4 GPUs
Tiempo (s)	4085.407	465.288	137.509
Individuos	1000	256	256
Generaciones	1000	1000	1000
Fitness	0.0345	0.028108	0.035662
Instancias de DE	1	4	4

Tabla 3: Resultados al replicar la instancia Tominaga de 2 genes y 50 intervalos de tiempo

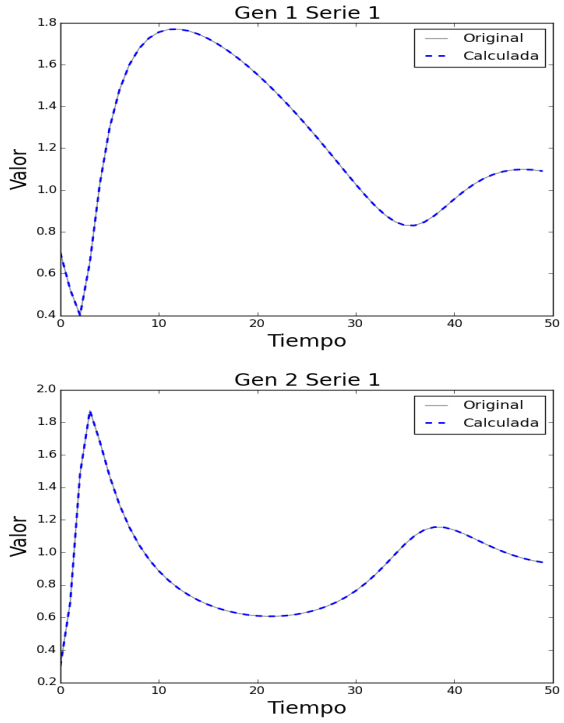


Figura 1: Niveles de expresión de ambos genes de la instancia Tominaga en los 50 intervalos de tiempo.

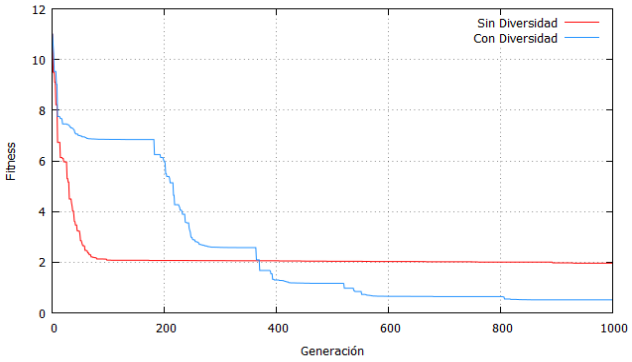


Figura 2: Gráfica de convergencia de la instancia de 2 genes de Tominaga. Los puntos que aparecen en la gráfica sin diversidad indican los puntos que satisfacen la condición dada para que fuera aplicada.

que utilizando un GPU tardo casi 10 veces menos de tiempo, con la mitad de individuos y ejecutando 4 procesos de evolución diferencial independientes en paralelo. Los valores de la instancia Tominaga para los 2 genes se muestran en la figura 1.

Esta configuración se eligió arbitrariamente ya que las capacidades de la tarjeta gráfica del equipo de desarrollo son muy conservadoras. Sin embargo, el clúster del equipo de pruebas, que cuenta con 4 GPUs tiene tarjetas específicamente diseñadas para cómputo paralelo.

4.2 Diversidad

La inclusión de diversidad evita el estancamiento en óptimos locales, la función renueva un porcentaje de la población cuando no se ha logrado reducir el fitness del mejor individuo (*BestSoFar*) de toda la población en k generaciones para permitir explorar en nuevos espacios de búsqueda.

En la gráfica se puede observar una comparación de la convergencia el fitness aplicando diversidad y sin ella, cabe mencionar que en esta implementación es importante permitir evolucionar a la población y no aplicar la diversidad al cabo de pocas generaciones para obtener buenos resultados.

4.3 Comparativa en GPUs

Se realizaron pruebas sobre la misma instancia Tominaga utilizando 100 generaciones y variando el tamaño de la población, los resultados obtenidos se muestran en la tabla 3.

Como se observa en la Tabla 1, con una población menor de 2048 individuos se logró una aceleración con respecto al tiempo requerido en CPU utilizando el equipo de desarrollo,

más allá de esta cantidad las limitaciones de la tarjeta gráfica comienzan a presentarse, sin embargo el fitness fue mucho mejor que el obtenido en el CPU con la misma cantidad de generaciones.

Es notable la aceleración obtenida al ejecutar la misma instancia en el equipo de pruebas, logrando una reducción de 2.8 veces con respecto a la misma instancia con los mismos parámetros en CPU. La implementación en CPU corresponde a una instancia sencilla de evolución diferencial, por lo que realiza una menor cantidad de procesos a comparación de la instancia de GPU que realiza cálculos para comparar a la población para aplicar diversidad y analizar la estructura de la red, estas herramientas son las que permiten reducir sustancialmente el fitness.

Hay que destacar la capacidad de utilizar equipo especializado (y una cantidad mayor de tarjetas) para obtener mejores resultados en menos tiempo, además de poder manejar configuraciones más robustas.

5. DISCUSIÓN

El uso de GPU puede ayudar a acelerar los procesos utilizados en la evolución diferencial gracias a su modelo de programación, lo cual hace especialmente útil este tipo de hardware para resolver procesos iterativos que requiere un gran poder de cómputo. Gracias a esto es posible reducir sustancialmente el tiempo requerido para obtener soluciones de buena calidad, tan buena como los obtenidos con métodos más comunes.

Una de las características más favorables es que al ser un lenguaje unificado, un algoritmo escrito para CUDA puede correr tanto en una tarjeta gráfica sencilla como en una especialmente diseñada para cómputo paralelo, aprovechando al máximo las capacidades del *Device* donde se esté ejecutando sin realizar modificaciones importantes.

La mayor limitación del uso de GPU es la memoria de la tarjeta gráfica. En las pruebas realizadas una instancia de 10 genes y 512 individuos puede ser ejecutada sin problemas en el equipo de desarrollo, pero la misma instancia utilizando 1024 individuos comienza a dar problemas de memoria en la tarjeta, sin embargo en el equipo de pruebas, se pueden correr instancias de 50 genes y 2048 individuos sin problema.

6. TRABAJO FUTURO

Extender las opciones de variantes de evolución diferencial, como *DE/best/2/bin*, *DE/rand/1/exp* etc, con la posibilidad de elegir la variante como un parámetro más del algoritmo.

La implementación de búsqueda local se encuentra en un estado primitivo, la idea principal sería poder combinar instancias independientes en cualquier momento de la evolución diferencial para modificar todas las poblaciones y compartiendo características totalmente agenas a las mismas.

Optimizar los parámetros en instancias de más de 20 genes, aunque es posible analizar este tamaño de instancias, es un proceso muy lento debido a las capacidades técnicas el equipo de desarrollo.

Buscar un reemplazo para el método numérico de Runge-Kutta como método de evaluación de la ecuación diferencial que calcula los niveles de expresión. Aunque es un método sumamente funcional y ya se encuentra paralelizado, en este método existe una dependencia del tiempo actual para el cálculo del siguiente tiempo por lo que no se pueden evaluar todos los tiempos a la vez como se pretendía.

7. REFERENCIAS

[1] M.˜O. Daisuke˜TOMINAGA. Nonlinear numerical optimization technique based on a genetic algorithm for inverse problems: Towards the inference of genetic networks. *DBLP*, January 1999.

[2] A.˜E. Eiben and J.˜E. Smith. Introduction to evolutionary computing. *SpringerVerlag*, 2003.

[3] D.˜C. M.-J.-B. S. L. C. S.˜R. Fabrizio F.˜Borelli, Raphael Y.˜Camargo. Accelerating gene regulatory networks inference through gpu/cuda programming. *IEEE*, June 2012.

[4] D.˜C. M. J. L. C.˜R. Fabrizio F˜Borelli, Raphael Y de˜Camargo. Gene regulatory networks inference using a multi-gpu exhaustive search algorithm. *Second IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS 2012)*, 14(18)(S5), February 2012.

[5] M.˜J. Flynn. Some computer organizations and their effectiveness. *IEEE Trans. Comput.*, 1972.

[6] E.˜R.-T. Luis E. Ramírez-Chavez, Carlos A. Coello˜Coello. A gpu-based implementation of differential evolution for solving the gene regulatory network model inference problem. 2011.

[7] N.˜Noman and H.˜Iba. Inference of gene regulatory networks using s-system and differential evolution. *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation (GECCO'2005),*, pages 439, 446, 2005.

[8] T.˜Schlitt and A.˜Brazma. Current approaches to gene regulatory network modelling. *BMC Bioinformatics*, 8(6)(S9), September 2007.

Consistent Transformations of Knowledge Bases in Answer Set Programming, Part I*

Juan C. Acosta G.†
A. University of Juarez
Cihuahua, Mexico
jguadarrama@gmail.com

ABSTRACT

One of the major and traditional topics of Artificial Intelligence over many years has been knowledge representation and reasoning. It has proved to be a strong theoretical framework for Logic Programming to manage dynamic knowledge bases. In this work, one of two parts, we go through current and some of those past proposals to update knowledge coded into Answer Set Programming knowledge bases, by analysing their features and identifying challenges to represent correct evolving knowledge.

Keywords

Belief Revision; Belief Update; Knowledge Bases

1. INTRODUCTION

A preliminary version of this paper has first appeared as a technical report in [2]. Here we revisit it and update it to current and past approaches, in an attempt to classify them and to come upon new challenges.

One of the major and traditional topics in Artificial Intelligence over the last years, knowledge representation and reasoning has proved to be a strong theoretical framework to manage knowledge bases. This particular topic has become more widely studied in administration of knowledge bases of intelligent (rational) agents, especially in situations of incomplete knowledge from a changing environment. This area of research is known in the literature as *belief change* and its relevance to program transformation is fundamental if we are interested in the resulting correct knowledge bases, especially if they represent critical systems requirements.

The history of semantics for updates[1] of logic programs (in our context, *formalisms for program transformations*) is rather long. Indeed, it begins in the days of some of the first versions of PROLOG with its commands assert and retract. However, sooner they began to yield conflicting information and other (perhaps unexpected) *side effects*. It was also time of research on databases with publications like [21], and in particular for *logical databases*[36, 20, 19]. Nevertheless, some of the first formalisms to carry out proper changes to *monotonic theories* have been originally studied by [6, 23, 29, 25, 27], whereas in the *non-monotonic side* by [26, 30, 28].

In particular, [22] formulated the *Stable Models Semantics* also refereed as *Answer Sets Semantics*, SM or simply ASP), and more concrete proposals arose within that framework, aimed at the problem of updating knowledge represented in ASP [38, 39, 42, 34, 35, 9, 16, 17, 14, 15, 32, 1].

In this work we introduce current and some of those past proposals to transform knowledge bases represented by logic programs. We point out features as well as some of their limitations to represent *correct evolving knowledge* and of course, *correct program transformations*. Nevertheless, this survey is just a small thread of a massive research over more than two decades, and is by no means exhaustive. It just takes into account those proposals that are the most relevant and of interest in our opinion.

The rest of our parer is divided into a very quick glimpse of basic background (Section 2) necessary to understand program transformations in ASP; the approaches to update logic programs are classified into several different categories (Sections 3–6) and a section for discussion and final remarks—Section 7. Each of the approaches show a few particular examples to illustrate their definitions, as well as common *observations* to show disadvantages and to compare with the others.

2. PRELIMINARIES

A main foundation of these proposals is the well-known Answer Sets Semantics, also known as Stable-Models Semantics. In this paper it is assumed, though, that the reader is familiar with basic notions of logic programming and (extended) disjunctive logic programs, DLP, EDLP, which are easily available in the literature.

2.1 Logic Programming and Answer Sets

*

†This project is mainly supported by The National Council for Science and Technology, CONACYT, and The Public Education Ministry, SEP, research grants. We are also grateful to The Autonomous University of Juarez. Thanks also to Michel Ramírez B. for his proofreading.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AI 2016 Chihuahua, Mexico

© 2016 ACM. ISBN 978-1-4503-2138-9…$15.00

DOI: 10.1145/1235

[1]For historical reasons, in this paper we call it update, although it's actually belief revision. A study on the difference was first introduced by [24].

As we represent knowledge by means of ASP programs for being one of the most studied and founded successful semantics to reason about incomplete (unknown) information, in the following we give a very-short description of Answer Sets Programming (ASP), which is identified with other names like *Stable Logic Programming* or *Stable Model Semantics* [22] and A-Prolog. Its formal language and some more notation are introduced from the literature as follows.

DEFINITION 1 (ASP LANGUAGE, $\mathcal{L}_{\mathsf{ASP}}$). *In the following $\mathcal{L}_{ASP}$ is a language of propositional logic with symbols: $a_0, a_1, \ldots$; connectives: "," (conjunction) and meta-connective ";"; disjunction $\vee$, also denoted as $|$; $\leftarrow$ (also denoted as $\rightarrow$); $\neg$ (default negation or weak negation, also denoted with the word* not*); "$\sim$" (strong negation, equally denoted as "$-$"). The propositional symbols are also called atoms or atomic propositions. A* literal *is an atom or a strong-negated atom. A* rule *ρ is an ordered pair $f_1(\rho) \leftarrow f_2(\rho)$, where $f_1(\rho)$ is a possibly-empty finite set of literals in disjunction and $f_2(\rho)$ a possibly-empty finite set of literals (or default-negated literals) in conjunction.*

The meaning we give the propositional constants is the same meaning than having an empty set in either component of a rule. Finally, a *logic program* (or just program) is a possibly empty finite set of rules.

With the notation just introduced in Definition 1, one may construct program clauses of several forms that are well known in the literature, such as Extended Logic Program (ELP), Extended Disjunctive Logic Program (EDLP), etc.

Informally, the semantics of such programs consists of reducing the general rules to rules without default negation "$\neg$", because the latter are universally well understood. For page limitation, we just skip the formal definition of such reduct, which can be easily found in the literature.

Formally, we say that a program is inconsistent if and only if it has no answer sets, and consistent otherwise.

DEFINITION 2 (EDLP). *An* extended disjunctive logic program *is a set of rules of form*

$$\ell_1 \vee \ell_2 \vee \ldots \vee \ell_l \leftarrow \ell_{l+1}, \ldots, \ell_m, \neg\ell_{m+1}, \ldots, \neg\ell_n \quad (1)$$

where ℓ_i is a literal and $0 \leq l \leq m \leq n$.

Naturally, an *extended logic program* (or ELP hereafter) is a finite set of rules of form (1) with $l = 1$; while an *integrity constraint* (also known in the literature as *strong constraint*) is a rule of form (1) with $l = 0$; while a *fact* is a rule of the same form with $l = m = n$. In particular, for a literal ℓ, the *complementary literal* is $\sim\ell$ and vice versa; for a set $\mathcal{M}$ of literals, $\sim\mathcal{M} = \{\sim\ell \mid \ell \in \mathcal{M}\}$, and $Lit_{\mathcal{M}}$ denotes the set $\mathcal{M} \cup \sim\mathcal{M}$; finally, a *signature* $\mathfrak{L}_{\mathcal{K}}$ is a finite set of literals occurring in $\mathcal{K}$. Additionally, given a set of literals $\mathcal{M} \subseteq \mathcal{A}$, the complement set $\overline{\mathcal{M}} = \mathcal{A} \setminus \mathcal{M}$.

The well-known *semantics of an EDLP* consists of reducing general rules to rules without default negation "$\neg$" because the latter can be interpreted in *classical logic* by means of the well-known *Herbrand models*. In particular, the reduced rules with no default negation Mon of a rule of the form (1) is

$$\ell_1 \vee \ell_2 \vee \ldots \vee \ell_l \leftarrow \ell_{l+1}, \ldots, \ell_m \quad (2)$$

where ℓ_i are literals and $0 \leq l \leq m$. This kind of rules is known in the literature as *monotonic counterpart* or *positive program*. Additionally, the monotonic counterpart of a set of rules is the set of the monotonic counterparts of its rules.

Now let us introduce the meaning of programs with both monotonic and nonmonotonic counterparts.

Suppose a finite ground program $\mathcal{K}$, consisting of clauses of form (1). For any set $\mathcal{S} \subseteq \mathfrak{L}_{\mathcal{K}}$, the *answer-sets reduct* $\mathcal{K}^{\mathcal{S}}$ corresponds to

$$\mathcal{K}^{\mathcal{S}} = \{\ell_1 \vee \ell_2 \vee \ldots \vee \ell_l \leftarrow \ell_{l+1}, \ldots, \ell_m \mid \{\ell_{m+1}, \ldots, \ell_n\} \cap \mathcal{S} = \emptyset\} \quad (3)$$

Stating $\mathcal{S}$ as a set of literals rather than atoms, makes one of the differences with *Stable-models* semantics.

Next, the meaning of a monotonic counterpart corresponds to its *minimal classical model* as follows.

DEFINITION 3 (MINIMAL CLOSURE, $\mathsf{Cn}(\mathcal{K})$). *Let $\mathcal{K}$ be a* positive extended disjunctive program *and $\mathfrak{L}_{\mathcal{K}}$ the* signature *(set of all ground literals) from $\mathcal{K}$. The set $\mathsf{Cn}(\mathcal{K})$ denotes the minimal subset of $\mathfrak{L}_{\mathcal{K}}$ where*

1. *for each ground clause $p_0 \vee p_1 \vee \cdots \vee p_l \leftarrow q_1, \ldots, q_m$ in $\mathcal{K}$, $q_1, \ldots, q_n \in \mathcal{S}$ implies $p_i \in \mathcal{S}$ for some $0 \leq i \leq l$; and for each ground clause of the form*

$$\bot \leftarrow q_1, \ldots, q_m \quad (4)$$

$\{q_1, \ldots, q_m\} \not\subseteq \mathcal{S}$.

2. *if $\mathcal{S}$ contains a pair of complementary literals, then $\mathcal{S} = \mathfrak{L}_{\mathcal{K}}$.*

Note that item (2) in this Definition 3 extends *Stable Models* by giving a meaning to *strong negation*.

Finally, an *answer set* of a given program $\mathcal{K}$ is a *minimal closure* of its reduct as following stated.

DEFINITION 4 (ANSWER SET). *Suppose $\mathcal{K}$ is a EDLP and $\mathcal{S}$ a set of literals. Then, $\mathcal{S}$ is an* answer set *of $\mathcal{K}$ if and only if $\mathcal{S} = \mathsf{Cn}(\mathcal{K}^{\mathcal{S}})$.*

Notice that all stable models can be viewed as *minimal Herbrand models* of a set of first-order sentences, but not the converse. Additionally, $\mathcal{S}$ is a *consistent answer set* of a given program $\mathcal{K}$ if it does not contain a complementary pair of literals.

Although we have introduced ASP as propositional (*ground*) programs, fixed *non-ground ASP-programs* of arbitrary *arity* are also considered in the same way than [13] do. Accordingly, non-ground ASP-programs with variables or constants as *arguments* can be seen as a simplified expressions of larger ground (propositional) ones without variables, where each *ground program* $\mathcal{K}$ is a set of its *ground rules* $\rho \in \mathcal{K}$. In addition, a ground rule is the set obtained by all possible substitutions of variables in ρ by constants occurring in $\mathcal{K}$ [13].

In general, ASP is the necessary background and main foundation that is common to all the approaches here presented. Yet another framework employed by a few of the approaches to update knowledge in ASP is called Generalized Answer Sets.

This is the basic background to understand the following approaches to update knowledge represented in ASP programs. So, let us begin with the different proposals.

3. EITER'S TEAM

To the best of our knowledge, [17] achieved the most complete survey of most known semantics for updates of logic programs, by gathering relevant postulates and principles from the literature. Their approach first appeared in [15] with a vast study of well-known and well-accepted postulates and properties, and later refined [17] and extended to be a main component in more general problems like agents [18] or preferences. They also implemented a solverthat is the main engine of an experimental graphical front end we have implemented[2].

In [17] they formulate a natural definition for updating logic program sequences on a restricted Answer Sets language by rejecting rules under a *causal rejection principle*. The principle is due to [9] that later, however, turned out to be counterintuitive, even to themselves see [18, 8, 32]. This "counter-intuition" comes from a strong dependency upon the syntax of programs, first noted by [18, 33, 37], and is further discussed in Section 4. In particular, the formula under which [17, 18] analyse and describe update properties is as follows.

Given an update sequence $(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$, with $2 \leq n$, over a set of atoms $\mathcal{A}$, assume $\mathcal{A}_\alpha$ as an extension of $\mathcal{A}$ by new pair-wise unique atoms $\mathsf{rej}(\rho)$ and α_i, for each rule ρ occurring in $\mathcal{K}$; each atom $\alpha \in \mathcal{A}$, and $1 \leq i \leq n$. An injective naming function $\mathsf{Name}(\cdot,\cdot)$ is also assumed, which assigns to each rule ρ in a program $\mathcal{K}_i$ a unique name, $\mathsf{Name}(\rho, \mathcal{K}_i)$, provided that $\mathsf{Name}(\rho, \mathcal{K}_i) \neq \mathsf{Name}(\rho', \mathcal{K}_j)$ whenever $i \neq j$. Finally, for a literal ℓ, they use ℓ_i to denote the result of replacing an atomic formula α of ℓ by α_i.

The intuitive idea of $\mathsf{rej}(\rho)$ is that of an atom that blocks (rejects or *inhibits*) a related rule ρ when the former is true, provided that there is another more recent rule ρ' with *conflicting information*.

Then [17] define the intended answer sets of an update sequence $(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$ in terms of the answer sets of $\mathcal{K}_\lhd = (\mathcal{K}_1 \lhd \cdots \lhd \mathcal{K}_n)$. In other words, the models are back expressed in the original alphabet by the intersection of them and the original atoms:

Definition 5 ([17]). *Given an update sequence,*

$$(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$$

over a set of atoms $\mathcal{A}$, then $\mathcal{S} \subseteq Lit_\mathcal{A}$ is an update answer set *of $(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$ if and only if $\mathcal{S} = \mathcal{S}' \cap \mathcal{A}$ for some answer set $\mathcal{S}'$ of $\mathcal{K}_\lhd = (\mathcal{K}_1 \lhd \cdots \lhd \mathcal{K}_n)$. The collection of all the update answer sets of $(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$ is denoted by $\mathcal{U}(\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_n)$.*

In addition to the declarative version, this semantics is also supported by a solver available both for downloading and running online[3], which is yet another valuable asset worth considering when comparing the approach with others.

Let us present the following example, inspired in [17], to illustrate their approach. Later we will reuse it with other approaches:

[2] The graphical online front end to [18] command-line solver may be found at: http://logic-lab.sourceforge.net/upd.html.

[3] We have installed an online version, which also provides a graphic-oriented interface on the server itself—http://logic-lab.sourceforge.net/upd.html. Obviously, no download or installation is necessary to run the latter version.

Example 1. *Suppose we have the following simple-but-illustrative two sets of* system requirements *(translated into ASP), $\langle \mathcal{K}_1, \mathcal{K}_2 \rangle$, representing the* initial and current knowledge *of an intelligent greenhouse, which acts autonomously under specific circumstances, where*

$$\begin{aligned}
\mathcal{K}_1 = \{ notify &\leftarrow night, \text{not } wSystem \\
night & \\
wPlants &\leftarrow wSystem \\
wSystem & \} \\
\mathcal{K}_2 = \{ \sim wSystem &\leftarrow blackout \\
blackout & \}
\end{aligned}$$

Program $\mathcal{K}_1$ might represent the following configuration:

- *Notify when it's night and there isn't evidence of the water system working.*
- *It's night now.*
- *Water the plants when the water system is working.*
- *The system is working now.*

The unique model of such requirements is

$$\{night, wPlants, wSystem\}.$$

Now suppose that the systems engineer needs to incorporate a new rule that states to not water plants when the ground is flood: $\sim wPlants \leftarrow gFlood$. By their definition, the transformed update program $\mathcal{K}_\lhd = (\mathcal{K}_1 \lhd \cdots \lhd \mathcal{K}_n)$ consists of the following rules—amongst the rest of the rules that we skip for page constraints:

$$\begin{aligned}
\cdots wSystem_1 &\leftarrow \text{not } \mathsf{rej}(\rho_4) \\
\sim wSystem_2 &\leftarrow blackout, \text{not } \mathsf{rej}(\rho_5) \cdots \\
\cdots \mathsf{rej}(\rho_4) &\leftarrow \text{not } wSystem_2 \cdots \\
notify_1 \leftarrow notify_2 & \quad notify \leftarrow notify_1 \cdots \\
wSystem_1 \leftarrow wSystem_2 & \quad wSystem \leftarrow wSystem_1 \cdots \\
\sim wSystem_2 \leftarrow \sim wSystem_3 & \quad \sim wSystem \leftarrow \sim wSystem_2 \\
blackout_2 \leftarrow blackout_3 & \quad blackout \leftarrow blackout_2
\end{aligned}$$

whose unique answer set is $\{notify_1, night, night_1, \mathsf{rej}(\rho_4), \sim wSystem_2, blackout, blackout_2, notify, \sim wSystem\}$ and its ***update answer*** *set is just:*
$\{night, blackout, notify, \sim wSystem\}$.

Let us complete Example 1, which is one of the major disadvantages the framework exhibits [35]:

Observation 1 (Continued Example 1). *Now let us consider Example 1 again and perform a* second update *to the sequence with program $\mathcal{K}_3 = \{\sim blackout\}$. Accordingly, the new answer set of the resulting update program is $\{wSystem_1, wSystem, night_1, night, wPlants_1, wPlants, \mathsf{rej}(\rho_6), \sim blackout_3, \sim blackout\}$.*

As a result, by Definition 5 the corresponding **update answer sets** *are*
$\mathcal{U}(\mathcal{K}_1, \mathcal{K}_2) = \{night, blackout, notify, \sim wSystem\}$ and
$\mathcal{U}(\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3) = \{wSystem, night, wPlants, \sim blackout\}$. However, this result does not coincide with common intuition, just because one of the possible models contradicts the latest fact ($\sim blackout$) stating that it is no longer happening!

On the other hand, the second model says that system is working, back again for no obvious reason, which is counterintuitive too.

Despite the satisfactory deep nice analysis they realise of known postulates and principles from the literature and their available sover, yet another major disadvantage of [17]'s approach has to do with syntactic and semantic contents, as illustrated by the following example inspired from [8] and modified by [37] that may produce *counterintuitive models*:

Observation 2. *Suppose an agent who believes that* when *it is* day *it is not* night *and vice versa, and that there are* stars *when it is* night *and when there are no* clouds. *Finally, that at the current moment it is a fact that there are no* stars. *This simple story may be coded*[4] *into a logic program* $\mathcal{K}_1$ *as follows:*

$$\mathcal{K}_1 = \{(day \leftarrow \text{not } night), \quad (night \leftarrow \text{not } day)$$
$$(stars \leftarrow night, \text{not } cloudy), \quad \sim stars \leftarrow \emptyset\}$$

whose unique answer set is $\{day, \sim stars\}$. *Later, the agent acquires new information stating that* stars *and* constls[5] *are the same thing, as coded in* $\mathcal{K}_2$. *As soon as the agent updates* $\mathcal{K}_1$ *with program*

$$\mathcal{K}_2 = \{(stars \leftarrow constls), \quad (constls \leftarrow stars)\}$$

the augmented alphabet *of the two programs contains only one new* extra atom *with respect to* $\mathcal{K}_1$*: constls. As the model of* $\mathcal{K}_2$ *is obviously the empty answer set, constls is considered* synonym *of stars by means of* $\mathcal{K}_2$*, and thus the update should not change the* original beliefs. *However, the update yields an* extra answer set *in some of the existing update semantics based on the* causal rejection principle *see [9, 17, 8]:* $\{stars, constls, night\}$, *which does not coincide with* common intuition.

The reason is that, although stars can not be true, introducing constls gives another possibility for stars to be true. Thus, the additional answer set is derived [37].

In general, these supplementary rules in the update are a conservative extension *[31] to* $\mathcal{K}_1$*: the original language is extended and all answer sets ought to be extensions of the old answer sets. In this specific situation, constls should be true if and only if stars is true.*

So, Observation 2 means that the proposed semantics is inappropriate to model the corresponding kind of problems.

Finally, updating knowledge through a sequence of logic programs does not seem to be natural in our opinion. For instance, consider Observation 1 again and try to perform a new update. Although the authors might have had other goals and intentions, their current approach does not allow to do that (i.e. you would have to "initialize" the sequence and append the new update to it.

To recap, [17] were *very* good at gathering relevant postulates and principles from the literature and at analysing them in terms of their proposal. Their approach, however, suffers from a few disadvantages for our interests owing to its reliance on the *causal rejection principle* (see [9]) and to its sequence-based approach.

[4] Notice that there are other ways to represent the story. The problem is, however, what to do in this particular situation, when the agent's original knowledge base runs across a redundant piece of information.

[5] i.e. *constellations*.

4. DYLP AND OTHER EARLY PROPOSALS

One of the earliest approaches to update logic programs through transformations appeared in the late 90's from [9, 10] that was extended to an interesting language called LUPS, by [11], to specify *explicit updates* in programs on a semantics that they called *Dynamic Logic Programming* or DyLP [9]. Some years later, though, [8] refined the latter, whom over the previous periods formulated a principle of rejection, also known as *causal rejection principle* [9, 17, 8] and [7].

Informally, the called *refined principle of rejection* consists in *rejecting rules* of previous and upcoming programs in an update sequence, whenever there are other conflicting rules at the current state.

To begin with, [8] motivation comes from a simple example to what they themselves called a *tautology* (a rule from which they expect no models):[6]

$$\text{not } p \leftarrow \text{not } p \tag{5}$$

One may verify that the rule alone produces no models in their semantics[7]—i.e. just the **empty model**, $\emptyset$, which actually *contains* all non-positive (default-negated) atoms—and updating knowledge with this rule should have no effect, an that is why they introduced the refined principle of rejection.

Additionally, [8] explain in a footnote what *tautology* means: A rule of the form $\ell \leftarrow f_2$ with $\ell \in f_2$, where ℓ is an atom (or a default-negated atom) and the body of a rule, respectively. Although the authors might have had other goals for their approach, this high dependency on syntax will prove to be one of their major disadvantages as a semantics for updates in our opinion, as explained along this paper.

Before introducing their proper definitions for their semantics, a special notation is necessary, which can be obtained from the literature and from [8].

Intuitively, a *refined interpretation of a DyLP* program is a dynamic stable model if the least model/s of the positive program that result/s from the difference of the rejected rules and the union of the default assumptions is the same than the union of the interpretation and the default-negated literals that do not appear in the interpretation.

The following definition introduces the model of the *transformed program* from the original dynamic logic program:

Definition 6 (Dynamic Stable Model, [8]). *Let* $\mathcal{P}$ *be a dynamic logic program and* $\mathcal{M}$ *an interpretation.* $\mathcal{M}$ *is a* dynamic stable model *of* $\mathcal{P}$ *if and only if*[8]

$$\overline{\mathcal{M}} = \mathit{least}(\rho(\mathcal{P}) \setminus \mathit{Rej}(\mathcal{P}, \mathcal{M}) \cup \mathit{Def}(\mathcal{P}, \mathcal{M}))$$
$$\mathit{Rej}(\mathcal{P}, \mathcal{M}) = \{\rho \mid \rho \in \mathcal{K}_i, \exists \rho' \in \mathcal{K}_j, i \leq j, \rho \bowtie \rho', \mathcal{M} \models f_2(\rho')\}$$
$$\mathit{Def}(\mathcal{P}, \mathcal{M}) = \{not_a \mid \nexists \rho \in \rho(\mathcal{P}), f_1(\rho) = a, \mathcal{M} \models f_2(\rho)\}$$

$\oplus_R$ *is the corresponding update operator.*

[6] Note that the kind of rule in (5) is invalid in $\mathcal{L}_{\mathsf{ASP}}$, and it also illustrates why strong negation "$\sim$" should not be a simple replacement to "not"in the head. Take for example, the program $\{\sim p \leftarrow \text{not } p\}$ whose unique **answer set** differs from the empty set: $\{\sim p\}$.

[7] The details are easily available in the literature, like [9].

[8] Notice that it seems they have missed the "_" and have typed "not a" rather than "not_a" in (6) from the original paper in [8]. Moreover, (6) should be a set of rules, rather than a set of literals, to be sound with the equality in (6)! In the rest of the section, it is assumed the former, as the authors do.

The intuition behind $\mathsf{Rej}(\cdot,\cdot)$ is the set of rules that are in conflict with both current and previous rules in the sequence. Moreover, $\mathsf{Def}(\cdot,\cdot)$ consists of the positive "default-negated" atoms that do not appear in the intended model.

This approach has had several implementations, including one for the original version before the so-called refined principle, and another for the principle itself. Additionally, LUPS is also implemented, which we consider a major asset, and the following list shows their respective locations:

- http://centria.di.fct.unl.pt/~jja/updates
- http://centria.di.fct.unl.pt/~banti/FedericoBantiHomepage/refdlp.htm

Now let us analyse an example inspired by [35], originally proposed by [9].

Observation 3. *Let us code the story introduced in Example 1 and Observation 1 as follows:* $\mathcal{P} = \mathcal{K}_1 \oplus_R \mathcal{K}_2 \oplus_R \mathcal{K}_3$, *where*

$$\begin{aligned}\mathcal{K}_1 = \{&(notify \leftarrow \text{not } wSystem), \quad (wPlants \leftarrow wSystem)\\ &(wSystem \leftarrow \emptyset)\},\\ \mathcal{K}_2 = \{&(blackout \leftarrow \emptyset), \quad (\text{not } wSystem \leftarrow blackout)\},\\ \mathcal{K}_3 = \{&\text{not } blackout \leftarrow \emptyset\}.\end{aligned}$$

$\mathcal{M} = \{wSystem, wPlants\}$ *is a* refined dynamic stable model *of the update sequence, since the unique rejected rule is*

$$blackout \leftarrow \emptyset$$

and the unique default is not_notify, where the least *model is*

$$\{wPlants, wSystem, not_blackout, not_notify\} = \overline{\mathcal{M}}.$$

In a similar example [35] argue, however, that the resulting model is counterintuitive *because, after the first update, the refined dynamic stable model was* $\{blackout, notify\}$ *and thus there is no reason to believe that once the power is restored, the system should be working back again! A similar counterintuitive result is given in Observation 1.*

In addition to that, there is still a simple example, first suggested by [17], that still causes counterintuitive results in this DyLP-semantics.

Observation 4. *Suppose an initial knowledge base* $\mathcal{K}_0 = \{(c \leftarrow r), (r \leftarrow \emptyset)\}$ *updated with* $\mathcal{K}_1 = \{\text{not } r \leftarrow \text{not } c\}$. *Firstly, the initial generalised stable* ***model*** *of* $\mathcal{K}_0$ *is* $\{c, r\}$, *and one would expect no further changes after the update because, according to [17], such an update should be irrelevant to the known fact* r *and the derived* c. *However, the update* $\mathcal{P} = \mathcal{K}_0 \oplus_R \mathcal{K}_1$ *has the* ***extra model*** $\mathcal{M} = \{not_c, not_r\}$ *because* $\overline{\mathcal{M}} = \{not_c, not_r\}$; $\mathsf{Rej}(\mathcal{P}, \mathcal{M}) = \{r \leftarrow \emptyset\}$; $\mathsf{Def}(\mathcal{P}, \mathcal{M}) = \{not_c\}$; *and* $\mathsf{least}[(\mathcal{K}_0 \cup \mathcal{K}_1) \setminus \mathsf{Rej}(\mathcal{P}, \mathcal{M}) \cup \mathsf{Def}(\mathcal{P}, \mathcal{M})] = \{not_c, not_r\} = \overline{\mathcal{M}}$.

Although [9] were some of the first researchers to formulate and implement a semantics for updates through program transformations, they still have quite a few disadvantages like the ones pointed out in this section: Firstly, for the particular syntax and semantics of their transformed generalised logic programs, which is a different variant of Stable-Models Semantics, or in other words, a *non-standard concept of* ASP [17]; secondly, for their *causal rejection principle* that produces the mentioned *counterintuitive results*. Finally, their sequence-of-knowledge-bases approach is in our opinion counterintuitive, as we have already commented at the end of Section 3.

5. MINIMAL CHANGES

In [34, 35] they propose three types of updates through program transformation, to which they call *inconsistency removal*, *view updates* and *theory updates*. Each of them correspond to a special case of updates and revision in the literature.

For page and comparison reasons, this paper is focused on theory updates, rather than other special cases such as making an inconsistent program consistent or differentiating between variant and *invariant knowledge*. As a result, this section does not include other types of belief changes, which are easily available in the literature.

Once the extended abductive program is normalised, its interpretations shall correspond to transformed *update programs* that consist of the rules of the original theory that don't belong to the *normalised abductive set*, merged with a new set of update rules, as specified in [35].

Next, this transformation of *update rules* takes part of a new transformation called update program of the normalised extended abductive program that is an intermediate EDLP. This intermediate program specification is as follows:

Definition 7 (Update Program, [35]). *Given an abductive program* $\langle P, \mathcal{A}_\alpha \rangle$, *its* update program, $\mathcal{UP}$, *is defined as an* EDLP *such that* $\mathcal{UP} = (P \setminus \mathcal{A}_\alpha) \cup \mathcal{UR}$.

Then the models of an *update program* denote the deletion of facts or rules from the original program in the pair.

Definition 8 (U-minimal Answer Sets, [35]). *An answer set* $\mathcal{S}$ *of* $\mathcal{UP}$ *is called* U-minimal *(U-MAS) if there is no answer set* $\mathcal{S}'$ *of* $\mathcal{UP}$ *such that* $\mathcal{S}' \cap \mathcal{UA} \subset \mathcal{S} \cap \mathcal{UA}$.

As a result, there is one or more new corresponding updated programs to the U-MAS.

5.1 Discussion

This abduction framework proves to have nice properties of a *syntactical minimal change* of rules in the original non-monotonic theory—Theory Updates Definition in [35]—by means of consistent interpretations of *hypothetical changes*. With this framework, [35] can perform particular kinds of updates and maintenance of knowledge-bases *consistency*, and can also present a vast analysis of disadvantages in other comparable approaches.

In general, [35]'s first goal is to provide an update semantics to compute their extended abduction. Secondly, they also characterise updates through the extended abduction, as they themselves state it. Consequently, the approach lacks of a proper analysis of more principles and postulates from the literature of program transformation or belief change. Additionally, they characterise different kinds of updates with their extended abduction, claiming that they can provide an algebra of rules deletion, besides the addition of them, to explain observations.

Let us recapitulate [35]'s approach in a few words. They construct their *update program* out of the normal abductive form of an *extended abductive program* $\langle P \cup Q, P \setminus Q \rangle$ whose models are *U-MAS's*, interpreted from an update program.

Last, the interpretations correspond to one (or more) new programs representing knowledge bases, derived from the addition and/or deletion of facts that the U-MAS's describe in turn.

In [35] they propose an example as an argument against other approaches—like [9, 17]—that brings back previous knowledge of the original theory. That is to say, their interpretation is that the TV turns itself on again and it is possible to watch it as well: $\{wSystem, wPlants, \sim blackout\}$, which does not coincide with their intuition. This is similar to our conclusions in Observation 1, where the system is working back again upon no good justification However, this argument seems to be too strong to generalise that all update semantics should behave accordingly, because [35] are differentiating *fluents* and *actions* in a language that does not have such an explicit difference.

Finally, there is a simple example that might represent another disadvantage of this approach.

Observation 5. *Suppose the initial knowledge base* $\mathcal{K} = \emptyset$ *updated by a simple fact* $\mathcal{K}_1 = \{x \leftarrow \emptyset\}$. *Following [35]'s framework, the answer sets of its update program is empty:* $\mathcal{UP} = (\emptyset \setminus \{x \leftarrow \emptyset\}) \cup \mathcal{UR}$, *where* $\mathcal{UR}$ *is clearly empty because the extended abductive program from the update pair has no abducibles:* $\mathcal{A}_\alpha = \emptyset \setminus \{x \leftarrow \emptyset\}$.

Moreover, although the authors present a deep analysis of their proposal and although it seems to be *robust*-enough for agent's *changing environment*, there is a lack of further and more *general properties* and solver that make it hard to compare with other alternate approaches. [42] pointed out that this approach is classified into a *syntax-based semantics*. As a result, it has no general semantic foundation that justifies its updates [42], and by interpreting the resulting knowledge bases with a given semantics might interfere with the ASP semantics that performs the update operation. It is clear that they justify their updates with an *extended abductive framework*, which is still a specific problem and then leaves the mentioned absence of update characterisation.

Finally, a minor disadvantage is that the approach is undefined to update a knowledge base with an inconsistency. [35] state that such a kind of update "makes no sense", which clearly does not mean that an agent or whoever is updating thie knowledge base will *never* come across an originally *inconsistent observation*. Nevertheless, they do consider cases where an *initial knowledge base* is originally inconsistent, and they identify such case as *inconsistency removal*. This method consists in updating an inconsistency knowledge base with an *empty update*.

6. ZHANG'S

An interesting proposal for updates through program transformation comes from [42], where the author identifies three types of problems to solve in an update process: *elimination of contradictory information*, *conflict resolution* and *syntactic representation*.

Additionally, one of the applications from that line is an interesting language [12] that is specialised in updates of agent *policies* and defined at the top of ASP. [12] specify such policies in terms of clauses with a predefined semi-imperative *syntactical structure*, as well as an initial *planning approach*.

However, owing to a specialisation the work has on *policies*, the programmer is restricted and obliged to use *reserved words* like "always", "implied by", "with absence", etc. which, besides constraining the domain to specific applications, it 'reduces' the language and has potentially different meanings in the *meta-language*. Nevertheless, they already have a fully fledged system, as they themselves mention it [12].

6.1 General View

In [42] he characterises program-transformation updates in terms of three main objectives: *contradiction elimination*, *conflict resolution* and *syntactic representation*. The first topic is one of the most obvious in semantics for updates, which should be real by preserving a *minimal-change principle* and a proper *justification*. On the other hand, *conflict resolution* has to do with potential future *contradictions* an update might yield because of the introduction of the two kinds of negations in logic programs—strong and default negation. Finally, once the process meets the two main goals, the author argues that a proper semantics should also *preserve* as many as possible of the original rules from the *updating knowledge base*.

In order to realise these three goals, [41] performs program update transformation by means of a called *prioritised logic program*. In an intuitive way, this kind of program consists in *preferring the latest update* to the *original knowledge base* including non-contradictory but *conflicting rules*. His approach definitions and plenty of examples are easily accessible in the literature, from which we also recommend [2].

A mandatory test is the example in Observation 2, which produces *counterintuitive results* in many of the existing semantics for updates. This is not the case for Zhang's semantics, whose expected answer set is just what intuition would tell us: $\{day, \sim stars\}$.

Despite this approach is well behaved, though, one of the counter-intuitive examples to [42]'s approach has to do with solving conflicts between rules, where most of the current semantics differ, as first pointed out by [17]:

Observation 6. *Suppose an initial knowledge base* $\mathcal{K}_0 = \{p \leftarrow \text{not } q\}$ *being updated by* $\mathcal{K}_1 = \{q \leftarrow \text{not } p\}$. *Its simple-fact update specification corresponds to*

$$U_{PLP}(\mathcal{S}_{\mathcal{K}_0}, \mathcal{K}_1) = (\mathcal{K}^*, \mathcal{N}, <),$$

where

Initial Knowledge: $i_0 : p \leftarrow \emptyset$

Inertial Rules:

$$\begin{aligned} i_1 &: \quad newp \leftarrow p, \text{not} \sim newp \\ i_2 &: \quad \sim newp \leftarrow \sim p, \text{not } newp \\ i_3 &: \quad newq \leftarrow q, \text{not} \sim newq \\ i_4 &: \quad \sim newq \leftarrow \sim q, \text{not } newq \end{aligned}$$

Update Rule $u_1 : newq \leftarrow \text{not } newp$

Preferences

$$\begin{aligned} N(i_1) < N(u_1) \qquad & N(i_2) < N(u_1) \\ N(i_3) < N(u_1) \qquad & N(i_4) < N(u_1) \end{aligned}$$

As a result, its unique answer set $S_{PLP}(U_{PLP}(\mathcal{S}_{\mathcal{K}_0}, \mathcal{K}_1)) = \{p\}$ *and none of its* $<$*-relations are used. Next, the minimal subset of* $\mathcal{K}_0$ *that is coherent with its answer set is just*

$\mathcal{K}_{(\mathcal{K}_0,\mathcal{K}_1)} = \mathcal{K}_0$. *Then, the update specification of* $\mathcal{K}_0$ *and* $\mathcal{K}_1$ *is* $U_{\diamond_Z}(\mathcal{K}_0, \mathcal{K}_1) = (\mathcal{K}_1 \cup \mathcal{K}_{(\mathcal{K}_0,\mathcal{K}_1)}, \mathcal{N}, <)$, *where*

$$(t : p \leftarrow \neg q), (u : q \leftarrow \neg p), (u < t).$$

Finally, its unique reduct $q \leftarrow \text{not } p$ *comes from*

$$(\mathcal{K}_1 \cup \mathcal{K}_{(\mathcal{K}_0,\mathcal{K}_1)}) \setminus \{t\}$$

defeating rule t. *Therefore, the* ***conclusion*** *of such an update is just* $\{q\}$, *what we would not expect.*

Last, besides not satisfying some of the principles already pointed out, one of the major drawbacks of this approach is being limited to only one update to a knowledge base. Namely, it is undefined for update sequences and for *successive updates*, which does not seem to lead to immediate practical use. Although [42] also suggests an extension to one of his earliest approaches in [40] to deal with multiple updates, his proposal still makes the same strong assumptions when deciding between multiple models, as in Observation 6.

7. CONCLUSIONS

This work is a critique of current and past proposals for update semantics, as an attempt to classify them and to bring out new challenges. This Part I starts with a set of early proposals. All of them update and revise knowledge by means of program transformations in ASP or similar semantics. They run from a wide range from *simple-fact one-step updates* of logic programs, up to unlimited updates either in a sequence or by an iterative process. Some of these works present a vast collection of postulates and principles for correct knowledge-base changes, and/or implementation. However, nearly all the proposals here introduced still present drawbacks either for being limited to only one update, or for relying on syntactical principles to change the original logic program that leads to *counterintuitive results*. Nearly all of the issues here pointed out have been surpassed by current approaches, like [4], where they have found nice properties and applications [5, 3]. On the other hand, for practical and commercial reasons, the proposals here presented are often called semantics for *updates*, although there are implicit operations in some of them that may suggest a different classification, say *semantics for belief revisions*.

8. ACKNOWLEDGMENTS

The author is very grateful to J. Dix and W. Jamroga for their many useful comments and proofreading some preliminary versions of this work.

9. REFERENCES

[1] J. C. Acosta-Guadarrama. Maintaining knowledge bases at the object level. In A. Gelbukh and A. F. Kuri Morales, editors, *Special Session of the 6th International MICAI Conference*, pages 3–13, Aguascalientes, Mexico, November 2007. IEEE Computer Society. ISBN: 978-0-7695-3124-3.

[2] J. C. Acosta-Guadarrama. A road map of updating in ASP. ISSN: 1860-8477 IfI-07-16, TU-Clausthal, Clausthal, Germany, December 2007. 35pp.

[3] J. C. Acosta-Guadarrama. Towards a logic-programming system to debug ASP knowledge bases. In M. Osorio, C. Zepeda, I. Olmos, J. L. Carballido, J. Arrazola, and C. Medina, editors, *7th Latin American Workshop on Non-Monotonic Reasoning 2011*, volume Vol-804, pages 3–12, Germany, November 2011. CEUR Workshop Proceedings.

[4] J. C. Acosta-Guadarrama. Towards a unified framework for declarative knowledge-change—principles and consistency. In M. Osorio, C. Zepeda, I. Olmos, J. L. Carballido, and C. Medina, editors, *8th Latin American Workshop on Non-Monotonic Reasoning 2012*, pages 51–62. Dirección de Fomento Editorial, Benemérita Universidad Autónoma de Puebla, 2012.

[5] J. C. Acosta-Guadarrama, R. Dávila Pérez, M. Osorio, and V. Saldivar. *Modeling Natural Language Metaphors with an Answer Set Programming Framework*, volume LNAI 8856 of *ISSN 0302-9743 ISBN 978-3-319-13646-2*, chapter IV, pages 28–36. Springer-Verlag, 2014.

[6] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, June 1985.

[7] J. J. Alferes, F. Banti, A. Brogi, and P. Hitzler. The well supported semantics for multidimensional dynamic logic programs. In C. Baral, G. Greco, N. Leone, and G. Terracina, editors, *Logic Programming and Nonmonotonic Reasoning*, volume 3662/2005 of *LNCS*, pages 356–368, Diamante, Italy, September 2005. Springer Berlin/Heidelberg.

[8] J. J. Alferes, F. Banti, A. Brogi, and J. A. Leite. The refined extension principle for semantics of dynamic logic programming. *Studia Logica*, 79(1):7–32, 2005.

[9] J. J. Alferes, J. A. Leite, L. M. Pereira, H. Przymusinska, and T. C. Przymusinski. Dynamic updates of non-monotonic knowledge bases. *Journal of Logic Programming*, 45(1–3):43–70, 1999.

[10] J. J. Alferes, J. A. Leite, L. M. Pereira, T. C. Przymusinski, and H. Przymusinska. Dynamic logic programming. In A. Cohn, L. Schubert, and S. Shapiro, editors, *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR-98)*, pages 98–111, San Francisco, June 2–5 1998. Morgan Kaufmann Publishers.

[11] J. J. Alferes, L. M. Pereira, H. Przymusinska, and T. C. Przymusinski. LUPS —A language for updating logic programs. *Artificial Intelligence*, 138(1–2):87–116, June 2002.

[12] V. F. Crescini and Y. Zhang. Policy updater: A system for dynamic access control. *International Journal of Information Security*, 5(3):145–165, 2005.

[13] E. Dantsin, T. Eiter, G. Gottlob, and A. Voronkov. Complexity and expressive power of logic programming. *ACM Computing Surveys*, 33(3):374–425, September 2001.

[14] T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. Considerations on updates of logic programs. In M. Ojeda-Aciego, I. P. de Guzmán, G. Brewka, and L. Moniz Pereira, editors, *Logics in Artificial Intelligence, European Workshop, JELIA 2000*, pages 2–20, Malaga, Spain, 2000. Springer Verlag.

[15] T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. On

updates of logic programs: Semantics and properties. Technical Report INFSYS RR-1843-00-08, TU Wien, Institute für Informationssysteme, 2000.

[16] T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. A framework for declarative update specifications in logic programs. In B. Nebel, editor, *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI*, volume I, pages 649–654, Seattle, Washington, 2001. Morgan Kaufmann.

[17] T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. On properties of update sequences based on causal rejection. *Theory and Practice of Logic Programming*, 2(6):711–767, 2002.

[18] T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. Reasoning about evolving nonmonotonic knowledge bases. *ACM Transactions on Computational Logic*, 6(2):389–440, 2005.

[19] R. Fagin. *Reasoning about Knowledge*. The MIT Press, Cambridge, Massachusetts, USA, 1995.

[20] R. Fagin, G. M. Kuper, J. D. Ullman, and M. Y. Vardi. Updating logical databases. *Advances in Computing Research*, 3:1–18, 1986.

[21] R. Fagin, J. D. Ullman, and M. Y. Vardi. On the semantics of updates in databases. In *PODS '83: Proceedings of the 2nd ACM SIGACT-SIGMOD symposium on Principles of database systems*, pages 352–365, New York, NY, USA, 1983. ACM Press.

[22] M. Gelfond and V. Lifschitz. The Stable Model Semantics for Logic Programming. In R. A. Kowalski and K. A. Bowen, editors, *Logic Programming, Proceedings of the Fifth International Conference and Symposium ICLP/SLP*, pages 1070–1080, Seattle, Washington, 1988. MIT Press.

[23] H. Katsuno and A. O. Mendelzon. A unified view of propositional knowledge base updates. In N. S. Sridharan, editor, *the 11th International Joint Conference on Artificial Intelligence, IJCAI-89*, pages 1413–1419, Detroit, Michigan, USA, 1989. Morgan Kaufmann.

[24] H. Katsuno and A. O. Mendelzon. On the difference between updating a knowledge base and revising it. In *KR'91*, pages 387–394, Cambridge, Massachusetts, USA, 1991. Morgan Kaufmann Publishers.

[25] H. Katsuno and A. O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, 1991.

[26] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, 44(1):167–207, 1990.

[27] D. Lehmann. Plausibility Logic. In E. Börger, G. Jäger, H. Kleine-Büning, and M. M. Richter, editors, *Computer Science Logic, 5th Workshop, CSL 91, Berne, Switzerland*, LNCS 626, pages 227–241, Berlin, 1992. Springer.

[28] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[29] D. Makinson. General theory of cumulative inference. In *Non-Monotonic Reasoning, 2nd International Workshop*, pages 1–18, Grassau, Federal Republic of Germany, 1988. Springer.

[30] D. Makinson. General Patterns in Nonmonotonic Reasoning. In D. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming, Nonmonotonic and Uncertain Reasoning*, volume 3, chapter 3, pages 35–110. Oxford University Press, Oxford, UK, 1994.

[31] M. Osorio, J. A. Navarro, and J. Arrazola. Equivalence in answer set programming. In A. Pettorossi, editor, *Logic Based Program Synthesis and Transformation: 11^{th} International Workshop; selected papers / LOPSTR 2001, Paphos, Cyprus*, LNCS 2372, pages 57–75, Paphos, Cyprus, November 2001. Springer-Verlag Berlin.

[32] M. Osorio and F. Zacarías. Irrelevance of syntax in updating answer set programs. In *Proceedings of the Fourth Mexican International Conference on Computer Science (ENC' 03) In Workshop on Logic and Agents*, Apizaco, Mexico, 2003. IEEE Computer Society.

[33] M. Osorio and F. Zacarías. On updates of logic programs: A properties-based approach. In *FoIKS*, pages 231–241, Wilhelminenburg Castle, Austria, 2004. Springer.

[34] C. Sakama and K. Inoue. Updating extended logic programs through abduction. In M. Gelfond, N. Leone, and G. Pfeifer, editors, *LPNMR '99: Proceedings of the 5th International Conference on Logic Programming and Nonmonotonic Reasoning*, volume 1730 of *LNCS*, pages 147–161, El Paso, Texas, USA, 1999. Springer-Verlag.

[35] C. Sakama and K. Inoue. An abductive framework for computing knowledge base updates. *Theory and Practice of Logic Programming*, 3(6):671–715, 2003.

[36] M. Winslett. *Updating logical databases*. Cambridge University Press, New York, NY, USA, 1990.

[37] F. Zacarías, M. Osorio, J. C. Acosta-Guadarrama, and J. Dix. Updates in Answer Set Programming Based on Structural Properties. In S. McIlraith, P. Peppas, and M. Thielscher, editors, *7th International Symposium on Logical Formalizations of Commonsense Reasoning*, pages 213–219, Corfu, Greece, May 2005. Fakultät Informatik, TU-Dresden, ISSN 1430-211X.

[38] Y. Zhang. On propositional knowledge base updates. *Australian Journal of Intelligent Information Processing Systems*, 2:20–29, 1995.

[39] Y. Zhang. The complexity of logic program updates. In *AI '01: Proceedings of the 14th Australian Joint Conference on Artificial Intelligence*, pages 631–642, London, UK, 2001. Springer-Verlag.

[40] Y. Zhang. Minimal change and maximal coherence for epistemic logic program updates. In *IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, pages 112–120, Acapulco, Mexico, 2003. Morgan Kaufmann.

[41] Y. Zhang. Two results for prioritized logic programming. *Theory and Practice of Logic Programming*, 3(2):223–242, March 2003.

[42] Y. Zhang. Logic program-based updates. *ACM Transactions on Computational Logic*, 7(3):421–472, July 2006.

Silla de Ruedas Autónoma e Inteligente que Avanza a Lado del Acompañante

Luis A. Jara-Lugo
Maestría en Mecatrónica, Instituto Tecnológico Superior de Cajeme
Ciudad Obregón Son. México
5216441412177
ljara@itesca.edu.mx

Fernando Orduña-Cabrera
Instituto Tecnológico Superior de Cajeme,
SEKAI.MX
Cd. Obregón Sonora, México
521644-4108650
forduna@itesca.edu.mx

Alberto Ramírez-Treviño
Instituto Tecnológico Superior de Cajeme
Ciudad Obregón Son. México
521644-4108650
aramirez@itesca.edu.mx

RESUMEN
La tecnología avanza cada día y por ello cada día hay nuevos métodos y nuevas tecnologías que facilitan/mejoran la calidad de vida de la gente. Considerando este último como prioridad presentamos un método de seguimiento autónomo al acompañante. Este método es aplicado en una silla de ruedas y permite controlar el avance de la silla, teniendo así una silla de avance autónomo que permite ir a lado al acompañante en el paseo.

1. INTRODUCCIÓN
A pesar de los avances tecnológicos que se han implementado para mejorar sillas de ruedas [1], las personas con limitaciones severas de movimientos (enfermedades musculares, accidentes, vejez, reposo medico etc.), siguen teniendo dificultad para trasladarse por sí mismos.

En el siglo XXI se han modificado las sillas de ruedas y han surgido diferentes tipos como lo es la silla de ruedas controlada por voz [2], la silla de ruedas controlada por la mente [3], la silla de ruedas que aprende de la experiencia [4], la sillas de ruedas con GPS [5], y la silla de ruedas controlada por los movimientos naturales de la cabeza [6]. Todos estos son avances que facilitan el movimiento, cabe señalar que cada uno de ellos lleva acabo procesos que requieren de cálculos complejos computacionalmente, lo que lleva tiempo respuesta. El método aquí presentado es un método en tiempo real basado en la implementación de hardware más reactivo que deliberativo, lo que nos permite un tiempo de respuesta y acción rápida.

En la actualidad el seguimiento de objetos se realiza mediante sensores ultrasónicos [7] o cámaras web [8]. El sensor ultrasónico es comúnmente utilizado en robots que con ayuda de un pequeño servo motor determina la localización de dicho objeto cuando el sensor lo detecta, mientras que la cámara utiliza algoritmos y técnicas de visión artificial para detectar diferentes tipos de objetos.

2. SISTEMA DE SENSORES
Las dimensiones ya estandarizadas de las sillas de ruedas son consideradas para la ubicación del hardware agregado a la silla, para ello se ha diseñado el hardware de adquisición de señales y se ubicó sobre el apoya brazos de la silla. La cual consiste en el uso de tres sensores ultrasónicos modelo HSR04 que trabaja con un voltaje de 5 volts [9] a una frecuencia de 40 Hz, su alcance de detección es de 4 metros como máximo y 2 centímetros como mínimo, teniendo un ángulo de apertura de 15 grados. Y trabaja con 5 volts.

Para encontrar la ubicación más propicia de los sensores se realizaron pruebas en donde se consideró la evaluación de la ubicación del acompañante a diferentes distancias y ángulos, hasta obtener resultados convincentes donde se determinó que la distancia que debe existir entre la persona acompañante y la silla de ruedas fuera de 60 centímetros, ya que es una distancia segura en caso de cualquier accidente que pudiera ocurrir, además es una distancia donde el acompañante puede alcanzar a reaccionar de manera rápida y también está lo suficientemente lejos como para lograr tener una buena interacción con el usuario. La silla de ruedas tiene sus apoya-brazos de 50 centímetros de largo, 8 centímetros de ancho y a una altura de 70 centímetros, por lo que fue necesario colocar los tres sensores a una distancia menor de 50 centímetros.

Después de una serie de pruebas realizadas se obtuvo que la distancia idónea entre sensor y sensor fuera de 22.5 centímetros, lo que da una distancia de 45 centímetros en total, ya que mientras más separados más grande es el área de detección de los sensores, dejando así un margen de 5 centímetros libres en la parte de enfrente del apoya brazos, por si el acompañante requiere sujetar la silla. Teniendo un sensor a la izquierda (sensor 1) otro a la mitad (sensor 2) y el otro a la derecha (sensor 3). Como se muestra en la figura 1.

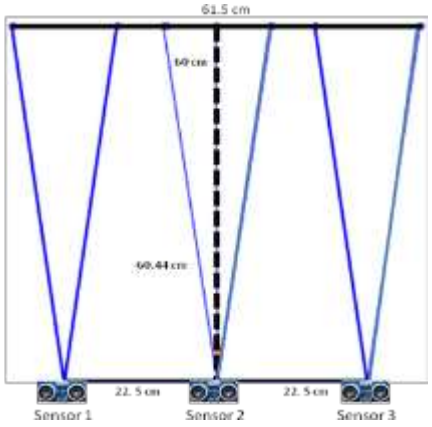


Figura 1. **Sensores ultrasónicos separados a una distancia de 22.5 centímetros entre ellos.**

También se ha considerado la velocidad promedio a la que camina una persona (0.92 m/seg), la cual es una velocidad rápida para un

paseo relajado. Para un paseo relajado el acompañante debe de caminar más lento (0.5 m/seg).

2.1 DETECCIÓN DE ÁREAS

El área donde el acompañante puede ser captado por cualquiera de los tres sensores es nombrada como "área de detección", esta debe de ser lo más amplia posible. Al evaluar la ubicación reflejada en la figura 1 no se obtuvieron resultados favorables, debido al rango de apertura en el sensado, (15 grados). Según como se muestra en la Figura 1, el largo del área de detección máxima es los 60 centímetros de distancia a la silla la cual es un área que nos genera solo 3 posibles áreas de ubicación del acompañante. Con estas características no es posible identificar comportamientos del acompañante como es el de dar vuelta a la izquierda y derecha.

Buscando una mejor detección del comportamiento del acompañante se experimentó modificando el ángulo del censado; quedando finalmente el "sensor 1" a un ángulo de 40 grados, el "sensor 2" a 90 grados y el "sensor 3" a 140 grados, así como muestra la figura 2.

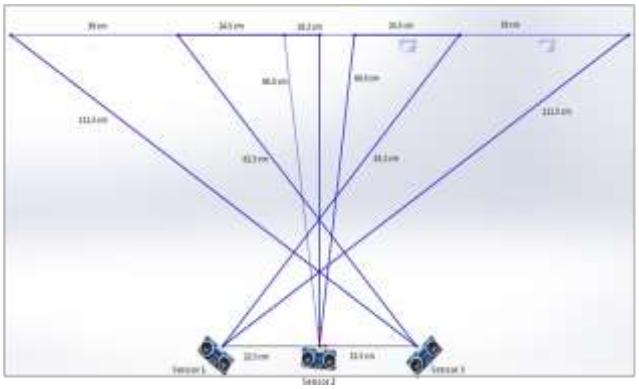


Figura 2. Sensores con el ángulo cerrado ("sensor 1" a 40°, "sensor2" a 90° y sensor 3 a 140°)

En la figura 2 se muestra como lo largo del área de detección a una distancia de 60 centímetros entre los sensores y el acompañante se extiende a 143.2 centímetros. Con las modificaciones se encontró el área óptima de ubicación del acompañante, la cual es de 65 centímetros, y también obtenemos tres áreas de ubicación. Se realizaron 1000 pruebas tomando en cuenta los movimientos de personas con diferentes características de peso y estatura. Las pruebas consistieron en caminar al frente y atrás, pararse frente al área de censado, girar a la derecha e izquierda. Para cada ejercicio hecho se consideraron las distancias arrojadas por cada sensor.

Se encontró que según la ubicación del acompañante los sensores marcan una distancia diferente, lo que permite hacer una identificación más detallada de la ubicación del acompañante como se muestra en la figura 3. Como se muestra en la figura son 11 áreas posibles de ubicación del acompañante, por ejemplo en el "área 1" el "sensor 1" es el único que puede detectar al acompañante con un rango de distancia de 80 a 100 centímetros, donde si el acompañante se mueve al "área 11" que se encuentra arriba de la uno lo detectara el mismo "sensor 1" con un rango de 100 a 120 centímetros. Una relación de rangos se muestra en la tabla 1.

La tabla 1 muestra los rangos de las distancias por cada área en centímetros y se utiliza el término FR para indicar que el sensor queda fuera de rango es decir que no detecta al acompañante en su área de censado.

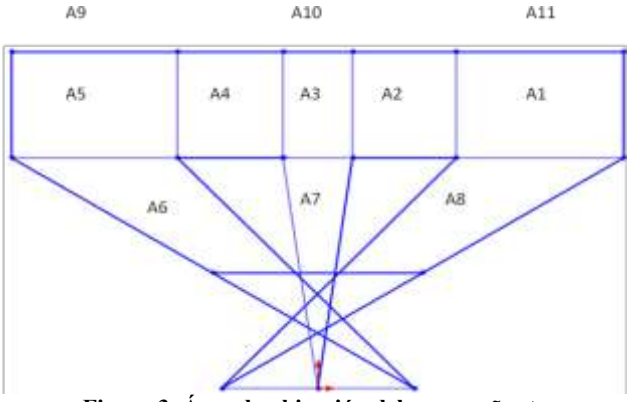


Figura 3. Área de ubicación del acompañante

Tabla 1. Rangos de las distancias de cada área

Área	Sensor 1	Sensor 2	Sensor 3
A1	80-100	FR	FR
A2	65-80	58-70	FR
A3	FR	50-70	FR
A4	FR	58-70	65-80
A5	FR	FR	80-100
A6	FR	FR	55-80
A7	20-40	20-40	20-40
A8	55-80	FR	FR
A9	FR	75-120	90-120
A10	FR	80-120	FR
A11	90-120	75-120	FR

También se presentan casos en los que 2 sensores se activan por la presencia del acompañante y otros casos los 3 sensores son activados como lo es el área 7, considerada como la ubicación más idónea.

2.2 COMPORTAMIENTOS DE LA SILLA SEGÚN UBICACIÓN DE ACOMPAÑANTE

De acuerdo a las áreas identificadas se determinaron los comportamientos o acciones que se deben llevar acabo y permitan que la silla se mantenga a lado del acompañante durante el paseo.

Tabla 2. Comportamientos por área

Área	Comportamiento
1	Frenar
2	Seguir
3	Seguir
4	Seguir
5	Acelerar
6	Frenar y doblar a la izquierda
7	Seguir y alertar al acompañante
8	Frenar
9	Acelerar y doblar a la derecha
10	Seguir constante acercándose a la derecha
11	Frenar

Considerando el área donde se encuentra ubicado el acompañante le corresponderá una acción o serie de acciones que permitan que la silla se mantenga a lado del acompañante; si el acompañante entra a un área donde el comportamiento de la silla es "seguir", la silla continuara a una velocidad constante de 0.5 m/seg. Si el acompañante llega a un área donde el comportamiento es "acelerar", la silla acelera hasta llegar a una velocidad de 0.75 m/seg, teniendo la opción de alcanzar al acompañante y obtener de nuevo la velocidad constante. Si el acompañante entra a un área donde el comportamiento de la silla es "frenar", la silla disminuirá su velocidad a 0.25 m/seg buscando ubicar nuevamente al acompañante en el área óptima. Si el acompañante no es detectado por ningún sensor la silla estará en alto total. Si el comportamiento es "doblar a la derecha" la silla va acelerar hasta los 0.75 m/seg para posteriormente doblar sus llantas 30 grados a la derecha. Si el comportamiento es "doblar a la izquierda" la silla va a frenar hasta los 0.25 m/seg para posteriormente doblar sus llantas 20 grados a la izquierda. Si el comportamiento es "alertar" la silla de ruedas simplemente indicará mediante una bocina al acompañante que se aleje. Si el comportamiento es "acercarse" la silla de ruedas a la velocidad promedio de 0.5 m/seg, girará las llantas 10 grados a la derecha y avanzará 50 centímetros y volverá a girar 10 grados hacia la izquierda y posteriormente avanzará.

2.3 AUTÓMATA DE DESPLAZAMIENTO

Cada una de las áreas de detección colinda con diferentes áreas y el acompañante es capaz de moverse y ser detectado por los sensores en cualquier combinación de dos áreas vecinas; algunas de estas combinaciones generan diferentes comportamientos donde solamente se toma en cuenta un área. Las combinaciones posibles se pueden mostrar en la Figura 4, donde si el acompañante es detectado en el área A1 las siguientes posibles ubicaciones son las áreas vecinas A8, A11 y A2.

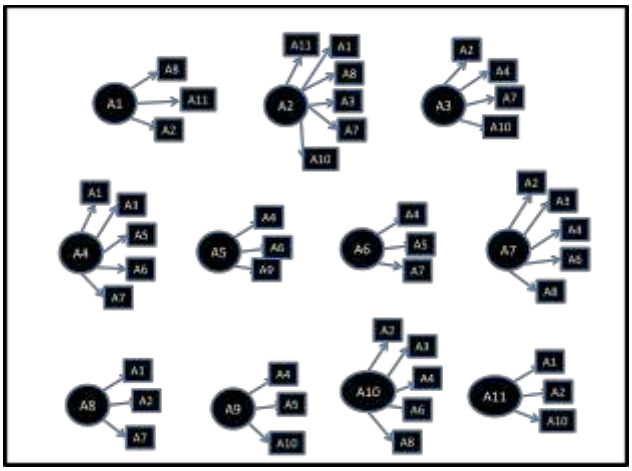


Figura 4. Autómata de desplazamiento del acompañante

Los comportamientos obtenidos de las combinaciones de los desplazamientos de áreas se muestran en la matriz de acciones y desplazamientos, como se muestra en la Tabla 3. Las filas indican el área donde es ubicado el acompañante, las columnas indican el área vecina a la que se puede desplazar el acompañante, y el número dentro de la celda indica el comportamiento de la silla en cada combinación. Por ejemplo si el acompañante está ubicado en el A1 sus posibles desplazamientos son sus áreas vecinas y la intersección de estas en la tabla 3 es la celda que alberga el comportamiento o acción que debe llevar acabo el controlador de la silla para ubicar al acompañante en el área idónea de seguimiento.

Tabla 3. Matriz de acciones y desplazamientos.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
A1	2	1	0	0	0	0	0	2	0	0	2
A2	2	0	1	0	0	0	7	2	0	6	2
A3	0	1	0	1	0	0	7	0	0	6	0
A4	0	0	1	1	3	4	7	0	5	6	0
A5	0	0	0	1	3	4	0	0	5	0	0
A6	0	0	0	1	1	4	7	0	0	0	0
A7	0	1	1	1	0	4	7	2	0	0	0
A8	2	0	0	0	0	0	7	2	0	0	0
A9	0	0	0	1	1	0	0	0	5	6	0
A10	2	1	1	1	3	0	0	0	5	6	2
A11	2	0	0	0	0	0	0	0	0	7	2

La relación de los comportamientos según su número en la tabla 3 se describen en la tabla 4, donde el comportamiento 1 indica la acción de Seguir, que indica continuar con el comportamiento que en ejecución, para este caso seguir adelante.

Tabla 4. Comportamientos según su número

Comportamiento	Acción
1	Seguir
2	Frenar
3	Acelerar
4	Vuelta a la izquierda
5	Vuelta a la derecha
6	Acercarse
7	Avisar

2.4 GIRO A LA DERECHA

Si el acompañante entra en un área donde el comportamiento es doblar a la derecha (según tabla 3, estando A4 o A5 desplazándose a V9) las llantas de la silla cambiaran su dirección 30 grados hacia la derecha y la silla va a acelerar tal como se muestra en la figura 5.1

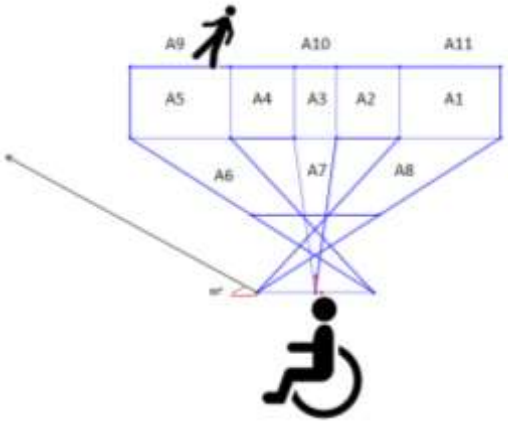


Figura 5.1

La silla de ruedas después de que recorre una distancia de 50 centímetros a la derecha ubica al acompañante en A4. Figura 5.2

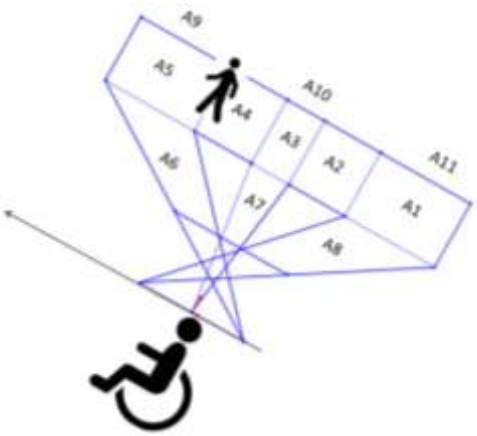


Figura 5.2

En el A4 el comportamiento de la silla es "seguir", pero si el acompañante sigue caminando hacia la derecha, dando una vuelta de 90 grados, el acompañante volverá al A9 y la tendrá que dar de nuevo vuelta a la derecha. Figura 5.3

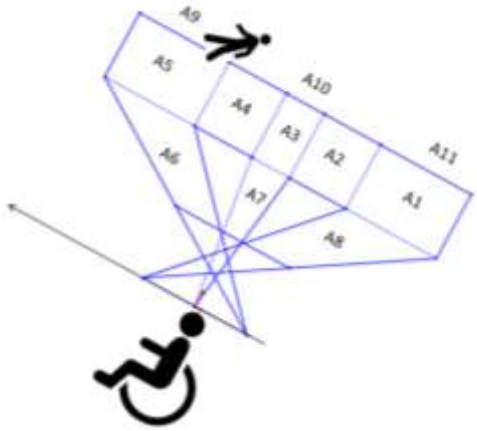


Figura 5.3

De nuevo la silla dobla 30 grados hacia la derecha y recorre una distancia de 50 centímetros. Logrando de nuevo ubicar al acompañante en A4. Tal como se muestra en la figura 5.4

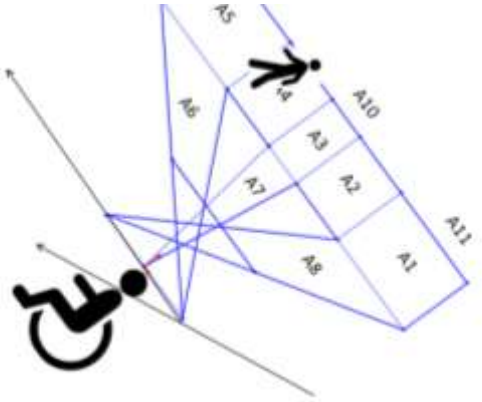


Figura 5.4

Por último si el acompañante sigue caminando hacia la misma dirección entrará a A5 y posteriormente al A9 donde automáticamente la silla dará su último movimiento de 30 grados y dar el recorrido de 50 centímetros, para quedar de frente al acompañante ubicándolo en el A7. Donde el comportamiento de la silla es seguir y avisar que está demasiado cerca para que el acompañante se acomode alejándose un poco de la silla. Figura 4.5.

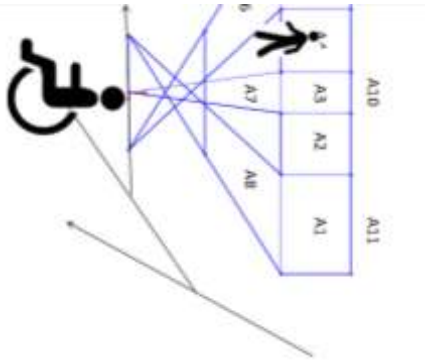


Figura 5.5

Siguiendo una dinámica parecida se ejecutan las diferentes acciones de vuelta a la izquierda.

3. CONCLUSIONES Y TRABAJO FUTURO

Como se ha mostrado en este trabajo es posible generar comandos de control para una silla de ruedas a partir de los movimientos de una persona acompañante, sin embargo esta debe de ubicarse a una distancia apropiada de la silla, lo que no siempre sucede, por lo que esto constituye un área de mejora.

Se demostró que el problema de seguimiento puede ser discretizado, lo que permite el uso de un autómata mejorando así el sistema de respuesta de la silla y facilitando la embebida en los microcontroladores.

Es de gran ayuda para el seguimiento automático, el determinar cada uno de los comportamientos de la silla en diferentes escenarios planteados, tomando en cuenta que el acompañante conoce las capacidades y movimientos de la silla.

Si el acompañante cambia la distancia entre él y la silla puede generar escenarios donde la silla puede perder la huella del acompañante. En un futuro se agregaran nuevos escenarios agregando el seguimiento mediante visión artificial y un algoritmo que compare ambos métodos, para que la silla tome una decisión idónea de una manera más inteligente en nuevos escenarios. Estos nuevos escenarios serán organizados siguiendo técnicas de indexación de la información [10, 11, 12] que nos permitan recuperar los casos que se presentan en entorno de la silla, estas técnicas debido a que son para implementarse en dominios continuos nos permitirán una reacción más eficiente y precisa, como se demuestra en [10, 11, 12]. Todo esto enfocado a obtener una silla inteligente seguidora del acompañante.

4. REFERENCIAS

[1] EL IMPACTO TECNOLÓGICO EN LAS PERSONAS CON DISCAPACIDAD. Ricardo A. Koon, María Eugenia de la Vega.2000
http://repositoriocdpd.net:8080/bitstream/handle/123456789/363/Pon_KoonRA_ImpactoTecnologicoPersonas_2000.pdf?sequence=1

[2] Silla de Ruedas Inteligente Controlada por Voz. Alcubierre, Minguez, Montesano Montano, Saz, Lleida. Primer Congreso Internacional de Domótica, Robótica y Teleasistencia para todos, 2005 Fundación ONCE.

https://webdiis.unizar.es/~jminguez/Silla%20de%20Ruedas%20Inteligente%20Controlada%20por%20Voz.pdf

[3] Developing Intelligent Wheelchairs for the Handicapped. Takashi Gomi and Ann Griffith. Kanata, Ontario, CANADA 1998.

http://cs.unc.edu/~welch/class/mobility/papers/yanco.pdf

[4] MIT Tech Talk- Finding its own way. David Chandler. Nokia and Microsoft 2008

http://news.mit.edu/2008/wheelchair-0919

[5] Towards Autonomous Wheelchair Systems in Urban Environments. Chao Gao, Michael Sands, and John R. Spletzer.2009
http://www.nrec.ri.cmu.edu/fsr09/papers/FSR2009_0064_5f4061377ab8e14454791a1877907bb3.pdf

[6] EMG-based Hands-Free Wheelchair Control with EOG Attention Shift Detection. Chun Sing Louis Tsui, Pei Jia, John Q. Gan, Huosheng Hu and Kui Yuan. International Conference on Robotics and Biomimetics December 15 -18, 2007, Sanya, China

http://cswww.essex.ac.uk/staff/hhu/Papers/ROBIO07-222.pdf

[7] Seguimiento de Objetos empleando Aforge.Net y Arduino. Sáenz Esqueda, José Armando; Sáenz López, Agustín; Reyes Espino, Vicente; Morales Salazar, Rafael. Revista de Arquitectura e Ingeniería, vol. 8, núm. 1, abril-, 2014

http://www.redalyc.org/pdf/1939/193931237002.pdf

[8] Design of Semi-Autonomous Robots for Competitive Robotics. P. Benavidez, C. Gleinser, A. Jaimes, J. Labrado, C. Riojas and Mo Jamshidi, Ph.D., Lutcher Brown Endowed Chair Department of Electrical and Computer Engineering The University of Texas at San Antonio

http://www.wacong.org/ace/Assets/doc/1569553971.pdf

[9] Ultrasonic Ranging Module HC - SR04. Data Sheet http://www.micropik.com/PDF/HCSR04.pdf

[10] A Dynamic Adaptive Framework for improving Case-Based Reasoning System Performance, Fernando Orduña-Cabrera, Ph. D Dissertation. Dept. of Comp Science, Univ. Politècnica de Catalunya, 2015.

[11] Using NIAR kd Trees to Improve the Case-Based Reasoning Retrieval Step, Fernando Orduña-Cabrera, Mexican International Conference on Artificial Intelligence, 314-325, Springer Berlin Heidelberg, 2013.

[12] Negotiation and learning in distributed MPC of large scale systems, Javalera Rincón, Valeria; Morcego Seix, Bernardo; Puig Cayuela, Vicenç (IEEE Press. Institute of Electrical and Electronics Engineers, 2010).

Multi-objective Optimization for Local Descriptors in Image Mosaicing

Cynthia B. Perez
Sonora Institute of Technology
Ciudad Obregón, Sonora, México
cynthia.perez@itson.edu.mx

Daniel Mendoza
Sonora Institute of Technology
Ciudad Obregón, Sonora, México

ABSTRACT

Image mosaicing through invariant features has rapidly emerged to become a main paradigm for the computer vision community due to its robustness against scene movement, speed in matching features and the ability to work without initialization. This paper refers to the problem of synthesizing descriptor operators for affine image transformations using a multi-objective genetic programming (MO-GP) framework. The descriptor operators represent mathematical expressions that would be used within a descriptor algorithm to characterize the description of local features. This research follows the idea that affine image transformations are presented in the process of creating mosaics; for this reason, matching through invariant features has turned into the best approach for this application. Thus, we expect to provide mathematical expressions that outperform the performance of traditional descriptor algorithms for affine transformations in an image mosaicing application.

Keywords

Genetic programming; Local descriptors; Multi-objective optimization; Image mosaicing

1. INTRODUCTION

Image mosaicing is an active area of research in computer vision, with applications in robotics, architecture, medicine, surveillance and industrial inspection, to mention but a few [12]. Image mosaicing is the process by which multiple images, that provide different views of a real-world scene, are combined to form a larger composite image called a mosaic.

Recently, there has been great progress in the use of invariant features for image mosaicing [1, 15]. These features are used to localize and match the same salient features in all images in order to create the corresponding mosaic. Thus, the basic idea on using local features consists on three principal stages. First, the detection of interest points, or regions, using specially designed operators that are covariant to a class of transformations [13]. Second, an invariant descriptor is computed for each detected point where the principal objective is to obtain a compact region description, usually in the form of a numerical vector, that captures the information about the local structure around each interest point. Finally, the image descriptors are matched in order to obtain the geometric relations between images.

Regarding features descriptors, the general consensus is that the Scale Invariant Feature Transform (SIFT), proposed by David Lowe [8], represents the state-of-the-art of local descriptor research. This algorithm provides a fast and reliable solution for matching the same candidate points that are present in two images. Hence, the SIFT algorithm has been widely used in different high-level vision tasks, including image mosaicing [4].

Therefore, a number of research works in the literature have attempted to optimize this descriptor in different ways [2, 11]. In this work, we are interested in developing an implementation of multi-objective genetic programming as the core platform to evolve mathematical operators that will replace the gradient magnitude used in the SIFT description process which can be extended to other local descriptors. Thus, the idea is to obtain an optimal affine-invariant operator for the SIFT descriptor.

The application of genetic programming (GP) to image related problems has received an increase level of attention in recent years. However, there is still a small number of relevant works that use MO-GP for computer vision problems. For instance, Krawiec [6] introduced a learning method of visual concepts for shape recognition; and Olague and Trujillo proposed en evolutionary design of interest point detectors through MO-GP [10]. On the other hand, in our previous work [11] we have shown how to optimize the SIFT descriptor using a canonical GP while keeping the DoG detector unchanged. The idea was to obtain scale-invariant-descriptor operators that could not-only replace the gradient magnitude within the SIFT descriptor, but also, that could outperform all other state-of-the-art descriptors. Thus, we correctly posed the problem of learning descriptor operators as mono-objective optimization, using the F-Measure proposed by Van-Rijsbergen [14], obtaining outstanding results. However, in this work, we pretend to explore a different optimization scenario by posing the problem of learning descriptor operators in terms of a multi-objective optimization given recall and precision as intrinsically conflicting objectives. Hence, we propose a MO optimization framework within genetic programming to automatically synthesize affine-invariant descriptor operators.

To summarize, the main contributions of this proposal are described as follows.

- This paper studies the problem of local descriptors from the standpoint of MO optimization using genetic programming, an approach that is unique in this field.
- The MO-GP algorithm could synthesize descriptor operators using an affine detector for an image mosaicing application, considering two widely used performance criteria: recall and precision.

- We would validate our proposal once the algorithm is implemented in a real-world problem such image mosaicing.

The remainder of this paper is organized as follows. Section 2 presents the performance evaluation of local descriptors and describes the optimization criteria considered in this paper. In Section 3 we describe the MO-GP methodology for learning SIFT descriptor operators. Section 4 presents an overview about our MO-GP algorithm proposal applied in image mosaicing. Finally, we conclude the paper in Section 5, with a summary and some final remarks.

2. PERFORMANCE EVALUATION OF LOCAL FEATURES

Local features are commonly employed in a wide range of real-world applications such as: object recognition, image matching, medicine, and with several commercial purposes due to their robustness, invariance and efficiency that they can provide. Local features consist on localizing interest points according to their position and scale within an image, followed by a description process for each image region, where the interest points are obtained; then, the result is a set of descriptors for each image. The purpose of local descriptors is to obtain a compact and reliable description that numerically captures information regarding the local shape and appearance around each interest point. Afterwards, each descriptor vector can be matched between different images if any local correspondences exist. It is expected that this matching should be invariant to image transformations such as rotations, scaling, illumination changes, affine distortion, eyefish lens deformation, to mention but a few.

The testbed of Mikolajczyk and Schmid [9] is widely accepted in the computer vision community as a standard performance evaluation for local descriptors. It contains images with different geometric and photometric transformations such as rotation, scaling, blur, illumination, JPEG compression and affine transformation. In this work, we are interested on using an image pair with affine transformation for learning and testing descriptor operators with the aim to apply these operators on an image mosaicing problem. Moreover, the test of local descriptors is based on the number of correct and false matches produced during the correspondence of descriptors between two images. The idea is to create a Recall vs 1-Precision curve that represents the descriptor performance for a type of image transformation. This curve is obtained by varying a threshold t in the matching process. Thus, Recall and 1-Precision metrics are defined as follows:

$$Recall = \frac{\#correct \quad matches}{\#correspondences} \ , \tag{1}$$

$$1{-}Precision = \frac{\#false \quad matches}{\#correct \quad matches \quad + \quad \#false \quad matches} \ , \tag{2}$$

Recall refers to the number of correctly matched regions with respect to the number of corresponding regions between each image pair like (I_A, I_B) using an overlap error. This overlap error measures how well the regions correspond under a homographic transformation according to the ratio of intersection and union of two regions R_{I_A} and R_{I_B}. On the other hand, 1-Precision refers to the number of falsely matched regions with respect to the total number of regions being matched. The false matches are computed from total matches minus the correct matches. Finally, the result of this test is a plot that allows us to visualize the performance curves of each descriptor where a perfect descriptor would give a recall equal to 1 for any precision.

3. LEARNING LOCAL DESCRIPTORS WITH MO-GP

In this section we describe the proposed implementation of a MO-GP algorithm that could automatically generate affine descriptor operators. In general, the goal is to explore a multi-objective optimization framework for learning affine descriptor operators. Hence, in order to understand our proposed methodology we define the basic concepts of genetic programming and multi-objective optimization.

Genetic Programming (GP) is an evolutionary computation methodology inspired from the principles of Darwinian evolution [5] with the aim to automatically create computer programs that learn through a user-defined function or fitness. A GP algorithm starts with an initial population of randomly generated programs. Normally, individual programs are represented using tree structures that are evaluated through the fitness function.

Furthermore, many real-world problems sometimes involve simultaneous optimization of several and often conflicting objectives due to the nature of the problem or due to a special interest in exploring a novel problem search space. Hence, multi-objective (MO) optimization using evolutionary algorithms (MOEAs), seem to be particularly well-suited for these kind of problems due to their inherent parallelism and their capability to exploit solutions in a multidimensional space, [3]. Moreover, it is possible to optimize several objectives at the same time by searching the set of optimal solutions that converge to the Pareto front. In MO optimization we have several objectives or functions (equally important) and decision variables in two complementary spaces. These two spaces are related by the mapping $\vec{f}: \mathcal{R}^n \rightarrow \mathcal{R}^k$ where a feasible region $\Omega \subseteq \mathcal{R}^n$ is defined by the constraints on the objective vector $\vec{f}(x) = [f_1(x), f_2(x), ..., f_k(x)]$. Formally, the optimality of a MO problem is defined as follows:

DEFINITION 3.1. *MO Problem: Optimize the vector function*

$$\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), ..., f_k(\vec{x})]^T, \tag{3}$$

finding the vector $\vec{x}^* = [\vec{x_1}^*, \vec{x_2}^*, ..., \vec{x_n}^*]^T$ *which will satisfy the following* r *inequality constraints*

$$g_i(\vec{x}) \geq 0 \quad where \quad i = 1, 2, ..., r \tag{4}$$

and the t *equality constraints*

$$h_i(\vec{x}) = 0 \quad where \quad i = 1, 2, ..., t \tag{5}$$

where the vector $\vec{x} = [x_1, x_2, ..., x_n]^T$ *is a vector of decision variables.*

DEFINITION 3.2. *Pareto Optimality: A point* $\vec{x}^* \in \Omega$ *(*$\Omega \subseteq \mathcal{R}^n$*) is* ***Pareto Optimal*** *if for every* $\vec{x} \in \Omega$ *and* $I = 1, 2, ..., k$ *either* $\forall_{i \in I}(f_i(\vec{x}) = f_i(\vec{x}^*))$ *or, there is at least one* $i \in I$ *such that*

$$f_i(\vec{x} > f_i(\vec{x}^*)) \tag{6}$$

DEFINITION 3.3. *Pareto Dominance: A vector* $\vec{u} = (u_1, u_2, ..., u_k)$ *is said to dominate* $\vec{v} = (v_1, v_2, ..., v_k))$ *(also denoted by* $\vec{u} \preceq \vec{v}$*) if and only if* u *is partially less than* v*, i.e.,* $\forall_i \in 1, 2, ..., k,\ u_i \leq v_i \wedge \exists_i \in \{1, 2, ..., k\} : u_i < v_i$

DEFINITION 3.4. *Pareto Optimal Set: The Pareto optimal set (*$\mathcal{P}^*$*) is defined for a given MO Problem* $\vec{f}(x)$*, as*

$$\mathcal{P}^* := \{x \in \Omega \mid \neg\exists x' \in \Omega \text{ that } \vec{f}(x') \preceq \vec{f}(x)\} \quad (7)$$

DEFINITION 3.5. *Pareto Front: The Pareto front (*$\mathcal{PF}^*$*) is defined for a given MO problem* $\vec{f}(x)$ *and Pareto optimal set* $\mathcal{P}^*$*, as*

$$\mathcal{PF}^* := \{\vec{u} = \vec{f} = (f_1(x), ..., f_k(x)) \mid x \in P^*\} \quad (8)$$

In this way, the evolutionary search is focused on providing an optimal set of non-dominated solutions. To incorporate MO principles within an evolutionary algorithm such as GP, two important adaptations are required. First, the evolutionary search algorithm has to evolve a Pareto front of non-dominated solutions with respect to multiple objectives, instead of evolving a single fittest solution. Second, fitness assignment should be based on Pareto Dominance relations. In this work, we pretend to use the Strength Pareto Evolutionary Algorithm 2 (SPEA2), which assigns the fitness values on the basis of dominance rank and count [16]. Furthermore, SPEA2 is a robust technique for approximating the Pareto set for MO problems. This technique has compared favorably with other MOEA algorithms and has been applied in several application domains [7, 10].

4. MO-GP ALGORITHM

The design of descriptor operators using a GP approach was previously proposed by Perez and Olague [11]. The idea was to synthetize mathematical operators obtained by a canonical GP that would replaced the gradient magnitude which is used as a weight in the histograms built by the SIFT descriptor algorithm. The evolutionary search was performed based on a mono-objective function based on the F-Measure metric. However, even though the F-Measure provides an excellent answer for the case of a mono-objective version of the local descriptor problem; we decide to explore the multi-objective approach using recall and 1-precision as the optimization criteria. Such an approach allows us to explore several equivalent solutions which are optimal in the Pareto sense, and also provides a framework to add new optimization criteria in future research. Moreover, in this work another goal is to evolve SIFT descriptor operators that are optimized for affine-covariant regions, which are common in real-world applications such as image mosaicing. Therefore, we propose a MO-GP algorithm to synthetize SIFT descriptor operators based on the SPEA2 technique for MO optimization [16], see Fig. 1. Our proposed evolutionary algorithm has three main differences with respect to our previous canonical GP algorithm: 1) we would include an affine-invariant detector (hessian-affine) instead of the scale-invariant detector (DoG) proposed by Lowe [8], 2) during learning, we would use an image pair with affine transformations instead of scale+rotation, and 3) we pose a MO optimization problem instead of mono-objective optimization.

4.1 Search Space

The mathematical operators, that will be produced by the MO-GP evolution will be made up of internal and leaf nodes which are composed of the function (F) and terminal (T) sets, respectively. These functions and terminals represent the problem search space as primitive elements; in such a way, the MO-GP will explore the search space in order to build the mathematical operators by combining the primitives in multiple possible ways. The function and terminal sets proposed in this work are similar to that used in our previous work [11] due to they were well-suited for evolving SIFT descriptor operators. Thus, the functions and terminal sets are defined as follows:

$$\begin{aligned} F = &\left\{+, |+|, |-|, *, \div, \sqrt{I_t}, \frac{I_t}{2}, log_2(I_t), D_x G_\sigma,\right. \\ &\left. D_y G_\sigma, G_\sigma, Sobel_x(I_t), Sobel_y(I_t), LoG(I_t), Median(I_t)\right\}, \\ T = &\{I, D_x(I), D_{xx}(I), D_{yy}(I), D_{xy}(I), D_y(I)\}, \end{aligned} \quad (9)$$

where I is the input image and I_t can be any of the terminals in T, as well as the output of any of the functions in F; D_u symbolizes the image derivatives along direction u then $D_u = I * G_{u(\sigma=1}$; G_σ are Gaussian smoothing filters with $\sigma = 1$ or 2; $D_u G_\sigma$ representing the derivative of a Gaussian filter with image blur σ. $Sobel_u$ is a Sobel Filter along direction u; LoG is a function for detecting edges using a Laplacian of Gaussian Filter; $Median$ is a 2-D Median Filter. In this way, GP would be able to create composite operators using these function and terminal sets which were carefully selected because its importance for achieving good results.

4.2 Optimization Criteria

The individuals in our proposed MO-GP algorithm would be evaluated using the recall and 1-precision metrics described in Equations 1 and 2. Hence, the optimization criteria are defined as follows:

$$F_1 = argmin\left\{\frac{1}{(\sum_{i=1}^{n} recall(i)/n)}\right\}, \quad (10)$$

$$F_2 = argmin\left\{\sum_{i=1}^{n} 1 - precision(i)/n\right\}, \quad (11)$$

where n is the number of thresholds for which $recall$ and $1 - precision$ were computed for a particular image pair; in this work, we use $n = 20$ following [9].

5. CONCLUSIONS

In this work a multi-objective genetic programming algorithm was described as an innovative method to automatically synthesize descriptor operators that would be robust to affine transformations. We believe that our proposed approach will successfully learn said affine descriptor operators using recall and 1-precision as optimization criteria. At this moment, we are working on the implementation of our proposed algorithm using images that present affine transformations. In summary, this work proposes new evolved descriptor operators for affine transformations using MO-GP

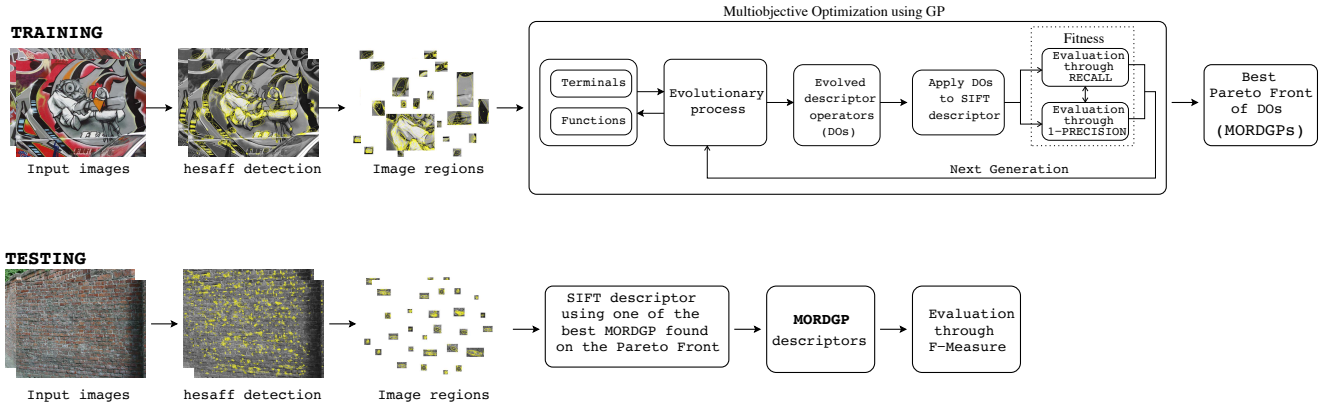


Figure 1: Overview of our MO-GP approach for learning affine-invariant descriptor operators.

where the principal advantage of using a multi-objective optimization framework is the incorporation of multiple objectives without significantly altering the core functionality of the search process. Furthermore, we pretend to test our proposed evolved descriptor on a real world application such image mosaicing. Finally, we believe that this approach could be extended and adapted to other real-world applications where specialized descriptor operators could be incorporated according to the requirements of the task.

6. REFERENCES

[1] M. Brown and D. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007.

[2] M. Brown and S. Winder. Discriminative learning of local image descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(13):43–57, 2011.

[3] C. Coello, D. Veldhuizen, and G. Lamont. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, 2007.

[4] M. Jiang, J. Hong, Q. Liao, and S. Huang. A sift-based method for image mosaic. *International Conference on Advanced Computer Theory and Engineering*, pages 423–427, 2010.

[5] J. Koza. *Genetic Programming*. The MIT Press, 1993.

[6] K. Krawiec. Generative learning of visual concepts using multiobjective genetic programming. *Pattern Recognition Letters*, 28(16):2385–2400, 2007.

[7] M. Lahanas, N. Milickovic, D. Baltas, and N. Zamboglou. Application of multiobjective evolutionary algorithms for dose optimization problems in brachytherapy. *International Conference on Evolutionary Multi-Criterion Optimization*, pages 574–587, 2001.

[8] D. Lowe. Object recognition from local scale-invariant features. *IEEE Conference on Computer Vision*, pages 1150–1157, 1999.

[9] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.

[10] G. Olague and L. Trujillo. Evolutionary-computer-assisted design of image operators that detect interest points using genetic programming. *Image and Vision Computing*, 29(7):484–498, 2011.

[11] C. B. Perez and G. Olague. Evolutionary learning of local descriptor operators for object recognition. *Genetic and Evolutionary Computation Conference*, pages 1051–1058, 2009.

[12] R. Szeliski. Image alignment and stitching: A tutorial. *Found. and Trends in Computer Graphics and Vision*, 2(1):1–104, 2006.

[13] T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: a survey. *Found. and Trends in Computer Graphics and Vision*, 3(3):177–280, 2008.

[14] Van-Rijsbergen. *Information retrieval*. Butterworth-Heinemann, 1979.

[15] J. Zaragoza, T.-J. Chin, M. Brown, and D. Suter. As-projective-as-possible image stitching with moving dlt. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1285–1298, 2014.

[16] E. Zitzler, M. Laumanns, and L. Thiele. Spea2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. *Evolutionary Methods for Design, Optimization and Control*, pages 19–26, 2002.

Captura inteligente de imágenes de actas electorales mediante dispositivos móviles

Alexis Plata M.*
Departamento de Eléctrica y Computación
Instituto de Ingeniería y tecnología, UACJ
Cd. Juárez, Chih, México
al150666@alumnos.uacj.mx

Víctor Morales R.†
Departamento de Eléctrica y Computación
Instituto de Ingeniería y tecnología, UACJ
Cd. Juárez, Chih, México
victor.morales@uacj.mx

RESUMEN

En el siguiente artículo se muestra cómo pueden ser corregidas las distorciones más visibles que se presentan durante la captura de imágenes de actas electorales, obtenidas por medio de teléfonos móviles. Dichas imágenes fueron procesadas aplicando diversas técnicas y filtros, tomando como principal la técnica de detección de líneas de Hough.

Palabras clave

procesamiento de imágenes; actas electorales; captura inteligente; teléfonos móviles

1. INTRODUCCIÓN

El trabajo que se presenta en este artículo forma parte de un proyecto de investigación que busca la transparencia en los procesos electorales a partir de auditorías, con el fin de otorgar credibilidad a las instituciones que manejan dichos procesos. El proyecto global Election Results Reliable Sampling (ERRS) es un sistema de verificación de resultados electorales que parte de la tecnología móvil, así como de técnicas de procesamiento de imágenes. El principal requisito para su implementación, consiste en la captura de imágenes de actas electorales por medio de teléfonos móviles. Los procedimientos siguientes parten de la incorporación de los métodos de procesamiento para eliminar las imperfecciones de las imágenes y clarificar los datos. En consiguiente, se comprueba la validez del acta a través del reconocimiento de patrones, se extrae la información contenida en las actas con técnicas de reconocimiento de caractéres, y finalmente, se aplican tecnologías criptográficas para mantener la integridad y veracidad de los datos, desde el origen hasta el final.

2. RECONOCIMIENTO DE PATRONES

Desde décadas pasadas, se han investigado y desarrollado algoritmos que permiten a una computadora imitar la habilidad del ser humano para describir y clasificar objetos a partir de imágenes. Esto lo han logrado gracias a las áreas de procesamiento de imágenes y reconocimiento de patrones.

El procesamiento de imágenes consta de métodos que mejoran el aspecto de una imagen digital y que permite resaltar áreas deseadas. Una imagen digital se representa en un plano espacial, apreciada como matriz, en donde cada valor en la matriz simboliza un nivel de color y cada elemento es llamado comúnmente pixel. Esta disciplina está relacionada con las matemáticas, la computación y el conocimiento de los órganos que ayudan a la visión del ser humano [1].

Por otro lado, el reconocimiento de patrones está orientado a los métodos de descripción y clasificación de objetos. Aunque existen 3 etapas básicas del reconocimiento de patrones en imágenes: preprocesamiento, extracción de características, y clasificación; Egmont-Petersen, Ridder y Handels [2] muestran los diferentes niveles en el que puede ser dividido, los cuales son:

Preprocesamiento/filtrado: Resultado de la imagen original modificada con las mismas dimensiones. Reducción de datos/Extracción de características: Extracción de componentes sustanciales de una imagen. Segmentación: Partición de una imagen en regiones. Detección o reconocimiento de objetos/Clasificación: Determina la posición, la orientación y escala de un objeto específico en una imagen y clasifica estos objetos.

La etapa de preprocesamiento es usada para eliminar el ruido que se genera al capturar una imagen. De la misma forma se puede usar para mejorarla con respecto al contraste, la calidad, distorción, entre otras características [3]. Asimismo, es un método que permite reducir las zonas oscuras de una imagen, preservando la información necesaria para una futura extracción de características [4].

La reducción de datos y la extracción de características son una parte importante del procesamiento de imágenes, ya que la compresión ayuda a almacenar y transmitir imágenes digitales. La extracción de características es utilizada para el reconocimiento de objetos, es decir qué tipo de características pertenecen a ese objeto. Las redes neuronales artificiales son empleadas para la compresión de una imagen y para la extracción de características [2].

La segmentación también es una parte significativa del procesamiento de imágenes. Rojas, Guerrero y Serrano [5] señalan que existen diferentes técnicas para la segmentación de imágenes y que estas se dividen en 4 grupos: basadas en valores del pixel, basadas en el área, basadas en borde, y basadas en la física. La técnica basada en el valor del pixel utiliza el histograma, es decir, se adquiere el histograma de la imagen, se analizan los valores máximos y los valores adyacentes. La técnica basada en el área se divide en 2 categorías, crecimiento de regiones y división, y combinación de regiones. La técnica basada en orillas se divide de igual forma en 2 categorías, técnicas locales y técnicas globales.

*Estudiante del programa de Maestría en Cómputo Aplicado.
†Prof. Investigador en áreas de voto electrónico y sistemas electorales

La detección de objetos básicamente consiste en localizar la posición, posible orientación y escala de un objeto en una imagen. Es la etapa encargada de asignar etiquetas a los objetos detectados. Las redes neuronales artificiales pueden ser entrenadas para reconocer diferentes objetos. Este reconocimiento puede estar basado en los datos de cada pixel o basado en ciertas características. Cuando se presenta por pixel, es realizado por un clasificador entrenado para reconocer el objeto completo de un patrón espacial de intensidades, aumentando la complejidad de manera exponencial de acuerdo al tamaño del objeto con el número de dimensión, 2D o 3D [2].

A continuación, se mencionará exclusivamente el trabajo realizado en la etapa de preprocesamiento de este proyecto, en el cual, se logró normalizar imágenes de actas de escrutinio, corrigiendo las distorsiones de perspectiva, variaciones de iluminación e intensidad.

3. PREPROCESAMIENTO

El preprocesamiento es el primer paso que se debe realizar en el reconocimiento de patrones, es decir, después de capturar una imagen con algún dispositivo electrónico. Normalmente, se utilizan dos maneras de capturar imágenes de documentos, una mediante cámaras (digitales, web o teléfonos móviles), y otra, a partir de escáneres.

La captura de documentos a través de teléfonos móviles se ha convertido en algo muy habitual, sin embargo, presenta algunas desventajas. Una de ellas, es al momento de adquirir la imagen, ésta puede sufrir de distorsiones de perspectiva, perdida de enfoque y/o variaciones de iluminación. A diferencia de las imágenes capturadas en teléfonos móviles, las escaneadas son adquiridas en un entorno controlado, en donde siempre existe el mismo tipo de luz, la misma posición, escala, etcétera, dando como resultado una imagen de buena calidad. Por otro lado, los algoritmos de reconocimiento de patrones utilizados por la mayoría de las aplicaciones asumen una imagen normalizada, en otras palabras, necesitan requisitos esecíficos para funcionar correctamente. Por ello, fue necesario desarrollar un método con el objetivo de eliminar las variaciones producidas por los problemas mencionados anteriormente.

El método consiste en detectar los bordes en la imagen, para después realizar la detección de posibles líneas, y así coincidir con los bordes del acta. Posteriormente, a partir de estas líneas se obtienen las intersecciones para definir las esquinas del acta, generando cuatro puntos. Una vez obtenidos estos puntos, se aplica una corrección de perspectiva con las medidas originales del acta. Obteniendo finalmente una imagen corregida. Además, se eliminó las variaciones de iluminación que se presentaban en la imagen del documento. En la figura 1, se observa el conjunto de imágenes que hace referencia a cada uno de los pasos que se llevaron a cabo.

3.1 Corrección de perspectiva

Fue necesario realizar un preprocesamiento de la imagen antes de corregir la perspectiva. La imagen original utilizada para esta explicación es la figura 1.A. Uno de los primeros pasos para lograrla fue aplicar un filtro de mediana. El filtro de mediana se utilizó para poder eliminar los elementos pequeños no necesarios para la detección de bordes en la imagen, por ejemplo caracteres, símbolos, firmas y escrituras a mano, y así conservar sólo las líneas más grandes en la imagen, ver figura 1.B. Éste es un filtro de orden estadístico

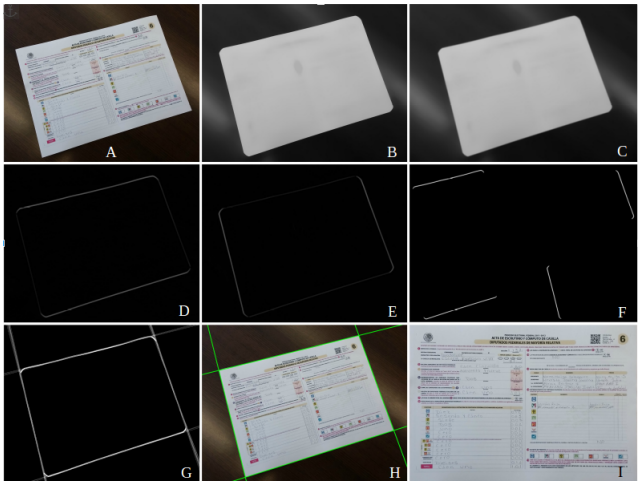


Figura 1: Imágenes del procedimiento realizado.

y es utilizado por su simplicidad y capacidad de preservar bordes. Dicho filtro trabaja en el dominio espacial, y se basa en utilizar un tamaño $W_m x W_n$, en donde W_m y W_n son números primos [6]. Dada un imagen f, la imagen filtrada g es definida por:

$$g(x,y) = median_{(s,t)\in S_{xy}}\{f(s,t)\} \tag{1}$$

donde (x, y) son las coordenadas del pixel localizado en el centro de una región S_{xy} definido por $W_m x W_n$, y (s, t) son las coordenadas de los pixeles pertenecientes a la región. Como resultado del filtro se obtendrá el pixel mediano de los datos contenidos en la región S_{xy}. Finalmente, se conseguirá la eliminación de los detalles finos, como líneas, esquinas, o distorciones en la imagen [6].

Después, se realizó un filtro de suavizado, basado en la media para disminuir la precisión y aumentar el rango de detección de bordes, y así facilitar la detección de líneas, ver figura 1.C. Este filtro de suavizado, es un filtro normalizado donde cada salida de un pixel es la media del kernel de sus vecinos, después de aplicado el proceso [7]. Donde dicho kernel esta establecido por:

$$K = \frac{1}{K_{ancho} \times K_{alto}} \cdot \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \tag{2}$$

Posteriormente, se siguió con la detección de bordes aplicando un filtro Sobel, se obtuvieron las diferencias de las intensidades de la imagen, lo que es equivalente a detectar los bordes. Ese proceso se hace por separado en dos orientaciones, vertical y horizontal, lo que permitió diferenciar la orientación de cada uno de los bordes, ver figura 1.D para detección horizontal y ver 1.E para detección vertical.

Cuando se habla de detección de bordes, se asume que estos pueden ser identificados cuando se busca el gradiente de una imagen, dicho de otra manera, es una discontinuidad direccional en función de la intensidad. Con lo anterior, si utilizamos la derivada del valor de la intensidad a través de la imagen, obtenemos los puntos donde la derivada tiene un

valor máximo y un valor mínimo, y así podemos localizar los bordes [8]. El gradiente es un vector, cuyos componentes son dos derivadas parciales, una a través de la dirección X y otra en dirección Y. Con esto se establece que:

$$\nabla f = \frac{\partial f}{\partial x_i}\vec{i} + \frac{\partial f}{\partial y_i}\vec{j} \tag{3}$$

La dirección del gradiente esta dado por:

$$\theta = \arctan \frac{\nabla f_y}{\nabla f_x} \tag{4}$$

Finalmente, la magnitud se establece como:

$$\|\nabla f\| = \sqrt{(\nabla f_x)^2 + (\nabla f_y)^2} \tag{5}$$

Las derivadas parciales necesitan ser calculadas para cada pixel en la imagen. Comúnmente, se utiliza un filtro de convolución de un área pequeña para hacer una aproximación. Deben de existir dos tipos de filtros de convolución, G_x para la dirección x, y G_y para la dirección y [9]. Los filtros utilizados son de tamaño 3×3, donde la ecuación 6 representa la detección en dirección horizontal y la ecuación 7 en vertical.

$$G(x) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot I \tag{6}$$

$$G(y) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \cdot I \tag{7}$$

Se decidió dividir la imagen en 4 secciones: superior, inferior, derecha e izquierda, con el fin de detectar específicamente las líneas con orientación horizontal o vertical, dependiendo de la sección. Después, se realizó una umbralización Otsu, que es adaptativa a la intensidad, para conservar los bordes de mayor intesidad en la imagen, y por consiguiente obtener una imagen binarizada, ver figura 1.F.

La umbralización es una técnica aplicada para la segmentación de imágenes. Tiene como objetivo clasificar una imagen en 2 tipos, las del objeto y las del fondo. Donde los pixeles del fondo comúnmente son asignados con un valor de 0, mientras que al objeto se le designa un valor de 255. La binarización Otsu es utilizada para umbralizar imágenes bimodales, es decir, cuyo histograma tenga dos picos, de tal manera que se calcula automáticamente el umbral del histograma de la imágen [10].

A continuación, se utilizó la Transformada de Hough para detectar posibles líneas en la imagen, es decir los bordes del documento. En la parte superior e inferior se buscan líneas horizontales, con un ángulo de 60 a 120 grados, mientras que en la parte derecha e izquierda se buscan líneas verticales, con un ángulo de 150 a 210 grados, ambos con una apertura de 60 grados, ver figura 1.G. A partir de cada uno de los grupos de líneas, pertenecientes a las diferentes secciones, se obtiene la media de ellas, dando como resultado sólo 4 líneas (superior, inferior, derecha e izquierda), ver figura 1.H.

La transformada de hough es una técnica utilizada para detectar líneas en imágenes. Se basa, en asignar cada pixel a un conjunto de líneas que pasen posiblemente a través de ese pixel. En donde, una línea puede ser representada en cartesiano como $y = mx + c$, o en su forma polar como $r = x\cos(\theta) + y\sin(\theta)$, siendo r la distancia perpendicular desde el origen hacia la línea, y θ el ángulo formado por la línea perpendicular [11][12].

Para cada pixel (x_0, y_0) se define un grupo de líneas pertenecientes a ese pixel:

$$r_\theta = x_0 \cdot \cos\theta + y_0 \cdot \sin\theta \tag{8}$$

Primeramente, por cada pixel en la imagen se evalua r para todos los ángulos pertenecientes a $\theta = \{0, 1, 2, \cdots, 180\}$. Posteriormente, se dibuja r en cada línea que pasa por el origen de coordenadas. Después, se le asigna un voto con valor de 1 a cada coincidencia del par (r, θ). Finalmente, el par con mayor número de votos, será la línea de intersección de los pixeles [13].

En el último paso de la correción de perspectiva, se calcularon las 4 intersecciones de las líneas que se obtuvieron con la transformada de Hough. Con las intersecciones resultantes y los puntos originales de la imágen del acta se pudo aplicar la correción de perspectiva, obteniendo así la imagen deseada, ver imagen final 1.I.

3.2 Mejora de color

Como ya se mencionó, las imágenes adquiridas por teléfonos móviles son suceptibles a variaciones de luz. Con la imagen corregida, se implementa un paso adicional para eliminar el efecto iluminante. Primeramente, se obtuvo la diferencia, dividiendo la imagen original entre la misma con un filtro de suavizado (mencionado anteriormente). Después, esta diferencia se eleva a una potencia n para incrementar el contraste, en donde los valores menores de 1 se harán más oscuros, y los mayores o iguales a 1 se quedarán blancos.

4. RESULTADOS

Se desarrolló una aplicación con el lenguaje de programación Java. También, se utilizó la librería Opencv para la implementación de los métodos mencionados anteriormente para la correción de perspectiva. Además, se utilizó un celular Samsung S5 con sistema operativo Android para la captura de imágenes de actas de escrutinio. Se capturaron 11 imágenes con diferentes orientaciones y fondos. Se puede observar la comparación en la tabla 1 en donde el tiempo aproximado de ejecución en una computadora con 4 núcleos fue de 3 segundos por imagen.

De los casos más significativos se mostró, cómo a pesar de tener imágenes con fondo de color similar al documento, e imágenes con bordes recortados, se obtuvo una imagen final rectificada. A continuación, se muestran algunos ejemplos de imágenes de los documentos, antes y después de aplicar la corrección de perspectiva y la corrección de intensidad de color, la primera de ellas es una imagen con fondo no uniforme e incompleta, mientras que la segunda se caracteriza por el fondo blanco, ver figura 2.

5. CONCLUSIONES

A partir del prototipo desarrollado en Java, se establece la aplicación automática de los filtros y transformadas necesarios para la rectificación de la imagen. De tal manera que se reconoce al método de Hough, en coalición con las demás técnicas (filtro de mediana, filtro de suavizado, y detección de bordes Sobel), como un procedimiento robusto para la detección de bordes en imágenes de documentos,

Tabla 1: Resultados obtenidos de la corrección de perspectiva

Tipo de imagen	Características	Rotación en ejes	Resolución	Cámara	Corrección
Tipo A	Fondo oscuro	z	3264x1836	SM-G900M	Exitosa
Tipo B	Fondo oscuro	x, y y z	3264x1836	SM-G900M	Exitosa
Tipo C	Fondo oscuro	x	3264x1836	SM-G900M	Exitosa
Tipo D	Fondo oscuro	x y z	3264x1836	SM-G900M	Exitosa
Tipo E	Fondo no uniforme	Sin rotación	3264x1836	SM-G900M	Exitosa
Tipo F	Fondo no uniforme	x	3264x1836	SM-G900M	Exitosa
Tipo G	Fondo no uniforme e incompleta	x, y y z	3264x1836	SM-G900M	Exitosa
Tipo H	Fondo no uniforme	x	3264x1836	SM-G900M	Exitosa
Tipo I	Fondo no uniforme e incompleta	x y y	3264x1836	SM-G900M	Exitosa
Tipo J	Fondo blanco	Sin rotación	3264x1836	SM-G900M	Exitosa
Tipo K	Fondo blanco	x, y y z	3264x1836	SM-G900M	Fallida



Figura 2: Imágenes con fondo no uniforme y fondo blanco, antes y después de la corrección.

consiguiendo satisfactoriamente la correción de perspectiva. Por otro lado, uno de los retos presentes, como se mostró en los resultados al adquirir una imagen con un fondo similar al del acta de escrutinio, en este caso blanco, y con una orientación en diferentes ejes, consiste en lograr la corrección de perspectiva. Dentro de las soluciones posibles se encuentra el rotar la imagen para disminuir la perspectiva del acta, o colocar el acta en un fondo distinto para establecer contraste de fondos. En un futuro, se planea establecer un método para evitar la pérdida de enfoque, de tal manera que se consiga una imagen lo suficientemente legible. También, se pretende migrar el prototipo desarrollado en Java al sistema operativo Android, debido a que corresponde a uno de los objetivos del proyecto, y se busca que el preprocesamiento se realice directamente en este sistema.

6. REFERENCIAS

[1] B.˜E. Ramírez, "Departamento de procesamiento de señales," Agosto 2006.

[2] M.˜Egmont-Petersen, D.˜d. Ridder, and H.˜Handels, "Image processing with neural networks, a review," *Pattern Recognition*, vol.˜35, no.˜10, pp.˜2279–2301, 2002.

[3] A.˜Ahmad, A.˜B. Mansoor, and R.˜Mumtaz, "Image processing and classification in diabetic retinopathy: A review," *2014 5th European Workshop on Visual Information Processing (EUVIP)*, pp.˜1–6, 2014.

[4] G.˜P. Teja and S.˜Ravi, "Face recognition using subspaces techniques," *2012 International Conference on Recent Trends In Information Technology (ICRTIT)*, pp.˜103–107, 2012.

[5] J.˜B. Rojas, M.˜Guerrero, and G.˜U. Serrano, "Segmentación de imágenes de color," *REVISTA MEXICANA DE FÍSICA*, vol.˜60, no.˜6, pp.˜579–587, 2004.

[6] H.˜Ibrahim, N.˜S. Pik˜Kong, and T.˜Foo˜Ng, "Simple adaptive median filter for the removal of impulse noise from highly corrupted images," *IEEE Transactions on Consumer Electronics*, vol.˜54, no.˜4, pp.˜1920–1927, 2008.

[7] OpenCV, "Smoothing Images." http://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html. [En línea; accesado 15-Septiembre-2016].

[8] O.˜R. Vincent and O.˜Folorunso, "A descriptive algorithm for sobel image edge detection," *Proceedings of Informing Science, IT Education Conference, InSITE*, 2009.

[9] W.˜Gao, L.˜Yang, X.˜Zhang, and H.˜Liu, "An improved sobel edge detection," *IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, vol.˜5, pp.˜67–71, 2010.

[10] C.˜Hima˜Bindu and K.˜Satya˜Prasad, "An efficient medical image segmentation using conventional otsu method," *International Journal of Advanced Science and Technology*, vol.˜38, pp.˜67–74, 2012.

[11] L.˜A. Fernandes and M.˜M. Oliveira, "Real-time line detection through an improved hough transform voting scheme," *IPattern Recognition*, vol.˜41, pp.˜299–314, 2008.

[12] OpenCV, "Hough Line Transform." hhttp://docs.opencv.org/2.4/doc/tutorials/imgproc/imgtrans/hough_lines/hough_lines.html. [En línea; accesado 15-Septiembre-2016].

[13] OpenCV, "Hough Line Transform." http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_houghlines/py_houghlines.html. [En línea; accesado 15-Septiembre-2016].

SVM for Inferring Drivers Distracted Behavior: A Feasibility Experiment

Marcela D. Rodríguez
Facultad de Ingeniería, UABC
Mexicali, B.C., México
marcerod@uabc.edu.mx

Antonio Ordorica
Facultad de Ingeniería, UABC
Mexicali, B.C., México
aordorica@uabc.edu.mx

José M. Ramírez
Facultad de Ingeniería, UABC
Mexicali, B.C., México
jramirez53@uabc.edu.mx

Luis A. Castro
Sonora Institute of Technology
Ciudad Obregón, México
luis.castro@acm.org

Jessica Beltran-Márquez
Sonora Institute of Technology
Ciudad Obregón, México
jessica.beltran@itson.edu.mx

Ángel G. Andrade
Facultad de Ingeniería, UABC
Mexicali, B.C., México
aandrade@uabc.edu.mx

ABSTRACT
Most methods for inferring driving distraction are based on visual features analysis of head pose, since it is a strong indication of driver distraction. In this work, we propose to use the inertial sensors embedded in smart glasses. To this aim, we collected data from five participants and conducted experiments to assess the feasibility of using support vector machines (SVM) to generate drivers' models for inferring their focus of attention. The results show that using the personalized training model renders an acceptable accuracy to identify particular car cabin's spots where drivers focus their attention (i.e. accuracy was greater than 50% and less than 81.44% for all subjects).

Keywords
Driving distraction; support vector machines; smart glasses;

1. INTRODUCTION
Driver distraction or inattention refers to the lack of attention to the driving tasks due to engagement in other tasks while driving. These other tasks can be any diversion of attention like cognitive, physical, or visual distractions that lead to performance degradation [1]. It is estimated that distractions cause 23% of crashes or near-crashes [2], and that could be reduced by 10-20% if driving behaviors are monitored or predicted [3,4]. Avoiding distracted driving has been of particular interest while using navigation artifacts [5]. Also, methods for detecting driver distraction and drowsiness have been explored, which are mainly based on the analysis of visual features of facial expressions, changes on head pose [4], and gaze behavior [6]. In particular, head pose (or head orientation) is a strong indicator of a driver's field of view and current focus of attention [7,8]. Therefore, it has been considered as an integral process for monitoring driver awareness [8]. For deploying these methods so far, the vehicle's cabin needs to be instrumented with cameras [7,8,9]. For instance, faceLAB uses two cameras to determine the head pose and eye gaze direction [9]. On the other hand, smart glasses have emerged as an interesting research and product platform for a wide range of wearable assistive systems due their sensing capabilities. The security regarding using smart glasses while driving has been questioned as they allow users to send text messages or share on social media via spoken commands. However, research has shown that they could improve safety for people who are at risk of operator fatigue (e.g. truck drivers), and that drivers using smartphones swerve longer and more frequently than those using Google Glass [10]. We propose the use of the inertial sensors embedded in smart glasses for obtaining head pose features to infer drivers' current focus of attention. To reach this aim, we used Support Vector Machines (SVM), a supervised learning technique that searches for an optimal hyperplane to separate data between defined classes [11,12]. In the following section, we explain the methods used for experimenting with the SVM so far. Section 3 presents our preliminary results, and Section 4 concludes and presents the possible future directions of this work.

2. METHODS
2.1 Data collection
We collected data from 5 subjects (mean age M=32; SD=11.20 years) through controlled experiments in near naturalistic conditions. Each subject participated in a session in which they conducted a set of tasks (see Table 1) in a car while wearing Google glasses. First, participants performed the tasks while the car was parked (P), and then while the car was being driven (D). The former experiment lasted a mean time M=5' 34'', SD=1'29'', and the later had a mean time of M=8' 34'', SD=2'55''. The collected data consisted of 9-tuple vectors i.e., accelerometer, gyroscope and orientation sensors (each along x, y and z axes respectively). The accelerometer's sampling frequency was set to the fastest through the Android API, around 15.5 Hz. All sessions were video recorded to obtain the ground truth by observing and labeling data. The aim of supervised learning techniques is to infer a function from labeled data [11,12]. We labeled the data using two classes related to the car's cabin locations where participants centered their attention by changing their head pose. As presented in Figure 1, these types of classes were: a) car's cabin spots and b) cabin's zones. Both of them are associated with tasks that drivers may conduct when their attention is focused on these spots and zones. For instance, a driver may focus his attention on the Comfort and Information zone (C&I_Zone which consisted of spot 5 and 6 according to Fig. 1) when changing the radio station (spot 5). Thus, we assessed the feasibility of identifying particular spots where drivers' focused their attention, which may help identify if he is conducting a secondary task.

2.2 Experiments
For generating the SVM model, we used WEKA [14] since it offers ways for automating the process of finding the optimal

parameters for the SVM classifier. Additionally, WEKA offers functions to normalize the numeric values in the given dataset.

We used the polykernel for the experiments conducted so far. Two models were trained, one using data from all participants (i.e. generalized model), and one using data of each participant (i.e. personalized model). In this paper, we report the personalized model, in which we used the 66-34% training modality (66% of data of each participant was used for training and 34% for validation). For the classifier training, we used the 9-tuple vectors containing the normalized values collected from the three sensors. For each experiment, we report the instances correctly classified or accuracy, precision and recall as classification performance measures [11,12].

Table 1. Experimental tasks which are reported to cause risks associated to distractions [13].

Secondary Tasks	Tasks conducted during experiment
Using cell phone	Dialing (quick keys), talking/listening, texting
Interacting with object	Looking at object, Moving object in vehicle, Reaching for object
Drinking	Lid no straw, Open container
Using comfort and information console	Turn climate on/off, adjust temp, CD-Insert/remove, Radio Adjustment

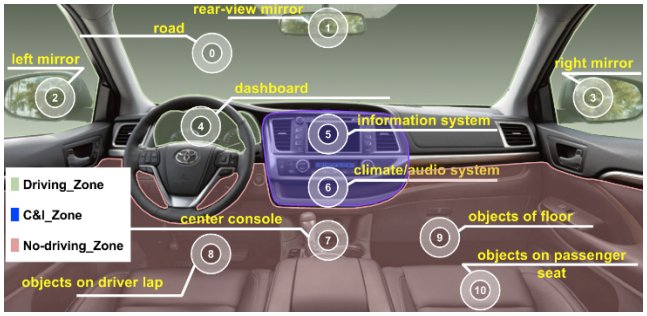


Figure 1. Cabin's spots and zones used as classifying classes.

3. RESULTS

We collected 21,621 9-tuple vectors from the experiment, for both, driving (D) and parked (P) setting conditions. We labeled 6,403 tuples (i.e., classes instances) for the driving condition as presented in Table 2. It shows that approximately 90% of the data belongs to spots located in the (Driving_Zone). We observed that spot0 (S0) was the predominant class, since it is the spot where the drivers' focus of attention remains most of the time. Therefore, we conducted experiments for observing changes on the classification accuracy by balancing the dataset through two different strategies: i) eliminating all spot0 instances, and ii) reducing spot0 class instances by eliminating those captured while performing the requested tasks during the driving session (see Table 1). Afterwards, we used the over-sampling method i.e., adding copies of instances from the under-represented classes. For instance, Table 3 shows that when reducing the number of spot0 instances for subject 2, the new over-represented class is spot1 (S1=281), consequently we even-up the rest of the classes to 281. While Table 4 shows that when removing all spot0 instances, the spot2 became the over-represented class (S2=256) for subject 3.

Reducing the quantity of spot0 instances resulted in an accuracy of 73.17% for subject 3, and 53.28% for subject 2. On the other side, accuracies increased when eliminating all spot0 instances, resulting in 79.18% for subject 3 and 57.48% for subject 2. The confusion matrices presented in Table 5 and 6, show that several instances tended to be misclassified when they correspond to spots located closely; for instance, S3 and S9 (see Table 5), which are on the right side of the car cabin; and S5 and S6 located in the C&I Zone (see Table 6).

4. CONCLUSIONS

Our results show that inferring driving distraction based on data collected from the inertial sensors of smart glasses is a viable solution. Also, using SVM for inferring the focus of attention renders acceptable accuracies measures. The results of our previous experiments reported in [15], show that using raw data (9-tuple vectors) is an appropriate choice for enabling smart glasses to execute a process for inferring inattention in real time. However, there exists clearly trade-offs between accuracy, and the effort taken for collecting and training data in real time conditions. We realized that the drivers' habits contributed to obtaining unbalanced datasets. For instance, Table 2 shows that subject 2, an older driver, stared more times at the rear-view mirror (281 S1 instances) and right mirror (S3 instances) than the rest of the participants. Subject 1 reported that he is not used to interact with the mobile phone while driving, and when conducting the tasks related to mobile phone usage, a great quantity of instances of S8 were obtained (i.e., he oriented his head more times to the No_Driving Zone than the rest of the subjects). Additionally, subjects interact with the comfort and information console (C&I Zone) as little as possible while driving. Further work is needed to efficiently infer distractions. We plan to conduct additional driving sessions with more users in order to collect a more balanced dataset. Additionally, we are considering reducing the granularity of spots classes, such as grouping spots 5 and 6 since they are close enough to be confused by the classifier.

Table 2. Data collected (DC) and data labeled (DL) obtained from each subject during the driving session.

Subject	Sample		Spots of Driving_Zone						Spots of C&I Zone			Spots of No_Driving_Zone				
	DC	DL	S0	S1	S2	S3	S4	Total DL	S5	S6	Total DL	S7	S8	S9	S10	Total DL
Subject1	3,443	1,877	1,064	122	434	17	16	**1,653**	89	20	**109**	8	97	7	3	**115**
Subject2	3,138	1,203	466	281	238	87	7	**1,079**	51	49	**100**	8	6	5	5	**24**
Subject3	2,768	1,351	933	2	256	26	56	**1,273**	30	24	**54**	13	0	4	7	**24**
Subject4	2,000	898	463	14	308	23	33	**841**	15	17	**32**	7	8	3	7	**25**
Subject5	2,042	1,074	489	37	384	38	0	**948**	36	21	**57**	18	23	9	19	**69**
Total	13,391	6,403	3,415	456	1,620	191	112	**5,794**	221	131	**352**	54	134	28	41	**257**

Table 3. Dataset after reducing the number of instances of the spot 0 (S0) class for subject 2 and subject 3

Subjects	Spots of Driving Zone					Total	Spots of C&I Zone		Total CIZone	Spots of No Driving Zone				Total
	S0	S1	S2	S3	S4		S5	S6		S7	S8	S9	S10	
subject2	225	281	238	87	7	**838**	51	49	**100**	8	6	5	5	**24**
subject3	306	2	256	26	56	**646**	30	24	**54**	13	0	4	7	**24**

Table 4. Dataset after eliminating all spot0 (S0) class instances for subject 2 and subject 3.

Subjects	Spots of Driving Zone				Total	Spots of C&I Zone		Total CIZone	Spots of No Driving Zone				Total
	S1	S2	S3	S4		S5	S6		S7	S8	S9	S10	
subject2	281	238	87	7	**613**	51	49	**100**	8	6	5	5	**24**
subject3	2	256	26	56	**340**	30	24	**54**	13	0	4	7	**24**

Table 5. Confusion matrix corresponding to the classification results for subject 3 when removing all spots 0.

Correctly Classified Instances 79.1826 %					Incorrectly Classified Instances 20.8174 %		Precision 0.815	Recall 0.792	
S4	S2	S3	S5	S6	S9	S7	S10	S1	<-- classified as
64	0	0	4	7	1	3	3	0	S4
8	**50**	7	5	9	2	0	2	0	S2
0	9	**61**	7	6	0	0	0	0	S3
3	0	5	**64**	5	0	0	0	0	S5
5	2	7	12	**54**	0	0	0	0	S6
0	0	30	0	0	**63**	0	0	0	S9
6	0	0	0	0	0	**85**	0	0	S7
15	0	0	0	0	0	0	**83**	0	S10
0	0	0	0	0	0	0	0	**96**	S1

Table 6. Confusion matrix corresponding to the classification results for subject 2 when removing all spots 0.

Correctly Classified Instances 57.4869 %					Incorrectly Classified Instances 42.5131 %				Precision 0.610	Recall 0.575
S1	S3	S2	S5	S7	S4	S6	S9	S8	S10	<--classified as
72	3	3	12	0	1	6	0	0	0	S1
67	**14**	2	21	0	3	2	2	0	0	S3
16	1	**40**	21	1	9	5	0	0	3	S2
0	3	5	**76**	1	11	3	0	2	1	S5
0	0	0	0	**67**	0	0	0	0	20	S7
0	0	0	0	0	**76**	0	0	13	0	S4
0	2	3	23	0	15	**42**	15	0	2	S6
0	0	0	23	22	0	0	**47**	0	0	S9
0	0	0	0	0	31	0	15	**23**	18	S8
0	0	0	0	0	0	0	0	0	**92**	S10

5. ACKNOWLEDGMENTS

We thank all the volunteers of the study, and authors 4 & 5 for their comments to improve this work

6. REFERENCES

[1] Sajan, S., Ray, G.G. Human Factors in Safe Driving - A Review of Literature on Systems Perspective, Distractions and Errors. In Proceedings of 2012 IEEE Global Humanitarian Technology Conference, 83-88

[2] Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. and Ramsey D.J. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. National Highway Traffic Safety Administration, USDOT, 2006.

[3] Bayly, M., Fildes, B., Regan, M., and Young K.. Review of crash effectiveness of intelligent transport system. Traffic Accident Causation in Europe (TRACE), 2007.

[4] Kang, H.B. Various Approaches for Driver and Driving Behavior Monitoring: A Review. In Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops (ICCVW '13), IEEE Computer Society, Washington, DC, 616-623.

[5] Tarqui, G., Castro, L. A., Favela, J. Reducing Drivers' Distractions in Phone-Based Navigation Assistants Using Landmarks. Ubiquitous Computing and Ambient Intelligence. UCAMI 2013, LNCS 8276, Springer, 342-349.

[6] Chamberlin, J. Smart glasses: Driver distraction or safety tool? March 2014, Vol 45, No. 3. Print version: page 12. Note available at: http://www.apa.org/monitor/2014/03/smart-glasses.aspx

[7] Zhang, L., Liu, F., and Tang J. 2015. Real-Time System for Driver Fatigue Detection by RGB-D Camera. ACM Trans. Intell. Syst. Technol. 6, 2, Article 22 (March 2015).

[8] Murphy-Chutorian, E., Manubhai, M. Head Pose Estimation and Augmented Reality Tracking: An Integrated System and Evaluation for Monitoring Driver Awareness. IEEE Trans. on Intelligent Transportation Systems. 2010, Vol. 11, No. 2, 300–311.

[9] Seeing machines Driver State Sensor. https://www.seeingmachines.com/solutions/

[10] Zhang, Y.F., Gao, X.Y., Zhu, J.Y., Zheng, W.L., Lu, B.L. A novel approach to driving fatigue detection using forehead EOG, Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference, 707 – 710

[11] Mohri, M., Rostamizadeh, A., Talwalkar, A. Foundations of Machine Learning. The MIT Press 2012.

[12] Mitchell, T.M. Machine Learning, 1st. Edition, Mc Graw Hill Higher Education, 1997.

[13] NHTSA. (2008, July). National Motor Vehicle Crash Causation Study Report to Congress. DOT HS 811 059. Washington, DC: National Highway Traffic Safety Administration.

[14] WEKA available at: https://sourceforge.net/projects/weka/

[15] Antonio Ordorica, Marcela D. Rodríguez, Luis A. Castro, Jessica Beltran. Support Vector Machines for Inferring Distracted Behavior of Drivers Wearing Smart Glasses. To be in the Springer Proceedings of Ubiquitous Computing and Ambient Intelligence, UCAMI, 2016.

Uso del Lenguaje Natural para la Búsqueda de Respuestas en un Sistema de Gestión de Conocimiento Académico

César E. Rose-Gómez
Tecnológico Nacional de México
Instituto Tecnológico de Hermosillo
Avenida Tecnológico S/N
Hermosillo, Sonora, México
(52) 662 2606500
crose@ith.mx

María T. Serna-Encinas
Tecnológico Nacional de México
Instituto Tecnológico de Hermosillo
Avenida Tecnológico S/N
Hermosillo, Sonora, México
(52) 662 2606500
tserna@ith.mx

Oscar M. Rodríguez-Elías
Tecnológico Nacional de México
Instituto Tecnológico de Hermosillo
Avenida Tecnológico S/N
Hermosillo, Sonora, México
(52) 662 2606500
omrodriguez@ith.mx

RESUMEN

La cantidad de información y conocimiento que se genera en un centro de investigación y docencia es considerable. Tradicionalmente los datos asociados al desarrollo de las actividades de los investigadores se mantienen en una base de datos en el mejor de los casos. Lo anterior permite tener una gestión de los datos para apoyar algunos procesos en la toma de decisiones de una manera limitada. En este artículo se presenta el uso de lenguaje natural para la búsqueda de respuestas como uno de los elementos de la recuperación de información en un sistema de gestión de conocimiento de una organización académica.

CCS Concepts

• Information systems ~ Information retrieval • Information systems ~ World Wide Web • Information systems ~ Data management systems.

Keywords

Búsqueda respuestas; gestión conocimiento; procesamiento lenguaje natural; ontología.

1. INTRODUCCIÓN

Actualmente en muchas organizaciones la cantidad de datos ha crecido de una manera considerable, con lo cual la generación de información y conocimiento es compleja y requiere del uso de diversas técnicas de las áreas de recuperación de información y de la inteligencia artificial. Entre estas organizaciones se encuentran también las universidades y los centros de investigación, entre otros.

Lo anterior ha motivado que en los actuales sistemas de información se mantengan diversos repositorios de datos, de información y tal vez de conocimiento. Los cuales pueden tener un formato estructurado, semi-estructurado o sin estructura, de tal manera que la recuperación de información o conocimiento plantea diversos escenarios para realizar tal tarea. Una definición de la recuperación de información es encontrar material (usualmente documentos) de una naturaleza no estructurada (usualmente texto) que satisfaga una necesidad de información desde una gran colección (usualmente almacenada en computadoras).

Los usuarios ante la necesidad de información plantean consultas a través de una interfaz que les permita recuperar información desde un repositorio que podría ser una base de datos, una base de documentos, un servidor web, entre otros. Los sistemas de recuperación tienen implementados los mecanismos para llevar a cabo las consultas de acuerdo al tipo de repositorio, asimismo, la interfaz esta diseñada de acuerdo a la usabilidad necesaria para simplificar su uso por parte del usuario.

Ésta interfaz puede ser gráfica o textual, lo cual permite plantear las consultas sin tener conocimiento técnico acerca de los lenguajes computacionales de consulta. Por ejemplo, en Google se tiene una interfaz muy sencilla donde se encuentra una caja de texto para escribir las palabras que indican el interés de la búsqueda, los usuarios considerados expertos solamente escriben las palabras necesarias para su búsqueda, sin embargo, habrá usuarios que escriban toda una frase u oración en lenguaje natural.

De tal manera que encontramos al menos dos problemas: 1) la falta de contexto en la consulta para obtener una respuesta mas certera y 2) la complejidad del lenguaje natural para construir una consulta que proporcione también una respuesta certera. En ambos casos encontramos que la semántica es un factor muy importante para obtener una respuesta certera.

En este articulo se presenta el uso de lenguaje natural para la búsqueda de respuestas en un sistema de gestión de conocimiento de una división de estudios de posgrado e investigación. En la sección 2 se muestran algunos trabajos relacionados. En la sección 3 se muestra el diseño del modelo de conocimiento, el cual se usa para el modelo de búsqueda de respuestas que se describe en la sección 4. Al final se encuentra los resultados obtenidos y la evaluación del prototipo.

2. TRABAJOS RELACIONADOS

El Procesamiento de Lenguaje Natural (PLN) se ha usado para la creación de sistemas de Recuperación de Información (RI), sistemas de Extracción de Información (EI) y obviamente, para sistemas de Búsqueda de Respuestas (BR). Actualmente existen diferentes grupos en el mundo que investigan sobre los sistemas BR. En 1992, nació TREC [1] con el fin de crear una comunidad de investigadores en sistemas de recuperación de información. Asimismo en Europa existe CLEF [2] con los mismos objetivos, pero considerando una diversidad de lenguas. Existen hoy en día una gran variedad de sistemas BR, entre los cuales se tiene a START [3], éste es un sistema de pregunta y respuesta basado en técnicas de recuperación de información para la Web; AnswerBus [4] realiza la recuperación a través de sentencias que pueden estar en inglés, alemán, francés, español, italiano y portugués. Las frases son buscadas en la Web mediante cinco buscadores y directorios como son: Google, Yahoo, WiseNut, Altavista y Yahoo News, los cuales extraen páginas Web que contienen respuestas potenciales. Ask Jeeves [5] mejor conocida como ask.com, es un motor de búsqueda para Internet con diferentes versiones para diferentes países.

Es importante destacar que la mayor parte de los trabajos acerca de sistemas BR son para la lengua inglesa y para recuperar páginas Web. Trabajos para el idioma español son realmente pocos, entre ellos se encuentra [6], que describe la arquitectura de un sistema que responde a preguntas de usuarios cuyas respuestas están en un corpus de más de ocho mil documentos que contienen resoluciones académicas y administrativas de una universidad.

En [7] se describe cómo las ontologías, como un componente esencial de la Web Semántica, pueden ser utilizadas en una organización de investigación para crear un lenguaje común para describir su conocimiento. La misma ontología se puede utilizar para gestionar proyectos, personas, documentos y productos. La ontología permite buscar tanto el contenido de conocimiento y expertos que están vinculados a diferentes temas, y así salvar la brecha entre el conocimiento tácito y explícito. La investigación presentada en [8] ha demostrado que la utilidad del paradigma de la web semántica radica en su capacidad para objetivar la vinculación del proceso de representaciones del conocimiento explícito debido que explica estos procesos como una estructura funcional de la lógica de significaciones que es fundamental para construir el conocimiento sobre la gestión de conocimiento en una organización académica.

Nuestra propuesta está enfocada en tres principales características: idioma español, uso de lenguajes semánticos para la recuperación y un dominio acotado de aplicación como lo es la División de Estudios de Posgrado e Investigación (DEPI) de una institución de educación superior tecnológica.

3. MODELO DE CONOCIMIENTO

En esta sección se presenta el modelo de conocimiento [9] usado para la búsqueda de respuestas, para el diseño de este modelo se ha usado la metodología KoFI [10]. La metodología KoFI se compone de cuatro etapas: Identificación de fuentes de conocimiento, identificación de tipos/temas de conocimiento, identificación de flujos de conocimiento, e identificación de problemas que afecten el flujo de conocimiento. Una vez que se ha llevado a cabo el análisis con las cuatro etapas de la metodología KoFI, se puede obtener un meta modelo.



Figura 1. Taxonomía parcial de la ontología DEPI.

Este meta modelo es la base para la implementación del modelo de conocimiento el cual considera un repositorio de datos y un repositorio de conocimiento, el repositorio de datos incluye una base de datos para los documentos y para los indicadores. Para el repositorio de conocimiento se ha diseñado una ontología usando la metodología Methontology [11], la ontología ha sido implementada con la herramienta Protégé [12], en la figura 1 se muestra parcialmente la taxonomía de la ontología.

Los datos y documentos se encuentran almacenados en una base de datos relacional. Del repositorio de indicadores estratégicos se recuperan datos, que son las instancias en la ontología. Para lo cual es necesario mapear los datos a la ontología usando un proceso de población de la ontología como se muestra esquemáticamente en la figura 2.

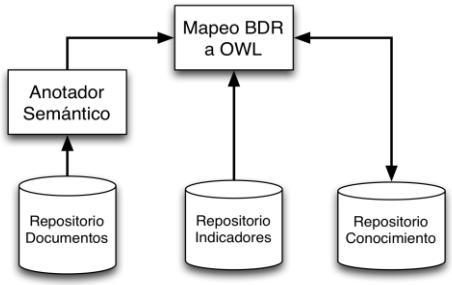


Figura 2. Procesos para la creación de instancias.

Adicionalmente, se usa un anotador semántico [13] para recuperar datos de documentos no estructurados que se encuentran en el repositorio de documentos. En la figura 3 se muestran parcialmente algunas instancias en la ontología.

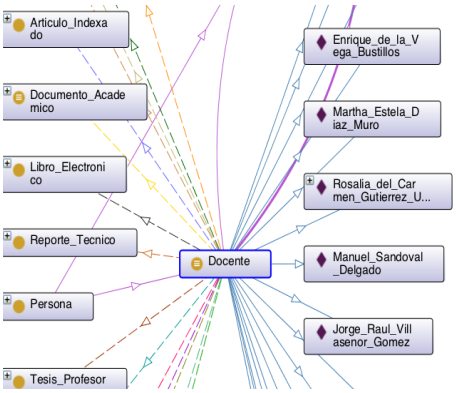


Figura 3. Conceptos e instancias en la ontología DEPI.

Este modelo de conocimiento permite realizar consultas usando el lenguaje SPARQL [14], SPARQL DL [15] o DL Query [16], así como realizar inferencias con diversos mecanismos de razonamiento, tales como Pellet [17], Fact++ [18], SWRL [19], entre otros.

4. BÚSQUEDA DE RESPUESTAS

4.1 Categorización de la Pregunta

En la figura 4 se muestran los componentes del proceso propuesto para la búsqueda de respuestas, como se puede apreciar el primer paso es la categorización de la pregunta.

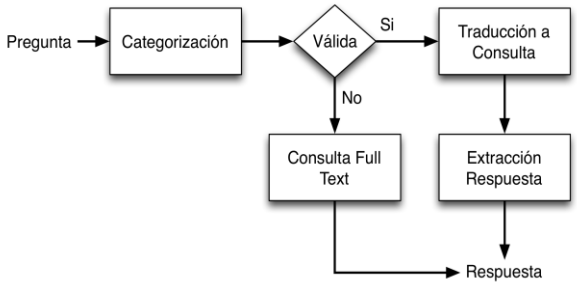


Figura 4. Proceso búsqueda de respuestas.

Para la categorización establecida en este trabajo, se ha usado como base la clasificación propuesta en [20]. La cual establece diferentes tipos de pregunta para los sistemas de respuesta automática, de ellos se consideraron solo los tipos que se muestran en la tabla 1 y su equivalencia usada en este trabajo.

Tabla 1. Comparativa tipo de pregunta

Propuesta [20]	Usadas en este trabajo
Sí / No	Booleana
Concretas	Concreta
Listas	Lista Objetos Lista Personas

Se consideró que esta clasificación de preguntas es suficiente para el dominio de aplicación previamente mencionado. Algunos ejemplos de las preguntas consideradas se muestran en la tabla 2.

La pregunta booleana es una pregunta cuyo resultado esperado es un valor booleano, esto es, "falso" o "verdadero". Se tienen de dos tipos: 1) las de membresía de un individuo que se encuentra en un conjunto o clase y 2) Las de herencia, esto es, una subclase pertenece a una clase.

En la pregunta concreta se espera un resultado que es un valor de un dato en concreto. Para la pregunta de lista de objetos se espera un resultado que sea la unión de valores concretos, este tipo de preguntas suelen ser más problemáticas que las anteriores, debido a que el sistema nunca puede saber la cardinalidad de la misma y no podría tampoco saber si existen contradicciones.

La pregunta de lista de personas es una variante de la pregunta de lista de objetos, también se espera un resultado que sea la unión de valores concretos, sólo que la lista del resultado es de personas.

4.2 Traducción al Lenguaje Semántico

Los sistema BR utilizan las técnicas típicas de PLN, que a veces pueden diferir del modelo básico del cuál se parte, pero que en esencia sirven de esquema para todos. Se inicia con la segmentación, que consiste en dividir el texto en frases y éstas en palabras. Una vez que se tienen las palabras se pueden categorizar y saber su aportación en cuanto a su significado.

Tabla 2. Clasificación de las preguntas

Tipo Pregunta	Pregunta	Ejemplos
Booleana	¿Es ….	¿Es César Rose profesor de posgrado? ¿Es César Rose investigador de posgrado? ¿Es un profesor un investigador?
Concreta	¿Cuál …. ¿Cómo ….	¿Cuál es el documento para solicitud de beca? ¿Cómo se llama el documento para solictud de beca? ¿Cuál es el proceso para el registro de proyectos?
Lista Objetos	¿Cuáles ….	¿Cuáles son los artículos de Germán? ¿César y José en cuales cursos han participado? ¿Cuáles publicaciones fueron hechas por Rosalía y Rose? ¿En cuales proyectos colabora María?
Lista Personas	¿Quién ….	¿Quién escribió acerca de semántica? ¿Quién ha impartido cursos de Minería de Datos? Personas que hayan escrito algún artículo Investigadores que hayan escrito algún artículo indexado

El análisis morfológico consiste en determinar la forma, clase o categoría gramatical de cada palabra de una oración. Algunos sistemas y herramientas permiten detectar entidades de tipo fechas, ciudades, personas, etc. Estas detecciones consideran varias palabras que se refieren a un mismo objeto. El etiquetado se basa en lematizar cada una de las palabras en la que se ha fraccionado el texto. El último paso es el análisis sintáctico. De todos los niveles de análisis, la sintaxis es a la que se ha dedicado mucho tiempo de investigación y al cual la lingüística le ha prestado mayor atención. La sintaxis permite conocer si la estructura gramatical de la oración es correcta.

En nuestro trabajo se usa el idioma español, de tal manera que se ha usado la herramienta FreeLing [21] que proporciona diversos servicios para el análisis del lenguaje. Entre estos, se encuentra la segmentación del texto, el análisis morfológico, el etiquetado gramatical, etc. De tal manera, que una vez escrita en la interfaz la frase, oración o pregunta, se procede a eliminar las palabras y puntaciones no deseados para obtener el etiquetado gramatical. En la figura 5 se muestra la obtención de las etiquetas Eagles para una sentencia usando FreeLing.

¿	Cuales	son	los	artículos	arbitrados	de	Germán	?
¿	cual	ser	el	artículo	arbitrar	de	germán	?
Fia	*PR0CP00*	*VSIP3P0*	*DA0MP0*	*NCMP000*	*VMP00PM*	*SP*	*NP00000*	*Fit*

Figura 5. Etiquetas Eagles usando FreeLing.

Esto nos posibilita para aplicar las reglas de sintaxis que fueron diseñadas y se deben cumplir para cada una de los tipos de preguntas. Para cada una de los tipos de pregunta, se diseñaron las reglas sintácticas que nos permiten verificar el tipo de pregunta y obtener las palabras de interés para realizar el mapeo hacia las consultas semánticas en lenguajes como SPARQL, SPARQL-DL y DL-Query. En este artículo por cuestiones de espacio se presenta sólo el caso de la pregunta de lista de objetos.

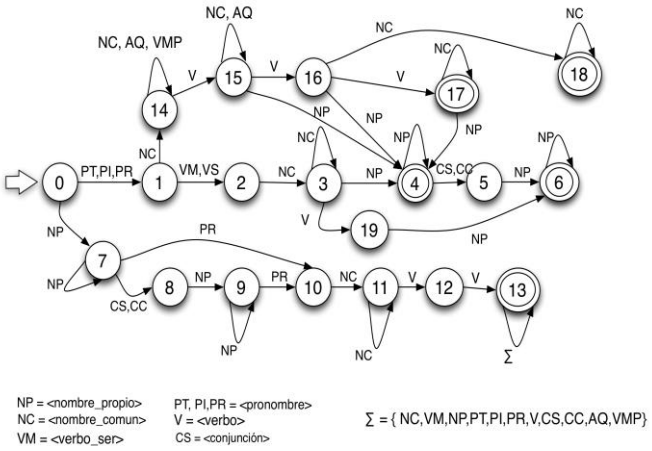


Figura 6. Autómata para la pregunta lista de objetos.

En la figura 6 se muestra el autómata que realiza la verificación de sintaxis para el reconocimiento de una pregunta de lista de objetos. El autómata tiene dos funcionalidades; la primera permite reconocer el tipo de pregunta a través de las etiquetas Eagles, esto se realiza una vez que se eliminan las palabras innecesarias y los símbolos de puntuación. Considere el ejemplo de la figura 5, sólo se usarían las siguientes etiquetas:

<PR0CP00 PR><VSIP3P0 VS><NCMP000 NC><VMP00PM VMP><NP00000 NP>

Se inicia en el estado 0 del autómata, considerando las anteriores etiquetas se llega al estado final 6, lo cual indica que se reconoce esta sentencia como pregunta del tipo de lista de objetos.

La segunda funcionalidad del autómata es obtener los sujetos y objetos usados en el mapeo a la consulta. Para el ejemplo de la figura 5 se obtiene como objetos a "artículos" y "arbitrados" y como sujeto a "Germán".

Lo anterior permite el mapeo de la cadena de texto a una forma lógica. Para este mapeo, se ha considerado que desde el punto de vista del análisis lingüístico, la categoría sintáctica y el contenido son las dimensiones centrales para la clasificación de la sentencia.

Las categorías sintácticas en este trabajo son: pregunta sí/no, pregunta cuáles, pregunta quiénes. Asimismo, se incluye un subtipo sintáctico que relaciona una palabra con el pronombre establecido en la sentencia. El contenido consiste de conceptos y relaciones de la ontología, así como el tipo ontológico siendo cuestionado. Para la siguiente pregunta de lista de objetos se tiene la forma que se muestra en la figura 7.

Pregunta: ¿Cuáles son los artículos arbitrados de Germán?

Para la siguiente pregunta de lista de objetos se tiene la forma que se muestra en la figura 8.

Pregunta: *¿Cuáles cursos fueron impartidos por los docentes?*

Forma	
Entrada	*¿Cuáles son los artículos arbitrados de Germán?*
Categoría Sintáctica	*Cuáles*
Subtipo Sintáctico	*artículos arbitrados*
Contenido	
Restricción	*Publicación*
Conceptos	*Docente* *Publicación*
Relaciones	*escribeArticuloArbitrado*

Figura 7. Forma pregunta lista de objetos.

Forma	
Entrada	*¿Cuáles cursos fueron impartidos por los docentes?*
Categoría Sintáctica	*Cuáles*
Subtipo Sintáctico	*cursos*
Contenido	
Restricción	*Curso*
Conceptos	*Docente* *Curso*
Relaciones	*imparteCursoPor*

Figura 8. Forma pregunta lista de objetos.

En la figura 9 se presenta el algoritmo principal, como se puede apreciar, muestra el proceso iniciando con el análisis de la sentencia en lenguaje natural usando las funcionalidades de FreeLing. Posteriormente se verifica el tipo de pregunta usando los autómatas y una vez clasificada el tipo de pregunta se ejecuta el correspondiente algoritmo de mapeo a la consulta semántica. En caso de no corresponder a uno de los tipos de pregunta, se

Algoritmo 1 General
Entrada: S sentencia en lenguaje natural
Salida: Resultado en texto
LS contiene los sujetos y *LO* los objetos al ejecutar el automata correspondiente
$L \leftarrow analisisFreeling$(S) donde L contiene la secuencia de etiquetas Eagles de S.
si $automataCLO$(L) es cierto **entonces**
  $Resultado \leftarrow algoritmoCLO$(LS,LO). {resultado lista de objetos}
si no, si $automataCLP$(L) es cierto **entonces**
  $Resultado \leftarrow algoritmoCLP$(LS,LO). {resultado lista de personas}
si no, si $automataCC$(L) es cierto **entonces**
  $Resultado \leftarrow algoritmoCC$(LS,LO).{resultado concreta}
si no, si $automataCB$(L) es cierto **entonces**
  $Resultado \leftarrow algoritmoCB$(LS,LO). {resultado si/no}
si no
  $Resultado \leftarrow algoritmoFullText$(LS,LO). {resultado full text}
fin si

Figura 9. Algoritmo principal.

realiza una consulta sintáctica usando el algoritmo de Full Text.

El algoritmo para la pregunta de lista de objetos se muestra en la figura 10. Como se estableció previamente, los objetos son nombres comunes asociados con la categoría sintáctica (pronombre relativo), la variable de la consulta es el subtipo sintáctico y es un objeto y los sujetos son nombres propios o nombres comunes.

Algoritmo 2 algoritmoCLO: Lista de Objetos
Entrada: LS lista de sujetos y LO lista de objetos
Salida: Resultado en texto
Resultado ←vacia.
si *docente*(LS) es cierto **entonces**
sujetoDocente ←true. {verifica sujeto como docente}
si no
sujetoAlumno ←true. {verifica sujeto como alumno}
fin si
propiedad ← *obtienePropiedad*(LO). {obtiene propiedad usando el subtipo sintactico y las reglas para determinar la propiedad correspondiente}
si *sujetoConNombrePropio*(LS) es cierto **entonces**
consulta ← *construyeConsultaSPARQL*(LS,LO,propiedad). {construye consulta en SPARQL usando los sujetos, objetos y la propiedad}
Resultado ← *ejecutaConsultaSPARQL*(consulta)
si no
consulta ← *construyeConsultaDLQuery*(LS,LO,propiedad). {construye consulta en DLQuery usando los sujetos y objetos}
Resultado ← *ejecutaConsultaDLQuery*(consulta)
fin si
devolver Resultado

Figura 10. Algoritmo Lista de Objetos.

A continuación se muestran los resultados de la construcción de las preguntas previamente mostradas, es importante señalar que la pregunta se puede mapear al lenguaje SPARQL, así como al lenguaje DL Query como se puede ver en el siguiente ejemplo.

Pregunta: ¿Cuáles son los artículos arbitrados de Germán?

Consulta SPARQL:

SELECT ?articulos_arbitrados

where {?a ?b ? articulos_arbitrados.

FILTER regex(str(?b),"escribeArticuloArbitrado","i")

FILTER regex(str(?a),"German","i")

}

Consulta DL Query:

articuloArbitradoEscritoPor some {German}

Pregunta: ¿Cuáles cursos fueron impartidos por los docentes?

Consulta DL Query: *(Curso and cursoImpartidoPor some Docente)*

Como se puede apreciar en el algoritmo CLO, las consultas se construyen en el lenguaje SPARQL o en el lenguaje DLQuery. En el algoritmo se muestra que la decisión de que lenguaje usar para construir la consulta se basa en la existencia de los sujetos con nombres propios. Adicionalmente, para esta decisión se considero la facilidad de la implementación de expresiones regulares para construir patrones de búsqueda de los nombres propios. Aún cuando se pueden construir expresiones regulares en DL Query, es más sencilla su implementación en SPARQL.

5. RESULTADOS

Para realizar la experimentación se implementó un prototipo basado en una arquitectura cliente-servidor, para el servidor web se usó Apache-Tomcat y se implementó en Java Server Page el código necesario para ejecutar los algoritmos implementados en Java para las diferentes tipos de preguntas previamente mencionadas. En la figura 11 se muestra la interfaz del prototipo con un ejemplo de pregunta de lista de objetos, en la figura 12 se muestra su correspondiente resultado.

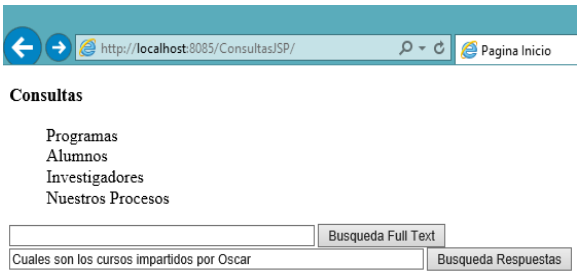


Figura 11. Interfaz prototipo para la pregunta.

Para el experimento se crearon instancias a partir de las bases de

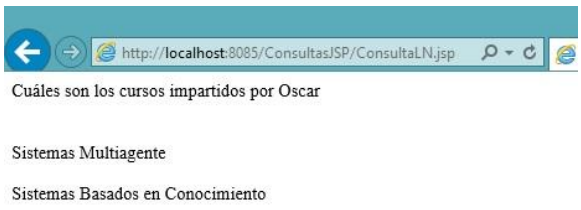


Figura 12. Resultado de la pregunta.

datos de indicadores y documentos. Se usaron los datos de 48 publicaciones académicas, 61 documentos administrativos y académicos, 17 profesores, 13 proyectos, 18 procesos administrativos y académicos y 8 cursos. Una vez creado el modelo de conocimiento se procedió a realizar las preguntas de los tipos anteriormente mencionados. La tabla 3 nos muestra la distribución de las consultas y los resultados obtenidos. Es importante destacar que no se hicieron consultas cuando correspondían a consultas que de antemano se conocía que no existía el individuo, por lo tanto no se obtuvieron consultas "No contestadas".

Por ejemplo la consulta: ¿Cuáles son los artículos de XYZ?. Ya que XYZ no se encuentra en la base de datos no se obtiene una respuesta, sin embargo, si se hubieran hecho este tipo de consultas sólo se incrementaría la cantidad de consultas pero el porcentaje de éxito se sigue manteniendo en el 100%, ya que sería correcta la respuesta.

Tabla 3. Resultados de la experimentación

Consultas	Booleanas	Concretas	Lista Objetos	Lista Personas	Total
Respuesta Correcta	8	11	15	22	56
Respuesta Incorrecta	0	0	0	0	0

6. DISCUSIÓN Y CONCLUSIONES

Una de las principales dificultades para comparar nuestra propuesta con otros sistemas BR es que nuestro sistema usa un dominio muy acotado y el resto usa la Web. Estos últimos han hecho su experimentación usando un corpus de documentos, por

ejemplo CLEF. En nuestro caso se utilizaron los repositorios del sistema de gestión de conocimiento de la DEPI.

La base de datos de indicadores se usó para generar las instancias de la ontología y para verificar en algunas preguntas los nombres propios de alumnos y profesores. Asimismo, la base datos de documentos se utilizó para la recuperación del documento. Es importante señalar que a través del sistema de gestión de documentos que es parte del sistema, se extrajo el corpus de texto de los documentos que se utiliza para su posterior recuperación.

Por otro lado, el uso de un dominio acotado permite construir sistemas de respuesta automática que puedan asegurar una buena precisión en las respuestas ya que el conjunto de posibilidades está restringido. Esto conlleva a obtener resultados más adecuados y concretos y evita que el usuario tenga que realizar un trabajo adicional para filtrar los resultados hasta encontrar lo que sea de su interés.

Otra ventaja de usar una interfaz de lenguaje natural permite a los usuarios interactuar con el sistema de gestión de conocimiento de manera sencilla, sin necesidad de contar con el aprendizaje de los lenguajes de consulta que se están utilizando para el acceso a los modelos de conocimiento, permitiendo al usuario a través del lenguaje natural una mayor flexibilidad con las consultas

7. REFERENCIAS

[1] TREC Web Site. Disponible online: http://trec.nist.gov/. [Consultado el 8 de Agosto de 2016].

[2] CLEF Web Site. Disponible online: http://www.clef-initiative.eu//. [Consultado el 8 de Agosto de 2016].

[3] START Web Site. Disponible online: http://start.csail.mit.edu/. [Consultado el 9 de Julio de 2015].

[4] Zhiping, Z. 2002. AnswerBus question answering system. In *Proceedings of the second international conference on Human Language Technology Research (HLT '02)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 399-404.

[5] Ask Jeeves Web Site. Disponible online: http://es.ask.com/. [Consultado el 5 de Julio de 2016).

[6] Cardoso, A. C., Bini, A., and Pérez Abelleira, M. A. 2015. Diseño de un sistema de búsqueda de respuestas para diversos tipos de preguntas. *In Argentine Symposium on Artificial Intelligence (ASAI 2015)-JAIIO 44* (Rosario, 2015).

[7] Bäck, A., Vainikainen, S., Södergård, C., & Juhola, H. (2003). Semantic Web Technologies in Knowledge Management. In ELPUB.

[8] Velázquez, P. C. (2014). Knowledge management, semantic web and logic of significances to create technologic skills for support e-learning methodologies. International Journal of Information and Education Technology, 4(1), 1.

[9] Zavala, A. 2013. *Modelo de conocimiento para un sistema de gestión de conocimiento de una División de Estudios de Posgrado e Investigación*, Tesis de Maestría, Instituto Tecnológico de Hermosillo.

[10] Rodríguez Elías O. 2007. *Metodología para el diseño de sistemas de administración del conocimiento: su aplicación en mantenimiento de software*. Tesis de Doctorado, Centro de Investigación Científica y de Educación Superior de Ensenada.

[11] Fernández-López, M., Gómez-Pérez, A., and Juristo, N. 1997. Methontology: from ontological art towards ontological engineering. In *AAAI Symposium on Ontological Engineering*, Stanford.

[12] Protégé Web Site. Disponible online: http://protege.stanford.edu/. [Consultado el 12 de Junio de 2016].

[13] Martínez, I. 2015. *Diseño de un Anotador Semántico Semiautomático de Documentos en un Dominio de Educación e Investigación Tecnológica*, Tesis de Maestría, Instituto Tecnológico de Hermosillo.

[14] SPARQL Web Site. Disponible online: http://www.w3.org/TR/rdf-sparql-query/. [Consultado el 8 de Agosto de 2016].

[15] SPARQL DL Web Site. Disponible online: http://www.derivo.de/en/resources/sparql-dl-api/. [Consultado el 8 de Agosto de 2016].

[16] DL Query Web Site. Disponible online: http://code.google.com/p/co-ode-owl-plugins/wiki/DLQuery. [Consultado el 8 de Agosto de 2016].

[17] Pellet Web Site. Disponible online: http://clarkparsia.com/pellet/. [Consultado el 8 de Noviembre de 2015].

[18] Fact++ Web Site. Disponible online: http://owl.man.ac.uk/factplusplus/, [Consultado el 8 de Agosto de 2016].

[19] SWRL Web Site. Disponible online: http://www.w3.org/Submission/SWRL/. [Consultado el 8 de Agosto de 2016].

[20] Celaá. D. 2010. *Sistema de respuesta automática basada en recursos Semánticos*. Universidad Carlos III de Madrid.

[21] Padró, L. 2012. Analizadores multilingües en freeling. *Linguamática*, vol. 3, no 2, p. 13-20.



Taller:
Knowledge Discovery and Cloud Computing Applications

Organizadores:
Víctor Jesús Sosa Sosa (CINVESTAV)
Maricela Claudia Bravo Contreras (UAM)
José Alejandro Reyes Ortiz (UAM)

An Approach for the Extraction and Visualization of Semantic Knowledge in Private Cloud Storage

Leyla Alvarez-Medina, Ivan Lopez-Arevalo, J.L. Gonzalez
Cinvestav Tamaulipas, México
lalvarez@tamps.cinvestav.mx
ilopez@tamps.cinvestav.mx
jgonzalez@tamps.cinvestav.mx

Ana B. Rios-Alvarado
Autonomous University of Tamaulipas
México
arios@uat.edu.mx

ABSTRACT

Cloud storage technology has becoming a cost-effective solution for organizations to manage their data in an efficient manner. The data accumulation over the time results in information repositories that would be considered as assets by organizations if data were analyzed. However, the information stored in private clouds are not usually analyzed by organizations. This avoid organizations obtaining knowledge and taking advantage of it in the improvement of data manage. This paper presents an approach for the extraction of semantic knowledge from cloud storage repositories as well as the visualization of the acquired knowledge. In this approach, the extraction knowledge is based on the topic detection from repositories of unstructured data (documents) stored in a private cloud storage, the extracted semantic knowledge is indexed in structured data, whereas an application, based on a topic index, enable organization to visualize the knowledge in the form of graphs of topics per cloud storage location. The implementation of proposed approach in a private cloud storage shows the feasibility of this approach to get and visualize semantic knowledge extracted from documents. This knowledge can be used as a criterion for any of the allocation of data in cloud storage systems, the categorization of users through profiles. It also could be useful for building distribution schemes based on topics and document search services.

Keywords

Semantic knowledge; Cloud storage; Topic detection; Graph Visualization

1. INTRODUCTION

The production of new data has grown in constant manner over the past few years [4] and it is expected to grow exponentially in the next years [5]. Cloud storage technology is a cost-effective solution for organizations to face up this trend in an efficient manner [4][5].

The outsourcing model associated to public cloud storage services comes at risk of issues such as vendor dependency [2], security [2] [14] and data mining [14] nevertheless. In order to take advantage of cloud computing benefits without suffering from the side-effects associated to the cloud storage services offered by public providers, organizations are building private cloud storage infrastructure [8] [13]. This technology has becoming a popular solution for organizations over the few past years.

In real scenarios, a data accumulation effect is produced when organizations stores data by using cloud services (either public or private) during large periods of time. This effect produces data repositories including information that would be considered as assets by the organizations if data were analyzed [9]. In fact, studies estimate that the percentage of information that could be useful if tagged and analyzed will be 40% by 2020 [5]. In this context, there is an opportunity for organizations to extract semantic knowledge of the information repositories of their private cloud infrastructure and take advantage of the it either in the improvement of data manage and making decisions processes.

For large repositories of documents it is a crucial task the detection of topics. The detection of topic of interest is a very important area for Information Retrieval (IR) that enables either researchers or people who make decisions to efficiently identify significant and representative topics and to identify trends from diverse documents. It is particularly useful when researchers modeling the behavior of social networks as well as customers and end-users of services. Although public cloud providers already earn a profit from analyzing information extracted from cloud storage repositories, in private clouds is not commonly that the information stored being analyzed by organizations.

This paper presents an approach for the extraction of semantic knowledge from cloud storage repositories as well as the visualization of the acquired semantic knowledge. In this approach, the extraction knowledge is based on the topic detection from repositories of unstructured data (documents) stored in a private cloud storage. The extracted semantic knowledge is organized in structured data, which enables our approach to display graphs of topics per cloud storage location.

In a private cloud storage, the implementation of the proposed approach includes a topic detector, an index manager and a visualization tool. The implementation was tested by using a set of abstracts from scientific articles from Med-

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

line[1], which were associated to a repository of PDF documents. This proof-of-concept shows the feasibility of our approach to get and visualize semantic knowledge extracted from documents stored in private repositories. We consider this knowledge can be used as a criterion for allocating data in cloud storage systems, the determination of user profiles or a base for distribution schemes based on topics and document search services. Figure 1 shows the main stages for this approach.

2. RELATED WORK

In data mining, proposals for the extraction of information from textual data aims to identify entities, events or key phrases and the relationships between them have been deeply studied in the past to reveal significant semantic information beyond a single set of words. Topic detection is a subtask from Topic Detection and Tracking (TDT), the clustering algorithm for text called CBC (Clustering By Committee) is an example of this scheme [11]. This algorithm manages a set of elements called representative committees with which it tries to discover centroid groups that describes unambiguously the members of a given class. CBC represents the elements (words) as a feature vector that represents the terms that appear in documents and their numeric value is the document term stat.

Traditional methods of text mining for topic detection are Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (pLSA), Latent Dirichlet Allocation (LDA) and topic correlation model. A scheme for measuring the similarity of texts were also proposed by [6], whereas a suite of algorithms for semantic document clustering were recently proposed by [10].

A method called *paragraph extension method* was proposed for the topic and subtopic detection [15]. In this method a document is considered as a set of consecutive paragraphs. Consecutive paragraphs are fusioned are considered as similar. Tthe number of similar words in the merged paragraphs are added to build scores of subtopics by using a corpus, which allows them to find words related to a secondary topic. The paragraph vectors are represented by subtopics instead of the composed words. The sub-theme of a paragraph are the most frequent in paragraph vector, whereas the theme of the document is the most common document subtopic. The main issue of this type of schemes is the computation time required for the topic detection.

In order to solve the problems of time efficiency in the detection of topics, an approach for topic detection inspired by the efficient methods of analysis and graph representation in a form of co-occurrence of relevant words was proposed by [12]. A hybrid analysis approach for relationships integrating semantic relations and relations of co-occurrence in the tasks of topic detection was proposed by [16]. This scheme integrates semantic information and information co-occurrence between terms for the detection of subjects. More specifically, the approach fused multiple relationships in a graph of terms by incorporating theories of discovery of ideas (Idea Discovery Theory) with modeling methods issues and on the extraction of semantic relationships based on LDA (Latent Dirichlet Allocation) to enrich the graph with the semantic information. However, although both schemes represent an efficient solution for the topic detection, they have no been implemented in a real cloud storage scenario.

3. AN APPROACH FOR THE EXTRACTION AND VISUALIZATION OF KNOWLEDGE

In this section we describe the components for topic detection, index manager and visualization, which are the main component in the approach for the extraction of visualization of semantic knowledge. We also describe the iteration of the implementation of our approach with a private cloud storage [7].

[1]National Library of Medicine journal citation database. http://www.ncbi.nlm.nih.gov/pubmed (visited on September 2016)

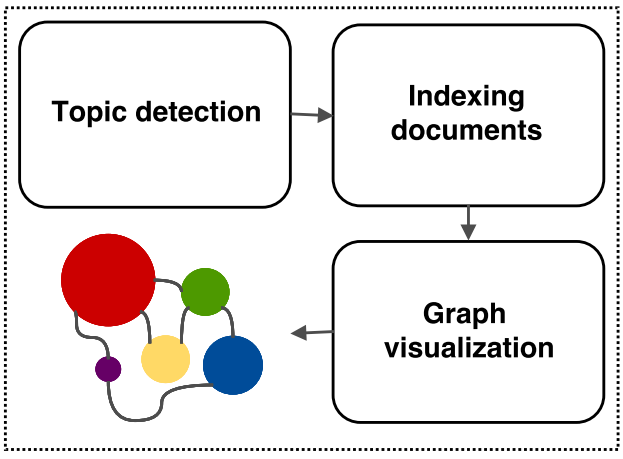


Figure 1: Methodology for extraction and visualization of semantic knowledge in private cloud storage.

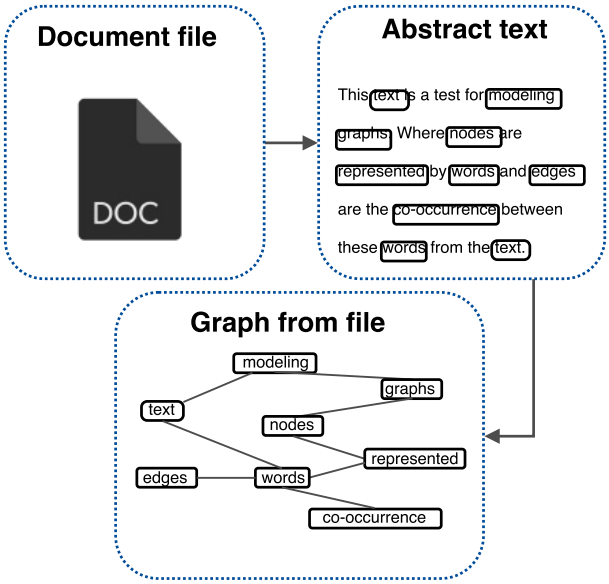


Figure 2: Representation of a document by using a graph model based on topic detection.

3.1 Topic detection and semantic aware component

This component was designed as a black box that receives a set of documents $D\{d_1..d_r\}$, identifies a set of keywords $k \in d_i$; as a result, each document d_i analyzed by this black box can be represented in a graph by $k \in d_i$ as nodes and the co-occurrences between words as edges. In this form, a document d_i that belongs to a set of documents stored in a cloud storage system can be processed by topic detection techniques and even more, displayed like a topic related system based on topics.

We implemented our black box for the detection of topics by using the *Keygraph algorithm* [12]; which enables us to analyze document repositories in reduced periods of time (seconds or even minutes depending on the size of the corpus to be analyzed).

Figure 2 shows the three procedure stages applied to each document into the black box to build a graph based on the information of a given document (d_i). As it can bee seen, the content of d_i (document file stage) is converted into plain text and an abstract is extracted from it (Abstract text stage). This abstract is preprocessed to remove stop-words from the text, which could affect the data mining technique. When the abstract has been cleaned, the remaining words are mapped into a graph structure making edges from their co-occurrence in text (Graph from file stage).

In the last stage a procedure in the black box chooses a $k \in d_i$ from the keywords graph as a representative topic for each document ($T \rightarrow d_i$). This is the information returned by this black box when is invoked by a component of the approach.

3.2 Index manager

When documents are assigned to a topic ($T \rightarrow d_i$), this information is delivered to a component so called *Index Manager*, which registers this map in a database; as a result, a graph of the analyzed documents $D\{d_1..d_r\}$ is built by using the maps managed by this component.

3.3 Semantic graphs generator

This component is in charge of creating a text file (R_{gf}) that reflects the knowledge of repository (R) by using the maps of $D\{d_1..d_r\}$ indexed by the Index Manager. This information can be processed by tools like Gephi [1] or the graph can be built by a Java program using a XML library or an API extended by Gephi.

3.4 Integrating components in a workflow

The workflow shown in Figure 3 depicts the interconection of the components of our approach. This workflow considers an example where a repository of documents stored in a cloud storage service (SkyCDS system) $D\{d_1..d_r\}$ is used by topic detection and semantic aware component to determine a topic per document (Abstract and Topic Detection). It also shows how this topic is sent to the Semantic graphs generator to create a text file (R_{gf}) (Build and display graph), which will used to display a knowledge graph (Graph visualization). For instance, Figure 4 shows a graph of topic-document relationship.

Coarse-grained and fine-grained visualization

We implemented a fine-tunning configuration tool that enables the users to assign colors to each topic and its keywords. All graphs are visualized using Fruchterman Reingold algorithm[3] to distribute the nodes. Topics are differentiated by having distinct colors.

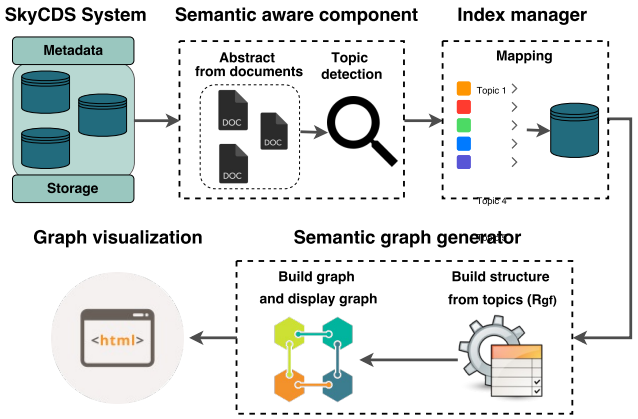


Figure 3: Workflow to build and display a semantic information on a graph.

This tool also enables the users to determine the semantic knowledge in different manners: *Fine-grained*. Shows a graph where each node represents a specific element. For instance, Figure 5 shows the knowledge extracted by showing keywords that conform the topics detected. That is fine-grained because it shows every single keyword and their relation to a topic.

Coarse-grained. Shows a visualization grouped such a zoom out of the extracted knowledge. An instance for coarse-grained is the Figure 6, grouping the keywords for a topic. It makes explicit change in volume of topics.

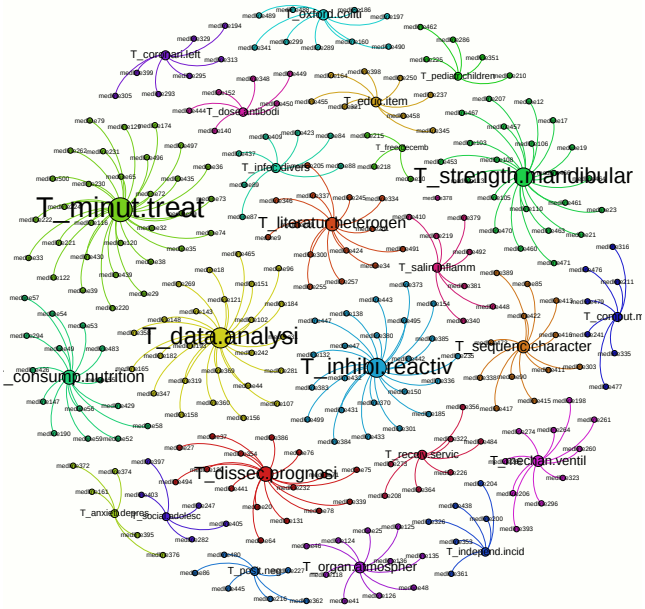


Figure 4: Medline documents assigned to topics detected in a bipartite graph. Documents are represented by small nodes and are linked to topics wich are showed as bigger nodes.

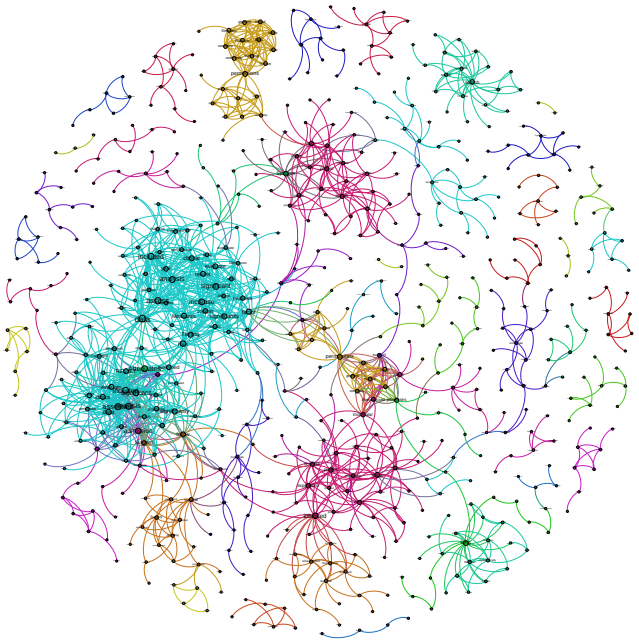


Figure 5: Keywords from detected topics on 500 Medline documents. Nodes are keywords and edges are co-occurrence between them. Each group color represents a topic

Figure 5 shows a graph of keywords and topics extracted from a set of 500 documents (obtained from MedLine) stored in SkyCDS [8] by using a modified version of *Keygraph algorithm*, which build the graph by an XML library. This graph also allows the owner of the repository to know the proportion of topics in its repository. There is another manner to display the volume of topics, by grouping the keywords, which is depicted in Figure 6.

Given that the detected topics are sets of related keywords. Each set (topic) is shown with a different color, just

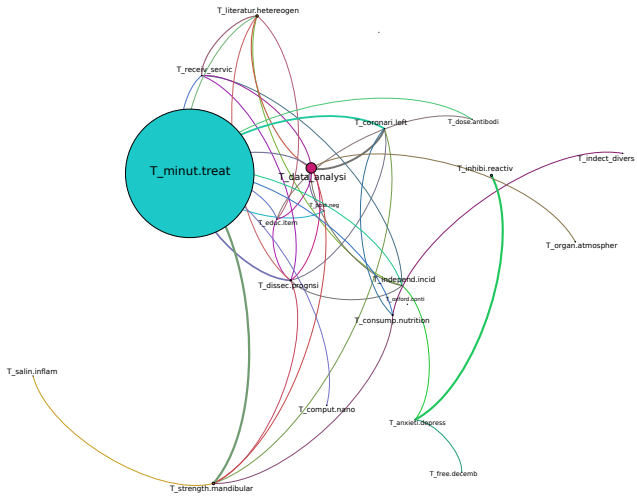


Figure 6: Keywords grouped from topics detected on 500 Medline documents stored in SkyCDS.

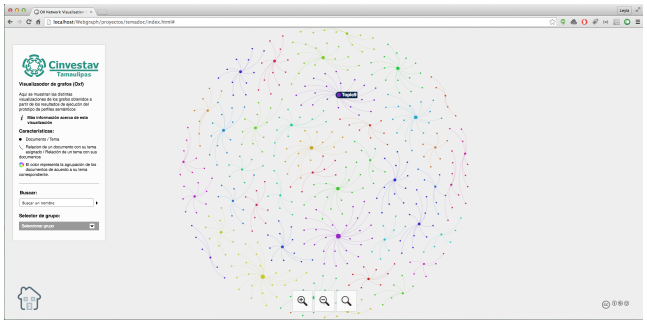


Figure 7: General display of the Webgraph. Topic and documents are nodes, edges are arcs.

as shown in Figure 5. The name of the topics can be simply observed in Figure 6 which is given by the structure `T_{keyword1}.{keyword2}`, where the name is taken from the first two keywords of the words list that forms the topic.

The approach also enables the administrator to visualize a structure formed by files stored in different cloud storage locations or topics detected and distributed per cloud storage location. This means the administrator could determine the load of topics per cloud storage location (for instance, google store, azure, amazon, etc) where documents were stored by a simple graph visualization.

4. THE WEBGRAPH VIEWER TOOL

For facilitating to the Administrator of cloud storage service the visualization of semantic knowledge by displaying information, we implemented a tool called *Webgraph viewer*. This tool shows in friendly manner on a web browser the results obtained by applying our method. This without the need for complex programs installed.

By using the *Sigma.js* library, a call is made to a JSON or gexf file to display the graph on the screen. The design of the Webgraph viewer was inspired on the *Interactive Visualizations* project made by the Oxford Internet Institute[2].

Each generated graph has a different structure so that the nodes represent words, documents or users respectively. Figure 7 shows a general view of Webgraph viewer, displaying a topic-document relationship graph, and a section to select the desired graph representation. This display is intended to be placed in the system controller storage management administration.

5. CONCLUSIONS

In this paper we presented the implementation of an approach for the extraction of semantic knowledge from cloud storage repositories as well as the visualization of the acquired knowledge. In this approach, the extraction knowledge is based on the topic detection from repositories of unstructured data (documents) stored in a private cloud storage. The extracted semantic knowledge is indexed in structured data, whereas an application, based on a topic index, enable organization to visualize the knowledge in the form of graphs of topics per cloud storage location.

The implementation of proposed approach in a private cloud storage shows the feasibility of this approach to get

[2]http://blogs.oii.ox.ac.uk/vis/ (Visited on September 2016)

and visualize semantic knowledge extracted from documents. This knowledge can be used as a criterion for any of the allocation of data in cloud storage systems, the categorization of users through profiles. It also could be useful for building distribution schemes based on topics and document search services.

6. REFERENCES

[1] M. Bastian, S. Heymann, M. Jacomy, et al. Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 2009.

[2] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina. Controlling data in the cloud: outsourcing computation without outsourcing control. In *CCSW '09*, pages 85–90, 2009.

[3] T. M. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Softw., Pract. Exper.*, 21(11):1129–1164, 1991.

[4] J. Gantz and D. Reinse. Extracting value from chaos. In *IDC-EMC2 report*, July 2011.

[5] J. Gantz and D. Reinsel. The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the Future*, pages 1–7, 2013.

[6] W. Gomaa and A. Fahmy. A survey of text similarity approaches. *International Journal of Computer Applications*, 68(13):13–18, 2013.

[7] J. Gonzalez, J. C. Perez, V. J. Sosa-Sosa, L. M. Sanchez, and B. Bergua. Skycds: A resilient content delivery service based on diversified cloud storage. *Simulation Modelling Practice and Theory*, 54:64–85, 2015.

[8] J. L. Gonzalez, J. Carretero Perez, V. Sosa-Sosa, J. F. Rodriguez Cardoso, and R. Marcelin-Jimenez. An approach for constructing private storage services as a unified fault-tolerant system. *The Journal of Systems and Software*, 86, 2013.

[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. Big data: The next frontier for innovation, competition, and productivity. 2011.

[10] M. P. Naik, H. B. Prajapati, and V. K. Dabhi. A survey on semantic document clustering. *Proceedings of 2015 IEEE, ICECCT 2015*, 2015.

[11] P. A. Pantel. Clustering by committee. Master's thesis, Citeseer, 2003.

[12] H. Sayyadi and L. Raschid. A Graph Analytical Approach for Topic Detection. *ACM Transactions on Internet Technology*, 13(2):1–23, 2013.

[13] S. Singh and T. Jangwal. Cost breakdown of public cloud computing and private cloud computing and security issues. *IJCSIT Vol 4, No 2, April 2012*, pages 17–31.

[14] R. H. Sloan and R. Warner. *Unauthorized Access: The Crisis in Online Privacy and Security*. CRC Press, first edit edition, 2013.

[15] M. Turan and C. Sönmez. Automatize Document Topic and Subtopic Detection with Support of a Corpus. *Procedia - Social and Behavioral Sciences*, 177(July 2014):169–177, 2015.

[16] C. Zhang, H. Wang, L. Cao, W. Wang, and F. Xu. A hybrid term-term relations analysis approach for topic detection. *Knowledge-Based Systems*, 2016.

Web Services for Controlling User's Environments Through Ontologies

Jose A. Bernal
Computer Science Department,
CINVESTAV-IPN
2508 Av. IPN, San Pedro Zacatenco
Mexico City, Mexico
jbernal@computacion.cs.cinvestav.mx

J. Rodriguez
Computer Science Department,
CINVESTAV-IPN
2508 Av. IPN, San Pedro Zacatenco
Mexico City, Mexico
rodriguez@cs.cinvestav.mx

ABSTRACT

A web service is becoming an essential component in the development of systems for environments with high human interaction such as in ubiquitous computing. In this field there is the need to have a higher understanding about the context in order to be capable of providing the most adequate services to users. The work developed proposes to modify the functioning of services in order to supply the right service according to the needs of the user and his location. This proposal presents a set of services for the interaction of the environment with the user, this proposal has been developed entirely with open source software and hardware.

CCS Concepts

•**Information systems** → **Web services;** *Web Ontology Language (OWL);* Web Services Description Language (WSDL); •**Computer systems organization** → **Sensor networks;** •**Theory of computation** → *Semantics and reasoning;* •**Hardware** → *Sensors and actuators; Sensor applications and deployments; Sensor devices and platforms; Wireless devices;*

Keywords

Wireless Sensor Networks, OWL, Hardware open source, Web Services, Sensors and Actuators

1. INTRODUCTION

In his article Mark Weiser [12] presents his vision of a world where computers are present in the everyday activities of users, in the background and interacting with them through daily use objects, in this way computers become invisible for users. The aim of ubiquitous computing is to improve the live of users by supplying useful information without requiring that user be distracted by the interaction with the system.

The evolution of technology in last years has made possible to implement systems following Weiser's vision. One of the components that has gained popularity is the wireless sensor networks (WSN) [7]. This kind of networks are used on several areas such as: e-health-care or industry, for example. In WSN the clusters of sensors can be deployed in a random way over the area to be sensed.

The use of web services allows to access in information from everywhere in the Internet giving as result the integration of that information. Another quality of WS is that they allow a direct interaction with applications through standardized protocols easing the link creation between the client (user) and the supplier (server), in this communication schema a server can serve several clients.

In other hand the use of ontologies has allowed to implement systems with a richest expressiveness that can be interpreted both by humans and by other systems by giving semantic meaning to data. In this way using ontologies is possible to define relations, properties and characteristics about data. Through the classes defined over the data defined and using inference rules to validate relations between data the ontology can produce information.

By combining the WS and ontologies the semantic web services (SWS) were created. The SWS use the protocols of WS and the semantic supplied by the ontologies. Giving the expressiveness supplied by the ontologies the SWS can supply more suitable information to the users according their requests and interests. This make easier the interaction between the users and the systems.

In this paper is presented a set of web services that through an ontology are capable to control the environment of a specific place and give information about persons and their environment.

The rest of this paper is structured as follows: in section 2 the state of the art about WSN and web services for ubiquitous computing are presented. In section 3 the application scenario and the implementation are presented. Section 4 presents the results and finally the conclusion are presented in section 5.

2. RELATED WORK

Several works have been developed where ontologies and WSN are integrated in order to enhance the information gathered from the environment. Combining the information from the environment with XML and OWL it is possible to give meaning to the data supplied to the user and making that data more useful.

Several approaches have been used in those works, some of them search for a better communication between hardware and software through the use of dynamic discovering of services [6], in this approach searches present several issues to be taken into account, for example the *a priori* acknowledge on the hardware, it's capacity for gathering data and

their characteristics. The but is that every single device found is added automatically into the system. The proposal of this authors is called Semantic Rubicon; this work allows the interaction of every device available in the network through the automatic discovering. The discovered services are stored in a table that is query in order to have a better interaction between devices and services.

One of the main qualities searched in WSN is the scalability moreover of an interaction with other type of systems. One option to this issue is to mount tiny web servers on embedded systems having sensors, these systems have few resources like computing power, and memory [10]. The communication is carried out using protocols like UDP or HTTP. These protocols are selected in order to use messages for communication over heterogeneous systems. The devices are free of communicating with any kind of devices in the system, even if they are just arriving into the system.

Developing smart everyday objects is a big step in developing ubiquitous systems, because those smart objects do not require that user should be distracted in order to interact with the device [1]]. In this proposal every day object, like coffee cups, are used as augmented objects. This kind of objects are equipped with sensors to detect if the object is moving and is capable of supplying his geographical location inside the office. A set of receivers are installed inside the building so it is possible to determine the location of users given that every object has a unique identifier. Given that each object belongs to a user, through the sensors and the identifier of the object is possible to know where the user is.

In spite of the progress in communication, computing, sensors and mobile devices, to implement a ubiquitous system there is always several issues to solve, mainly the heterogeneity in devices and software. In order to face this heterogeneity, a "federation of ubiquitous systems" has been proposed. This federation is a collection of intelligent nodes where each node is composed of a set of heterogeneous hardware and software devices. A framework is responsible of the interaction between the nodes it allows to users to carry out a task in one node and leave the current node to continue in other node without losing data.

A Web Service is a program that can be accessed remotely supplying a service in a direct way or through third components [11]]. For this is needed to use notations using (Web Service Description Language (WSDL) and XML to publish the service in a register, allowing it's discovering, it's invocation or composition. In WSDL 2.0 the man types of notations include: Interface notation, where an element of the interface can be noted using a referenced model to supply the link to some concept in the interface.

The semantic web has come to give to the Web the ability to produce information that can be interpreted both by humans as by computers, by giving meaning to the data and allowing that both users and computers work in a collaborative way. In a semantic system the computers must have access to well structured information collections and the set of rules to be used in order to produce automated reasoning. One of the main challenges in the semantic web is to create a language that could be interpreted by any knowledge representation system that could be used in the Internet.

The term ontology has been adopted by the IA community and is used to defining the relations between terms [2][8]]. Moreover, the ontology must define the inference rules and the taxonomy. The taxonomy defines the classes and the relations between them. One of the principal uses of semantic web is to collect information from different sources in web to process such information and exchange it with other systems.

The semantic Web is considered as the application of techniques for representing knowledge in a distributed environment [3][4][5]and it has been used to produce enhanced descriptions of web services known as Semantic Web Services.

3. SCENARIO

Humans always has been receiving information coming from the environment. The information produced can have different nature so that it is necessary to get and process all this information in order to be presented to the user in the most suitable way, and to improve the understanding and usability. As it can be seen the information gathered from the environment must be filtered in order to be useful for the user according his location.

The figure 1 represents the whole system layers, it's important to explain briefly each layer of the system. The first layer is called "Application layer" and it's located at the top of the architecture here is where the end user application must be deployed, means that here can be developed different applications that invoke Web services developed to this system.

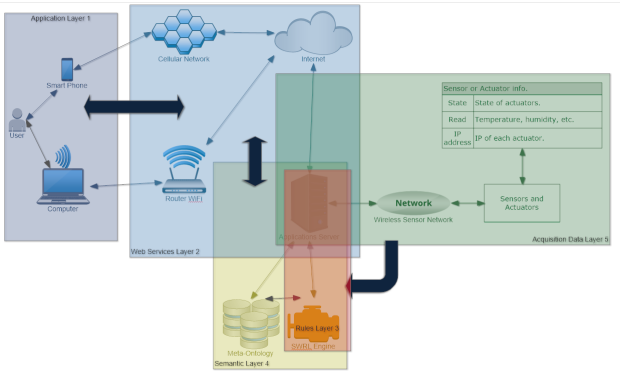


Figure 1: Architecture of the prototype system

The second layer belongs to web services and here is where the applications server maintains the services online, and they can be invoked from any location with a Internet connection, in this layer the Apache Tomcat server runs.

In the third layer is executed the rules engine SWRL this engine can recognize inconsistencies in the ontology, even has the capacity to create new knowledge once a rule or set of rules are executed, for this prototype only are needed the queries to get information and realize some actions.

For the penultimate layer is needed a section a little bit

bigger, which It will be taken up again later, as a introduction to this layer is important to say that, here is where the meta-ontology get the associations with others ontologies even the individuals inside the ontologies are modeled based on the devices that exists in the environment, by example, sensors, actuators, etc.

Last but not least the layer of Acquisition data, in this layer can be found the wireless sensor network and the daemon program that scans a set of IP address looking for "individuals" that exists in the meta-ontology, a individual is a instance created inside the ontology to model something of interest, for our prototype some examples can be found in the next images.

As a last point we present figure 2 here can be observed the

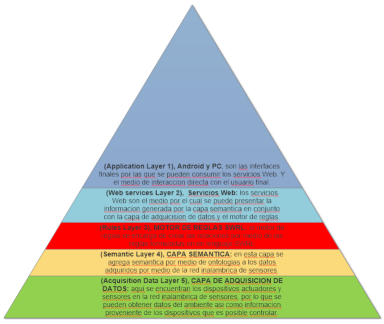


Figure 2: pyramid of the proposed architecture

organization of the layers and some characteristics, also it's going to be important explain a little bit deeper the base of knowledge for the layers two, three and four and they will resumed below

3.1 Semantic Layer

An ontology has been defined for this proposal, this is composed of three sub-ontologies. The first of them is used to store information about physical spaces, as buildings, offices, etc. The information stored about these spaces are: responsible of space, devices, type and number of sensors and actuators, location. The second sub-ontology is responsible of store information about devices, such as: mobile devices, sensors, actuators, displays, air conditioned, etc. The third sub-ontology is used for storing information about people such as: name, roll, position, id, etc.

3.2 Wireless sensor network

A sensor network has been implemented using Arduino. Arduino allows the prototyping of devices with Wi-Fi connection at low cost. The network uses Arduino MEGA 2560, this model has been selected because his memory of 256KB, his low energy consumption. Different sensors have been used in this network, mainly for temperature and humidity, light intensity and presence. The actuators are devices that can be activated through a power unit [9].

3.3 Web Services

Once the data have been gathered and processed into the ontology, the information can be presented to the user through web services. Web services have been selected because they ease the consumption of the information through heterogeneous platforms. The web services were implemented using Apache CFX.

4. IMPLEMENTATION

This proposal was implemented Java and the OWLAPI framework for the *.owl* files of the ontology. Once the ontology and the sub-ontologies are loaded the semantic relations can be inferred for example, what devices are in the office of professor X, or ask if professor Y is in the hall B. It is possible to activate devices, for example the air conditioned, by using web services for asking to the ontology about the state of the device and sending the command to the actuator to turn on the air conditioned.

The information gathered from sensors through the network is stored directly into the ontology by use a daemon program that scans a set of IP address and get the information from the different devices in the WSN, the daemon program run all time in the server as a background process and scans the WSN every ninety seconds. After the information for individuals has been supplied directly to the meta-ontology the SWRL engine is executed in order to get the information that is sent to the user through the web services, using the web services it is possible to activate the actuators in the network, through a daemon that reads the information in the ontology and compares against the preferences of users using a second web service.

As mentioned above, the sensors/actuators network has been implemented using Arduino Mega 256. Everyone of them has been added with different characteristics, in order to make them capable of get different information. An example of these extended characteristics is the capability of identify users through RFID using the exclusive ID of the user. The RFID is used too as a presence sensor, given that it is possible to use it as a counter of users in the same space, see Fig.3. Fig.4 shows the response that returns from the

```
<html><head></head>
<body>
<a id="H254"> H254 </a>
<a id="hasNumberOfPeople"> 3 </a>
<a class="Sensor" id="RFIDReader"> RFID</a>
<a name="RFIDReader" id="idCardRead">
    00000000 </a>
<a name="RFIDReader" id="idNumberRead"></a>
<a class="Cubicle" id="H254"> H254 </a>
<a name="H254" id="hasNumberOfPeople"> 0 </
    a>
</body></html>
```

Figure 3: Information of users given by RFID

lightweight web server running in the arduino, this kind of response is used to feed the otologies, the temperature, humidity and the intensity of light, the set of the actuators, and the state of devices (on/off), and the IP address of the device being controlled by the arduino. through instructions sent from the daemon program that compares the data from the ontologies with the user preferences set from the fourth web service. As it can be seen the information about devices, sensors and actuators is gathered using HTML, this allows that the ontology could be fed directly from the in-

```
<html><head></head>
<body>
<a class="Sensor" id="DHT11A">DHT11A</a>
<a name="DHT11A" id="hasTempValueA"> 29.00
    </a>
<a name="DHT11A" id="hasHumidityValueA">
    38.00 </a>
<a class="Sensor" id="DHT11B">DHT11B</a>
<a name="DHT11B" id="hasTempValueB"> 24.00
    </a>
<a name="DHT11B" id="hasHumidityValueB">
    30.00 </a>
<a class="Sensor" id="LightA">LightA</a>
<a name="LightA" id="hasLightValueA">
    125.00 </a>
<a class="Sensor" id="LightB">LightB</a>
<a name="LightB" id="hasLightValueB">
    117.00 </a>
<a class="Actuator" id="Lum1">Lum1</a>
<a name="Lum1" id="hasStateLum1"> 0 </a>
<a name="Lum1" id="hasIPAddressLum1">
    192.168.2.20 </a>
<a class="Actuator" id="Fan1">Fan1</a>
<a name="Fan1" id="hasStateFan1"> 0 </a>
<a name="Fan1" id="hasIPAddressFan1">
    192.168.2.20 </a>
</body></html>
```

Figure 4: Getting information from sensors and actuators

formation of the web services, few additional processing is required. The arduino responses showed in fig 3 and fig 4 are the response of the light-weight Web server that is running in each arduino, the web service is very simple and can be used as a HTTP server, to get advantage of HTTP characteristics we used the tag <a> that allow define attributes, events and destination or origin of a link.

The use of tag <a> is to define attributes and get a easier identification of the different devices or sensors that can be read or controlled, by example *DHT11A* here the tag <a> allows to create a attribute of type <class> and this is associated to the class "Sensor" in the meta-ontology, then the *id* is the individual in the class Sensor, for * 29.00 * "name" is used to identify a individual and *id* is the "dataproperty" in the ontology that stores the value that the sensor read from the environment, in case of a sensor or actuator has more values or properties we can define more different attributes by use tag <a>, by the way the characteristics of tag <a> get more easier put the data gathered form the WSN to the meta-onotlogy.

An Apache server was mounted and the implemented services can be queried in a directly way in the ontology. Fig.5 shows an example of a WSDL file with the services using Web Services Description Language. At time to invoke a service, a SWRL rule is executed to make inferences on the ontology and get the information required, as example Fig. 6 shows the rule to get information about actuators in a specific place. This query returns the number of people, the IP addresses of devices and their state and the name of the specific place.

The average temperature, humidity and light intensity are obtained too. The average is calculated in case of existence several sensors for the same parameter in the same place. Every service has a similar structure to their purpose and the information is obtained when executing their associated rule. Another service uses the name of a person as parameter and returns information about the environment where that person is, Fig.7 shows the results for the rule associated to this service. This information is supplied by the ontology and the rules engine, as it can be seen the information is presented with their type and labels, they are presented in text because this information is going to be used for an additional application in the future. The third service takes

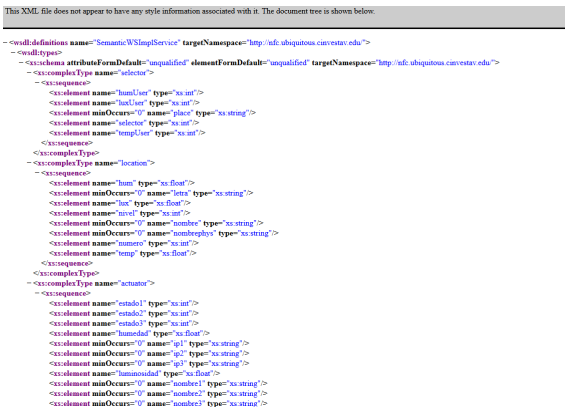


Figure 5: Example of WS that presents information about a person and his environment

the name of a person as parameter and as result it returns the information about the environment of the person and information about the actuators that can be controlled. This information is used by a daemon program that compares the information from the service against the preferences in the profile of the user and keeps the conditions required for the user in that environment.

At this point only works with the preferences of only one user, in the next update to the prototype is expected add the characteristic to manage more users and their preferences. The last service takes as parameter the operation mode of the daemon program, means operation mode of the daemon program and it's selected by the user where "0 = Manual" and "1 = Automatic",also returns the temperature, the humidity and the intensity of light of the location. These two last services are used for specifying their preferences without needing to modify code in the application, this configuration can be made too from any devices connected in an external network, by example Internet.

5. CONCLUSIONS

This paper shows that is possible to build semantic systems capable of gathering the information from the environment, and present that information in a friendly form to make it useful for human users or computers. The information supplied for the meta-ontology is intended to be useful for the user and in this way keep the path about the requests of users. The aim of this proposal is to minimize

```
SQWRLResult Q3 = queryEngine.runSQWRLQuery(
    "Q3", "ontologydevices:Actuator(?a)^
    ontologydevices:itIsLocatedIn(?a,?loc)^
    ontologyphysicalspace:hasNumberOfPeople
    (?loc,?numpeople)^ontologydevices:
    hasIPAddress(?a,?ip)^ontologydevices:
    hasState(?a,?state)^
    ontologyphysicalspace:hasName(?loc,?
    name)^ontologydevices:hasNameDevice(?a
    ,?namedev)^ontologyinstitution:hasTemp
    (?loc,?temp)^ontologyinstitution:hasHum
    (?loc,?hum)␣^␣ontologyinstitution:
    hasLight(?loc,?lux)->sqwrl:select(?
    numpeople,?ip,?state,?name,?namedev)^
    sqwrl:avg(?temp)^sqwrl:avg(?hum)^sqwrl:
    avg(?lux)");
```

Figure 6: Getting information about actuators in a specific place.

```
<soapenv:Envelope xmlns:soapenv="http://
    schemas.xmlsoap.org/soap/envelope/"
    xmlns:nfc="http://nfc.ubiquitous.
    cinvestav.edu/">
<soapenv:Header/>
<soapenv:Body>
<nfc:name>Jose Abraham</nfc:name>
</soapenv:Body>
</soapenv:Envelope>
----------------Response----------------
<soap:Envelope xmlns:soap="http://schemas.
    xmlsoap.org/soap/envelope/">
<soap:Body>
<ns2:location xmlns:ns2="http://nfc.
    ubiquitous.cinvestav.edu/">
<hum>25.0</hum>
<letra>H</letra>
<lux>190.0</lux>
<nivel>2</nivel>
<nombre>Jose Abraham</nombre>
<nombrephys>Sistemas de informacion
    inteligentes</nombrephys>
<numero>54</numero>
<temp>20.0</temp>
</ns2:location>
</soap:Body>
</soap:Envelope>
```

Figure 7: Information about a person and his environment

the distraction of users and at the same time supply useful and easy to understand information for user and computers. The use of web services ease the exploitation of the information in the ontology. And make easier to query about the environmental information of a place or to find information about somebody. Work is to be continued mainly about the use of time and courses schedule.

As some contributions that can be mentioned are: the use of lightweight web servers running on embedded devices to get data from almost any sensor and even create actuator devices that can be controlled through these small server. As another contribution is the architecture proposed as a system made-up using different technologies even adding semantics and the option to grow to long term. Also the use of the Web technologies helped to connect through different infrastructures for achieve a ubiquitous architecture.

6. REFERENCES

[1] M. Beigl, H.-W. Gellersen, and A. Schmidt. Mediacups: experience with design and use of computer-augmented everyday artefacts. *Computer Networks*, 35(4):401–409, 2001.

[2] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):34–43, 2001.

[3] G. Borriello and R. Want. Embedded computation meets the World Wide Web. *Communications of the ACM*, 43(5):59–66, 2000.

[4] R. Cáceres, A. Friday, R. Caceres, and A. Friday. Ubicomp Systems at 20: Progress, Opportunities, and Challenges. *IEEE Pervasive Computing*, 11(1):14–21, 2012.

[5] M. Jung, J. Weidinger, W. Kastner, and A. Olivieri. Building Automation and Smart Cities: An Integration Approach Based on a Service-Oriented Architecture. *2013 27th International Conference on Advanced Information Networking and Applications Workshops*, pages 1361–1367, 2013.

[6] T. Kindberg and A. Fox. System Software for Ubiquitous Computing. *Pervasive Computing*, 1(1):70–81, 2002.

[7] L. Nachabe, M. Girod-Genet, and B. El Hassan. Unified Data Model for Wireless Sensor Network. *IEEE Sensors Journal*, 15(7):3657–3667, 2015.

[8] K. Park, Y. Kim, and J. Chang. Semantic reasoning with contextual ontologies on sensor cloud environment. *International Journal of Distributed Sensor Networks*, 2014, 2014.

[9] M. Rashid H. Electrónica de Potencia. Circuitos, Dispositivos y Aplicaciones. *Electrónica de Potencia. Circuitos, Dispositivos y Aplicaciones*, pages 1–878, 2004.

[10] Z. Sheng, H. Wang, C. Yin, X. Hu, S. Yang, and V. C. M. Leung. Lightweight Management of Resource-Constrained Sensor Devices in Internet of Things. *IEEE Internet of Things Journal*, 2(5):402–411, 2015.

[11] H. H. Wang, N. Gibbins, T. Payne, and A. Patelli. A survey of semantic web services formalisms. In *Proceedings - 2013 9th International Conference on Semantics, Knowledge and Grids, SKG 2013*, pages 135–142, 2013.

[12] M. Weisser. The Computer for the 21st Century. *Sci Am*, 265(3):94–104, 1988.

An ABC-based Approach for Web Services Clustering

Maricela Bravo
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
mcbc@azc.uam.mx

Roman Mora
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
rmg@azc.uam.mx

Roberto Alcántara
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
raar@azc.uam.mx

Iván Flores
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.

ABSTRACT
In this paper, we describe a hybrid approach for Web services clustering based on the bio-inspired algorithm of Artificial Bee Colony (ABC), K-means and a consensus method. Clustering starts by calculating eight semantic similarity measures between all Web services to produce similarity matrixes that are used as input to the hybrid algorithm that after execution returns clusters of services. Evaluation of current approach is done by executing a series of experiments with three collections of Web services. Experiments were statistically normalized and evaluated.

CCS Concepts
CCS→ Computing methodologies→ Machine learning→ Machine learning approaches→ Bio-inspired approaches→ Artificial life

Keywords
Web Services Clustering; Artificial Bee Colony Optimization; Bio-inspired algorithm.

1. INTRODUCTION
Web services clustering is the task of identifying the most important characteristics from the descriptions of Web services and determine to which class or group they belong. The implementation of mechanisms for automatic clustering of Web services is important because it facilitates the automatic organization of large collections of Web services according to any criterion selected. Getting large collections of organized Web services is a permanent objective of public Web services directories, only through a well-organized directory the automatic search, selection, substitution and invocation of any Web service will be possible.

In recent years there has been an increasing interest on bio-inspired algorithms to solve complex computing problems such as clustering. Some examples of these are genetic algorithms, evolutionary algorithms and algorithms that imitate the behavior of some animal species, etc. Living beings like animals and plants exhibit extraordinary, complex and fascinating behaviors trying to find the best solution to solve their problems and maintain the perfect balance between its components and their environment.

This is the main idea of computing inspired by nature, that is, to imitate the steps that nature has developed and adapt them to solve a problem, making a bio-inspired algorithm [1]. There are many bio-inspired algorithms that mimic nature to solve optimization problems [2].

The use of conventional methods for solving optimization problems requires enormous efforts and computational resources; they tend to fail as the problem size increases [3]. This is the main reason for the implementation of an algorithm that exploits the collective intelligence; this is a characteristic that the algorithm Artificial Bee Colony (ABC) poses and represents a good alternative for Web services clustering.

This paper describes an innovative approach for the clustering of Web services using a hybrid algorithm based on the ABC algorithm, K-means and consensus. For evaluation of the proposed method a collection of 1027 public Web services were used and the stages of processing, classification, testing and analysis were performed. The algorithm receives as input a collection of eight semantic similarity matrices, which were calculated using data extracted from Web service descriptions.

2. GENERAL ARQUITECTURE
The work reported in this paper is part of a research project of greater scope whose objective is to design and implement a directory of Web services using efficient methods, fully automated, and allowing the organization, composition and classification of Web services with a semantic approach. Figure 1 show the overall architecture of the semantic directory, which is composed of modules each with a specific purpose, in particular, this work contributed to the module of classification and grouping of Web services.

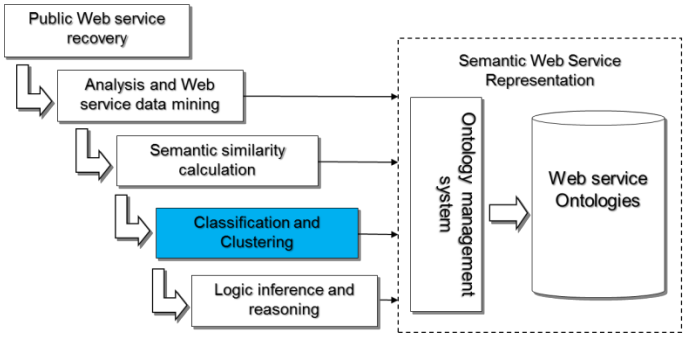


Figure 1. Semantic Web service directory architecture.

a) Public Web services recovery. The descriptions of Web services are searched and retrieved from the network by using autonomous programs that seek to recover resources available on the Web (also known as Web crawlers or spiders).

b) Analysis and Web service data mining. Recovered services are processed through programs that analyze the file and extract the relevant information services. The specific service information is stored in ontologies that represent Web services.

c) Semantic similarities. The extracted data is used for calculating the semantic similarity using eight similarity measures.

d) Classification and clustering. The classification and clustering of Web services is performed using different algorithms, for example, support vector machines, neural networks, ant colony optimization, among others. In particular in this paper a hybrid method for Web services clustering is reported derived from the algorithm Artificial Bee Colony (ABC).

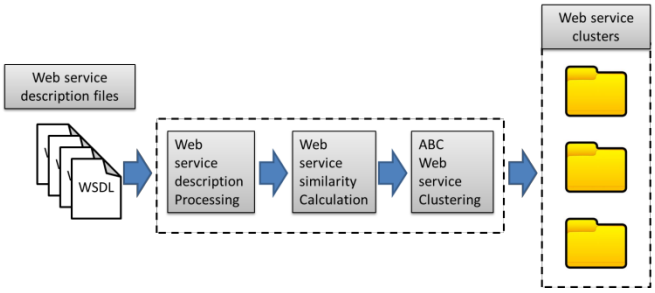


Figure 2. Architecture of the clustering process.

Figure 2 shows the general structure of the proposed solution, which consists of the following elements:

a. Collection of local Web services. It consists of a large collection of Web services, which are used as input to the clustering process, Web services of this collection were previously retrieved and stored locally by Web spiders.

b. Web services processing module. This module performs the task of processing all the information extracted from web services descriptions for further grouping.

c. Module for Web services clustering using the hybrid algorithm. This module determines how many groups should be generated, determines the centers of the groups, and groups the services according to the centers generated.

d. Repository of Web services grouped. This module contains all the web services organized in groups, the aim of this module is to obtain an optimal clustering scheme.

3. ABC ALGORITHM DESCRIPTION

The main idea of the Artificial Bee Colony (ABC) algorithm is to mimic the behavior of bees by searching and collecting food performing very specific tasks such as:

a. Exploration. Task carried out by unemployed bees to search and locate new sources of food.

b. Exploitation. Task carried out by the worker bees in a food source.

c. Recruitment. Task carried out by unemployed bees with worker bees to exploit a food source.

d. Abandonment. It happens when a better source of food is found.

The ABC algorithm is composed of two groups of bees: worker bees and bee waiting. Waiting bees are further subdivided into: scout bees and spectators bees, these two kinds of bees (worker and spectators) have the characteristic to choose and adjust by experience (memory) the food source. In the algorithm ABC is assumed that each position of the food source is equivalent to a feasible solution and the amount of food that is collected from the food source represents the quality associated with the solution [2]. Figure 3 summarizes the algorithm ABC.

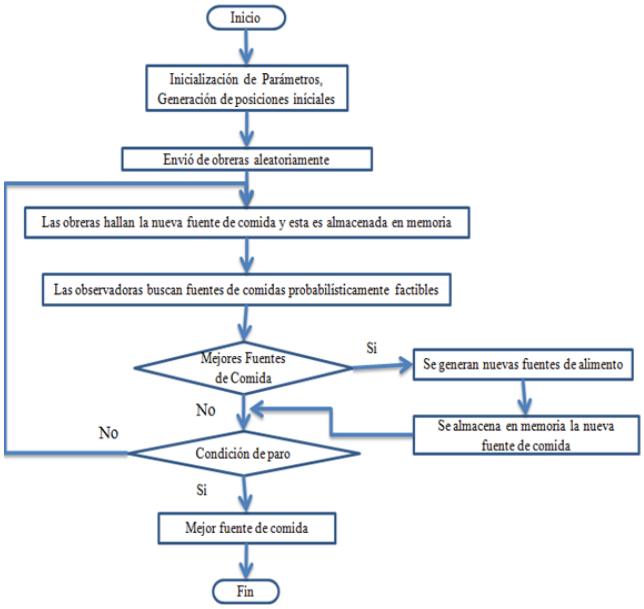
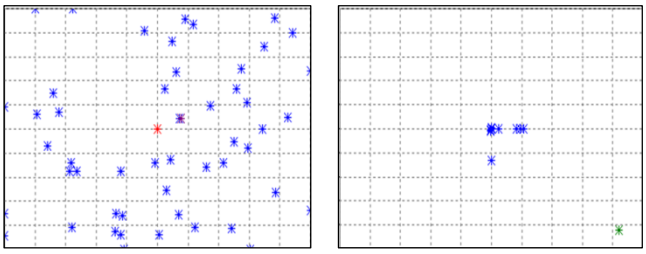


Figure 3. ABC algorithm.

Figure 4 shows the behavior of the ABC algorithm, where the blue color represents the worker bees, the green represents the scout bees, the purple color represents the best solution obtained, the red color represents the desired optimum. The graph on the left side (a) shows how the ABC algorithm begins positioning randomly workers; the right side (b) shows the result after a series of executions that the feasible solution is reached.



(a) **(b)**

Figure 4. Behavior of the ABC algorithm.

The combination of other clustering methods such as the consensus method allows the bio-inspired algorithm the creation of groups that are generated from two groups (minimum value of allowed groups) to the number of Web services that are used from 2 (maximum value of groups that can be generated). For the case of K-means method, the algorithm selects elements within groups generated before turning on centers which can measure the similarity; also the hybrid algorithm uses heuristic methods to generate random numbers that the algorithm uses for calculation in generating groups.

The solution is presented as an input vector of size *n* (number of Web services to group) where the element in the vector, return the

group to which it belongs. The objective function of the hybrid algorithm is:

$$\text{Min} \sum_{i=1}^{C} \sum_{\substack{x \in ci \\ y \in ci}} d(x_i, y_i)$$

Where

d: Distance

y_i: Center of the *i* group

x_i: One of the services included in the *i* group. No group may be empty and intersection between groups cannot exist.

The first step performed by the hybrid algorithm is the filtration of the 8 matrices containing information similarities of Web services using equation (1), discarding values exceeding the permitted limits and established by the same metrics, new matrices are generated filtered with 95% accuracy in measurements.

$$\mathcal{X} - 1.96 \frac{\sigma}{\sqrt{N}} \leq \mu \leq \mathcal{X} + 1.96 \frac{\sigma}{\sqrt{N}} \qquad (1)$$

Where

$\mathcal{X}$: Average matrix

1.96: Constant

σ: Standard deviation

N: An element of the similarity matrix

μ: Average similarity

Then filtered data is stored in an average matrix (food sources) discarding positions containing no information or 0. All values that were filtered were averaged and accepted as feasible values that contribute information to the algorithm.

Then the average similarity matrix is read for the first time to generate a series of arrays (bees) which will have information stored, for example the generated groups, selected centers, etc. (see Figure 5).

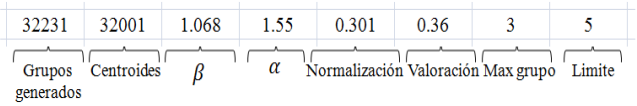


Figure 5. Information of the vector of generated groups.

Figure 5 shows the structure of a solution generated by the proposed hybrid algorithm. The information registered in the vector consists of:

Max group. The proposed algorithm does not require the user to indicate the number of groups to generate, as the algorithm iterates it determines the number of groups to be generated based on the results obtained by consensus.

Initially, the i-th bee generates a random number of clusters to generate, based on a discrete uniform distribution with limits 2 to N / 2; in subsequent iterations of the algorithm, the i-th bee produces a random number γ (based on a normal distribution of the weighted average variance of Max group determined by the colony in the previous iteration), then a simple rounding be applied to γ.

Then, the *i-th* bee must determine the centroid of the γ groups involved in clustering. The sub-vector centroid is formed by *N* integer elements; where *xijk* is zero if it is not considered a centroid for any group, if that *xijk* = a implies that the *j-th* Web service is centroid the *k-th* group. Initially these values are assigned randomly, and then using equation (2) sub-vector values are obtained.

$$X'_{new} = \text{round}\,(X'_i - \phi(X'_i - X'_s)) \qquad (2)$$

Where

X'_{new}: New generated vector

X'_i: First vector generated by the algorithm

ϕ: Random number between 0 and 1.

It is important to note, that vector X'_{new} is repaired with the objective that only *k* elements are different from zero and all groups have a different centroid.

The sub-vector of generated groups is constructed based on the sub-vector centroid; zero elements are replaced by the *k* value which is the closest centroid to the Web service.

β is the sum of the similarity between centroids of the groups while α is the sum of the similarity between members of each group to the corresponding centroid. The objective of the algorithm is to minimize β and maximize α simultaneously.

For each of the bees the value of *f (x)* = 1/β + α, and is intended that this value of *f (x)* as large as possible. The value of *f (x)* is stored in the vector solution.

Normalization implies that $n_i = \frac{f(x_i)}{\sum_{i=1}^{N} f(x_i)}$ element will determine the quality of the food source found. A bee will abandon its current food source if another food source explored in the neighborhood is better.

Limit is a counter indicating the number of iterations that the bee has been in the current food source, when it comes down to a value *g*, and as long as the bee is not in the best food source found, the *i- th* bee is forced to abandon its current food source.

4. EXPERIMENTATION

For evaluation purposes, three collections of 50, 647 and 1027 Web services respectively were used for clustering. For the creation of the similarity matrix a set of eight different semantic similarity metrics were used for each of the collections.

The determination of the parameters used by the algorithm performed by brute force, resulting in use 10 and the value of "phi" (φ) was set to 0.8, the limit value that makes each vector from the outset was established with the value of 10.

To characterize the behavior of the algorithm 20 experiments were executed using 100, 200 and 500 iterations of the algorithm. For each of the executions the best value of *f (x)* found by bees was determined. Subsequently, a statistical analysis of the results was performed.

Tables 1, 2 and 3 show the average values of the f(x) for each of the collections and iterations of the algorithm.

Table 1. Average results of the instances with 100 iterations

Matrix size	*f(x)*
50	0.4631
647	0.4457
1027	0.4414

Table 2. Average results of the instances with 200 iterations

Matrix size	*f(x)*
50	0.5011
647	0.4152
1027	0.4782

Table 3. Average results of the instances with 500 iterations

Matrix size	*f(x)*
50	0.5169
647	0.4228
1027	0.4542

Based on the above tables, it can be said that for the instance of 50 services the best values are 500 iterations, while for instance 647 services the best values are obtained with 100 iterations. Finally, for instance 1027 it is obtained with 200 iterations. Tables 4 and 5 show results obtained with the best settings. In Table 4 the best, worst, average, variance and standard deviation of *f (x)* about 20 executions of the algorithm are shown.

Table 4. Results with the best settings obtained

Collection -Iterations	Best	Worst	Average	Variance	Deviation
50-100	0.58936208	0.29274988	0.45901481	0.00593354	0.0770295
50-200	0.54853117	0.44904535	0.50052478	0.00088301	0.02971549
50-500	0.57163535	0.46554965	0.51820976	0.00090081	0.03001343
647-100	0.54947584	0.2153186	0.44401451	0.01027133	0.10134758
647-200	0.54778745	0.20257995	0.40923046	0.01554565	0.12468218
647-500	0.54881541	0.22262872	0.41911288	0.01280278	0.11314939
1027-100	0.55531806	0.27378002	0.44488931	0.00847871	0.09207991
1027-200	0.56846422	0.28919387	0.47529505	0.00772078	0.08786795
1027-500	0.66112025	0.20302145	0.4674717	0.00977301	0.09885856

Table 5. Results with the best settings obtained

Collection-Iterations	Best	Worst	Average	Variance	Deviation
50-100	2	10	4.31578947	6.22105263	2.49420381
50-200	2	4	3	0.89210526	0.94451324
50-500	2	6	3.94736842	2.36578947	1.53811231
647-100	3	49	18.0526316	221.207895	14.8730594
647-200	2	67	18.9473684	305.831579	17.488041
647-500	2	73	23.2105263	426.368421	20.6486905
1027-100	2	110	31.2631579	1066.82895	32.6623475
1027-200	2	69	22.1052632	355.884211	18.8648936
1027-500	2	123	31	1179.21053	34.3396349

Figures 6 and 7 show a comparison of algorithm results for 100, 200 and 500 iterations, respectively; the collection of 1027 service instances was used. It can be seen that the best values of *f (x)* are produced with 500 iterations; however, the result of the number of groups involved is more stable with 200 iterations.

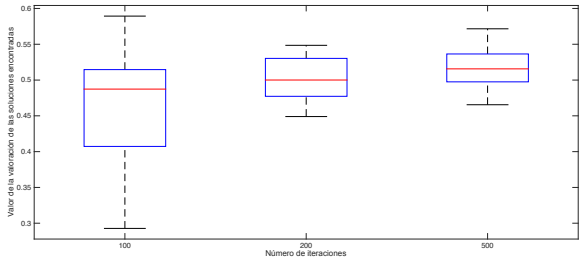


Figure 6. Similarity comparison between the best solutions found with 1027 services.

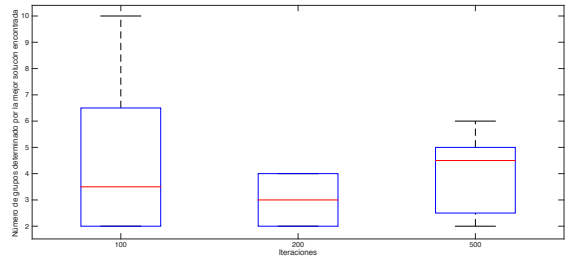


Figure 7. Comparison of the number of groups in the best case for 1027 services.

5. RELATED WORK

Web service clustering and classification has been largely studied from different perspectives ranging from statistical approaches, stochastic and variants. It has also been addressed using innovative approaches in the field of semantic web services using ontologies and inference engines. However, little has been done to address the problem by means of hybrid algorithms, meta-heuristics and biologically inspired algorithms. This section provides a review of the work related to the issue of classification and grouping of Web services as well as some studies reporting the use of biologically inspired algorithms applied in Web services is presented.

Dong in 2004 [4] introduced a clustering approach to search for Web services. This search consisted of two main stages: the user provided key words into a search engine service, the services returned by the search engine, the system of grouping extracted semantic concepts of natural language descriptions in the descriptions of Web services. Using the co-occurrence of terms in the names of the input and output parameters and the names of the operations, calculating similarities was made to employ agglomerative clustering algorithm and classify in terms meaningful concepts. By combining the original keywords and

concepts drawn from descriptions of the services the similarity of two services are performed at the level of concepts. From the standpoint of use of concepts, this approach is similar to that described in this paper, however, the clustering algorithm that was used differs significantly.

In 2007 Arbramowicz et. al [5] proposed an architecture for filtering and grouping of Web services. This approach is considered within the research topic of semantic Web services, since the author used the services described in OWL-S language (Web Ontology Language for Web Services). Filtering services is based on profiles representing users and application information. To improve filtering services, the use of a clustering method is proposed. The purpose of this approach was to support the process of reconciliation between services, saving time and improving execution refinement of stored data.

6. CONCLUSIONS

The meta-heuristics represent a powerful tool to solve optimization problems in different application domains. In this paper an adaptation of the ABC algorithm was presented to propose a novel approach for Web services clustering. The ABC algorithm has shown good results for clustering services, further modifications are possible, for example measuring semantic similarities and the incorporation of more information extracted from the description of services and data types. As a future other bio-inspired algorithms for sorting and grouping of Web services will be implemented.

The text should be in two 8.45 cm (3.33") columns with a .83 cm (.33") gutter.

7. REFERENCES

[1] D. Floreano, C. Mattiussi, Bio-Inspired Artificial Intelligence Theories, Methods and Technologies, The MIT Press Cambridge, Massachusetts London, Englan, 2008.

[2] B. Xing, J. Wen, Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms, Springer, Volume 62, 2014.

[3] S. Binitha, S S. Sathya, A Survey of Bio inspired Optimization Algorithms, International Journal of Soft Computing and Engineering, Volume-2, Issue-2, May 2012.

[4] X. Dong, A. Halevy, J. Madhavan, E. Nemes and J. Zhang. Similarity Search for Web services. In Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004.

[5] W. Abramowicz, K. Haniewicz, M. Kaczmarek, D. Zyskowski. Architecture for Web services filtering and clustering. In Internet and Web Applications and Services. 2007. 18-18.

Addressing Clarity and Coherence during Ontology Construction

Maricela Bravo
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
52 55 53189532 Ext. 149
mcbc@azc.uam.mx

José A. Reyes-Ortiz
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
52 55 53189532 Ext. 142
jaro@azc.uam.mx

Roberto Alcántara
Universidad Autónoma Metropolitana
Av. San Pablo #180, Col. Reynosa Tamaulipas, México D.F.
52 55 53189040
raar@azc.uam.mx

José Rodríguez
CINVESTAV-IPN
Av. Instituto Politécnico Nacional No. 2508, Colonia San Pedro Zacatenco, México D.F.
52 55 5747 3800 Ext. 6563
rodriguez@cinvestav.m

ABSTRACT
This paper describes a methodology for ontology construction from scratch that addresses the design principles of clarity and coherence. Axioms are defined in the entire ontology as facts asserted by the ontology engineer. Many methodologies and tools have been reported in literature, but little attention has been paid to the implementation of specific methods and techniques to achieve clarity and coherence in the resulting ontology. In this paper the design and construction of ontologies is primarily focused on the definition of axioms that specify the membership of individuals to concepts avoiding misinterpretations. The objective is the creation of consistent ontologies as an objective from the beginning of the design process.

CCS Concepts
Computing methodologies → Artificial intelligence → Knowledge representation and reasoning → Ontology engineering

Keywords
Ontology design; design principles; ontology coherence.

1. INTRODUCTION
Ontology design is the process of selecting and applying methods, techniques and principles with the objective of producing an ontology or ontology system. Ontology design encompasses different methodologies, techniques and design principles, which have been proposed to support the efficient design of ontologies. A common concern in achieving efficiency during ontology design is to build ontologies that are usable, reusable and maintainable. A key factor to achieve efficiency during ontology design is the principle of coherence. However, traditional ontology design approaches have paid little attention on the importance of consistency checking.

There are some reported works addressing ontology coherence. However, their main approach consists of dividing a large pre-existing ontology into consistent ontology modules or extracting a consistent segment or piece of a large ontology. In this paper the design of ontologies is primarily focused on coherence, considering the creation of consistent ontology modules as an objective from the beginning of the design process.

2. RELATED WORK
In this section the work related with the topic of ontology design considering mainly the aspect of design principles is presented. Some of the most popular ontology design methodologies addressing the aspect of clarity, coherence and modularization are described.

Design principles are guidelines that orient the ontology construction to produce reusable and shareable ontologies. According with Gruber [1] design principles, represent objective criteria that guide and evaluate ontology design. Gruber defined the following set of design principles: clarity, coherence, extendibility, minimal encoding bias, and minimal ontological commitment. The approach reported in this paper addresses particularly clarity and coherence design principles.

Clarity design principle. According with [1], ontology should communicate the intended meaning of defined terms. Definitions should be objective. Definitions should be stated in formal axioms, and a complete definition (defined by necessary and sufficient conditions) is preferred over a partial definition (defined by only necessary or sufficient conditions).

Coherence design principle. This principle is also referred as soundness or consistency. Coherence specifies that ontology definitions should be individually sound and should not contradict each other [2].

According with Gómez-Pérez [3], a methodology is composed of methods, techniques, processes and activities. An overview of the most cited and popular ontology design methodologies are described in this sub-section. The analysis of ontology building methodologies focuses specifically on reviewing whether methodologies provide methods and techniques for building ontologies systems (networked ontologies). We also review the issue of modularity of ontologies, seen as a design principle for all phases of the methodologies.

Lenat and Guha [4] described in 1990 the methodology they followed to build the Cyc ontology at the Microelectronics and Computer Technology Corporation. The Cyc ontology represents an important attempt to codify a large amount of common sense knowledge. Authors describe a method composed of three processes to build the ontology: manual extraction and codification of knowledge, computer-human aided extraction and codification of knowledge and full computer aided codification of knowledge. This methodology is not intended to build ontology systems nor the incorporation of modularity as a design principle.

Uschold and King presented a methodology for developing and evaluating ontologies in 1995 [5], later refined by Uschold and Grüninger in 1996 [6]. This methodology included the following stages to build ontologies: identify the purpose, build the ontology - ontology capture, ontology coding and integration of existing ontologies-, evaluation, and documentation. They incorporated a scoping process during the ontology capture stage which includes grouping. Grouping was defined as the task of structuring terms

into work areas corresponding to naturally arising sub-groups. In this sense, grouping is the closest concept to the ontology modularization concept presented in this paper. However, the Uschold and colleagues methodology does not cover integration guidelines for networked ontologies or ontology systems.

In 1995, Grüninger and Fox [7] described the process of building an ontology for the Toronto Virtual Enterprise (TOVE) modeling project. The TOVE ontology was constructed with the objective of representing a common sense enterprise model. The methodology defined the following procedure: Describe a motivating scenario, define the competency questions, define the terminology of the ontology (objects, attributes and relations), and specify the definitions and constraints on the terminology specification, and evaluation of the ontology by completeness theorems. Even though that authors describe a two-ontology example (composed of Activity and Organization ontologies), they do not provide any guideline for ontology modularization nor its application to a networked ontology.

In 1996, Berneras et.al [8] presented a methodology for building ontologies under the project KACTUS. The main objective of authors when proposing this methodology was to reuse existing ontologies and evaluate the feasibility of knowledge reuse in complex systems. Their methodology stated the following general processes: development of the list of requirements (specification of the application), identification of the general terms to generate a preliminary model, refinement and structuring the ontology, and search for pre-existing ontologies to reuse them. The KACTUS methodology main objective is very similar to the approach presented in this paper, which is to promote ontology reuse throughout the ontology design process, but incorporating the design principle of ontology modularization.

The METHONTOLOGY framework was presented for the first time by Gómez-Pérez in 1996 [9]. It was later refined and detailed by Fernández et.al in 1997 [10], 1998 [11], and 1999 [12] to facilitate the construction of ontologies. METHONTOLOGY defines a development process and life cycle which consists of the following activities: specification, conceptualization, formalization, implementation, and maintenance. Of special interest is the conceptualization activity which sets the following tasks: to build the glossary of terms, to build concept taxonomies, to build binary relation diagrams, to build the concept dictionary, to describe in detail each binary relation, to describe in detail each instance attribute, and to describe in detail each constant.

CommonKADS [13] is an important methodology for knowledge-based systems construction, not specifically tailored for the design and construction of ontologies. However, CommonKADS is a methodology which includes the concept of model components for reuse. This methodology recognizes the need for reusing knowledge-model elements or a combination of them, considering that large parts of a given model are not specific and may re-occur in other domains and/or tasks. CommonKADS methodology includes the following steps for knowledge-based systems: knowledge identification, consisting of getting familiarized with information sources, glossary and scenarios, and identifying model components for reuse; knowledge specification, aims at constructing a knowledge specification of the knowledge model by choosing the task template and the initial domain conceptualization; and knowledge refinement, which validates the knowledge model and completes the knowledge bases.

The NEON methodology [14] is based on the use of ontology design patterns (ODP). This methodology establishes the reutilization of ontologies from a public ontology repository and a set of known ontology design patterns to integrate them by means of a re-engineering process. The general steps defined in this methodology are: 1) identify requirements, 2) identify available design patterns, 3) divide and transform the selected problem into partial problems, 4) match selected partial problems with ontology design patterns, 5) select the design pattern, 6) apply selected patterns to make a composition, 7) evaluate partial designs solutions, and 8) integrate partial solutions.

According with Gangemi and Presutti [15] a design pattern provides a modeling solution to a frequent ontology design problem. Authors define Content Ontology Design Patterns as small ontologies that mediate between problem types and design solutions, and are used as modeling components to the extent that a new ontology can be built from the composition of multiple components.

The incorporation of ontology design patterns is a good approach whenever the set of required solutions exist in a public available repository. However, many times ontology design requires the implementation of tailored constructs for complex engineering systems and heterogeneous domains. Even more, ontology reuse requires the adaptation of the preexisting ontology design solution to the specific application needs. The notion of ontology design pattern is close to the concept of modularized ontology design, in the sense of reusability of ontology modules and composition of new ontologies based on a set of initial components. In particular, this work stresses the modularized ontology design as a design principle, taking into consideration that there may not be a frequent ontology design solution for every design problem; however, the resulting ontology should maintain the modularization criteria during its design.

This paper promotes the reutilization of ontologies achieving clarity and coherence during ontology design and implementation ontology from the scratch. In this paper ontologies are seen as reusable modules not as general design patterns. The idea behind this approach is that the owners of resulting ontology modules can reuse them inside the enterprise for more applications, being the designers of the structures of ontologies they will be able to reuse them more easily instead of searching in a given repository for general design patterns.

3. PROPOSED METHODOLOGY

In this section, an ontology design methodology is presented; its main goal is to assist ontology developers by providing methods and techniques to support ontology system construction considering the coherence design principle.

The main characteristics of this methodology are:

a. Ontology design is domain-oriented
b. Promotes reutilization by modularized ontology design.
c. Resulting ontologies are coherent.
d. The methodology is iterative and incremental.
e. Resulting ontologies are evaluated by means of competency questions.

This proposed methodology is defined to work with teams integrated with a group of knowledge domain experts, a group of computer programmers and analysts with experience implementing

ontologies and applications that exploit ontologies; and one or two ontology engineers. Figure 1 show the methodology divided into four main phases.

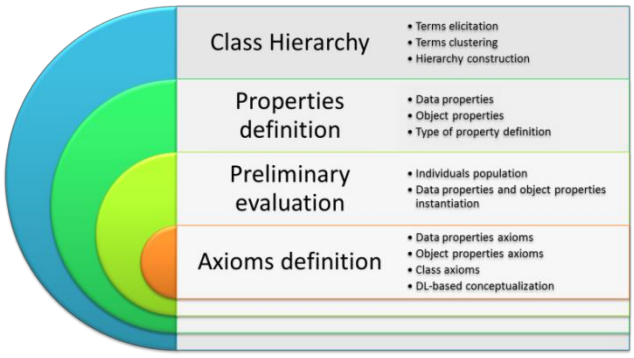


Figure 1. Ontology design methodology.

3.1 Case Study

Considering a traditional academic institution in which professors are teaching subjects to students in classrooms, pre-graduate students are developing their thesis, there is a chief for each department directing and supervising administrative activities; there are academic coordinators attending student academic issues, aided with the support of administrative staff (secretaries, janitors, etc.), and visitors who come for various reasons.

In the following sub-sections, we describe the steps executed to show the feasibility of the proposed methodology.

3.1.1 Class Hierarchy

During this step, the initial set of terms is acquired and the general class structure is defined using is-A relationships. The following activities are performed:

- ✓ **Terms elicitation**. From the motivating scenario, the following concepts arise as an initial list of terms: institution, professor, subject, student, classroom, department, chief, coordinator, staff, and visitor.
- ✓ **Terms clustering**. Using as input the list of terms, the clustering consists of putting together all domain-related terms and assigning a unique class identifier with its name in singular.
- ✓ **Class hierarchy construction**. For each group of terms hierarchical relations should be defined, adding more concepts into the hierarchy if necessary. The class hierarchy of the Person class (Figure 2) shows three additional classes defined to complete the general concepts of employee and two sub classes: Academic and Administrative employees. Verification of the correct definition of hierarchical relations is done by enunciating their derived meanings, for this case the following meanings are possible derivations: a student is a person; a professor is an academic, is an employee and is a person, etc.

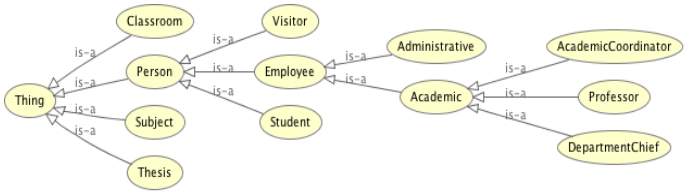


Figure 1. Asserted class hierarchy

3.1.2 Definition of Properties

During this step, data characteristics are defined and inter-classes relationships are defined. The following activities are carried out.

- ✓ **Data properties definition**. Data properties represent attributes or characteristics of each concept. In a hierarchy of concepts, first define characteristics that are common to all concepts. Common characteristics should be defined at the correct superclass level. Then for each concept, it is important to define at least one distinctive characteristic in order to facilitate the reasoner task of classifying individuals. Figure 3 shows the data properties defined for the *AcademicUAM* ontology.

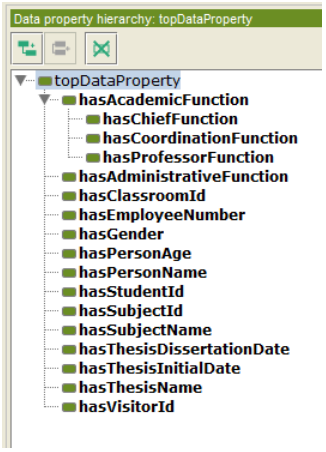


Figure 2. Data properties of the *Person* class hierarchy

- ✓ **Object properties definition**. Object property relations are those relations that link semantically concepts into an ontology. Some of the object properties that were defined for this scenario are shown in Fig. 4. Object property relations are not hierarchical relations; it is a frequent mistake during ontology design the confusion between these types of relations. The ontology engineer together with domain experts should enunciate these relations to identify the type of relationship.

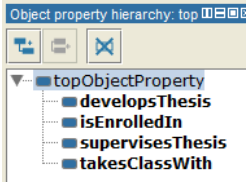


Figure 3. List of object properties defined for the *AcademicUAM* ontology

- ✓ **Type of properties definition**. For each data property and object property it is very important to decide if the relation is: functional, inverse functional, transitive, symmetric, reflexive, etc.

3.1.3 Preliminary Ontology Evaluation

The objective of this step is to evaluate the structure, relationships and usability of the ontology by instantiating few individual examples for each of the classes defined, only individuals for the classes at the end of the hierarchy should be created. The following activities are realized:

- ✓ Individual's instantiation
- ✓ Data properties instantiation
- ✓ Object properties instantiation

As a result of this preliminary evaluation, some data properties or object properties may have been eliminated or redefined, or some new properties may have been added to the definitions. This kind of modification is frequent during ontology population, when the ontology engineer realizes if assertions will be useful.

3.1.4 Definition of Axioms

The purpose of this step is to define the minimal set of characteristics that each individual should fulfill to be a member of a class or concept. These definitions are made through equivalences. The following activities are realized:

- ✓ **Data properties axioms definition**. For each data property, its restrictions are defined: cardinality restrictions, existential and universal restrictions, value restrictions. For instance, Figure 4 shows two restrictions for the *hasPersonName* data property, meaning that a Person should have exactly one name and its range should be some string.

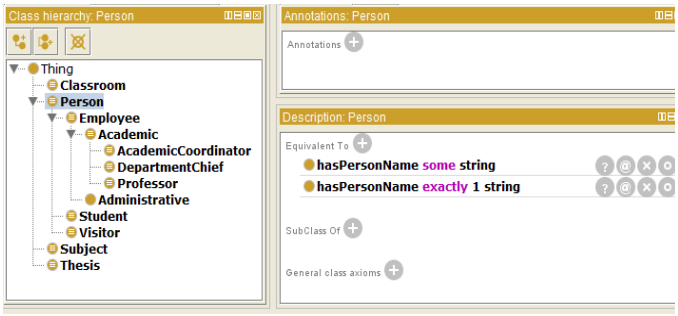


Figure 4. Data properties axiom definition for the class Person.

- ✓ **Object property axioms definition**. For each object property, its restrictions are identified and defined. For instance, to establish that an individual from the class Person has parent only from the class Person the following axiom is used in the definition of Person: Person and (hasParent only Person).
- ✓ **Class axioms definition**. For each concept or class into the ontology, the knowledge engineer should decide if classes are disjoint or not. This is a very good moment to create the core well defined classes by means of equalities. For instance: Which are the necessary and sufficient characteristics that define a concept Person?
- ✓ **DL-based conceptualization**. Ontology Conceptualization is the process by which all classes (concepts) are explicitly defined. In order to clearly define concepts, all common characteristics are normally inherited, but a minimal set of specific and distinguishable characteristics must be identified to unequivocally define each concept.

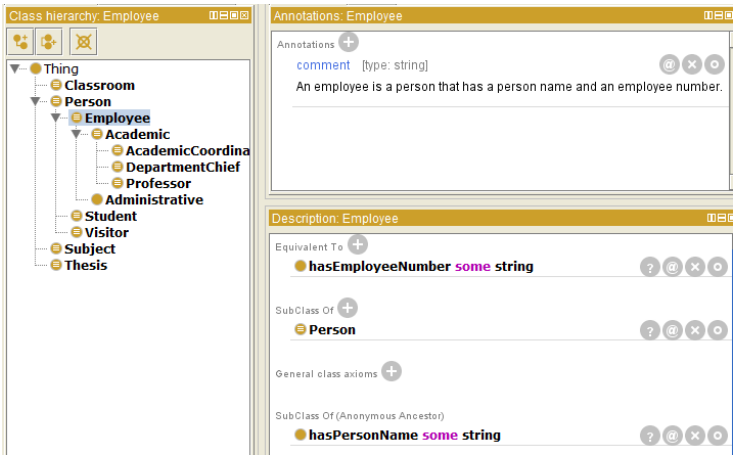


Figure 5. Informal and formal definition of an *Employee*.

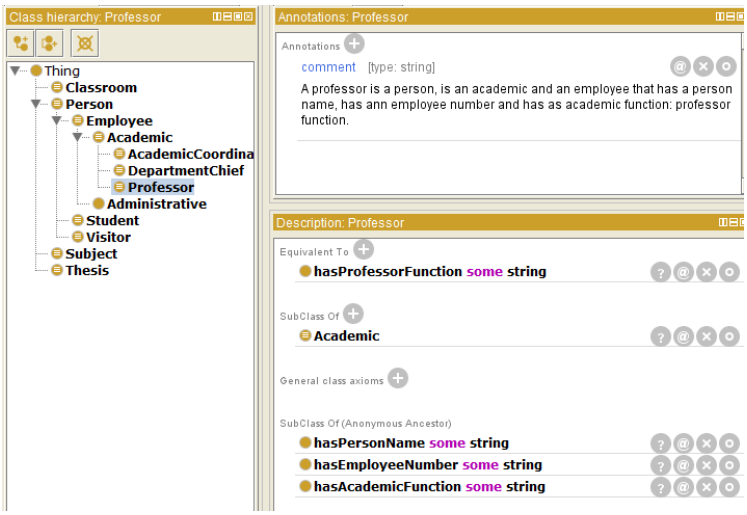


Figure. 6. Informal and formal definition of a Professor.

Table 1 shows the class or concept, the human definition and the implemented ontology definition.

Table 1. Ontology definitions

Class or concept	Human definition	Ontology definition
Student	A student is a person that has a person name, and	Class: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Student>

	has a student id.	**EquivalentTo**: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#hasStudentId> some xsd:string SubClassOf: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Person>
Professor	A professor is a person, is an academic and an employee that has a person name, has an employee number and has as academic function: professor function.	Class: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Professor> Annotations: rdfs:comment "A professor is a person, is an academic and an employee that has a person name, has an employee number and has as academic function: professor function."^^xsd:string **EquivalentTo**: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#hasProfessorFunction> some xsd:string SubClassOf: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Academic>
Employee	An employee is a persona that has a person name and has an employee number.	Class: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Employee> **EquivalentTo**: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#hasEmployeeNumber> some xsd:string SubClassOf: <http://www.semanticweb.org/marielabravo/ontologies/2013/1/AcademicUAM#Person>

4. ONTOLOGY EVALUATION

Ontology system consistency checking is executed to verify that none of the class definitions and axioms has logical contradictions, or the individual's instantiated into the ontology. This final activity consists of executing the reasoning tasks of taxonomy classification, compute inferred types, and consistency checking. As result, the ontology accomplishes the following design principles:

a. *Clarity design principle*. We assured that no incorrect interpretation by defining only equalities in axiomatic class definitions.
b. *Coherence design principle*. After execution of reasoner Fact++, individuals were correctly classified. For instance: professor Ricardo Lopez was classified as member of the Professor class.

The most important design principles were considered and verified through protégé tools such as Fact++ reasoned and DL-query tool.

5. CONCLUSIONS

In this paper a methodology for ontology design was presented, this methodology supports the construction of ontology systems from scratch.

The proposed methodology is holistic because it comprises from the design of classes to the integration an evaluation of the overall ontology, the principal design criteria is the coherence of ontologies from the design assuring the definition of modules with reduced complexity that can be reused separately or in an individual way.

It is based on well-studied concept such as Clarity, Coherence, Modularity and classic methodologies for designing ontologies, the presented example shows that this approach allows the production of consistent ontologies.

6. REFERENCES

[1] Gruber, Thomas R. "Toward principles for the design of ontologies used for knowledge sharing?" International Journal of Human-Computer Studies 43.5 (1995): 907-928.

[2] Morbach, Jan, Andreas Wiesner, and Wolfgang Marquardt. "OntoCAPE—A (re) usable ontology for computer-aided process engineering." Computers & Chemical Engineering 33.10 (2009): 1546-1556..

[3] Gómez-Pérez, A. (1999). Ontological engineering: A state of the art. Expert Update: Knowledge Based Systems and Applied Artificial Intelligence, 2(3), 33-43.

[4] Lenat, D. B., & Guha, R. V. (1989). Building large knowledge-based systems; representation and inference in the Cyc project. Addison-Wesley Longman Publishing Co., Inc..

[5] Uschold, M., & King, M. (1995). Towards a methodology for building ontologies (pp. 15-30). Edinburgh: Artificial Intelligence Applications Institute, University of Edinburgh.

[6] Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. Knowledge engineering review, 11(2), 93-136.

[7] Grüninger, M., & Fox, M. S. (1995). Methodology for the Design and Evaluation of Ontologies.

[8] Bernaras, A., Laresgoiti, I., & Corera, J. (1996). Building and Reusing Ontologies for Electrical Network Applications'. In ECAI (pp. 298-302). PITMAN.

[9] Gómez-Pérez, A., Fernández, M., & Vicente, A. D. (1996). Towards a method to conceptualize domain ontologies.

[10] Fernández, M.; Gómez-Pérez, A.; Juristo, N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Symposium on Ontological Engineering of AAAI. Stanford (California). March 1997.

[11] Gómez Pérez, A. Knowledge Sharing and Reuse. In J. Liebowitz (Editor) Handbook of Expert Systems. CRC. 1998.

[12] M. Fernández-López, A. Gómez-Pérez, A. Pazos-Sierra, J. Pazos-Sierra, Building a chemical ontology using METHONTOLOGY and the ontology design environment,

IEEE Intelligent Systems & their applications 4 (1)(1999) 37–46.

[13] Schreiber, G. (2000). Knowledge engineering and management: the CommonKADS methodology. MIT press.

[14] Suárez-Figueroa, M. C. (2010). NeOn Methodology for building ontology networks: specification, scheduling and reuse (Doctoral dissertation, Informatica).

[15] Gangemi, A., & Presutti, V. (2009). Ontology design patterns. In Handbook on ontologies (pp. 221-243). Springer Berlin Heidelberg.

Semantic Analyzer for Spanish, Using Ontologies

Alma Delia Cuevas Rasgado
Universidad Autónoma del Estado de México (UAEM)
Av.Jardín Zumpango s/n, Tejocote
Texcoco, Edo de México, 56265
almadeliacuevas@gmail.com

Yedid Erandini Niño Membrillo
Universidad Autónoma del Estado de México (UAEM)
Av.Jardín Zumpango s/n, Tejocote
Texcoco, Edo de México, 56265
yeninom@uaemex.mx

Farid García Lamont
Universidad Autónoma del Estado de México (UAEM)
Av.Jardín Zumpango S/N El Tejocote
Texcoco, Edo de México, 56265
fglamont@yahoo.com.mx

ABSTRACT
This paper reports the development and implementation of a semantic analyzer that helps to build successfully an ontology or semantic network a common definition of ontology is an explicit and formal specification of a shared conceptualization that represents a descriptive text, it analyzing texts in Spanish.

The semantic analyzer extracts data from a database consisting of ontology relations, concepts and values stored according to their meaning. The base ontology starts with quantitative data manually fed by a user. The analyzer responds "true" if a sentence is semantically related or "false" if not. During its development we have designed an ontology and we have built a prototype. It has obtained better results than parsers that build a tree disregarding semantics.

CCS Concepts
• Theory of Computation Semantics and reasoning→ Program Semantics →Denotational Semantics.

Keywords
Natural language processing, ontology, knowledge representation, semantic analysis.

1. INTRODUCTION
Analyze a text of a particular topic and convert it to a structure understandable by computers it is not a trivial task. Partly because in the conversion process it should check grammar, semantics and pragmatic, at least. This is one of the challenges of this project.

With the reduction in prices of computing capacity, with access to a large number of documents, and with cheaper storage of large volumes of information, it is possible (and desirable) use a computer to analyze relevant information in documents that are written in natural language, specifically in Spanish. Current analyzers use mainly syntax, i.e., they obey the grammar rules of Spanish language, but they do not take into account it's semantic.

For this reason, our task is to build an analyzer that considers both the grammatical and the semantic information. Other challenge is the automatic extraction of knowledge from those texts. Tools needed for this conversion are taggers, syntactic, morphological and semantic analyzers.

The inherent characteristic of Natural language is its ambiguity, which is solved by people using common sense, but this is not part of computers. For example, "*la puerta es amarilla*" {the door is yellow} (to facilitate reading this article, the authors (not our software) have added English explanations in {brackets}). In this case the door can refer to either a car door or a house's door, but how can the computer know this? This paper shows a semantic analysis tool that allows the computer to solve some of these problems. It is limited to descriptive texts. It has been tested in texts related to transports.

The analyzer identifies information and finds its definition, common uses and perhaps models of transports. In this paper we only describe the semantic analyzer. It starts its operation by analyzing three inputs: concept, relation and concept. For example: "door material plastic" contains the following elements: noun (door), relation (material) and noun (plastic).

To analyze a text, or to find documents, it is important to know the parts of a sentence, to interpret its meaning as close as possible to the reality. A morphological and syntactic parser such as Freeling [1] makes a limited recognition, since it does not use semantic knowledge.

Currently most information searchers use keywords to locate information and return a list of text documents that mention those words and the user will have to find your answer from these documents. However search engines do not solve the part of semantics, unlike our proposal. For example, if we search with Google the sentence "Mowers are also used for", the search engine will return to us 110, 000 results. It contains information such as: 1) what kind of lubricant is use; 2) polypropylene bags; 3) polypropylene bags, best price, quality and service; 4) plastic bag (Wikipedia), etc. If you look at most of the links, only Wikipedia provides an answer like: "The plastic bag is an object used to carry small quantities of goods." But the answer is somewhat incomplete.

The semantic analyzer shown in this paper is used prior to the construction of the dependency tree [1] or prior to searching information on web pages. It searches in an ontology (where concepts are related by meaning) the grammatical elements of speech (noun, verb, preposition, adjectives, etc.) and their semantic relationships, to be able to build the tree or the answer that the user expects about his or her request, using syntactic and semantic information. For example, leaves are relates to trees, rather than to bags or plastic.

Spanish also suffers from ambiguity (words that are morphologically similar but semantically different), e.g. "*Toma el gato que está dentro del carro …* " {Take the cat is in the car} in Spanish "*gato*" {cat} is an animal or "*gato*" is a {mechanical lever to change a tire} because both concepts can be in the car, other example that has ambiguity in Spanish and English: "*Juan es un alto official"* {John is a high officer} may mean that John is tall, or that he is a senior officer. Other works [2], [3] solve this by semantic proximity.

Our solution uses frames [4], which represent better the semantics of a word. A frame is a description of an object, process, action or situation. For example the car is a frame that describes the scenario in which events occurs (travel, driving, rest, etc.) and its usual parts (seat, piston, pneumatic jack, mechanical lever, motor, internal combustion engines, etc.). With frames, ontology is built to keep the semantic relations between two concepts. Some projects (FrameNet, [5]) had using them.

Ontology consists of relationships or links (representing actions, typically verbs, and properties) between nodes (representing objects, typically nouns). Both nodes and links represent concepts. We use the OM notation [6] to represent the ontology, see Table 1, because its relationships can also be nodes. OM is a notation in similar manner to XML, see figure 1. For a more formal definition of ontology, refer to [6].

The creation of frames is intended to guide the analyzer to provide the semantics of words. To obtain the meaning, it has an ontology that considers concepts and their relationships, synonymy, property inheritance and the gloss of the concept. Based on the above, it is necessary that the semantic analyzer check the concepts and relationships between them to get the correct meaning.

```
<concept>vehiculo de ruedas
    <Language>Spanish<word>vehiculo de ruedas, vehiculo automovil</word> <Language>
    <subset> vehiculo</subset>
    <relation>tiene = ruedas </relation>
    <concept>vehiculo Autopropulsado
        <Language>Spanish<word>vehiculo Autopropulsado</word> </Language>
        <subset> vehiculo de ruedas</subset>
        <concept>vehiculo de motor
            <Language>Spanish<word>vehiculo de motor</word> </Language>
            <subset> vehiculo Autopropulsado</subset>
            <concept>carro
                <Language>Spanish<word>carro, Auto, Automovil</word> </Language>
                <subset> vehiculo de motor</subset>
                <relation> cuenta-con = color</relation>
                <relation> cuenta-con = ruedas</relation>
                <concept>Jetta
                    <Language>Spanish<word>crossfox</word> </Language>
                    <subset> carro</subset>
                    <concept>carro de Juan
                        <Language>Spanish<word>carro de Juan</word> </Language>
                        <member> Jetta</member>
                        <part>puerta de carro de Juan</part>
                        <relation>color = amarillo </relation>
                    </concept>
                </concept>
```

Figure 1. Fragment of an ontology in OM Notation.

Figure 1 shows the concept "*vehículo de ruedas"* {vehicle with wheels}. It contains the subset "*vehículo autopropulsado"* {self-propelled vehicle} that has the subset "*vehículo de motor"* {motor vehicle}, which has as subset "carro" {car}. The set "*carro"* {car} has relations "*cuenta-con"* = color {it has = color}, "*cuenta-con = rueda"* {it has = wheel}.

The structure of the ontology is created according to WordNet. WordNet is a semantic database which its synset are connected which each other but the user must to find the relation between the concept, i.e. car is related between tourist or driver. This semantic relation is not shows for WordNet. Our semantic module can does it.

In the OM notation, synonyms appear inside <word> inside <Language>. Thus, "*vehiculo de ruedas"* {vehicle with wheels} y "*vehículo automóvil"* {vehicle automobile} denote the same concept (they are synonyms). That is, the concepts represented in ontology have no language; the concept "Domestic quadruped animal that barks" is the same in English and in Spanish. However, to understand sentences written in Spanish, we must link the frames of ontology with their corresponding Spanish words.

Table 1. Labels of OM Notation used to represent ontologies

SYNTAXIS	MEANING
<concept> concept </concept>	Represents the name of the concept or idea
<language> idiom </language>	Represents the language in which the words are defined
<word> words </word>	Represents the words that describe the concept, i.e. its synonyms.
<subset> subset </subset>	Represents a subset of the concept (when it is a set)
<relation>data = value </relation>	Represents the relationship indicated by data and its value
<member> member </member>	Represents a member of the concept
<part>part </part>	Represents a part of the concept

Semantic analyzer does the following:

1. It gets a sentence as input: "*la olla de color verde*" {the pot of green color} this sentence is sends a tagger which gives to each word a label. The sentence (la (DA0FS0) *olla* (NCFS000) *color* (NCMS000) *verde* (AQ0CS0)) is coverted by adding EAGLES tag to identify each word, e.g. ("*olla"* is a noun N, common C feminine F of singular S).

2. It identifies the triple: "*color*(*olla*, *verde*)" {*color*(*pot*,*green)}*. It searches it in the ontology (details below).

3. Measuring the semantic proximity between elements of the triple, its consistency [7], returning true (phrase is semantically correct) or false. Other examples of semantic proximity in the process of disambiguation of prepositions can be found in [8].

The method fails if there are no frames corresponding to the words that are analyzed. This may seems impractical (to build many frames in order to do something useful), but it is essential. How could the computer (or the reader) know the meaning of "*carro"* {car} if you have not told or seen it before? However, once it is placed in the ontology (using frames), its meaning is clear. See Figure 1.

Another important effort is achieved by placing in the ontology the frames of the relations as closed to the root "*entidad*" {entity} as possible. For example, instead of saying that dogs can drink milk ("*perro"* {dog} "*bebe"* {drink} "*leche"* {milk}), we can ask: Who else can drink? What else can be drank?, and generalize the passive and active agent, placing drink in the concept animal (animals can perform the action of drinking) and using fluid instead of milk (a fluid can be drank). Active agent who start the action, passive agent who receive the action.

In the ontology defining dog as a descendant of animal, and milk as a descendant of liquid, our semantic tool "knows" that the dog can drink milk because it is defined in its frame dog. Thus, if we introduce the phrase "*el perro negro toma leche caliente"* {black dog drinking hot milk}, the semantic relation it will evaluate it to true, because physical objects (an animal is a physical object) can have color, and physical objects can have be hot.

An important objective of our semantic analyzer is to combine it with Freeling (or other parser) to obtain a correct analysis of sentences of text tools, and other descriptive texts. The organization of the paper is as follows: Section 2 presents the supporting theoretical framework. Section 3 contains the algorithm of the analyzer. Section 4 shows the results. The concluding remarks are in Section 5. Finally, references are shown in Section 6.

2. THEORETICAL FRAMEWORK

A central feature of any system that can be considered smart from the human been perspective (for Turing), is the ability to communicate with people in their own language [9]. A key aspect to consider is that most of human knowledge is encoded in natural language, so it is necessary to build programs that use the same skill and naturalness of human beings, for instance using Artificial Intelligence tools [10].

2.1 Natural Language Processing

Natural language processing is used to understand the meaning of a text. It presents at least three problems: 1) lexical ambiguities, 2) structural ambiguity and 3) grammatical relations. It requires a semantic processing that analyzes a statement breaking it down into its structural elements, in order to check their semantic validity, while seeking the correct semantic relations to represent the information contained in the sentence.

2.2 Knowledge Representation

Knowledge representation is an area of artificial intelligence whose main objective is to represent knowledge in a way that facilitates searching it and making deductions. The idea is that the computer receives phrases and sentences across of a text and it can be understood for the computer. It finds their morphological, syntactic and semantic relations, much as persons do [11]. Therefore, the conceptual representation of knowledge is highly relevant because it allows deductions from a concept. There are various techniques to represent knowledge such as frames, rules, semantic networks and ontologies. We prefer to use ontologies containing frames, since they can represent a wide semantic spectrum. Our OM notation (Figure 1) reflects that.

In the OM notation, a concept can have more than one predecessor, allowing both to inherit properties to the concept. For example, Mexico is a country but it is also an emerging market.

3. RELATED WORK

Semantic analysis is concerned with solving various problems in natural language processing. These tasks include the word sense disambiguation (WSD), disambiguation of prepositions, resolving anaphora and extraction of relationships between words [12]. Therefore, the construction of a complete semantic analyzer is very complicated, since it must solve all the problems that are not solved by lexical or syntactic rules, i.e., those that require knowledge of the context.

For the English language, there are many more works, not so for Spanish; Panchenko in [13] shows a method for extracting semantic relationships between words using the k-nearest neighbor's algorithm (KNN) and two measures of semantic similarity, about summaries of Wikipedia. The procedure receives a set of terms like: {alligator, animal, building, house, telephone} and should generate output that shows the possible relationships, which in this case are: {{alligator, animal}, {building, house}}. They do not state the relationship between terms, i.e. between alligator and the animal. Our project goes beyond this, establishing not only a relationship but describing it.

Melcuk in [14] describes an automatic procedure to identify lexical patterns which represent semantic relationships between concepts in an online encyclopedia. Some concepts of the existing entries are extracted from Wikipedia, and searched in WordNet in order to determine its meaning. Subsequently, other words are searched within these entries, to relate them to the concept and in this way, create a broader context. A context pattern is created. Once created, it looks for these patterns in other entries of Wikipedia, in an attempt to generalize. Finally, it is verified whether the generalized patterns contain relations which initially did not exist in WordNet and if so, they are added.

Agrawal and Kakde in [15] describe a method for semantic analysis of natural language queries using domain ontology about "trains," as well as the design of a semantic analyzer. The structure of the ontology contains concepts that describe the set of properties associated with a concept, which when linked to other concepts, define the relationships between them. The ontology has three main concepts: objects (train); mental (concession) and social (department). It also has events that are actions or situations that occur over a period of time and properties used to define concepts of the ontology. This method works with a triplet which is processed to obtain the semantics using an ontology, as we propose in this article, but the difference is that we work to detail the properties of the concept, being able to obtain the semantics of the parts of a concept.

There are few studies that consider semantics as an important resource to structure the information. Most of them work only with syntax. For this reasons we think that our algorithm would improve searches and ontological representations and deductions.

4. METHOD

This proposal provides a method of semantic data analysis provided as a triple "relationship (object, value)" which is sent to the ontology to search the same meaning of the input data. The purpose of an analysis of words to look for them in the knowledge base represented by an ontology is to provide the semantics of the specific information it sent.

The meaning of what is searched in the triple can occur in different ways:

1. It can be inside of the concept, such as "*Color* (*Carro de Juan*, null)" where the "color" relationship is associated with the color yellow (yellow in this case could be the value that is expected to null).

2. To find the triple "*Color*(*Carro de Juan*, null)", the concept that contains: "*Automóvil de Juan"* {John's automobile} would be as a synonym for "*Carro de Juan*".

3. Another way to find semantics is through inheritance of properties where a property is not associated with the concept itself but the general concept (hypernym), for example wheels is associated car but if not specified in this to some predecessor inherited carriage (e.g. land transport).

4. Getting the gloss of the concept. In our example the car's definition is obtained when the triple does not find the relationship and value e.g. "null (*Carro*, null)".

Note that in all cases, you can get the semantics of what you want to always know and when ontology is provided with it, if the ontology concepts and sought relationships are presented.

This section shows with an example how the semantic analyzer works.

The Semantic Analyzer takes a phrase, use the EAGLES tagger to identify the triple: has (car, door), in text: "*carro(cuenta-con, puerta)"* {car has door} and returns "true" if they are semantically related or "false" if they are not. The cases of the analyzer are as follows:

Case 1. Search the first noun, "*carro"* {car}, in the ontology. It is found among the relations of car. In the relations of car searches: "*cuenta-con rueda"* {it has wheel}. If found, it returns "true".

Case 2. Look for the first noun, "*carro"* {car} in the ontology**.** It finds it. Then it searches the relations of car, for the relation "*cuenta-con puerta"* {it has wheel}. It does not exist. Reaching the ancestor concepts of car, it searches in them the relation "*cuenta-con puerta"* {it has door}. If it is found in one of them (for instance, in its immediate predecessor, "*vehículo de motor"* {motor vehicle}), it infers that: if motor vehicle has door and car is a subset of motor vehicle, then "car has door" (by inheritance). Thus, it returns "true".

Case 3. Search for the noun "*carro"* in the ontology. It finds it. See if it has the relation "*cuenta-con puerta"*. It is doesn't, but it finds the relation "*cuenta con portezuela"* {has small door}. It searches the concept "*cuenta-con"* {has} in the ontology, finds it and searches its synonyms "*posee"* {possess}. Search the noun "*portezuela"* {small door} in the ontology. It exists. It browses over its synonyms for the word "*puerta"* {door}. It finds it. This case applies to "*cuenta-con"* {has} with its synonym "*posee"* {possess} and "*portezuela"* {small door} with its synonym "*puerta"* {door}; it returns true. If "*cuenta-con"* did not have synonyms, or if "*portezuela"* did not have them, it would return "false."

Case 4. The concept "*carro"* {car} is not in the ontology, but the concept "*automóvil"* {automobile} is there, and it is a synonym of "*carro"*. Our analyzer sees if "*automóvil"* has the relation "*tiene puerta"* {has door}. It finds that relation; consequently, it returns "true".

Case 5. The relationship "*cuenta-con puerta"* {has door} does not appear in the concept "*carro"* in the ontology. Our analyzer seeks the frame "*cuenta con"* in the ontology. It finds it. What are its active agents, those concepts that can have the relation, "*cuenta-con"*? In this case it is "*artefacto"* {artifact}. Is this the same as "*carro"*? It is not. Is it an ancestor of "*carro"*? Yes, it is. Thus, our analyzer returns "true".

Case 6. The concept "*carro"* is not in the ontology, but its synonym "*automóvil"*. Nevertheless, the concept "*automóvil"* does not have the relation "*cuenta-con"*. Our analyzer cannot find the relation that links "*puerta"* {door} with "carro". It finds that "*automóvil corre velozmente"* {automobile runs fast}. It returns "false", since "automobile runs fast" and "car has door" have not the same meaning.

Case 7. There is not the concept "*carro"* in the ontology, nor any of its synonyms. The frame contains neither "*cuenta con puerta"*, nor synonyms of "*cuenta con"*, nor synonyms of "*puerta"*. It returns "false".

In search for a match, the algorithm, if necessary, searches all predecessors of a concept, from closest to farthest. The semantic analyzer is part of build a semantic tree, but it don´t build the semantic tree, as shown in Figure 2. The semantic analyzer performs the seven steps above, returning true if the sentence is semantically correct (it makes sense). Then ontology tree is constructed using a text-to-ontology procedure no described in this paper.

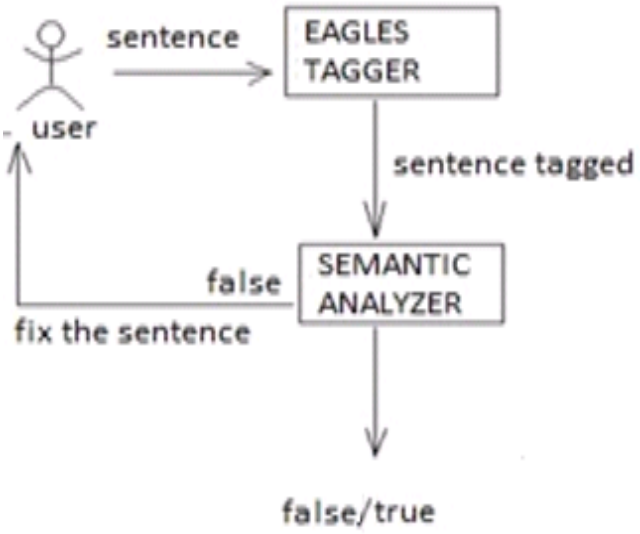


Figure 2. Construction of the semantic tree of a document, showing the role of our semantic analyzer.

5. TEST AND RESULTS

To confirm that the natural language text has semantics with the theme of this work, a set of tests were performed from the input data in the form of triple relation (object, value) for the semantics of the data entered. Note, which in the triple can enter the word null, depending on whether you want to retrieve the relationship, the object, the value or the gloss of the concept in question. The relationship can be a verb or a noun; the object is a noun and the value is a noun. The triple obtained is sent to the ontology to search the same meaning.

The tests were performed using the template R(O,V) where: R is the relation, O is the concept having the relation, and V is the value of the relation. For example, "*bebe(gato, leche*)" {drinks

(cat, milk)}. The relation "*bebe leche"* {drinks milk} is defined in concept "*gato"* {cat}. Our frames are adapted from the corresponding concepts in WordNet 2.1. Thus, for instance, in example 1, the definition for car concept is: "wheeled motor vehicle usually propelled by an internal combustion engine".

Figure 3 shows the triple "*tiene(carro, puerta*)" {has(car, door)}, the semantic analyzer searches ontology the concept car and its relationship, return "true".

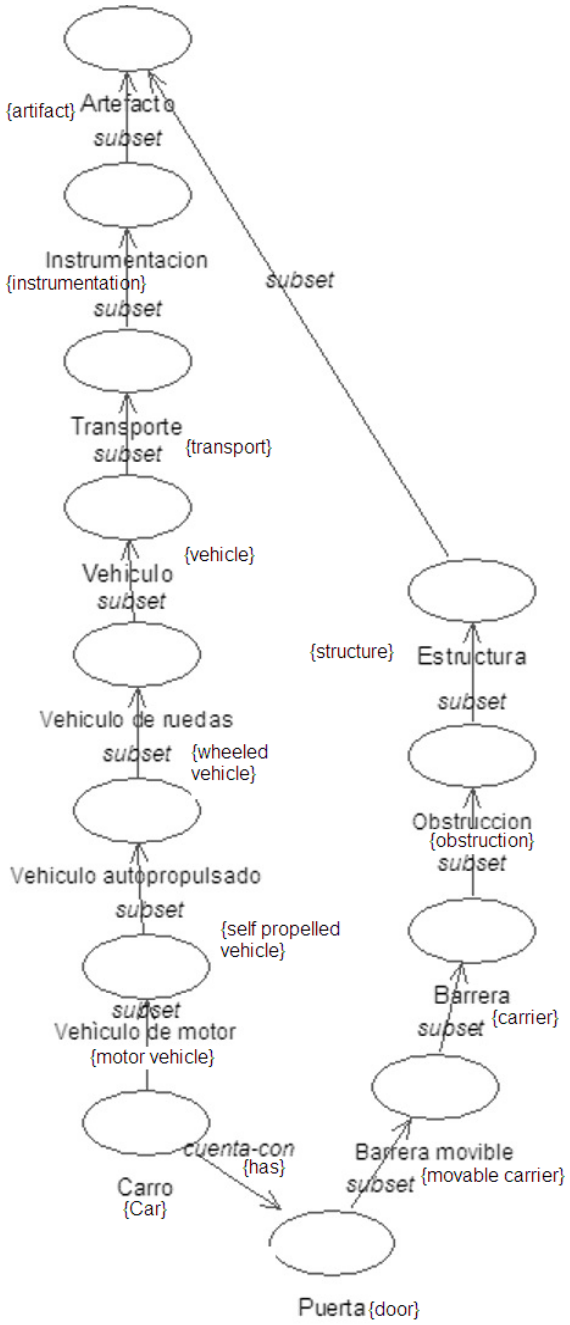


Figure 3. Concept *Carro* {car} and its relation *cuenta-con*{has door}.

Example 2, see Figure 4, the parser analyzes the triple: "*cuenta-con(carro, puerta*)" {has, car, door}. The concept "*carro"* {car} is in the ontology, but "*carro"* {car} does not have the relation "*cuenta-con"* {has}. This relation exists for "*vehículo de motor"* {motor vehicle}, an ancestor of "*carro*" {car}. The analyzer returns "true".

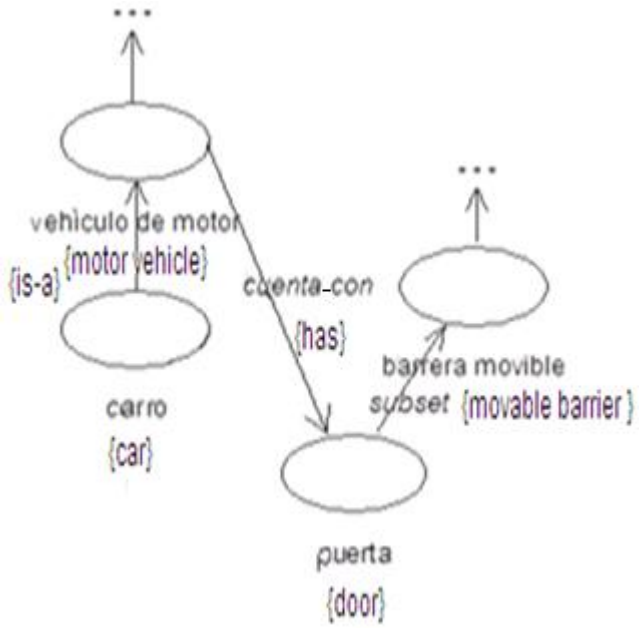


Figure 4. Analyzing the triple "(*cuenta-con(carro, puerta*)" {has(car, door)}, the relation "*cuenta*-con"{has} is found in the predecessor of "carro" {car}, thus car inherits that relation

Example 3. When analyzing the triple "*cuenta-con(carro, puerta*)" {has, car, door}, Figure 5, the concept "*carro"* {car} lacks the relation "*cuenta-con puerta"* {has door} but it has, instead, the relation "*tiene portezuela"* {has small door}. The frame "*tiene"* {has} has as synonym "*cuenta con"* {has}. The frame "*portezuela"* {small door} has as synonym "*puerta"* {door}. The analyzer responds "true".

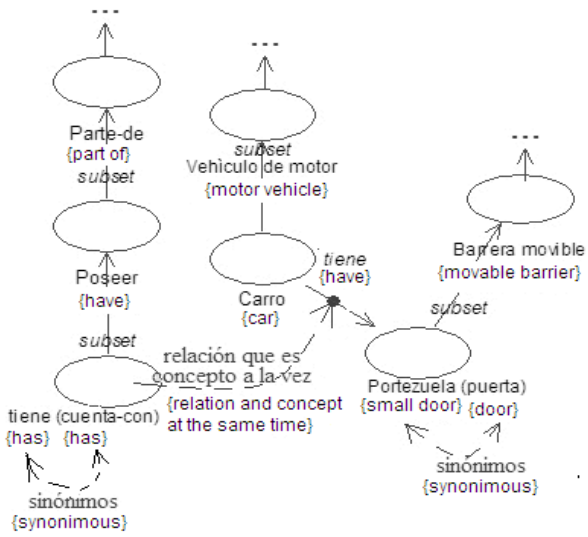


Figure 5. Show that {has, car, door} is the same as {has, car, small door}.

This is so because of the synonym relation between "*cuenta-con"* and "*tiene"*, and between "*puerta"* and "*portezuela" {small door}*. In OM notation, synonyms are defined by adding words between the mark <word>, as explained in Figure 1.

Example 4. The triple to analyze is *cuenta-con(carro, puerta)* {has (car, door)}. Figure 6 shows that "*carro"* {car} does not have the relation "*cuenta-con"* {has}, but "*automóvil"* {automobile} (which does have the relation "*cuenta-con"* {has}) is synonym of "*carro"* {car}. The analyzer returns "true".

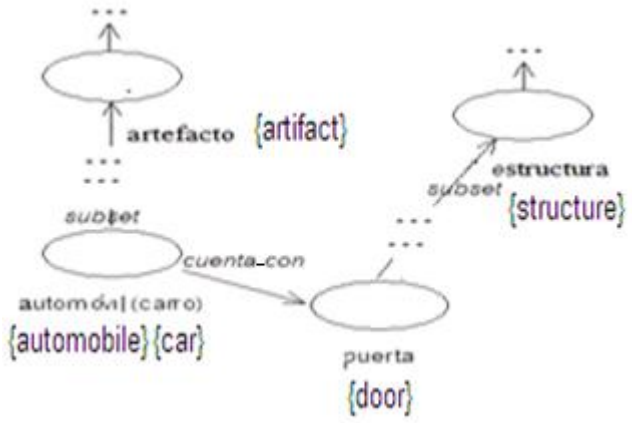


Figure 6. carro {car} is synonymous with automóvil {automobile} (all synonymous are inside of brackets).

Example 5. The triple is "*cuenta-con(carro, puerta)"*. Figure 7. The analyzer does not find the relation "*cuenta-con"* {has} in concept "*carro"* {car}. Neither its predecessor, "*vehículo de motor"* {motor vehicle} has it. Then, it searches that the active agent of relation "*cuenta-con"* {has} is "*artefacto"* {artifact}, and the passive agent (the concept receiving the consequences of the relation) is "*estructura"* {structure}. These are the ancestors of "*carro"* {car} and "*puerta"* {door}, respectively. Thus, there exist a semantic relation, and the analyzer return "true".

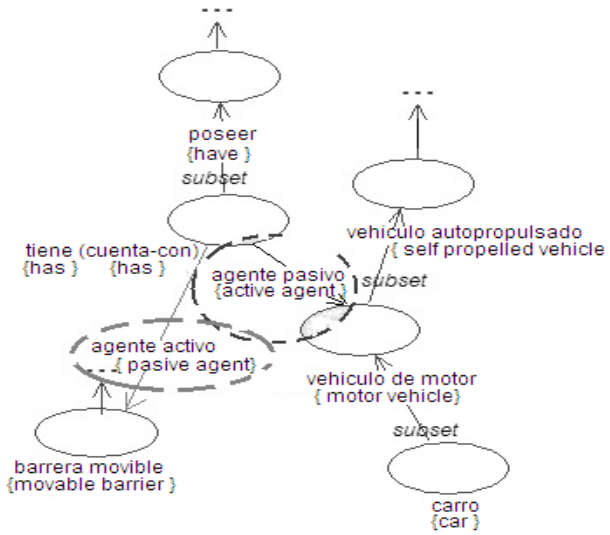


Figure 7. When there is no direct relation between "*cuenta-con puerta*" {has} and concept {car}, the relation is found with the analyzing of the active and passive agents of relation {has}

Example 6. The input is the triple "*cuenta-con(carro, puerta*) {has(car, door)}. "*cuenta-con"* {has} is not in "*carro"* {car}, nor in its ancestors. Neither the ancestors of "*cuenta-con"* {has}. The only relation of car is "*tipo de combustible"*=gasoline. Our analyzer returns "false". See Figure 8.

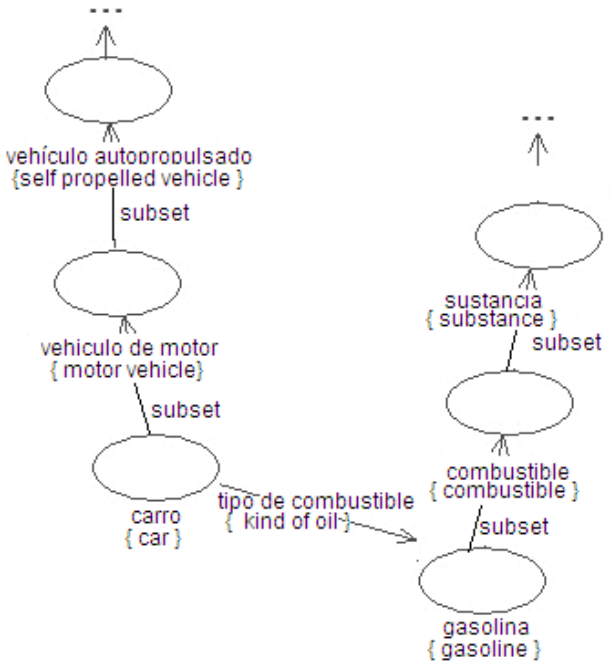


Figure 8. There is no relation between {has door} and {kind of fuel = gasoline}

An example of a fragment of code that uses the semantic analyzer is shown in Figure 9:

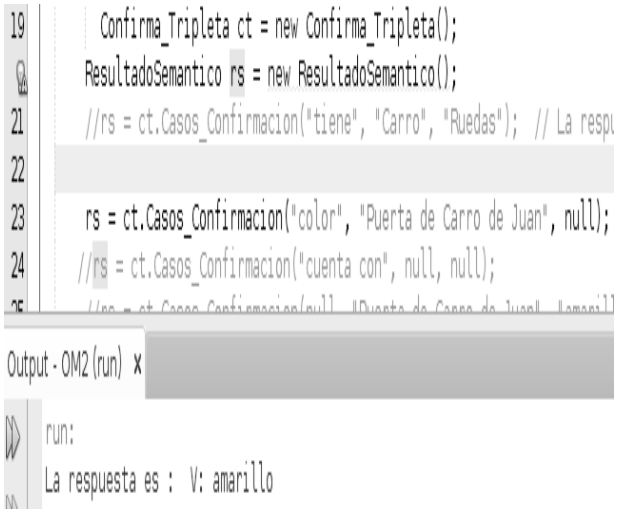


Figure 9. The rs object uses "*casos_confirmacion*" method to search semantic relation between "*color*" and "*puerta de carro de juan*" {door of Juan´s car}. The result is "*Amarillo*" {yellow}, that it to say "the color of the door Juan´s car is yellow"

The Figure 10 shows a fragment of code about "Casos_Confirmacion" method that solves semantic relation of a triple. The steps are the follow: first, ontology is loaded, we can load two or more ontologies but for this test we load an unique ontology. All ours ontologies have the ONT extension.

The semantic analyzer can return the closer value semantic between words. It value is obtained adding the number using the Confusion algorithm. The semantic analyzer uses confusion [16] to solve this.

Function CONF(r, s), called the absolute confusion, computer the confusion that occurs when object r is used instead of object s, more details in [16].

```
public String Casos_Confirmacion(String o,String r,  String v)
{   //String cadetemp;
    Lista ontoAcargar = new Lista();
    String archivos[] = { "OntoPrueba.ONT" };
    LoadOntology lo = new LoadOntology( archivos );
    Concepto co = null;
    String respuesta = null;
    StringBuilder sb_cadenas = new StringBuilder("");
    //funciona que el valor sea una cadena o sea un concepto. (pendiente
    // ----- carga la onto y busca la respuesta ---------
    ontoAcargar = lo.getOntologies();
    ontoAcargar.rewind();
    Concepto cRelacion = null;
    while(ontoAcargar.next())
    {
        Ontology onto = (Ontology) ontoAcargar.currentNodo().getDato();
        onto.getOntology();
        HashTableOntology hto = onto.getHashTable();
        float valorCercaniaSemantica = 0.0f;
        String re ;
        if (o!=null)
```

Figure 10. A fragment of code that describes a part of semantic analyzer, to calculate the closer semantic.

The development of the semantic analyzer led to the realization of various activities such as: the definition of an ontology (set of interrelated elements, like a network of neurons in the human), creating frames (definition of scenarios, contexts or domains) and obtaining the meaning (semantics).

This resulting of semantic analyzer is for a machine can to apply in another module, to answer complex questions from human being for example or for a robot can to find alive humans in a build after a disaster natural.

6. FINAL THOUGHTS

The analyzer is not a program to solve noun ambiguous it is a program that to find semantic relations between three words. It runs after an anaphora-solver [17] and a prepositions solver have preprocessed the Spanish descriptive texts. The base ontology must contain the appropriate relations (verbs) and nodes (nouns). The cases that are solves here cannot be solved using only grammar and syntactic information, because the semantics relations are necessary. This is the main reason why other analyzers have failed in solving this problem.

7. REFERENCES

[1] Padró, L. and Stanilovsky, E. 2012. Freeling 3.0: Toward wider multilinguality". *In Proceedings of the Eight International Conference on Language Resources and Evaluation* (Istambul, Turkey May 23-25, 2012). LREC'12, 2473-2479.

[2] McInnes, T. and Pedersen, T.2013. Evaluating Measures of Semantic Similarity and Relatedness to Disambiguate Terms in Biomedical Text. *Journal of Biomedical Informatics*. Elsiever Science. 6 (San Diego, USA, December), 1116-1124. DOI = http://dx.doi.org/10.1016/j.jbi.2013.08.008

[3] Colorado, F. 2008. *Mapeando palabras a conceptos*. M. Sc. Thesis. In Spanish. CIC-IPN.

[4] Minsky, M. 1974. *A Framework for Representing Knowledge*. Memo. 306MIT. AI Laboratory.

[5] Collin, B. 2008. FrameNet, Present and Future. In *Proceedings of the First International Conference on Global Interoperability for Language Resources*. (Hong Kong: City University), 12-17.

[6] Cuevas, A., and Guzmán, A. 2010. Knowledge accumulation through automatic merging of ontologies. *Expert Systems with Applications*, Elsevier Editorial. 37, 3. System, USA. ISSN: 0957-41741991-2005.

[7] Guzmán, A., Cuevas, A., and Jiménez, A. 2011. The centroid or consensus of a set of objects with qualitative attributes. *Expert systems with applications*. Elsevier Editorial. 38. 4908-4919.

[8] Villanueva, D., Cuevas, A., Juárez, O., and Guzmán, A. (2013) "Using frames to disambiguate propositions". *Expert Systems with applications*, Elsevier Editorial. 40, 2. (February 2013). 598-610. DOI= http://dx.doi.org/10.1016/j.eswa.2012.07.061.

[9] Aliseda, A. 2013. ¿Inteligencia Mecánica? La pregunta de Alan Turing. *Ciencia. Revista Mexicana de la Academia de Ciencias*, 64, 4. ISSN 1405-6550. In Spanish. 10-17.

[10] Pino, R., Gómez, A., and de Abajo, N. 2001. *Introducción a la Inteligencia Artificial: Sistemas expertos, redes neuronales artificiales y computación evolutiva*. Universidad de Oviedo. ISBN 84-8317-249-6. 106.

[11] Davis, R., Shrobe, R., and Szolovits, P. (1993). What is a knowledge Representation? *Artificial Intelligence Magazine*. 14, 1. 17-33.

[12] López, Y. 2012. *System for extracting and representing knowledge from descriptive texts*. M. Sc. Thesis. In Spanish. CIC-IPN

[13] Panchenko, A. 2013. Similarity measures for semantic relation extraction. Doctoral Thesis. Université Catholique de Louvain and Bauman Moscow State Technical University.

[14] Melcuk, I. 1996. Lexical functions: *A tool for the description of lexical relations in a lexicon*. *In Lexical functions in lexicography and natural language processing*. John Benjamin, (Amsterdam, Philadelphia, 1996). 37–102.

[15] Agrawal, A., and Kakde, O. 2013. Semantic Analysis of Natural Language Queries Using Domain Ontology for Information Access from Database. *International Journal of Intelligent systems and Aplications* . (December, 2013), 81-90. .DOI= DOI: 10.5815/ijisa.2013.12.07

[16] Levach, K., and Guzman A. 2007. Hierarchy as a New Data Type for qualitative Values. Journal Expert Systems with Applications. 32, 3.

[17] Toledo, G., Valtierra, R., Guzmán, A., Cuevas, R., and Martínez, S. 2014. AnaPro, tool for identification and resolution of direct anaphora in Spanish. *Journal of Applied Research and Technology*. 12, 1. (February, 2014), 14-40

Traducer of Letters of the Mexican Sign Language to Text Using Leap Motion

Gil Díaz
CINVESTAV-IPN
Mexico city, Mexico
gdiaz@computacion.cs.cinvestav.mx

José Rodríguez
CINVESTAV-IPN
Mexico city, Mexico
rodriguez@cs.cinvestav.mx

ABSTRACT

Computers are hard to manipulate for people with some kind of handicap, in the case of people with oral communication issues, they need to learn a sign language and their relatives too, but this is a solution for their closer ambiance. People that does not know the sign language cannot communicate with somebody that have oral communications issues. In this paper is presented a software tool that makes manual gesture recognition of the Mexican Sign Language through the use of neural networks and leap motion sensor.

Keywords

Artificial neural network, sign language, hand gesture recognition, leap motion

1. INTRODUCTION

Using the technology to help people with some kind of handicap has been a recurrent subject research in areas as Ambiance Intelligence. The use of modern devices in the everyday activities of people gives as result that new tools are being produced that could help people in their daily activities.

A special case of handicapped people are those with communication oral issues, those people use a sign language (SL) to comnicate with other persons. This kind of languages, moreover that must be learned by the person with the communication issue, they must be learned too by their relatives and people in his nearest ambiance [4]. However, mute people can not communicate with people that does not know the SL [5]. Nevertheless, the SL is not universal, in the same way that there are several oral languages there are too several SL and these SLs are not compatible between them. For this reason, in Mexico the Mexican Sign Language is used. According to the National Institute of Statistics and Geography (INEGI) in Mexico there are 400,000 people with oral communication issues [6].

In this paper is presented a leap motion based traducer for the Mexican Sign Language that takes the pattern recognized for the device and represents his meaning in Mexican Spanish. This paper is structured as follows: section 2 presents the related works developed with different devices and pattern recognition mechanism. In section 3 is explained how this proposal were developed. Results are presented in section 4 and conclusions in section 5.

2. RELATED WORK

Several works have been developed for hand gesture recognition, nevertheless the technological development in last years gives as result that better sensors with enhanced capabilities for precision and sensing such as leap motion, kinect and others that help in mencioned task.

Guna and Grega [15] implemented a test for the leap-motion to evaluate his functioning, reliability and precision. They planned a dynamic and a static scenario. In the static one they used a plastic model of the hand. The aim of using this model is keeping the same gesture moving the hand through 37 locations in the reading field of the sensor. With this test the locations with better sensing by the sensor were founded.

In the dynamic test a tool in a shape of v were used to simulate two fingers to be sensed by the sensor keeping always the same distance between them. For this test the results show that the best recognitions of the hand are in areas closest to the sensor. Inconsistences were found in the performance of sensor according to the variation in the sensing frequency.

A.S.Elons y Menna Ahmed [3], present a system where the data produced by the leap motion sensor are used for training a three layer perceptron neural network. The input layer get the data from the vector generated by the sensor, this vector contains the position of fingers. The middle layer is responsible of the learning, and the output layer gives the number of patterns used to test the system. This proposal get the best percentage in recognition of gestures using the distance between the fingers for training the multilayer perceptron. This distance gives invariable measures in function the positon of the hand according the sensor.

Ching-Hua Chuan and Eric Regina [2] developed a pilot system with 26 moving signs of the alphabet using several input data produced by the leap motion as: strength grip (to determine if hand is open or closed), the normal vector of the hand palm, speed of the moving hand palm, position of the palm, distance from every finger to the thumb, and size and direction of every finger and the speed of displacement of every phalanx of fingers.

For every finger 4 characteristics were calculated: the extended distance (from the fingertip to the center of the hand-palm), fingertip-inclination projection (i.e. the projection of the direction vector over the normal vector of the hand-palm), X-order, this is the order of the finger over the x-y plane according the other fingers), angle this is the angle between finger overt the plane x-z. To make the recognition

of the alphabet the k-nearest neighbors algorithm with different values for k and different functions for the kernel was used.

Giulio Marin y Fabio Dominio [10], used the leap-motion sensor and Kinect to implement the recognition of a set of 10 static gestures of the American Sign Language. They used the leap-motion sensor to calculate new data: angle for every finger, distance between fingers and the inclination angle of fingers. The Kinect sensor were used to get two sets of data. The first with the distances between points in the hand and the center of the hand-palm. The second set represents the curve of the hand´s contour. With all these data 5 vectors were generated for the recognition of gestures using a multi-class vector support machine. The tests were conducted first using the sensors in an individual way and then working together sharing the data: Using the distance between the fingers, the leap-motion sensor gets 76% of success in recognitions of gestures. Using the angle between fingers 74.2%, using the inclination angle of fingers 73% and, putting together the 3 parameters the success in recognition were 81%. With the Kinect sensor using the distance between a point in the hand and the hand-palm the success in gesture recognition were 65%. Using the curve of hand´s contour the success rate were of 87.3. Using both parameters the success rate were 89.7%.

Prashant G. Ahire y Kshitija B. Tilekar [1], implemented a voice-gesture bidirectional traducer using the Indian Signal Language. To implement the voice conversion, they used the Speech To Text Google´s API. Once the written word was identified, this is searched in a database of gestures and an animation is reproduced in the screen. In the translate from gestures to voice a Web-Cam is used. The video is decomposed in a sequence of images then noise is removed from the images. The regions of interest are searched using a method of region growth. Once the region of interest was identified a the gesture is searched over the database using an approach based on correlation. After the gesture has been identified the translation to text is performed using the Google´s Text to Speech API to generate the corresponding audio.

M.Mohandes y S. Aliyu [11] developed a system of recognition for the Arabic Signal Language using the leap-motion sensor. First they get the gestures and extracted their characteristics to create vectors with them. For the classification of gestures, they used a vector support machine and hidden Markov models. The aim of tis project was try to identify the gesture before it would be completed for the user. The test were carried out with the 50%, 70%, 80%, and 100% of sample of the gesture. The results show that this proposal gets 90% of success in recognition of the gesture and with times lower than 0.5 of a second.

3. DEVELOPMENT

3.1 Leap Motion Sensor

The leap motion sensor is composed of two infrared cameras and three infrared emitting light diodes, these five elements allow the sensing in 3D of both hands getting more specific information of phalanx, finger tips, size of finger, hand-palm center, angles according the axes of the sensor, speed in motion, etc.[3, 16, 15]. With all these capabilities the leap-motion sensor has been used in several fields as: control of robots, personal computers, games, gesture recognitions and sign language processing, etc. According to the manufacturer, the leap-motion sensor can follows the hands with a rate of 200 frames by second, moreover the sensor has a sensing field of 150 grades with a volume of 8 feet [7], see figure 1.



Figure 1: Sensed field of leap motion sensor.

3.2 Data Selection

The data to be used in this proposal are the set of angles formed by the fingers when a gesture is produced. Several angles are produced with phalanxes when the gesture is produced, the distance between phalanxes and the angle between the phalanxes and the normal vector of the hand palm. Figure 2 shows the structure of the hand, to produce a gesture the phalanxes must be set in different positions and angles. Different types of angles are formed, between the phalanxes in the finger and separation between fingers.

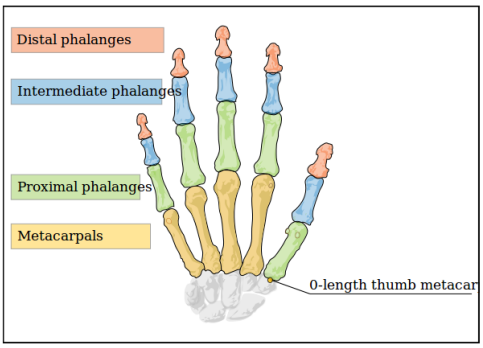


Figure 2: Hands Morphology.

In this proposal three the set of data are used, the first is composed by the angles formed by the phalanx see figure 3, the second has the angles of separation between the proximal phalanxes, see figure 4. And the last set is composed of the angles between phalanxes and the normal vector of the hand, this vector is generated by the leap-motion, see figure 5.

3.3 Structure of the neural network

To implement the recognition of the Mexican Sign Language a multilayer perceptron neural network is used. Three layer compose this neural network, each layer is composed of a set of neurons that are interconnected with neurons in another layer [13]. The neurons of the middle layer are connected with the neurons in the input and output layers. The neurons in the input layer don´t make any processing on the information, they supply the data obtained by the sensor to the neurons in the middle layer. The middle layer, called the hidden layer, is composed of neurons that perform the

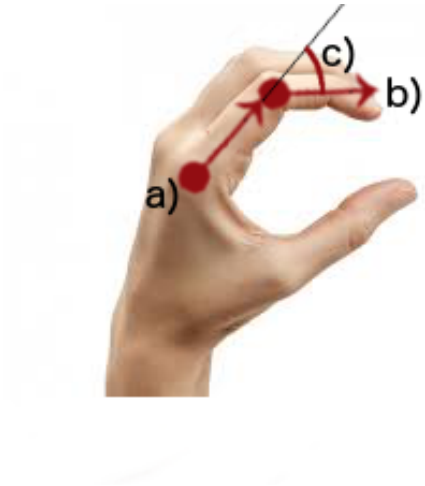


Figure 3: Vector of proximal phalanx, of medium phalanx and the angle between them

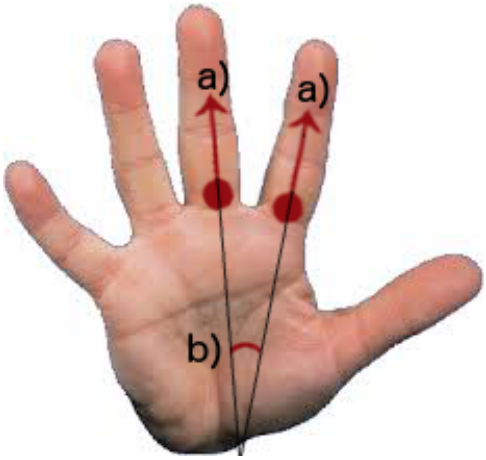


Figure 4: Angle between proximal phalanxes

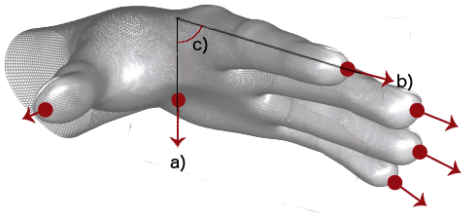


Figure 5: Normal vector, vector of the phalanx and the angle between both vectors

transformation of the data sent by the input layer into signals that the output layer is going to interpret as answer of the neural network. The output layer is composed of the output neurons, the set of neurons in this layer represents the answer of the neural network [12, 8]. The training of the neural network was carried out using a backpropagation and a stochastic descendent gradient algorithm, several training and validation iterations were executed. In every training iteration a random sub-dataset was selected as sample, the validation set was always the same during training. The sigmoid function selected as the activation function for every neuron and the error evaluation was estimated using the quadratic average error[14, 9].

The input layer is composed of 38 neurons one for every angle used to classify the gestures. The number of neurons in the output layer is defined of by the number of gestures that the neural network can classify [14,15,16,17]. The Mexican Sign Alphabet is composed of 27 letters, where 21 are represented with static gestures and 6 with dynamic gesture. In this proposal only the 21 static letters are processed. The letters where classed into three groups, according to the position of the hand-palm into the sensing field: in face of the sensor, on his side (with the little finger as the nearest to the sensor), and with the hand-palm in opposite orientation to the sensor see figure 6. Every group of gestures is classed by a specific neural network, it means that 3 neural networks were implemented. The neural network that classify the gestures with the hand-palm facing the sensors classes 11 letters: a, b, d, e, l, r, s, t, u, v, w y z. The neural network that classify the gestures with little finger near to the sensor classes 5 letters: c,f,i,o,p. The neural network that classify the gestures with the hand-palm in opposite direction to the sensor classes 5 letters: y,g,h,m,n.

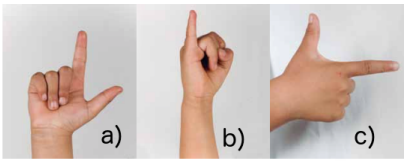


Figure 6: Different positions of the hand: in face of the sensor a), the little finger close to the sensor, and the hand-plan against the sensor.

For the gesture recognition first a pre-processing is carried out to identify the position of the hand-palm. Then the neural network is selected and the data obtained from the sensor.

4. RESULTS

Three neural networks were implemented, for the training of the neural networks 1100 samples were used, every sample is composed of 38 angles of the phalanx. 770 (70%) of these samples and 330 (30%) were used as validation data. Given that 21 gestures were classed, 23100 samples were used. Every neural network was entrained and validated wit samples from the same user. The success rate in classification of the neural networks was 99.5%. Then 4 additional users were selected with different sizes in hands, every one of them makes 10 gestures and every gesture was sensed 5 times. Every neural network was trained an validated with data samples from the same user. The training parameters were modified to find the best results. The learning rate selected was 0.1 and 0.5, the number of neurons in the middle layer varies between 30 and 120 neurons. Additional tests were carried out with training subsets of 10%, 20% and 30% of the data training set.

Next are presented the experiments used to determine the number of neurons in he middle layer. The results presented were obtained using the neural neurons capable of recognize 11 gestures.

Table 1 shows the results in training using a subset of 10% of training data. The best results were obtained using a small learning rate. With a learning rate of 0.1 two times the 100% in success for gesture recognition were obtained, the first case was with 90 neurons in the middle layer and the second case with 100 neurons.

Table 1: Results of the training with a subset of 10%

	50	60	70	80	90	100	110	120
0.1	90.9	91.9	95.1	94.3	100.	90.9	100.	93.7
0.2	81.8	78.8	79.0	83.6	81.8	81.8	81.8	83.0
0.3	63.6	63.6	68.6	90.9	81.8	79.3	81.8	90.9
0.4	56.5	58.1	63.6	78.0	63.6	70.8	72.7	78.0
0.5	45.4	63.6	72.7	72.7	72.7	81.8	72.7	77.4

Table 2 shows the results using a subset o training data of 20%. Using a learning rate of 01 the 100% in gesture recognition was obtained 4 times, and with a learning rate was only 1 time. With this subset the 110% in gesture recognition was achieved using middle layers with 70, 90, 100, 110, and 120 neurons, as can be seen in table 2.

Table 2: Results of the training with a subset of 20%

	50	60	70	80	90	100	110	120
0.1	96.6	84.9	100.	95.5	100.	100.	100.	96.9
0.2	85.2	70.8	89.5	90.9	92.3	94.5	94.0	100.
0.3	78.4	72.7	90.9	90.9	89.1	83.0	81.8	88.7
0.4	67.8	72.7	77.6	81.8	77.6	87.9	81.8	87.7
0.5	67.4	81.8	72.7	78.4	72.7	81.8	80.6	81.8

Finally, with the 30% de training data. The 100% of success on gesture recognitions was obtained seven times, five with a learning rate of 0.1. One with a learning rate of 0.2 and another with 0.3. Using 110 neurons in the middle layer three times the 100% was obtained, and one time with middle layers using: 70, 80, 90 and 120 neurons.

Table 3: Results of the training with a subset of 30%

	50	60	70	80	90	100	110	120
0.1	98.0	97.8	100.	100.	100.	99.3	100.	100.
0.2	81.8	84.9	90.9	96.8	99.0	90.6	100.	93.3
0.3	81.7	81.8	82.6	90.9	89.9	90.0	100.	86.6
0.4	72.7	78.9	72.7	81.8	90.9	90.9	90.9	90.9
0.5	73.3	72.7	72.7	81.8	81.8	81.8	90.9	90.9

he same process was followed for the other two neural networks where the 100% in gesture recognition success was obtained too using different learning rates. The main criteria in selecting the number of neurons for the middle layer was the computational cost, so the number of neurons selected was 30.

After the training the system was tested with 4 users having different sizes of hand. Everyone of users made 10 gestures and every gesture was read 5 times by the application, this is 200 reads were made by the sensor, 20 of them were wrong classed, this means a 90% of success en gesture recognition. In the case of gesture wrong classed, it was observed that the gesture was recognized in place of another close gesture with similar position of the hand, see figure 7. The wrong classed gestures are presented in table 4.

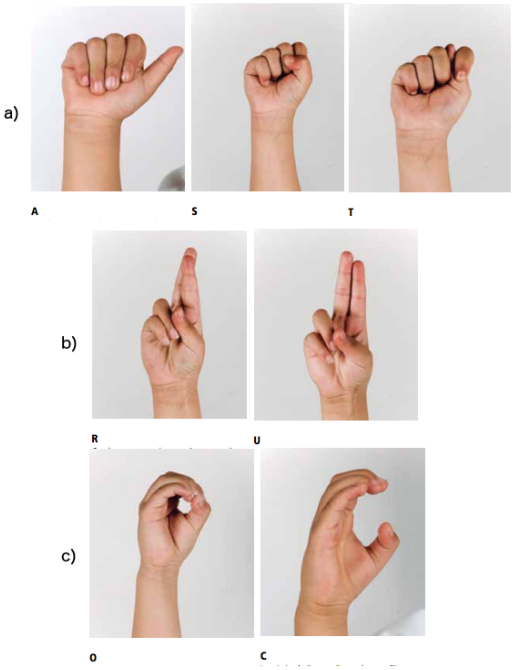


Figure 7: Gestures miss recognized

Table 4: Letters miss recognized.

Gesture	Answer	Number of errors
A	S	2
S	T	1
S	A	1
T	A	3
T	S	3
O	C	3
C	O	2
R	U	3
U	R	2

5. CONCLUSIONS

In this paper is presented a tool for the recognition of the gestures of the Mexican Sign Language, this proposal uses the data supplied by the leap motion. The data set is composed of angles between the fingers, between the phalanxes of the finger and the angle of the normal vector of the handpalm. With these data the implemented neural networks give 100% of success. However, the biggest number of gesture to classify the hardest to get the 100% on success. For this reason, the gestures were classed into three sub-sets.

The sets of data show that this is an approximation that gives good enough results given that the classification of ges-

tures get several times the 100% of success with hands with different size. However, some of the errors were produced by the data supplied for the sensor, sometimes the sign produced by the infrared cameras is limited and affected by the position of fingers hiding other fingers. The success rate is near to the 100% this allowing to think that a good generalization of the data is possible. The classification using different sizes of hand, gives too a good result.

6. REFERENCES

[1] P. G. Ahire, K. B. Tilekar, T. A. Jawake, and P. B. Warale. Two way communicator between deaf and dumb people and normal people. In *Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on*, pages 641–644, Feb 2015.

[2] C. G. Ching-Hua Chuan, Erick Regina. American sign language recognition using leap motion sensor. In *2014 13th International Conference on Machine Learning and Applications (ICMLA)*, pages 541–544, 2014.

[3] A. S. Elons, M. Ahmed, H. Shedid, and M. F. Tolba. Arabic sign language recognition using leap motion sensor. In *Computer Engineering Systems (ICCES), 2014 9th International Conference on*, pages 368–373, Dec 2014.

[4] P. Gupta, G. Joshi, and M. Dutta. Comparative analysis of movement and tracking techniques for indian sign language recognition. In *2015 Fifth International Conference on Advanced Computing Communication Technologies*, pages 90–95, Feb 2015.

[5] http://www.assh.org/handcare/hand-arm anatomy. July 2016.

[6] http://www.inegi.org.mx. July 2016.

[7] http://www.leapmotion.com. July 2016.

[8] D. Kriesel. *A Brief Introduction to Neural Networks*. 2007.

[9] B. KrÃűse, B. Krose, P. van der Smagt, and P. Smagt. *An introduction to Neural Networks*. 1993.

[10] G. Marin, F. Dominio, and P. Zanuttigh. Hand gesture recognition with leap motion and kinect devices. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1565–1569, Oct 2014.

[11] M. Mohandes, S. Aliyu, and M. Deriche. Arabic sign language recognition using the leap motion controller. In *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, pages 960–965, June 2014.

[12] M. A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.

[13] R. Rojas. *Neural Networks : A Systematic Introduction*. Springer, 1996.

[14] J. L. G. Rosa. *Artificial Neural Networks - Architectures and Applications*. InTech, 2013.

[15] J. G. G. J. M. P. S. Tomažič and J. Sodnik. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors*, 14(2):3702, 2014.

[16] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13(5):6380, 2013.

An Ontological Model for Business Process Domain

Tania Guerrero-Melendez
Faculty of Engineering and Science
Autonomous University of Tamaulipas
Victoria, Tamaulipas, Mexico, 87000
(52) 835-3181718
tyguerre@docentes.uat.edu.mx

Edgar Tello-Leal
Faculty of Engineering and Science
Autonomous University of Tamaulipas
Victoria, Tamaulipas, Mexico, 87000
(52) 835-3181718
etello@uat.edu.mx

Luis de la Cruz-Llopis
Telematics Engineering Departartment
Polytechnic University of Catalonia
Barcelona Spain, 08034
(34) 93- 4016014
luis.delacruz@entel.upc.edu

ABSTRACT
Ontology is one of the semantic technology standards that can make sharing knowledge easier. This paper describes a technological solution which focuses on contributing to the management of procedures for the business process environment. The technological solution proposed consists in the use of an ontological model developed in the business process domain with the purpose of optimizing the interaction between two or more organizations. In this paper, the use of a Business Process Ontology is suggested in order to solve some technological and organizational heterogeneity problems existing among participants, involved in an inter-organizational collaboration.

CCS Concepts
• Applied computing → Business process management.

Keywords
Ontology; Business Process; BPMN; BPM.

1. INTRODUCTION
Collaboration between organizations and the exchange of information have a relevant place in this increasingly globalized world, enabling an increase in the use of technological tools to obtain and share the knowledge that belongs to such organizations. This information shared represents a fundamental component in order to reach the business goals previously established since it facilitates the opportune business decision-making. For this reason, organizations invest resources in the search of competitive advantages that facilitate decision making to their partner organizations. The existence of technological and organizational heterogeneity among organizations stimulates the use of tools to manage information efficiently where ontologies represent an option to be used. The ontologies are knowledge representation systems on a specific domain in order to obtain a formal representation of concepts and the relationship between those concepts [1] [2] [3]. Berners-Lee in [4] explains that an ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules; therefore, an ontology allows to share a common perspective about information of interest and makes possible the reuse of knowledge for different domains. In particular, the knowledge domain addressed in this work is the business process environment; specifically, this paper describes the construction of an ontology based on the BPMN language [5].

The term ontology has been used in the field of knowledge for a long time [6] but, currently the use of ontologies has become popular with the Semantic Web, in which information is given well-defined meaning and is based on knowledge representation rules and assertions [7]. Ontology is used to define a punctual vocabulary and relationships for modeling the business process domain.

This paper describes a technological solution focused on contributing in management procedures for the business process environment. The technological solution proposed consists in the use of an ontological model developed in the business process domain with the purpose of optimizing the interaction between two or more organizations. Then, the construction of an ontology derived from the BPMN language specifications, which is the standard for modeling business processes, is described. The Ontology will be used in a knowledge-based information system, this system will be developed in order to allow making decisions about business processes that are executed based on process instances available.

The present work is structured as follows: section 2 presents a basic theoretical framework where Ontology, Business Process and BPMN topics are introduced. In section 3, the methodology used in the Ontology building process will be described. In section 4, some results are presented. Finally, conclusions will be shown in section 5.

2. BACKGROUND
In this section, a brief background is presented in order to establish a theoretical framework regarding the development of ontology.

The term ontology can be defined in different ways; in computer and information science, an ontology is defined as a "*technical term denoting an artifact that is designed for a purpose, which is to enable to the modeling of knowledge about some domain, real or imagined"* [8]. In other words, an ontology is a formal and explicit definition of elements and their properties and relationships. An ontology is composed of classes or concepts, the relationships that represent the interaction between classes (forming a domain taxonomy), instances used to represent objects of a class (individuals) in an ontology and, axioms that provide expressivity to the ontology [9].

The Business Process Management (BPM) is the discipline that combines knowledge from information technology and knowledge from management sciences and applies this to operational business processes [11]. The BPM consists of

executing a life cycle of continuous improvement to processes of organizations, which involves performing several stages for each process: analysis/evaluation, design, implementation, and execution. BPM can also be seen as the application of a set of technologies, methods, software tools, and standards that analyze, design, implement and manage organizational processes explicitly in order to automate them by integrating people and software applications an organization, as well as integrate, manage and successfully execute collaborative business processes among organizations [12]. The concept of a process model (business process diagram) is fundamental in BPM. A process model aims to capture the different ways in which a case (i.e., process instance) can be handled [11]. A business process is a set of activities done in coordination with a technical and organizational environment where such activities lead to the realization of business goals [12].

In the business process design trying to separate the process logic of the application logic, in a way that the business process can be automated through a business process management system. One of the main languages of notations to model operational business processes is the business process management and notation (BPMN) [5]. BPMN is a graphic notation that describes the logical steps in a business process [5]. This notation is specially addressed to coordinate the sequence of processes and messages that flow among participants in different activities. BPMN is a widely used language in business process modeling that describes a business process in a Business Process Diagram (BPD). BPMN is an Object Management Group (OMG) specification [13]; the most recent version of BPMN was published on December 2013. In the BPMN version 2.0.2. the specification document mentions that "BPMN was developed to provide a notation that is readily understandable by all business users, from business analysts that create the initial draft of the processes, to technical developers responsible *for implementing the technology that will perform those process, and finally to business people who will manage and monitor those processes"* [5].

3. RESEARCH METHODOLOGY

After specifying the knowledge domain and purpose of the ontology, the ontological model development started based on the following three methodology phases (see Figure 1):

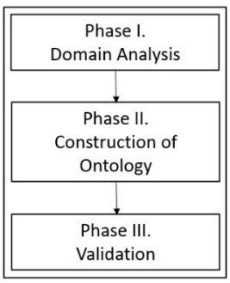


Figure 1. Research methodology

Phase I. In this step, an analysis of BPMN Language is done as ontology knowledge domain. The BPMN is an Object Management Group (OMG) specification, widely adopted as the graphical language for business process modeling. A business process is described in a Business Process Diagram and the BPMN Specification provides a detailed description of the graphical elements that compose this kind of diagrams. On the other hand, this specification also provides a detailed description of all the language elements, their properties, and attributes.

Phase II. In this step, the process followed to build the ontology and the tools used for this purpose are described. Protégé is a free, open source ontology editor useful to build knowledge-based solutions in diverse areas [14] and, to construct the Ontology the Protégé 5.0-beta version was used.

The ontology developing process involves the formalization of the BPMN 2.0.2 specification and the Ontology construction. The first step to develop an ontology consists in obtaining a set of basic concepts that define a specific domain, for that it is necessary to start by defining the taxonomy using the following sub-steps:

- Identify and classify all the BPMN language elements; every element classified is modeled as a class in the ontology
- Define the class hierarchy
- Represent formally the attributes and properties of each element modeled
- Create instances

Finally, the taxonomy that initially had been composed by classes and properties is enriched through axioms.

Phase III. Once built the ontology, it is necessary to evaluate in order to make sure that the ontology model created in this work is correct. One way to validate the ontology accuracy is using Pellet reasoner implemented in Protégé; Pellet provides functionalities to check consistency of ontology, classify taxonomy and check links [15]. On the other side, the precision of knowledge domain is a task that must be performed by a highly skilled person in the Business Process domain.

4. RESULT

As previously mentioned, the ontology model was supported by the analysis of BPMN 2.0.2 Specification and all the elements were modeled, therefore, the ontology model developed is composed by 161 classes hierarchically ordered. Under the root class *Thing* the ontology contains 26 concepts considering super-classes and 135 classes derived, in Figure 2, the main classes (superclasses) of the ontology are shown; one of this superclasses is the BaseElement class and their taxonomy includes the following 31 subclasses:

Artifacts, Assignment, Auditing, CategoryValue, ComplexBehaviourDefinition, ConversatiojnAssociation, CoversationLink, ConversationNode, CorrelationKey, CorrelationPropertyBinding, CorrelationPropertyRetrievalExpression, CorrelationSubscription, DataAssociation, DataState, Definitions, Documentation, Expression, ExtensionAttributeValue, FlowElement, InsputSet, ItemAwareElement, MessageFlowAssociation, Monitoring, Operation,OutputSet, Participant, ParticipantAssociation, Relationship, ResourceParameter, ResourceParameterBinding, and *RootElement;* 15 of these classes have more subclasses derived, an example of this is class: FlowElement. The definition of classes, class hierarchy and relationships follows the structure of BPMN ver.2.0.2.

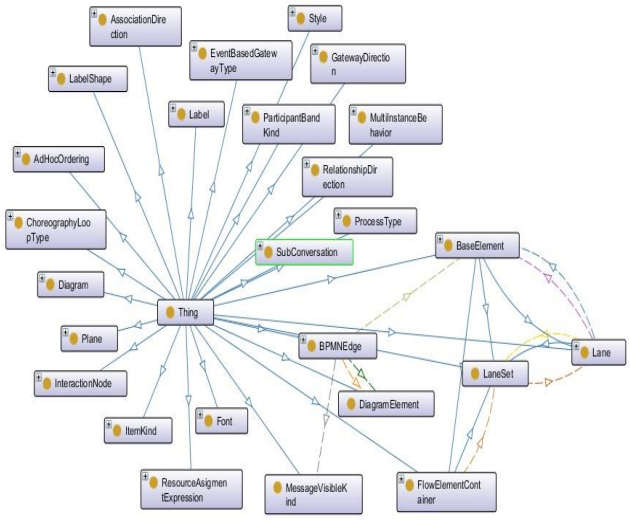


Figure 2. Ontology superclasses

Superclasses are composed by other derived classes, and have relation with the rest of the taxonomy in a different manner. An example of this is presented with class:Activity. An Activity represents the work performed within a business process and could be compound or atomic. The compound type means that its work is broken down into a finer level of detail, that is, a compound activity contains more than one sub-process; on the other side, atomic activities contain only one task. Activity is a subclass derived from FlowNode class who derives from class FlowElement and FlowElement, finally it is derived from BaseElement superclass. In Figure 3 the structure of class Activity is shown with a caption of the schema presented by Protégé tool. This taxonomy class includes nine main subclasses and from this, 39 subclasses are derived. Figure 4 shows the relationship of BaseElement taxonomy where it is possible to observe the class hierarchy in a graphical form.

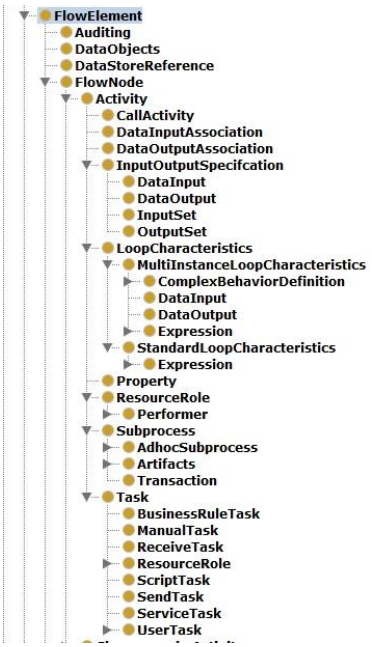


Figure 3. Hierarchy of class Activity

Essentially, the ontology is composed by a taxonomy of business processes concepts, the existing relationships and a set of axioms. This taxonomy is enriched by relationships and attributes of each element formalized like *data property* or *object property*.

In Table 1 a summary of ontology metrics is shown, where a number of logical axioms (1618) are considerably compared to the amount of classes available (161). A partial structure of the ontology model can be seen in figures 2, 3 and 4. In this same table, the instances loaded in the ontology are shown, the individuals come from different business processes modeled with the BPMN language and used in a real environment.

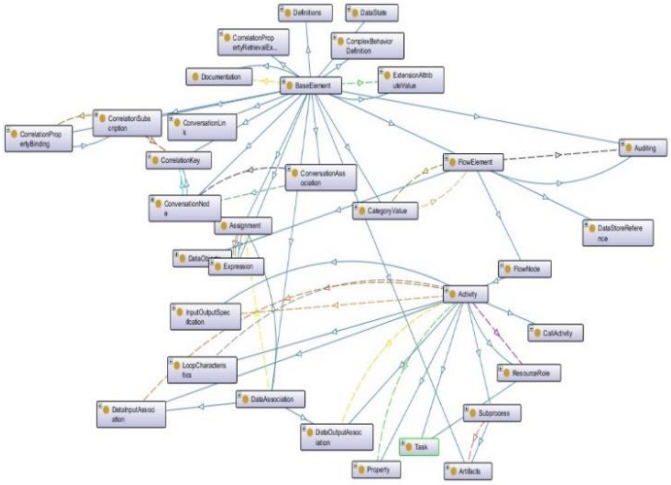


Figure 4. Taxonomy of BaseElement class

Table 1. Ontology metrics

Axiom	2179
Logical axiom count	1618
Class count	161
Object property	233
Data property	133
Individuals	34

The scope of our work includes the Business Process Modeling and Notation formal representation inside an ontology, making possible the management of different business processes to any business environment; the only feature that must comply with this process is to be correctly defined under the BPMN specification. The three phases methodology used cover the analysis, development, and validation of this ontology; the consistency and accuracy of the ontology were checked using the Pellet reasoner of Protégé and no logical errors or conflicts were found. With this last validation phase, we ensure that the ontology itself does not contain any contradiction.

5. CONCLUSION

Considering the use of information technologies in the business process domain, this paper presents an ontology in the Business Process domain. The ontology represents information in a manner that it can also be used by machines, not just to display it, but also to automate, integrate and reuse. The ontology development was supported by the analysis of BPMN Specification, exactly in accordance with BPMN language ver.2.0.2. The ontology

contains all elements that support the language specifications and is comprised by 161 classes, additionally, a detailed formalization of properties and conditions of the modeled elements is provided through 233 *object properties* and 133 *data properties* to represent the relationships among classes and the attributes of each element formalized, allowing to keep instances of private (Orchestration), publics (Choreographies) and collaborative business process. Also, this allows establishing guidelines for performing the combination among BPMN elements and valid BPD.

The ontology presented in this paper is part of a knowledge-based information system (in development). This system will support decision-making concerning the business process to be executed by organizations involved in collaboration.

6. ACKNOWLEDGMENTS

Our thanks to the Autonomous University of Tamaulipas and the Government of Mexico for the support given by the PRODEP-SEP DSA / 103.5 / 16/5907 scholarship.

7. REFERENCES

[1] S. Staab and R. Studer, Handbook on Ontologies, Springer, 2004.

[2] R. Gil and M. Martin-Bautista, "A novel integrated knowledge support system based on ontology learning: Model specification and case study," *Knowledge-based Systems,* no. 36, pp. 340-352, 2012.

[3] M. L. Caliusco and G. Stegmayer, "Semantic Web Technologies and Artificial Neural Networks for Intelligent Web Knowledge Source Discovery," in *Emergent web intelligence: Advanced semantic technologies* , London, Springer, 2010, pp. 17-36.

[4] T. Berners-Lee, J. Hendler and O. Lassila, "The semantic web," *Scientific American,* 2001.

[5] Object Management Group-OMG, "OMG-BPMN," 09 Diciembre 2013. [Online]. Available: http://www.omg.org/spec/BPMN/2.0.2/PDF/. [Accessed agosto 2015].

[6] T. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition,* vol. 5, pp. 199-220, 1993.

[7] E.-M. Kalogeraki, S. Theocharis, D. Apostolou, G. Tsihrintzis and T. Panayiotopoulos, "Semantic Concepts in BPMN 2.0," in *The 5th International Conference on Information, Intelligence, Systems and Applications,*, Chania Crete, Greece, 2014.

[8] L. Liu and T. Özsu, Encyclopedia of Database Systems, Springer-Verlag, 2008.

[9] O. Corcho, M. Fernández-López and A. Gómez-Pérez, "Ontological Engineering: Principles, Methods, Tools and Languages," in *Ontologies for Software Engineering and Software Technology*, Springer Berlin Heidelberg, 2006, pp. 1-48.

[10] M. Weske, Business Process Management. Concepts, Languajes, Architectures., 2 ed. ed., Berlin: Springer, 2012.

[11] W. M. van der Aalst, "Business Process Management: A Comprehensive survey," *ISRN Software Engineering,* vol. 2013, p. 37, 2013.

[12] E. Tello-Leal, O. Chiotti and P. D. Villarreal, "Software Agents for Management Dynamic Inter-Organizational Collaborations," *IEEE Latin America Transactions,* vol. 12, no. 2, pp. 330-341, 2014.

[13] Object Management Group - OMG, "OMG," [Online]. Available: www.omg.org. [Accessed Agosto 2015].

[14] University of Stanford, "Protégé," Stanford Center for Biomedical Informatics Research, 2016. [Online]. Available: http://protege.stanford.edu/. [Accessed mayo 2016].

[15] Pellet, "Pellet," [Online]. Available: http://clarkparsia.com/pellet. [Accessed marzo 2016].

Text Mining System for Information Extraction and Classification of Scientific Documents

Ana B. Rios-Alvarado
Autonomous University of Tamaulipas
Ciudad Victoria
Mexico
arios@uat.edu.mx

Edgar Tello-Leal
Autonomous University of Tamaulipas
Ciudad Victoria
Mexico
etello@uat.edu.mx

Fabian A. Ramirez-Tovar
Autonomous University of Tamaulipas
Ciudad Victoria
Mexico
a2121010017@uat.edu.mx

ABSTRACT

The quantity of digital text documents generated as a result of scientific research is increasing, and current information systems for the management of these documents are based on structured or semi-structured data. However, these systems do not consider the use of the information contained in the document text, i.e. unstructured information. Therefore, information systems that enable information extraction from unstructured information are required. In this paper, a text mining system that allows information extraction and document classification of the domain of Computer Science is proposed. On one hand, the information extraction process is comprised of the phases of extracting title, author extraction, extraction of keywords, and construction of abstract. For which purpose an approach Named Entity Recognition (NER) in combination with the taxonomy of the ACM is used. On the other hand, the document classification process uses keywords previously extracted to relate the text document with a theme of ACM taxonomy. Finally, the proposal is tested with a corpus of 300 documents using precision and recall measurements, allowing to recognize the efficiency of the proposed approach.

CCS Concepts

•**Information systems** → **Data mining;**

Keywords

Named Entity Recognition; Text Mining; Information Extraction; Natural Language Processing; Document Classification

1. INTRODUCTION

Currently, a large amount of information is generated in digital format. Most of this information is unstructured, because of that their management is complicated for information systems as well as humans. Therefore, new methods to help filter and structure the most relevant information are required. The text classification and information extraction are examples of tasks that allow you to perform these processes.

The techniques and tools in the area of Natural Language Processing (NLP) are essential to execute the tasks of classification and information extraction (IE) from unstructured information. These techniques facilitate human-computer communication through human language. The NLP is responsible for developing computational techniques that allow this communication, and such techniques should be able to process human language, spoken or written, using some type of linguistic knowledge. The term "Named Entity" has been used widely in the field of IE, question answering (QA), and various fields of applications of NLP [8].

The Named Entity Recognition (NER) is essential in the field of IE, allowing to obtain specific information from text documents, this extracted information relates to persons, places, organizations, dates, numerical quantities, etc. [2]. NER in the area of Information Retrieval (IR) can extend the search space. The text summarization and identifying sentences tasks are usually linked with the identification of named entities contained in the text [4].

In this paper, we describe the development of a software system based on Web technology that allows extraction and classification of information in digital text documents of the Computer Science domain. The proposed software system allows the named entity recognition and extracting information such as title, author, keywords, and the construction of a summary using as a base the keywords identified. In addition, enables the classification of text documents using the ACM[1] taxonomy. Finally, the proposal is tested with a corpus of 300 text documents using precision and recall measurements, allowing to recognize the efficiency of the proposed approach.

The rest of this paper will be structured as follows: Section 2 describes the related works. Section 3 introduces the methods for extraction and classification of text documents. Section 4 will show the evaluation of the information extraction and document classification processes, through precision and recall measurements. Finally, Section 5 will present the conclusion and future work.

2. RELATED WORK

In [6] describes a system for classifying documents semantically from an unstructured texts. The associated classes are drama, comedy, horror and science fiction; taking these classes as a limit for the system. The proposed method processes the contents of electronic documents in PDF format, extracting nouns, and applying text mining techniques to find keywords. Then, associate the document to one of the aforementioned genres. In [3] describes a process for identifying user profiles. The specific features of a user are defined to search for keywords in your profile, from which patterns are detected, and this can separate people into groups. Further, these features allow to create a dictionary, accumulating

[1]Association for Computing Machinery

the frequency of occurrence of these features, in order to associate the profile to a group.

In [1] two methods for the classification of documents applying text mining are proposed. The first method works for classifies documents from highly frequent keywords. The second method also implements the frequency of keywords but applies a process of several iterations in order to find the best classification. In [7] presents a system based on semantic rules and linguistic markers, where using summaries and curricular records can identify and represent individuals for ontology population task. The information extracted in the ontological model is useful for tasks such as information retrieval in the academic domain.

3. EXTRACTION AND CLASSIFICATION SYSTEM

The development of a system based on Web technology that allows the information extraction (author, title, keywords and abstract), and automatic document classification of the domain of computer science is proposed. In addition, the system will load a document and consult all of the documents associated with a theme.

The functional requirements identified for the system are:

- The application must receive an input document.
- The application must allow access to classified documents.
- The application must allow to analyze a new document.
- The application must classify the document in a topic.
- The application must show each previously classified document.

In the analysis phase of system development was determined that the tasks to be performed:

- Defragment the contents of the documents in sentences and words.
- Remove punctuation and stop words.
- Clean the document (join divided by the end of the line words) to identify keywords.
- Identify keywords from nouns and generate a list.

According to the functionalities that the system must support, the following processes are proposed. For which purpose, the Standford CoreNLP[2] library containing the feature of Name Entity Recognition (NER)[3] and a grammatical tagger[4] is used.

- *Title Search*. To identify the title, first the contents of the document is separated by lines and subsequently processed by the recognizer named entities, taking the first line in which there are names of people or organizations. We consider that the line predecessor to the author's name or affiliation contains the title of the document
- *Author Search*. To identify the author considers that the first line in the document containing the name of a person, can be labeled as the author's name. For which purpose, the document is processed by separating each line, and identifying the named entities.
- *Keyword Search*. To identify the keywords, each identified sentence is processed with a grammatical tagger, when each word has its label (noun, adjective, subject, etc.), the words with a label substantive (NN or NNS) are selected. Subsequently, are searched in the database, the keywords of the ACM taxonomy, and are compared with those identified nouns, counting the occurrence of the word. Finally, a list is generated, which is filtered to exclude keywords that are not in the sentences of the document, eliminating repeated and organize them (the list is sorted from highest to lowest, where the higher repeated more times) according to the content of the document.
- *Summary Generate*. Once you have identified the keywords, then are obtained all sentences containing the keyword group from the database. After recovering sentence list, are filtered to remove repeated, and are ordered according to times that repeat (the most often repeated is the one that occupies the first position and so on).
- *Document Classify*. To classify a document, the list of keywords found through the *Search Keywords* process is used, then are searched in the database the themes that belong keywords, generating a list. This list is filtered to remove the repeated themes, selecting the most repeated theme (this theme is the one that contains the higher number of keywords). Then, this theme is associated with the document.

Additionally, all documents are applied the processes of: extract the plain text, separated by lines, join the words at the end of the line, and remove stop words.

4. RESULTS

When new prototypes are designed, it is appropriate to evaluate them in order to see its behavior with respect to the benchmarks, and compare them with other existing technologies. This also applies for the extraction of information [5], where from the application software implemented the results are obtained and evaluated.

Therefore, in this section, the results obtained with the proposed software system are presented. The evaluation for the extraction of information is displayed, the *precision* for extraction of the title, author, keywords, and abstract is obtained. In addition, the results of *recall* for the classification of documents by subject are presented.

The precision (P) is calculated with the following formula:

$$P = \frac{RelevantDocumentsRetrieved}{RetrievedDocuments}$$

where $RelevantDocumentsRetrieved$ is equal to the total number of correct documents obtained by the application, and $RetrievedDocuments$ is the total number of documents used for testing. It is considered correct if the item is well detected: title, author, keywords or lists of sentences for the summary, according to the evaluation of a human.

[2]http://stanfordnlp.github.io/CoreNLP/
[3]http://nlp.stanford.edu/software/CRF-NER.shtml
[4]http://nlp.stanford.edu/software/lex-parser.shtml

Moreover, *recall* will be determined in the process of document classification. The formula to measure recall (R) is:

$$R = \frac{RelevantDocumentsRetrieved}{RelevantDocuments}$$

where $RelevantDocumentsRetrieved$ is equal to documents that were classified on the correct topic, and $Relevant-Documents$ is the total number of documents that belong to that topic.

4.1 Title extraction

Identifying the title was obtained through the *title search* process. The proposed method searches the first lines of the document, because when analyzing the documents manually was observed that the title is found commonly before the authors. Since the documents contained this feature, the line where the authors are found is identified, then the lines that are found before authors are stored (since the title of the document is the element that is being analyzed. Table 1 the precision value obtained from the set of 300 documents for testing are shown.

Table 1: the Precision value in the extraction of title

Total Documents	Correct Documents	Precision
300	298	0.99333

One can observe in the results presented in Table 1, the possibility of failure is present in the identification of the title. However, this problem is mainly due to the use of Stanford CoreNLP library, because it does not have an extensive catalog for NER, which means that if can not be found a named entity in the selected lines. Then, in the title field of the document is stored with a null or empty value.

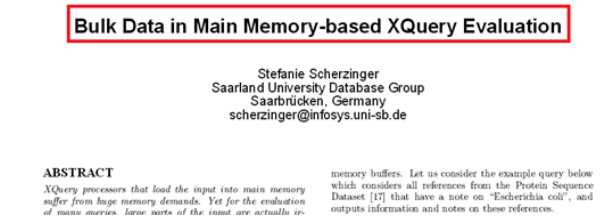


Figure 1: Title identified in text document



Figure 2: Title extracted from the text document using the application

In Figure 1 a digital document is displayed with the "title" element highlighted, where the title matches the title extracted by the application, which is shown in Figure 2. In Figure 2 the title tab is shown, which contains the title identified by the application, such title appears in a text box. In case the title identified is correct, the user must press the "Save" button to continue navigating within the application. In the case of an error in identifying the title, the user can edit and then press the "Save" button.

4.2 Author extraction

The method for extracting author consists in reading the first few lines of the document being analyzed, in search of a line containing a named entity with the "PERSON" label, if the label is found, then the value of the label is stored, and all lines that do not fulfill this condition are discarded. Table 2 shows the value of precision obtained by applying this method to search authors in a set of 300 documents.

Table 2: the Precision value in the author extraction

Total Documents	Correct Documents	Precision
300	258	0.86

In search of authors we observed some errors in the task of identifying authors, because the library used does not have a wide range of entities named in the English vocabulary or with unusual names in the Occident. But sometimes it recognizes names of other origins. Because of this, a higher failure rate is obtained in the analyzed documents.

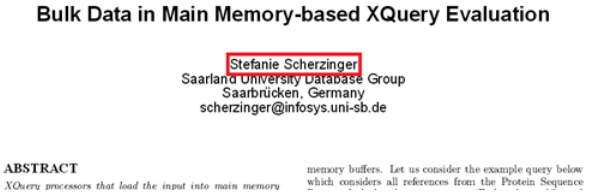


Figure 3: Author identified in text document



Figure 4: Author extracted from the document using the software application

In Figure 4 is shown the author tab that includes a text box with data extracted from the document, which matches the value of the author element contained in the digital document Figure 3.

4.3 Keyword extraction

Table 3 shows the precision with which the keywords are obtained. To determine the keywords, firstly the nouns are identified in the document, then these nouns are compared with the keywords of the taxonomy of ACM.

Table 3: The Precision in the keyword extraction

Total Documents	Correct Documents	Precision
300	297	0.99

In this experiment the probability of failure is minimal because the keywords are discovered according to the frequency of nouns, which are compared with the list of the ACM Taxonomy, taking all that are equal to that noun, to create a list of keywords. This list is filtered to show the keywords that really are in the document, discarding those that are not in the context of the document being analyzed.

When performing this process of extracting information can be found faults but not by the logic of the application but by the format of the input document, which in some cases to be written by hand and scanned to be stored, becomes unreadable content. Therefore, the application does not find matches and follows its normal flow.

In Figure 5 shows a list of phrases, which are related to nouns found in the document, and the application returns them as keywords. Next, the keywords are sorted by their relevance to the first that appears is the one that has greater importance because it is repeated more times and so on, by the number of times it is repeated during the search process. In another example, in Figure 6 an ordered list of keywords is shown (in the keywords tab), which is generated based on the nouns found when analyzing the document.

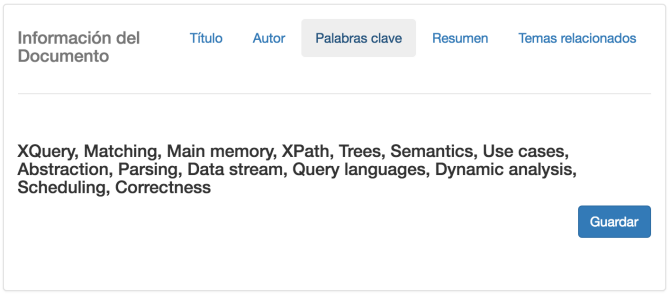


Figure 5: Phrases extracted from the text document

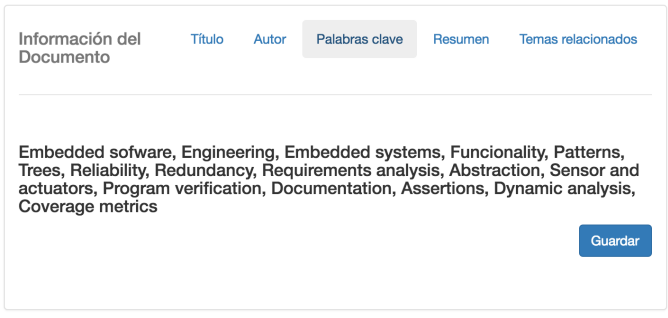


Figure 6: Keywords extracted from the text document

4.4 Extracting Summary

To perform the extraction summary process the frequency of each sentence is used (becoming ordained from highest to lowest), considering the top ten sentences. The entire list is displayed in the case of finding fewer to 10. Table 4 shows a high value of success in extracting the summary. However, there is the probability of failure that may be because there were no keywords in the document that are similar to the list of keywords of the ACM taxonomy, or can not be read the contents of the document.

Table 4: Precision obtained in the extraction summary

Total Documents	Correct Documents	Precision
300	297	0.99

Given a large number of sentences that can contain a document, we decided to display only sentences that were most relevant to that document, using as input the keyword list to extraction summary process, as shown in Figure 7.

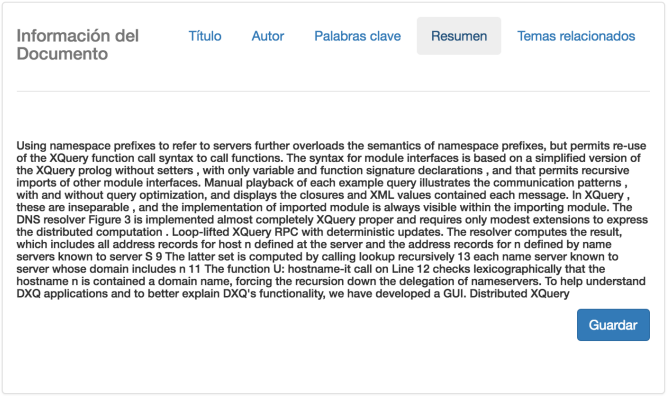


Figure 7: Summary generated from the text document

In Figure 7 shows the summary constructed from previously extracted keywords (see Figure 6). Also, shown in this Figure as the sentences are sorted according to the number of times repeated or frequency of occurrence, where the first sentence presented a higher frequency.

4.5 Document Classification

The classification of documents is carried out by the results obtained in the related theme extraction. When the processed document generates the necessary data to create a record (*index card*) with key data (title, author, keywords, summary), this index card associates the document to a theme with the highest number of keywords correctly identified, then the record is classified and stored in the database. The identification of themes is then performed from the keyword list obtained in the process *Search Keywords*, because each keyword is associated with a theme, according to the results of precision it may be observed a high success rate (see Table 5) in assigning a related theme.

Table 5: Precision obtained in the extraction of themes

Total Documents	Correct Documents	Precision
300	297	0.99

The software application has a *Related Topics* tab (see Figure 8), where the theme to which is assigned the document

is displayed. We consider only the theme most frequently at the time of processing the document.



Figure 8: Theme assigned to a document through the system

Next, two examples of the classification of documents are described, in which recall is evaluated in order to determine whether classified documents are correct for a particular topic.

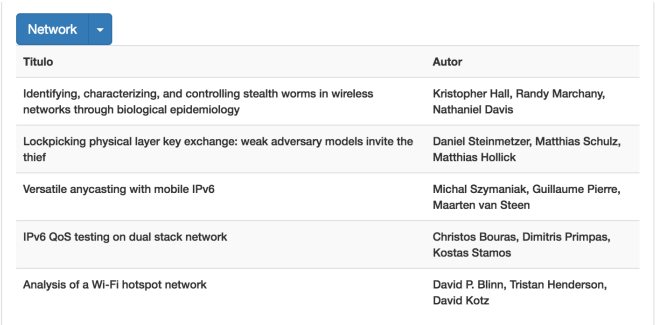


Figure 9: Classified documents for the "Network" theme

In the first example shown in Figure 9, the classification to theme "networks" is displayed, where for this nine theme documents were classified, of which eight are correct, ie relate to this theme. In Table 6 shows the result of *recall* measurement.

Table 6: The Recall value for the "Network" theme

Total Documents	Correct Documents	Recall
9	8	0.88888

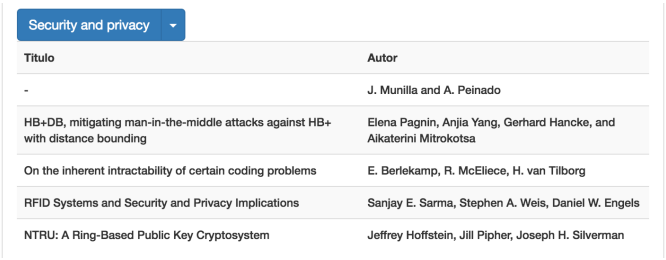


Figure 10: Classified documents for the "Security and Privacy" theme

In the second example displayed in Figure 10, the classification of documents to the theme "Security and Privacy" "is analyzed, which has eight documents, of which eight are correctly classified for this subject. In Table 7 shows the result of *recall* measurement for the "Security and Privacy" theme.

Table 7: The Recall value for the "Security and Privacy" theme

Total Documents	Correct Documents	Recall
8	8	1

5. CONCLUSIONS

This paper contributes a software application that extracts basic information from a digital document in PDF format. The basic information to get this application are the title, author, keywords and abstract from a text. Due to the difficulty of processing natural language text written, Natural Language Processing tools that allows the Named Entity Recognition (NER) and part-of-speech tagging were incorporated.

Additionally, the proposed software system allows the classification of text documents by extracting themes and their association with the themes of the ACM taxonomy for the Computer Science domain. The results obtained in the evaluation (precision and recall) of the tasks of information extraction and document classification enable considering the implementation of the proposed approach in other domains, and in a bigger corpus of text documents.

6. REFERENCES

[1] Coyotl-Morales, R. Clasificación automática de textos considerando el estilo de redacción. Master's thesis, Instituto Nacional de Astrofisica Optica y Electrónica, Puebla, Mexico, 2007.

[2] Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., Petrak, J., and Bontcheva, K. Analysis of named entity recognition and linking for tweets. *Information Processing & Management 51*, 2 (2015), 32 – 49.

[3] Espinoza, P., Vilariño, D., Pinto, D., Tovar, M., and Beltrán, B. Identificación de perfiles de usuario. *Research in Computing Science 95* (2015), 89–98.

[4] Konkol, M., Brychcín, T., and Konopík, M. Latent semantics in named entity recognition. *Expert Systems with Applications 42*, 7 (2015), 3470 – 3479.

[5] Morato, J., Sánchez-Cuadrado, S., Ruiz-Robles, A., and Moreiro-González, J.-A. Visualización y recuperación de información en la web semántica. *El profesional de la información 23*, 3 (2014).

[6] Pacheco-Luz, E., Trujillo-Romero, F., and Juárez-López, G. Clasificación semántica de textos no estructurados mediante un enfoque evolutivo. *Research in Computing Science 95* (2015), 49–59.

[7] Reyes-Ortiz, J., Bravo, M., Herrera, O., and Gudiño, A. Poblado automático de ontologías de perfiles académicos a partir de textos en español. *Research in Computing Science 95* (2015), 159–170.

[8] Wibawa, A., and Purwarianti, A. Indonesian named-entity recognition for 15 classes using ensemble supervised learning. *Procedia Computer Science 81* (2016), 221 – 228.

SEXWISE: An IBM Watson-Powered Mobile Application to Promote Sexual Education.

Miguel Royo-León
Department of Computer Science, The University of Texas at El Paso
500 W. University Ave.
El Paso, Texas, 79968.
mroyoleon@miners.utep.edu

Laura M. Rodriguez
Department of Computer Science, The University of Texas at El Paso
500 W. University Ave.
El Paso, Texas, 79968.
lmrodriguez@miners.utep.edu

Paola Gallardo
Department of Computer Science, The University of Texas at El Paso
500 W. University Ave.
El Paso, Texas, 79968.
pgallardo2@miners.utep.edu

Timothy Gonzalez
Department of Computer Science, The University of Texas at El Paso
500 W. University Ave.
El Paso, Texas, 79968.
trgonzalez@miners.utep.edu

Rosalyn Negron
Department of Public Health, City of El Paso
5115 El Paso Dr.
El Paso, Texas, 79905.
negronr@elpasotexas.gov

Diego Aguirre
Department of Computer Science, The University of Texas at El Paso
500 W. University Ave.
El Paso, Texas, 79968.
daguirre6@miners.utep.edu

ABSTRACT

Adolescents and young adults are particularly at-risk of acquiring sexually transmitted diseases or having unplanned pregnancies. They use the Internet and mobile technologies constantly for different purposes, one of which is searching information concerning sexual health.

Cognitive computing systems try to mimic the thinking process of humans. A famous success case is IBM Watson™, which uses several Artificial Intelligence techniques to *understand* the information contained in a document corpus (holding MS Word, PDF and web page documents), to answer questions.

SEXWISE is a Java-based, Android mobile phone application prototype which employs IBM Watson™'s cognitive capabilities to answer questions concerning sexuality. Watson was fed with trustworthy, officially-sanctioned and openly-accessible documents concerning sexuality and sexual health, and trained by asking it over 800 sexuality-related questions. SEXWISE also queries Google Maps™ to find health care professionals in the vicinity.

CCS Concepts

•**Information systems** → **Question answering;** •**Human centered computing** → *Mobile phones;* •**Applied computing** → Health informatics;

Keywords

Sexual Education, Mobile Application, Cognitive Computing, Question Answering Systems, IBM Watson.

1. INTRODUCTION

Adolescents are a population segment particularly vulnerable to sexual risks due to their ongoing cognitive, behavioural, emotional and social development [24]. The 2015 Sexually Transmitted Disease Surveillance Report from the Centers for Disease Control and Prevention (CDC), records some alarming statistics:

- Young people on the age range of 15-24 in the United States acquire half of the new Sexually Transmitted Diseases (STDs) registered.
- 1 of 4 sexually-active adolescents has an STD [5].

At the local level, the Teen Birth Rate (defined as the amount of births for every 1000 female adolescents) for the City of El Paso, Texas, is well over the national average (See Fig. 1).

The youth population constantly uses the Internet. As of 2010, 93% of teens (ages 12-17) and young adults (18-29) used the Internet [26]. In 2015, 92% of the teens reported using it on a daily basis, and 24% connected *"almost constantly"* [25]. Nearly $\frac{3}{4}$ of these teens either own or have access to a smartphone [25].

Thanks to the Internet, they gain access to educational material [28] and are able to use an anonymous way to find general health information, particularly about sexual health. The CDC considers important targeting young people through technology, to ensure their sexual health [24]. Sadly, a problem is that adolescents have difficulty in judging the quality of the information [28].

As young adults, the authors of this paper have witnessed the nuisances that this age group faces concerning sexuality, having experienced taboos, myths spread mouth to mouth, lack of trust in figures of authority (parents, guardians, teacher), and inability to discern the validity of the information.

Cognitive computing is a novel trend in the computing world. These systems read the information contained in documents and, by applying Artificial Intelligence processes, answer questions about the information contained in the documents [1]. Since there is trustworthy information about sexuality available in open-access documents, these files can

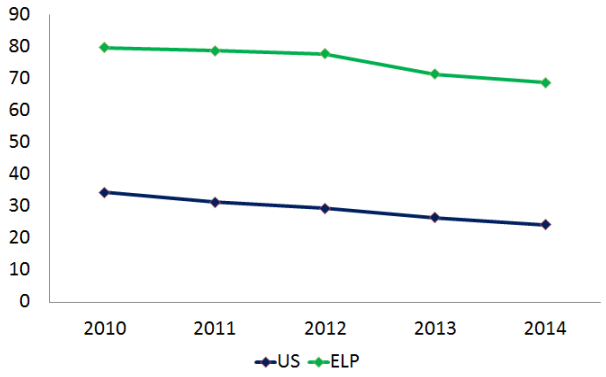


Figure 1: Teen Birth Rates for the United States and the City of El Paso. Graph created with data from [32], [6], [7], [8], [9], and [10].

be fed into cognitive systems to answer questions from users.

2. BACKGROUND

2.1 Structured / Unstructured Information

Information can be classified as structured and unstructured. **Structured information** consists in well-defined data with explicit semantics. For example, a relational data base has an established structure to store data, with explicit semantics. On the other hand, **unstructured information's** semantics are usually implicit and need to be derived from background information and through inferences. Human communication, for example, occurs in an unstructured way: we use background information to understand, and we can infer [12].

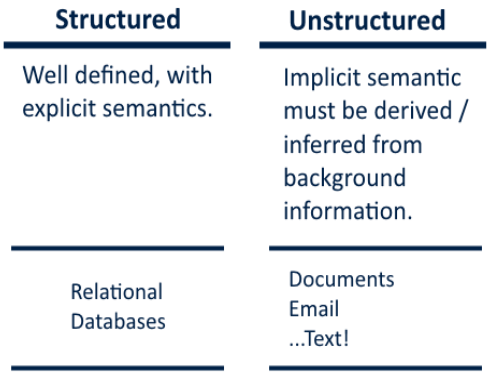


Figure 2: Structured vs. Unstructured Information: Definition and Examples.

2.2 Cognitive Computing

Cognitive computing systems try to mimic the thinking process of humans in computers, to extend what either humans or machines could do on their own, learning and interacting naturally. These systems aim to be helpful to human experts in decision making, by penetrating complex and unstructured big data [1].

2.3 IBM Watson™

IBM Watson™ is a cognitive system that became famous for participating in the Jeopardy! game show during 2011, competing against, and defeating, some of the greatest human players of the show [23]. Watson navigates through unstructured information in text passages, understands questions in natural language, and produces an answer to the question from the textual information [14]. This problem is called **open-domain question answering**, and has been studied since the 1970s [30]. Scientists working on this area use different **Artificial Intelligence (AI)** techniques from different AI areas, such as:

- Natural Language Processing (NLP).
- Knowledge Representation and Reasoning (KR&R).
- Machine Learning (ML).
- Human-Computer Interaction (HCI).

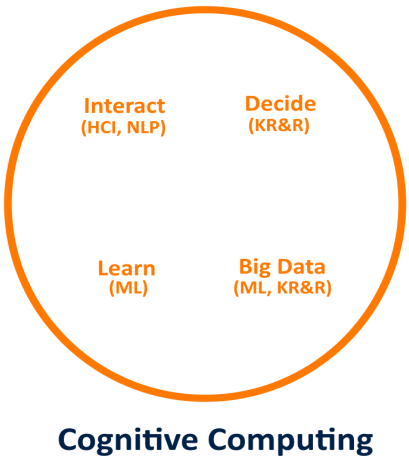


Figure 3: Cognitive Computing system characteristics and AI techniques employed.

2.3.1 *IBM Watson™'s Applications*

Watson is used for technical / scientific purposes, and with creative approaches. Let us review two examples of each:

- **IBM Watson for Oncology.** The Memorial Sloan Kettering Cancer Center in New York trained Watson as if they where teaching a medical or surgical oncologist [16] with the objective to help doctors enhance decision making in cancer clinical matters [13]. The corpus contains their own literature, over 290 medical journals, over 200 books, and 12 million text pages [20].
- **IBM Watson for Cyber Security.** The cyber security industry will need 1.5 million employees by 2020 [22]. Even if this quota is covered, this personnel will need to reach a certain skill level very quickly. Watson for Cyber Security has been trained to understand security research, attack pattern recognition, and evidence of attacks and threats [21]. The corpus contains the IBM X-Force research library, detailing 8

million spam/phishing attacks, and over 100,000 vulnerabilities. IBM plans to add about 15,000 security documents (including threat intelligence reports, cyber crime strategies, and threat databases) monthly [21].

Two examples in which Watson is being used for creative tasks are:

- **IBM Chef Watson.** Watson helps cooks creating original recipes through flavor compound algorithms [18]. It receives ingredients and figures out others that you might use with your dish [4]. The corpus contains data sets with regional / cultural knowledge and molecular / food pairing theories. After selecting the ingredients, Watson mixes them and cross references the flavour compounds with people's likes (or dislikes) [19].

- **Cognitive Dress.** On March, 2016, model Karolina Kurkova wore a dress designed by the British design studio Marchesa and IBM Watson. The dress was designed considering five sentiments and questioning Watson, fed with two data sets: a color psychology data set and a collection of runway dresses images [31].

2.3.2 *How does IBM Watson™ work?*

Humans understand and take decisions by observing the evidence at hand, interpreting the situation to generate hypotheses, evaluating these hypotheses by considering the evidence, and finally taking a decision and acting accordingly to the best hypothesis analyzed [15].

Watson tries to mimic this behavior. Evidence is **fed** to Watson as documents. It will understand the usage of the words and language of the domain. These documents must have been **curated** by experts, discarding irrelevant, outdated or poorly regarded information; even formatting the information correctly can affect the results that Watson will provide.

Thereafter, Watson will **ingest and pre-process** the document corpus to create indexes and other meta data. Finally, Watson must be **trained** by asking questions/answer pairs: the question is captured manually and the correct answer is selected amongst those provided by Watson [1].

For the version employed in this project, Watson can't be retro fed: we only can upload new questions directly through the web interface, not programmatically. Therefore, the applications must keep track of questions, in order to feed them to Watson.

When a question is sent to Watson, it will perform the following process, also illustrated in Figure 4:

1. Question Analysis: Identify the parts of speech of the question, determine its type and what is being asked, and generate different possible interpretations of the question.

2. Hypothesis Generation: Generate different hypotheses for the different interpretations, and generate a large number of possible answers for them.

3. Hypothesis and Evidence Scoring: Search for both positive and negative evidence for every hypothesis, and evaluate it.

4. Final Merging and Ranking: Considering that the type of evidence can be better suited for certain types of questions, calculate the weighted evidence scores (a degree to which evidence is useful or not) and rank the answers, to finally estimate the confidence on the answer [15].

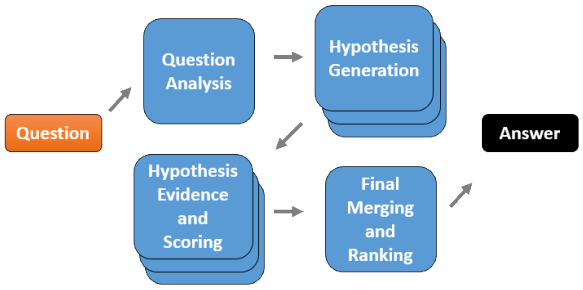


Figure 4: The DeepQA methodology, simplifed form [11].

This DeepQA (Deep Question Answering) approach is based in massive parallelism (to produce answers for different hypotheses), a pervasive confidence estimation, using many expert components, and integrating both shallow and deep knowledge. Watson employs over 100 different techniques in Natural Language analysis, source identification, hypothesis generation, evidence scoring, and hypothesis ranking and merging. IBM considers more important than the techniques, how they where combined and how they allow overlapping to contribute in speed, confidence or accuracy [11].

The Watson architecture was inspired in a great part by the structure of the Jeopardy! game, in which a player selects a question from a category and tries to answer it to win money [12]. After gathering evidence about the possible answers, Watson will rank the confidence for the answers, and give the answers for which a higher confidence has been found. There is a varying confidence threshold, defined by the precision with which questions have been answered and the amount of questions, that will define if Watson should risk and try to answer the question fast [11]. Since the Jeopardy! game is a betting game in some phases, an additional step is to evaluate the amount of money to be bet, depending on the confidence the system has on the answer.

3. METHODOLOGY

The IBM Watson™ service is provided directly by IBM from their own infrastructure. This instance reportedly has a previous version of the Watson system, and is provided for academic usage and testing [1].

This particular Watson instance was shared amongst different development teams, therefore different corpora were allocated during the development of this project. Only one corpus could be deployed at a certain time, so the querying application could be asking its questions to a different corpus that the one it was intended to.

IBM Watson™ was accessed through the IBM Watson Experience Manager™, a web-based platform used to create and manage the user projects. It allows performing tasks such as creating a corpus, adding documents to this corpus,

ingesting and pre-processing the corpus for usage, and creating training questions about the information within the documents. The interface is simple and intuitive, allowing users to easily work with it.

The ingestion time for the corpus neared 10 minutes, but we did not measure this time in a systematic manner. The corpus must be re-ingested every time that documents are added or updated.

Questions are entered to the system through the Expert Training tool, which will allow to select the answer to the question and keep track of the approval or rejection of the questions, the user that loaded them, etc. There are three options to select the answer of a question:

1. Match a Similar Question: Associates related questions that contain different terms, phrases, idioms, etc, in clusters.
2. Match an Answer: When there are either no questions or no related questions, the user should select an answer that matches the question. Watson indicates how relevant it considers that the answers are for an specific question.
3. Specify an Answer: When neither of the previous options has provided a valid answer, the user may still select directly the text section or sections that better answer the question [17].

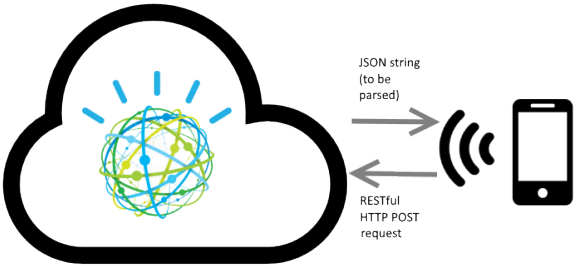


Figure 5: IBM Watson and SEXWISE Interaction.

The understanding that Watson obtains from the questions decreases if a user needs to employ the second and third options. The question-answer pairs establish the *ground-truth*. After Watson executes training runs to this ground-truth, it will create machine learning models [17].

The Watson instance also allows programmatic access to the trained corpus through the Question Answer Application Programming Interface (QAAPI), a RESTful web service [17] that receives HTTP POST requests and delivers a JSON string [2]. A computer or mobile application can query the corpus, and obtain an answer on JSON format, which must be parsed to extract the answer and other information about the documents related [2], as shown on Figure 6.

The process to create our cognitive application was:

1. Watson: Access IBM Watson™ and:
   - Create the project and the document corpus.
   - Load the documents to be used for question answering to the corpus.
   - Add the training questions and link them to the answers within the loaded documents.
   - Deploy the corpus for querying.
2. Application: Create the mobile application:
   - Develop the basic interface for the mobile application.
   - Configure the connection to IBM Watson™ instance.
   - Develop further characteristics within the application.

Fig. 7 shows two screen shots of the IBM Watson's Experience Manager.

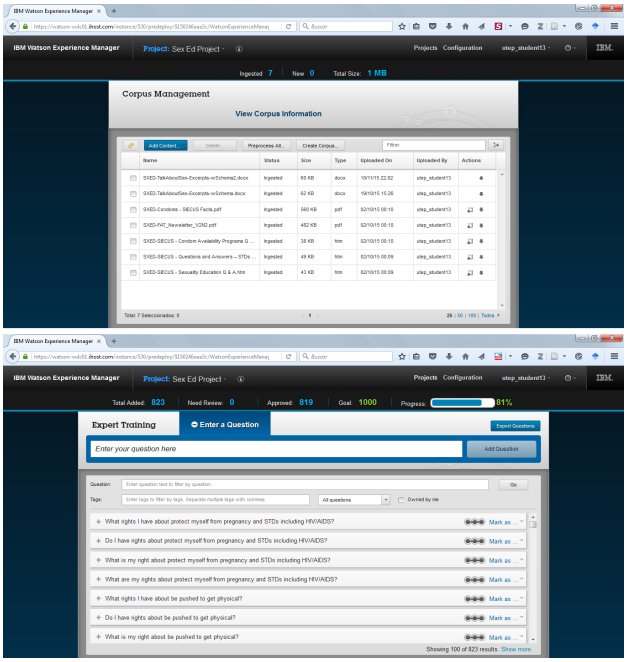


Figure 6: IBM Watson™: Document Upload and Corpus Training.

4. RESULTS

SEXWISE is a **prototype mobile application** developed to help young people in learning about sexuality, pregnancy and other topics related to a healthy sex life. It consists of two parts:

- The **document corpus**, allocated in the IBM Watson instance, whose service is provided directly by IBM under an academic agreement. Watson is key to provide appropriate answers to the questions of the users.
- The **SEXWISE mobile application** for the Android operating system (version 4.2.2 and over). The application can be executed over about 85% of the Android phone market share [3]. It was developed using the Java programming language version 1.7.0, with the Android Studio IDE version 1.3.2.

4.1 Corpus

4.1.1 Development

We developed our document corpus with **openly- accessible and officially-approved documents**, so users can trust the information and review the sources from where it was originated. All the documents in our corpus where obtained from the *Sexuality Information and Education Council of the United States* (SIECUS), an organization concerned in providing information and education about sexuality, and sexual and reproductive health [29].

The corpus contains 6 documents: 2 PDF, 3 HTML and 1 DOCX text files, comprising about 1 megabyte. The topics covered with these documents are:

- Sexuality and Behavior: Deciding, Communication, Safety, Abuse.
- STDs and HIV/AIDS.
- Sexual organs.
- Contra conceptive methods.
- Right to: Information, Expression, Deciding, Saying "No".

4.1.2 Curation

Most of the documents where processed by Watson without issue, excepting the DOCX. In order to help Watson to better capture the information of this specific document, we experimented by giving a hierarchical structure to the information. Microsoft Word allows assigning a certain paragraph type or hierarchy to the text. Watson was much more capable to understand the text as answers after the curation process.

4.1.3 Training

We trained Watson with **819 questions** related to the information from the documents, using all three system options for question-answering. During our experiments, we found that we could relate words within the questions by using the logical connective OR, allowing Watson to understand that the answers could apply for either word. This characteristic was tested for *slang* terms, given that youth commonly use them to name the sexual organs. Our corpus became capable of giving appropriate information, even though the user could use a slang word to refer to a sexual organ.

4.2 Mobile Application

We considered two main characteristics for the application prototype: answer sexuality questions and provide information about the source document from which Watson obtained the answer, and search for medical and psychological professionals in the area. A third characteristic, a user profile with personal information, was also added. Figure 7 shows screen shots of the implementation on the SEXWISE Mobile App.

SEXWISE has three tabs:

- The Watson tab, which performs two tasks:
  - Question answering interaction with Watson.
  - Display metadata (data whose primary purpose is to describe, define or annotate other data that accompanies it [27]) about the document containing the answer provided. We supply the name of the original document, a description of it, the original web address where the document is available, its origin organization name and description of the organization.
- Pros tab, with two functions:
  - Sexuality FAQ. Quick answers about sexuality.
  - Find Professionals. Launches a search on Google Maps for health professionals (medical doctors and psychologists) in the user's area.
- Profile tab. A customizable user profile for data such as Age, Sex, Sexual Orientation and Location.

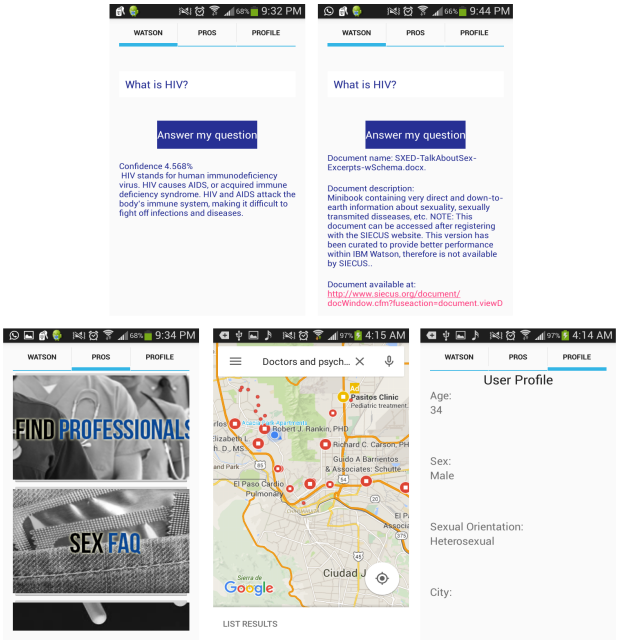


Figure 7: SEXWISE: Application Tabs.

Watson's answering time in delivering answers to the application was apparently seamless, or at least not noticeable. We would need to modify our prototype and save data in order to ascertain a deeper analysis on the answering time.

5. CONCLUSION

SEXWISE is a fully-working prototype mobile application to answers questions about sexuality, thanks to a corpus of documents and the question-answering power of IBM Watson™. SEXWISE also queries Google Maps™ to find health professionals in the vicinity.

We believe that these characteristics are appealing to young users, which may employ the application to gain a better understanding of sexuality and the responsibilities related to it.

6. FUTURE WORK

Having completed the prototype, the next step would be entering a testing phase to evaluate the application, the corpus and training questions. There is no established plan for the testing phase.

Another interest is to test the accuracy of the answers that Watson provides, by bringing the app to a group of domain experts, that may also provide feedback concerning the information and the documents that we may use.

Since Watson cannot be retrofed automatically, we would implement a question record, in order to find questions for which Watson could not provide a correct answer or no answer at all. We would use these questions as a new set of training questions.

Finally, we want to use the information from the user's profile to build the questions, and obtain appropriate answers. This would require performing further training on the corpus, to obtain personalized answers according on the user's age, gender and sexual orientation.

7. ACKNOWLEDGMENTS

We wish to acknowledge IBM, for granting access to Watson through the Department of Computer Science of The University of Texas at El Paso (UTEP). Also, to Silvia Martinez from the Healthy Miner Sex Positive Peer Education Program of UTEP, for pointing us to trustworthy information resources. Finally, to Schéhréazade Feddal and Séverine Pillado from the United Nations Educational, Scientific and Cultural Organization (UNESCO), for very kindly providing information and documentation related to sexual education. M. Royo-Leon was supported by the Fulbright-Garcia Robles scholarship, the mexican CONACyT (CVU No. 596739), the Government of the State of Chihuahua, Mexico through the Special Scholarship Program GobEdo-UTEP, and the School of Engineering of the Autonomous University of Chihuahua.

8. ADDITIONAL AUTHORS

Olac Fuentes (The University of Texas at El Paso, email: `ofuentes@utep.edu`).

9. REFERENCES

[1] D. Aguirre. Cognitive computing. Personal Collection of Diego Aguirre, 2015.

[2] D. Aguirre. JSON - Watson - Tutorial. Personal Collection of Diego Aguirre, 2015.

[3] Android Developers. Android developers dashboards, 2016.

[4] R. Bilow. How IBM's Chef Watson actually works. *Bon Appetit*, 2014.

[5] Centers for Disease Control and Prevention and others. Sexually transmitted disease surveillance, 2014. *Atlanta, GA: US Department of Health and Human Services*, 2015.

[6] Country Health Rankings and Roadmaps. 2010 Texas Data, 2010.

[7] Country Health Rankings and Roadmaps. 2011 Texas Data, 2011.

[8] Country Health Rankings and Roadmaps. 2012 Texas Data, 2012.

[9] Country Health Rankings and Roadmaps. 2013 Texas Data, 2013.

[10] Country Health Rankings and Roadmaps. 2014 Texas Data, 2014.

[11] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. A. Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, J. Prager, et al. Building Watson: An overview of the DeepQA project. *AI magazine*, 31(3):59–79, 2010.

[12] D. A. Ferrucci. Introduction to This is Watson. *IBM Journal of Research and Development*, 56(3.4):1–1, 2012.

[13] J. Fu, A. Gucalp, M. G. Zauderer, A. S. Epstein, M. G. Kris, J. Keesing, A. Caroline, M. Megerian, T. Eggebraaten, R. DeLima, et al. Steps in developing Watson for Oncology, a decision support system to assist physicians choosing first-line metastatic breast cancer (MBC) therapies: Improved performance with machine learning. In *ASCO Annual Meeting Proceedings*, volume 33, page 566, 2015.

[14] R. High. The era of cognitive systems: An inside look at IBM Watson and how it works. *IBM Corporation, Redbooks*, 2012.

[15] IBM. IBM Watson: How it works, 2014.

[16] IBM. MSK collaborates with IBM to teach Watson in the Fight Against Cancer., 2014.

[17] IBM. *IBM Watson Ecosystem Academic Partners, Getting Started Guide. v. 1.9.* IBM, 2015.

[18] IBM. IBM Chef Watson., 2016.

[19] IBM. IBM Watson Cognitive Cooking Fact Sheet., 2016.

[20] IBM. IBM Watson for Oncology., 2016.

[21] IBM. IBM Watson to Tackle Cybercrime., 2016.

[22] IBM Security. White paper: Cognitive security. Technical report, IBM, April 2016.

[23] J. Jackson. IBM Watson vanquishes human Jeopardy foes. *PC World*, 17, 2011.

[24] R. Kachur, J. Mesnick, N. Liddon, C. Kapsimalis, M. Habel, C. David-Ferdon, K. Brown, K. Gloppen, H. Tevendale, D. Gelaude, et al. Adolescents, technology and reducing risk for HIV, STDs and pregnancy. *Atlanta, GA: Centers for Disease Control and Prevention*, 2013.

[25] A. Lenhart, M. Duggan, A. Perrin, R. Stepler, H. Rainie, and K. Parker. *Teens, social media & technology overview 2015*. Pew Research Center, 2015.

[26] A. Lenhart, K. Purcell, A. Smith, and K. Zickuhr. Social media & mobile internet use among teens and young adults. millennials. *Pew internet & American life project*, 2010.

[27] P. M. Nadkarni. What is metadata? In *Metadata-driven Software Systems in Biomedicine*, pages 1–16. Springer, 2011.

[28] K. Purcell, L. Rainie, A. Heaps, J. Buchanan, L. Friedrich, A. Jacklin, C. Chen, and K. Zickuhr. How teens do research in the digital world. *Pew Internet & American Life Project*, 2012.

[29] Sexuality Information and Education Council of the United States. SIECUS Web Page, 2015.

[30] R. F. Simmons. Natural language question-answering systems: 1969. *Communications of the ACM*, 13(1):15–30, 1970.

[31] L. Stinson. IBM's Watson Helped Design Karolina Kurkova's Light-Up Dress for the Met Gala., 2016.

[32] U.S. Department of Health and Human Services. Trends in teen pregnancy and child bearing, 2016.

VM-into-PM: A Ball-into-bins scheme for Virtual Machine Allocation in Private Cloud

Víctor J. Sosa-Sosa, J.L. Gonzalez, Mario Gomez-Rodriguez
Cinvestav-Tamaulipas
Cd. Victoria, México
vjsosa,jgonzalez,mgomez@tamps.cinvestav.mx

Edwin Aldana-Bobadilla
Conacyt-Cinvestav, Tamaulipas
Cd. Victoria, México
ealdana@tamps.cinvestav.mx

ABSTRACT

This paper presents VM-into-PM, a scheme for the allocation of virtual machines (VM) into physical machines (PM) for private clouds. VM-into-PM includes a VM distributor emulator for algorithms based on the balls-into-bins metaphor; where the *balls* represent VMs running operating system and a set of applications, whereas the *bins* represent PMs chosen to receive balls. This scheme also includes an algorithm called *SortedUF* that distributes balls into a pool of bins by using utilization factors (UF), which represents the consumed portion of a bin. SortedUF calculates the size of the balls using a set of VM performance metrics and determines the most suitable bin for this ball according to its UF. In this procedure, the algorithm organizes the bins as an array sorted in ascending order and chooses the bin with lowest UF; as a result, the algorithm always chooses the bin with more available resources in the pool reducing the probability of overload in PMs. We connected VM-into-PM to a cloud monitoring tool to receive performance metrics of benchmarking applications and implemented in it a set of traditional ball-into-bin algorithms. An experimental evaluation was conducted with that implementation, which revealed the feasibility of VM-into-PM when testing large scale scenarios. It also shown that SortedUF produced a fair distribution of VMs to PMs in the cloud and produced less VM allocation errors as well as better load balancing than traditional allocation algorithms for all the studied scenarios.

Keywords

Virtual Machine Allocation; Balls-into-bins; Cloud monitoring; Emulation; Private cloud;

1. INTRODUCTION

Cloud computing is a cost-effective technology that enables organizations to face up IT issues such as management of software, platform and infrastructure in an efficient manner.

The elasticity, on-demand and pay-as-you-go model associated to this technology represents incentives for organizations to adopt it as IT solution. The on-demand feature refers to the ability to launch as many virtual machines (VM) into a pool of physical machines (PM) as required by an organization, whereas elasticity refers to the possibility to either increase or reduce the VM launched in a cloud depending on the needs of an organization. The pay-as-you-go model adapts the payment performed by organizations to the cloud service providers to the usage of the VMs launched into the cloud.

Nevertheless, delegating control over data and even metadata in this outsourcing model could result in a functional dependency with the service provider as it is shown in [5]. In this context, private cloud computing and storage has become a safety alternative for organizations to take advantage from traditional cloud computing benefits and to avoid suffering from side-effects of cloud provider dependence [6].

Private cloud computing enables organization to recover control over data and IT procedures but it also recovers the responsibility of the cloud infrastructure administration to the organization, which includes a key task as the allocation of VMs into PMs. This task, in private cloud environments, is traditionally performed by using a round-robin distribution [1]. This scheme could produce an unfair VM distribution that could affect the cloud performance in either the long term or large scale. In order to improve flexibility of cloud computing infrastructure administration, the VM allocation in cloud infrastructure managers as Open Stack currently includes a manual VM allocation scheme [2], which however is not a suitable solution for large scale.

Theoretical proposals exploring VM mapping allocation based on metrics weight schemes [7], dynamic round-robin versions [14] and genetic algorithms[15] focused on consolidation to safe energy costs have been proposed. These schemes are not quite suitable for making on-line decisions handling large scale scenarios where the performance represents the key metric for organizations. In this context, probabilistic schemes based on balls-falling-into-bins [13][4] [16] seem to be more suitable for on-line allocation because of their reduced computing requirements.

This paper presents VM-into-PM: an emulator of virtual machines allocation into physical machines for private clouds. VM-into-PM has been designed to deploy balls-into-bins allocation models, which considers the *balls* as applications installed in VMs, whereas the *bins* represent PMs chosen to receive balls (VMs running applications).

VM-into-PM includes a front-end that receives VM allocation requests, a cloud monitoring tool to extract performance metrics from PM included into the bin pool and a ball generator that creates balls for each application and calculates the size of each ball. In this emulator, a ball distributor called balls-into-balls allocator is in charge of choosing bins for balls and indexing the bin-ball maps, which can be used by a resource manager for launching VMs in the cloud.

As an additional contribution of this paper, we also have

designed a ball-falling-into-bin allocation strategy we called *SortedUF*, which chooses PMs for VMs by using the size of the balls and the bin capabilities as criteria. This strategy chooses the bin with the lowest load from a bin pool for each ball allocation; as a result, a fair distribution of VM per PM is achieved in the cloud.

We developed allocation schemes such as round-robin[14], Random[13], two choices[12] and our own strategy SortedUF. We implemented these algorithms as VM distributor in our VM-into-PM emulator. We conducted, by using this emulator, a comparative experimental evaluation of the efficiency of the allocation algorithms included into the VM-into-PM.

The evaluation revealed the feasibility of VM-into-PM for testing large scale scenarios. It also shown that SortedUF produced a fair distribution of VMs to PMs in the cloud and produced less VM allocation errors as well as better load balancing in studied scenarios than traditional allocation algorithms.

2. RELATED WORK

The workload distribution in cloud computing environments is basically managed in two step mechanisms. The first one is based on balancers that distributes load in clusters of data centers and in a second step allocators algorithms choose the PM best suitable to receive the load of a given VM [10].

The load balancing schemes are based on bio-inspired [11][3] and genetic[17] algorithms as well as dynamic versions of probabilistic allocation schemes (round-robin[14], randomized, etc). These are not realistic VM allocation solutions for large scale or on-demand scenarios based on making on-line decisions as calculation training and learning off-line stages are required and they were designed for solving the load balancing step.

In the case of allocation schemes, theoretical proposals exploring VM mapping allocation based on metrics weight schemes [7], dynamic round-robin versions [14] and genetic algorithms[15] focused on consolidation to safe energy costs have been proposed. These schemes are not quite suitable for making on-line decisions handling large scale scenarios where the performance represents the key metric for organizations. In this context, probabilistic schemes based on balls-falling-into-bins [13][4][16] seem to be more suitable for on-line allocation because of the reduced computing requirements of these schemes.

In real scenarios, the allocation of virtual machines (VM) into physical machines (PM) in private cloud environments is traditionally performed by using a round-robin distribution [1]. This scheme could produce load unbalancing affecting the cloud performance in either the long term or the large scale. In order for private cloud administrators to decide the best suitable PM for their VMs, the VM allocation in cloud builders as Open Stack currently can be performed in a manual manner [2], which however this is not a suitable solution for large scale scenarios.

3. THE VM-INTO-PM COMPONENTS

VM-into-PM has been designed as a modular scheme for designers to evaluate VM allocation algorithms/strategies.

Figure 1 depicts the modular VM-into-PM design. As it can be seen, VM-into-PM emulator includes a front-end that

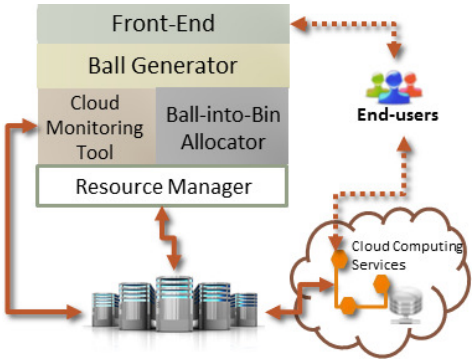


Figure 1: A conceptual representation of VM-into-PM emulator architecture

receives user requests for VMs. The user requests include a set of metrics to be taken into account in the allocation operation (applications requirements). The front-end component sends the use requests to the next component called *ball generator*. This component is in charge of determining the ball size by using a weight function that receives the performance metrics that have been defined by the users.

In order to preserve modularity, the weight function can be changed by a another one defined by the designers of allocation algorithms. A component called *ball-into-bin allocator* receives the ball size defined by the ball generator and retries the features of bins by performing calls to a *cloud monitoring tool* [8].

When the ball-into-bin allocator has retrieved the information about the ball and the bins in cloud pool, it determines the most suitable bin to receive that ball by using an allocation algorithm (round-robin, random, two-choices, etc), which produces registers about the mapping of a VM into PM. These maps enable VM-into-PM to perform tasks such as allocation and location of VMs into a cloud. At this point, the Resource Manager can launch the VM by using the maps produced by the allocator component and to send a notification to the users through the front-end. This component has not been implemented for this paper and the ball-into-bin allocator only returns the maps in a log file by now.

3.1 SortedUF: An allocation algorithm based on utilization factors

We designed an allocation algorithm based on utilization factors (UF), which represent the portion of bin resources consumed by the balls already allocated into it. A UF of a Bin *bi* can be calculated by the sum of consumed resources of a given metric by the balls already allocated (*Aballs*) as shown in the following formula:

$$UF_{bi} = (\sum_{r=0}^{Aballs} MetricCon)/MetricCap_{bi} \quad (1)$$

Where $MetricCon$ represents the consumption of a given performance metric of *bi* ($MetricCon \in bi$). $MetricCap_{bi}$ represents the whole capability of *bi* for that metric, whereas *Aballs* represents the number of balls allocated into a *bi*. For instance, if our metric of interest is the RAM of a bin, then $MetricCap_{bi}$ represents the whole RAM in *bi* and $MetricCon$ represents the consumed RAM in *bi*.

Algorithm 1: SortedUF

Input: $AppF$: Features of Applications; ω; $BinF$: Vector of Bins; n=number of bins in pool;
Output: $CBin$: Chosen bin
$Ball_Size = Create(AppF, \omega)$;
For $(i = 0; i < n; i++)$;
{
$UFArray[_i] = Av_{bi} + Ball_Size$;
}
$BinLowsetUF = SortUF(UFArray)$;
if $BinLowsetUF > 1$ **then**
    Return (Overload);
    Exit;
else
    $ChosenBin = BinLowsetUF$;
    Return ($ChosenBin$);
end

Nevertheless, some metrics could be more important for some application than others. We thus have included a weight parameter in the UF calculation to include a set of metrics in the calculation of UFs.

The following formula allows SortedUF algorithm to calculate the UF per bin taking a set of metrics into account:

$$UF_{bi} = \sum_{i=0}^{\omega}(\sum_{j=0}^{Aballs} MetricCon_{i,j})/MetricCap_{bi} \quad (2)$$

Where ω represents the number of metrics to be taken into account in the calculation of the UF of the bins.

In both formulas when $UF > 1$ the bin is overloaded and this is the situation our algorithm is trying to avoid in each ball allocation.

In both cases, the percentage of each *bi* available that the algorithm can use for an allocation operation is

$$Av = 1 - UF_{bi} whenever UF < 1 \quad (3)$$

In the case of UF including a set of metrics, this UF is managed as a vector of UF per metric; as a result, the allocator could assign a weight to each metric and make decisions in function of the importance of each metric for each application arriving to the allocator.

The SortedUF algorithm shown in algorithm 1 receives as input the features of an application running in a VM (ball) to be allocated in a bin. It creates a priority queue that registers the UFs of all bins in the pool. This priority queue is sorted in ascending order with respect to Av of the bins. When it comes to launch a new ball, the function SortUF obtains the $UF < 1$ with the lowest UF and the ball will be always assigned to the less-loaded bin in the priority queue; that is, the bin with more available resources Av.

When there are no bins with $UF < 1$ in the priority queue, the algorithm reports PM saturation to the Front-end component.

4. METHODOLOGY

We conducted an experimental evaluation by running a set of tests in a controlled environment. In this environment, we configured the ball generator to receive a workload of ball requests by using a sequential synthetic trace of 40, 400 and 4000 applications from a trace generated in advance by benchmarking applications.

We also configured the cloud monitoring tool to extract four metrics (cpu, memory, network, storage) from a pool of 10, 100 and 1000 bins; as a result, three scenarios were emulated:

Table 1: Features of the bins

Metric	Capability
CPU	100%
RAM	1 Gigabyte
HD	= 15GB
Network	100 Mbit/s

1. *Sce1:* Emulating the distribution of 40 balls into 10 bins.
2. *Sce2:* Emulating the distribution of 400 balls into 100 bins.
3. *Sce3:* Emulating the distribution of 4000 balls into 1000 bins.

With regard to the PMs, it is assumed that all of them are homogeneous and have the characteristics shown in Table 1. Notice that CPU metric was extracted on by the cloud monitoring tool as this metric is expressed in a consume percentage (CPU steal).

4.1 VM-into-PM Implementation

The four allocation algorithms implemented in the VM-into-PM emulator, the emulator and the cloud monitoring tool were developed in Java and tested in Open Stack kilo version.

4.2 Configurations

The ball-into-bin allocation of the emulator was configured with the following allocation algorithms.

- *RoundRobin*: This algorithm calculates the module between the ball number and the total number of bins in the cluster to obtain a bin where the ball will be allocated.
- *RandSChoice*: This algorithms selects a bin randomly with a uniform distribution and inserts the ball into the bin. This paradigm produces a maximum load of: $m/n + sqrt((mlogn)/n)$ where: $m \geqq nlogn$ [9] (m represents the balls and n the bins)
- *Rand2Choices* (Two Choices): This algorithm is based on Multiple-choice paradigm [12] and selects r bins randomly with a uniform distribution, with $r \geqq 2$. It inserts the ball into the less-loaded bin. The load is independent of the number of balls and produces a maximum load $m/n + (loglogn)/(logd)$ where: $m \geqq nlogn$ [9]
- *SortedUF*: This configuration represents our algorithm based on utilization factors.

4.3 Metrics

We defined two basic metrics for evaluating our proof of concept:

- *Fairness*: This metric represents the portion consumed by applications distributed into the bin pool by each configuration per scenario.

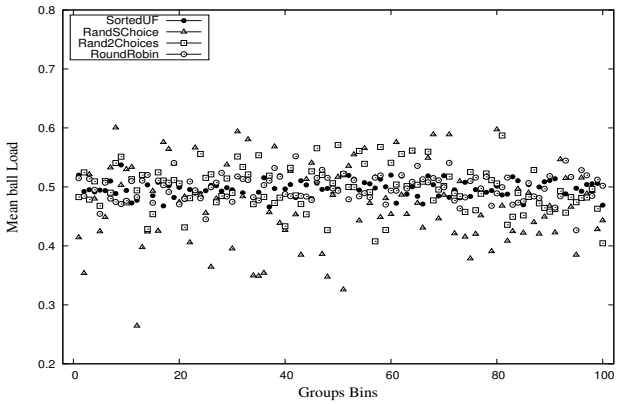


Figure 2: The fairness degree for all configurations (Sce3)

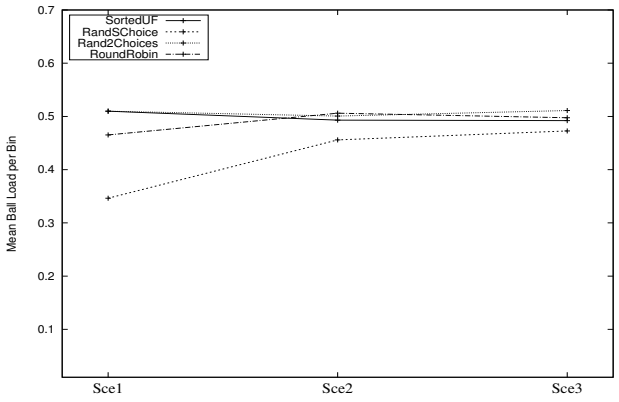


Figure 3: The average percentage of difference between ideal load and load produced by studied configurations for three scenarios

- *Allocation Error Rate*: This metric represents the allocation error rate produced by the tested configurations. An allocation error means that the allocator sent a workload Application to a bin which was overloaded.

Each emulation was launched 31 times for normalizing the average obtained per metric for each configuration. The VM were not launched in the cloud as only load balancing and allocation metrics were considered in this study.

5. EXPERIMENTAL EVALUATION

In this section, we describe the results obtained for the fairness and allocation error rate metrics for configurations tested in each studied scenario.

5.1 Fairness analysis

The fairness degree is the first metric we analyzes in this section, which is key to measure the allocation efficiency. The fairness degree is measured by the portion of bin capabilities consumed by the balls allocated into it. The fairness in the ball distribution to the the bins produces a direct impact on the cloud performance [16].

Figure 2 shows the fairness degree for the studied configurations in the third scenario (Sce3). The vertical axis shows the average portion of consumed resources by groups of 10 bins, whereas the horizontal one shows the bin pool; as a result, each point in graph represents the consumption for 1000 bins shown in groups of 10.

The basic idea is: the more close to the mean usage per bin (.5 in this graph), the more fairness degree shown by an algorithm. As it can seen, RandSChoice algorithm offered the lowest fairness degree followed by Rand2Choices, which improves its fairness when increasing the number of bins in the pool nevertheless. The fairness produced by Round-robin was close to SortedUF, which shown the best fairness values for this and all studied scenarios.

Figure 3 shows the fairness degree per configuration for three scenarios. As it can be seen, a similar behavior to Figure 2 is observed.

Figure 4 shows the fairness degree achieved by the all the studied configurations (bars) for all the scenarios (horizontal axis) as well as the amount of variation or dispersion of the measurements for each configuration. We observed that the randomized algorithms shown more difference and variation

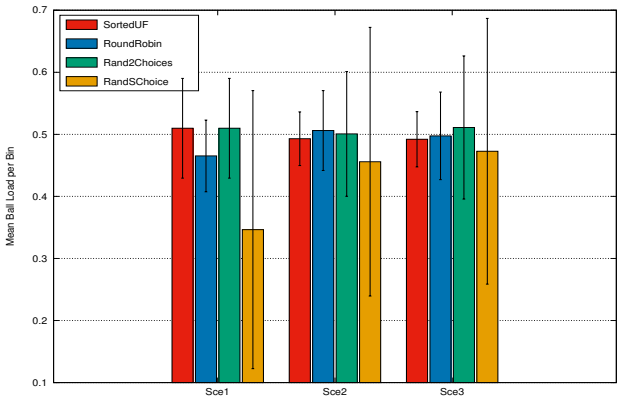


Figure 4: The fairness degree per configuration for three scenarios

from the ideal mean usage per bin (.5). than Round-robin and SortedUF configurations. SortedUF shown a difference and variation stabilized for large number of bins in comparison with Round-robin.

In addition, Round-robin can not offer a fair ball distribution in scenarios where the size of balls and bins is heterogeneous [16] as this algorithm was designed for homogeneous environments[14]. In turn, SortedUF can manage the heterogeneity of balls and bins by using weights as shown by emulations performed with homogeneous and heterogeneous size balls. This improves the flexibility offered by Round-robin strategy.

5.2 Error allocation analysis

The VM allocation errors result in overload of some performance metrics of PMs, which is an error observed by the end-users.

Figure 5 shows the number of allocation of errors (Vertical axis) per studied scenario. The randomized algorithms produced allocations that resulted in metrics overloaded in PMs.

In large scale scenarios as Sce3, Rand2Choices improves the error rate produced by RandSChoice and SortedUF has no shown errors in the three emulated scenarios.

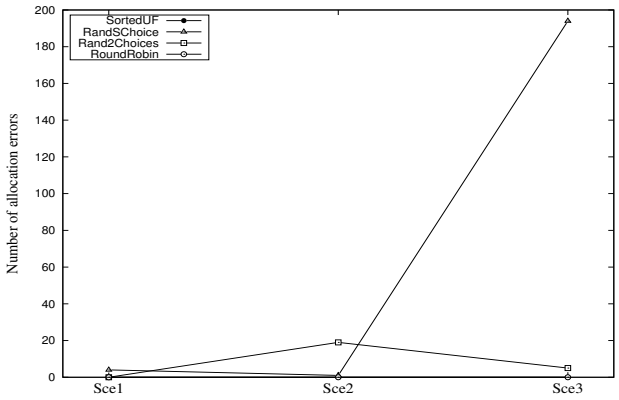


Figure 5: The number of allocation producing overload in metrics of PMs

6. CONCLUSIONS

This paper presented VM-into-PM: a scheme for the allocation of virtual machines (VM) into physical machines (PM) for private clouds. VM-into-PM includes an emulator of VM distributors based on a balls-into-bins metaphor; where the *balls* represent applications installed in VMs, whereas the *bins* represent PMs chosen to receive balls (VMs running applications).

VM-into-PM also includes an algorithm called *SortedUF*, which chooses PMs for VMs by using the size of the balls and the bin capabilities as criteria. SortedUF builds balls for applications and chooses bins to receive them. The *SortedUF* algorithm chooses bins with lowest load from a sorted array of bins; as a result, a fair distribution of VM per PM is achieved in the cloud when VM-into-PM emulator distributing balls into bins.

We connected VM-into-PM to a cloud monitoring tool to receive performance metrics of benchmarking applications and implemented traditional ball-into-bins algorithms in it. An experimental evaluation was conducted with that implementation and revealed the feasibility of VM-into-PM emulator for testing large scale scenarios. A comparative evaluation also shown that SortedUF algorithm does not produce allocation errors and yielded a better fair distribution of VMs to PMs in the cloud than traditional allocaiton algorithms, which produced VM allocation errors.

7. FUTURE AND ONGOING WORK

We are currently working on the integration of VM-into-PM in an implementation of a making-decision architecture for the automatic allocation of Virtual machines in private clouds built with Open Stack. We will conduct a comparative performance study between consolidation and load balancing schemes in private cloud management.

8. REFERENCES

[1] Opensack kilo. *https://www.openstack.org/software/kilo/*.

[2] Opensack mitaka. *https://www.openstack.org/software/mitaka/*.

[3] B. Balusamy, J. Sridhar, D. Dhamodaran, and P. V. Krishna. Bio-inspired algorithms for cloud computing: A review. *Int. J. Innov. Comput. Appl.*, 6(3/4):181–202, Nov. 2015.

[4] P. Berenbrink, A. Brinkmann, T. Friedetzky, and L. Nagel. Balls into non-uniform bins. *Journal of Parallel and Distributed Computing*, 74(2):2065 – 2076, 2014.

[5] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina. Controlling data in the cloud: outsourcing computation without outsourcing control. In *Proceedings of the 2009 ACM workshop on Cloud computing security*, CCSW '09, pages 85–90, 2009.

[6] D. C. P. B. S. Emerson Network Power Independently conducted by Ponemon Institute LLC. Cost of data center outages, January 2016.

[7] A. E. Ezugwu, S. M. Buhari, and S. B. Junaidu. Virtual machine allocation in cloud computing environment. *Int. J. Cloud Appl. Comput.*, 3(2):47–60, 2013.

[8] J. L. G. J. A. Perez-Espinoza, Victor J. Sosa-Sosa. A distributed architecture for monitoring private clouds. In *DEXA Workshops*, pages 186–190, 2015.

[9] V. R. U. W. J. Y. John MacCormick, Nicholas Murphy and L. Zhou. Kinesis: A new approach to replica placement in distributed storage systems. *Trans. Storage 4, 4, Article 11, 28 pages.*, 2009.

[10] R. Lee and B. Jeng. Load-balancing tactics in cloud. In *In proc. International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), IEEE*, pages 447–454, 2011.

[11] X. G. Z. G. D. Y. . W. D. Li, K. Cloud task scheduling based on load balancing ant colony optimization. *Sixth Annual ChinaGrid Conference*, 2011.

[12] M. Mitzenmacher. The power of two choices in randomized load balancing. *IEEE Trans. Parallel Distrib. Syst.*, 12(10):1094–1104, 2001.

[13] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, New York, NY, USA, 2005.

[14] D. M. . L. P. M. S. Mohanty, R. Design and performance evaluation of a new proposed fittest job first dynamic round robin (fjf- drr) scheduling algorithm. *International Journal of Computer Information Systems, Vol. 2, No. 2*, pages 23–93, 2011.

[15] N. Quang-Hung, P. D. Nien, N. H. Nam, N. H. Tuong, and N. Thoai. A genetic algorithm for power-aware virtual machine allocation in private cloud. *CoRR*, abs/1302.4519, 2013.

[16] M. Quezada-Naquid, R. Marcelín-Jiménez, J. L. Gonzalez-Compeán, and J. C. Perez. Rs-pooling: an adaptive data distribution strategy for fault-tolerant and large-scale storage systems. *The Journal of Supercomputing*, 72(2):417–437, 2016.

[17] C. Zhao, S. Zhang, Q. Liu, J. Xie, and J. Hu. Independent tasks scheduling based on genetic algorithm in cloud computing. In *Proceedings of the 5th International Conference on Wireless Communications, Networking and Mobile Computing*, pages 5548–5551, 2009.

Dos Aproximaciones para Expandir Consultas en un SRIB Usando Sinónimos*

Ana Laura Lezama Sánchez
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
yumita1102@gmail.com

Mireya Tovar Vidal
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
mireyatovar@gmail.com

Darnes Vilariño Ayala
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
dvilarinoayala@gmail.com

RESUMEN

En el presente trabajo presentamos una aproximación de expansión de consultas por sinónimos para un Sistema de Recuperación de Información Booleano (SRIB). En [12] se presenta una primera aproximación de expansión de consultas para el SRIB, extrayendo los sinónimos de WordNet, de los conceptos o consultas completos, dichos conceptos fueron extraídos de 4 ontología de dominio utilizadas en este trabajo. En [13] se presenta una segunda aproximación para la expansión de consultas para el mismo SRIB, pero esta vez extrayendo los sinónimos de las palabras que conforman al concepto original, para después realizar nuevas consultas con los sinónimos extraídos y los conceptos originales que componen a la consulta. Ahora en esta propuesta, se divide las palabras de los conceptos y se realiza un tratamiento independiente que posteriormente son buscados en WordNet, la propuesta obtiene resultados satisfactorios al compararlo con el SRI Booleano sin expansión.

Palabras clave

Expansión de consultas; sistema de recuperación de información booleano; ontologías.

1. INTRODUCCIÓN

La Recuperación de Información *(RI)*, nace ante la creciente cantidad de información en texto, utilizada en diferentes ramas y hasta en la vida diaria, ya que se debe de acceder a ella de manera precisa y rápida, y es por ellos que [9], definen la *(RI)* como representación, almacenamiento, organización y acceso a los elementos de información necesaria. Para [2] los Sistemas de Recuperación de Información son aquellos que son capaces de almacenar, recuperar y mantener la infomación. En este contexto la información puede componerse de texto, imágenes, audio, video y otros objetos multimedia. En este trabajo, la información utilizada en los experimentos es de tipo texto. Un sistema de RI que tiene como modelo de recuperación el modelo booleano, es donde una vez indizados los documentos, y ante la pregunta de un usuario, existen dos objetos, la entidad documento que se representa mediante un conjunto de términos de indización que expresan los diversos temas contenidos en el documento y el objeto pregunta, que se representa también por otro conjunto de términos de indización, pero en este caso combinados estos términos entre sí mediante operadores booleanos (AND, OR, NOT).

En esta investigación se parte de un Sistema de Recuperación de Información Booleano que permite recuperar documentos de un corpus de dominio, asociados a cada concepto y relaciones de una ontología de dominio. Las palabras de tales conceptos y relaciones son utilizadas para formar las consultas que se emplean en la entrada a dicho sistema, todo esto con la finalidad de mejorar los índices de precisión de un SRIB que no realiza expansión.

Este artículo se encuentra estructurado de la siguiente manera, en la sección 2, se dan a conocer algunos trabajos de diversos autores relacionados con la RI con expansión de las consultas, en la sección 3 se describen los pasos que sigue el algoritmo desarrollado para lograr la expansión de las consultas, en la sección 4 se muestran los resultados obtenidos con el algoritmo desarrollado, y en 5 las conclusiones a las que hemos llegado después de que el algoritmo proporcionara los resultados, y la bibliografía que sustenta esta investigación.

2. ESTADO DEL ARTE

En [3] se presenta un algoritmo y su variación para la realización de la expansión de consultas en un Sistema de Recuperación de Información . Dicho algoritmo y su variación son evaluados por la similud coseno y el algoritmo de Roccio, que es el algoritmo propuesto por los autores para llevar a cabo la expansión de las consultas, donde se destacan elementos como el conjunto de documentos evaluados como relevantes, un conjunto de documentos evaluados como no relevantes y una consulta expresada como vector de términos con pesos.

*Esta investigación es parcialmente apoyada por el proyecto PRODEP-SEP ID 00570 (EXB-792) DSA/103.5/15/10854, por el proyecto ID 00570 VIEP-BUAP. Apoyado por el Fondo Sectorial de Investigación para la Educación, proyecto Conacyt 257357.

En [1] se diseñó un Sistemas de Recuperación de Información basado en ontologías de dominio, también realiza la expansión de consultas de forma automática sin la intervención del usuario, esto lo logra mediante el uso de ontología de dominio, el propósito de dicha ontología será añadir nuevos términos que permitan proveer un contexto más preciso para la búsqueda y obtener mejores resultados en la relevancia de los documentos recuperados.

En [8] crearon un modelo que permitió la selección y generación de un conjunto de consultas reformuladas para el proceso de búsqueda siguiendo las condiciones del usuario en la consulta original, con ello el sistema es capaz de generar un conjunto de consultas a partir del intercambio de términos de la consulta original por otros obtenidos de una fuente de conocimiento u ontología. Utilizan la plataforma BUDI, que se centra en la implementación de soluciones de recuperación de información basadas en técnicas como lógica, entre otras.

En [4] proponen un Sistema de Recuperación de Información aplicado a contextos de documentación médica, que utiliza las principales métricas de similitud semántica. Dentro del estado del arte realizan un estudio de las diferentes técnicas de expansión de consultas en el campo Biomédico. Durante el procesado de las consultas, cada consulta o tópico, es procesada mediante parametrización, hecha con la herramienta MetaMap, que se encarga de descomponer la cadena o frase original en cadenas sencillas denominadas frases que caracterizan síntomas, partes del cuerpo, etc., y por cada frase creada se obtienen diferentes conceptos que los representan, y por ende cada cadena o frase se puede descomponer en varios términos y cada uno puede ser representado por diferentes conceptos biomédicos, dichos conceptos biomédicos de las frases son combinados entre sí dando paso a subfrases por cada frase, realizando de esta manera la expansión de la consulta.

En [7] se realiza el estudio, implementación y experimentación de métodos que automaticen el proceso de extracción de información trascendente de documentos, utilizándola para expandir o reformular consultas, de esta forma se pretende mejorar la precisión de los resultados asociados, tales métodos serán integrados en la herramienta Ontology Annotator y enfocados a la fuente de datos de PubMed, estudian diferenes modelos de expansión de consultas, pero el método que utilizan en su implementación es el método de expansión de consultas por el modelo de espacio vectorial.

En [10] se propone una aproximación de expansión de consultas, usando una ontología del dominio del mercado bursátil de Madrid. Mostraron una prueba de concepto de expansión de la consulta usando una ontología pero los resultados no permiten asegurar la utilidad del enfoque, para ello sería necesario escalar el tamaño de la ontología y aplicarlo a un problema de recuperación de información financiera.

3. ALGORITMO PROPUESTO

Se proponen dos nuevos algoritmos de expansión de consultas, con el uso de sinónimos extraídos desde WordNet, pero en este caso, lo hemos clasificado como primera y segunda triangulación. Para ambas triangulaciones, se realiza un preprocesamiento, donde se obtienen las nuevas consultas, y esas se ingresan a WordNet para extraer todos los sinónimos correspondientes. Las consultas estan formadas por las palabras que integran a los conceptos extraídos de las ontologías de dominio. A continuación se describe el algoritmo propuesto con sus dos variantes:

1. Extracción de conceptos y relaciones de cada ontología de dominio utilizando Jena[1]. Denominamos concepto original, al concepto extraído de la ontología de dominio.

2. Tipos de triangulación:

   (a) Primera triangulación
   Una vez obtenidos los conceptos de cada ontología de dominio, se ingresan los conceptos originales de cada ontología para realizar el primer triangulado, y dar paso al algoritmo que nos permitirá llevar a cabo la primera triangulación.

      i. Se fracciono el concepto original, después esas pequeñas consultas extraerán sinónimos desde WordNet [5], por ejemplo, para nuestra consulta pal_1 pal_2 pal_3 pal_4, su primera triangulación quedaría de la siguiente forma:
      A. pal_1
      B. pal_2 pal_3 pal_4
      C. pal_1 pal_2
      D. pal_3 pal_4
      E. pal_1 pal_2 pal_3
      F. pal_4

   (b) Segunda triangulación
   Igual que en la primera triangulación se ingresan los conceptos originales de cada ontología para realizar ahora el segundo triangulado, y dar paso al algoritmo que nos permitirá llevar a cabo la segunda triangulación.

      i. Separaramos los pares de palabras o palabra según sea el caso en longitud del concepto original que serán los que obtendrán sinónimos en WordNet [5], y el resto que ingresará el SRIB de forma normal, dicha triangulación se obtiene de tal manera que el concepto original quede fraccionado en pequeños trozos de palabras, por ejemplo para la misma consulta pal_1 pal_2 pal_3 pal_4, su segunda triangulación quedaría de la siguiente forma:
      A. $*pal_1$ pal_2
      B. pal_3 pal_4
      C. pal_1
      D. $*pal_2$ pal_3
      E. pal_4
      F. pal_1 pal_2
      G. $*pal_3$ pal_4
      En este caso los asteriscos, le indican al algoritmo que pares o que palabra son los que extraerán sinónimos de WordNet, ya que como explicabamos anteriormente no todas las pequeñas consultas, de esta segunda triangulación, que hemos obtenido al triangular, extraerán si es que cuentan con ellos, sinónimos de WordNet [5].

   (c) Extracción de los sinónimos con WordNet para las dos formas de triangular el concepto original.

[1] https://jena.apache.org/

3. Preprocesamiento del corpus de dominio, de los conceptos, de las relaciones, de los sinónimos y de los resultados sin sinónimos, para el caso de la segunda triangulación. Esta etapa incluye las siguientes acciones:
   (a) División del corpus en líneas.
   (b) Eliminación de símbolos especiales, números y palabras cerradas.
   (c) Aplicación de un lematizador, en particular se utiliza el algoritmo de Porter [6].
4. Formación de consultas. Existen cuatro tipos de consultas para la aproximación propuesta
   (a) Consultas formadas con las palabras del concepto.
   (b) Consultas formadas con los sinónimos de la o las palabras del primer triangulado.
   (c) Consultas formadas con los sinónimos de la o las palabras con asterisco del segundo triangulado y palabras del concepto original que complementan las palabras asignadas para la búsqueda de sinónimos
   (d) Consultas formadas con las palabras del concepto que forman la relación semántica.

   (a) Primera triangulación
      i. Aplicación del Sistema de Recuperación de Información Booleno (SRIB) para conceptos, sin expansión.
      ii. Aplicación del Sistema de Recuperación de Información Booleno (SRIB) para los sinónimos de los conceptos, con expansión.
      iii. Mezcla o unión de los resultados obtenidos (posting) por el SRIB para resultados de la aproximación y con sinónimos de los dos pasos anteriores.
      iv. Aplicación del operador AND para la consulta que incluye los dos conceptos que forman la relación semántica. El operador AND realiza la intersección de las líneas que integran los posting de ambos conceptos que forman la relación semántica.

   (a) Segunda triangulación
      i. Aplicación del Sistema de Recuperación de Información Booleno (SRIB) para conceptos, sin expansión.
      ii. Aplicación del Sistema de Recuperación de Información Booleno (SRIB) para los sinónimos de los conceptos, con expansión, sin expansión con resultados obtenidos durante la expansión y los que no ingresaron a WordNet, pero forman parte de la triangulación.
      iii. Mezcla o unión de los resultados obtenidos (posting) por el SRIB para resultados de la aproximación y con sinónimos de los dos pasos anteriores.
      iv. Aplicación del operador AND para la consulta que incluye los dos conceptos que forman la relación semántica. El operador AND realiza la intersección de las líneas que integran los posting de ambos conceptos que forman la relación semántica.

Tabla 1: Conjunto de datos

Dominio	Ontología			Corpus de referencia			
	C	T	NT	D	T	V	O
AI	276	205	61	8	11,370	1,510	475
SCORM	1,461	1,038	759	36	1,621	34,497	1,325
OIL	48	37	-	577	546,118	10,290,107	168,554
Turismo	963	1,016	-	1,801	877,519	32,931	36,505

En el caso de la evaluación de los resultados obtenidos, se utilizan las Ecuaciones 1 y 2 para medir la precisión a nivel de conceptos y relaciones.

$$P_C = \frac{Conceptos\ recuperados}{Total\ conceptos} \quad (1)$$

$$P_R = \frac{Relaciones\ recuperadas}{Total\ relaciones} \quad (2)$$

Donde: *Conceptos recuperados* es el total de conceptos obtenidos por el SRIB, y el *Total conceptos* es el total de conceptos existentes en la ontología de dominio. En el caso de *Relaciones recuperadas* se evalúa por separado las relaciones taxonómicas y las relaciones no taxonómicas (para más información ver [11]). El *Total relaciones* corresponden al total de relaciones de cada tipo existentes en la ontología de dominio evaluadas de manera independiente.

4. RESULTADOS

En esta sección se presenta el conjunto de datos (ver 4.1) utilizado para los experimentos realizados, así como los resultados experimentales obtenidos (ver 4.2) con el algoritmo propuesto.

4.1 Conjunto de datos

En la Tabla 1 se presenta el número de conceptos (C), el total de relaciones taxonómicas (T) y el total de relaciones no taxonómicas (NT) de las ontologías evaluadas. También se incluye el número de documentos (D), el número de tokens (T), la cantidad de vocabulario (V), y el número de oraciones. Todos los datos descritos anteriormente, fueron obtenidos de cada ontología de dominio utilizada, los dominios utilizados en los experimentos son Inteligencia Artificial (IA), Aprendizaje e-Learning (SCORM) [14], ontología del dominio de Petróleo (OIL), y Turismo (Turismo), por lo tanto hacemos mención de que se tratan de ontologías ya existentes en la literatura.

4.2 Resultados experimentales

A continuación se presentan los resultados experimentales obtenidos por las dos formas de triangular las consultas, y su comparación, es decir, resultados del Sistema de Recuperación de Información Booleno (SRIB) sin expansión de consultas y del Sistema de Recuperación Información Booleano (SRIB) con expansión de consultas, para la aproximación presentada.

4.2.1 Primera Triangulación

En la Tabla 2 se presentan los resultados experimentales obtenidos por la primera triangulación propuesta y su comparación, es decir, resultados del Sistema de Recuperación de Información Booleno (SRIB) sin expansión de consultas y del Sistema de Recuperación Información Booleano (SRIB) con expansión de consultas. Los resultados obtenidos por

Tabla 2: Resultados de la primera triangulación de la aproximación propuesta, del Sistema de Recuperación Booleano con expansión y sin expansión, para el caso de los conceptos.

Dominio	Ontología							Corpus			
	CO	C	F	P	CE	FE	PE	O	OE	DI	%
IA	276	274	2	0.992	275	1	0.996	1,994	3,720	1,726	86.55
SCORM	1,461	1,434	27	0.981	1,440	21	0.985	23,406	54,634	31,228	133.41
OIL	48	48	0	1.00	48	0	1.00	232,603	303,441	70,838	30.45
Turismo	963	682	281	0.708	797	166	0.827	86,353	251,215	164,862	190.91

Tabla 3: Resultados de la primera triangulación de la aproximación propuesta, del Sistema de Recuperación Booleano con expansión para el caso de las relaciones taxonómicas de cada ontología de dominio.

Dominio	Ontología							SRI			
	RT	RR	F	P	RRE	FE	PE	O	OE	DI	%
IA	205	205	0	1.00	205	0	1.00	782	967	185	23.65
SCORM	1,038	1,002	36	0.965	1,004	34	0.967	10,640	13,351	2,711	25.47
OIL	37	32	5	0.864	32	5	0.864	12,696	13,616	920	7.24
Turismo	1,016	291	725	0.286	454	562	0.446	5,606	22,837	17, 231	307.36

ambos algoritmos, para el caso de los conceptos, se muestran en la Tabla 2 para cada ontología revisada (Dominio). También se muestra el total de conceptos extraídos de la ontología (CO), los conceptos recuperados por el SRIB sin expansión (C), los conceptos que no obtuvieron líneas asociadas (F) y la precisión (P); los conceptos recuperados por el SRIB con expansión (CE), los conceptos que no logró recuperar el SRIB con expansión (FE) y la precisión obtenida (PE). Además, en la tabla se incluye la cantidad de oraciones obtenidas por el SRIB sin expandir (O), con expansión (OE), la diferencia del número de líneas recuperadas con expansión y sin ella (DI) y el porcentaje de incremento (%).

En base a los resultados obtenidos para los conceptos, se observa que en los casos de los dominios de SCORM y Turismo principalmente, se incrementó el número de conceptos recuperados que los que se recuperan con el SRIB sin expansión. Además, la cantidad de oraciones que contienen los sinónimos del concepto incrementa la cantidad de líneas u oraciones asociadas a cada concepto de las ontologías, esto ocurre para cada dominio. El porcentaje de incremento de la información recuperada en número de líneas, por el SRIB con expansión es mayor al 30%, lo que indica que el concepto puede ser representado en el corpus por su sinónimo correspondiente y que esta información es adicional a la presentada por el SRIB sin expansión. En la Tabla 3 se presentan los resultados obtenidos por ambos Sistemas de Recuperación de Información con expansión y sin ella, para relaciones de tipo taxonómicas de cada ontología de dominio. La columna RT corresponde al total de relaciones taxonómicas incluidas en la ontología de dominio correspondiente. La columna RR es el total de relaciones recuperadas con información del SRI sin expansión y con expansión RRE. La columna correspondiente a F es la diferencia de las relaciones recuperadas por el SRI booleano sin expansión y con expansión (FE). La precisión del sistema sin expansión (P) y con expansión (PE). También se incluye la cantidad de oraciones recuperadas en total por el SRIB sin expansión (O) y con expansión (OE) para este tipo de relaciones, la diferencia obtenida (DI) y el porcentaje de la diferencia (%).

En base a los resultados obtenidos se observa que el número de relaciones de tipo taxonómicas de las tres primeras ontologás se mantienen por los dos algoritmos diseñados, pero en el caso de la ontología de Turismo el número de relaciones se incrementa de 291 a 454 esto indica que existen relaciones en el corpus que sólo se pueden encontrar por su correspondiente sinónimo y al SRIB sin expansión no le es posible encontrarlo exactamente. También, la cantidad de oraciones asociadas a los SRIB con expansión se incrementa para las cuatro ontologías y más aún para la ontología de Turismo, reforzando nuevamente la existencia de los sinónimos de las relaciones encontradas en el corpus.

En el caso de las relaciones tipo no taxonómicas, que sólo las ontologías IA y SCORM tienen, se observa que la cantidad de relaciones recuperadas es la misma para ambos sistemas. La columna RNT corresponde, al número de relaciones no taxonómicas incluídas en cada ontología de dominio correspondiente. Las columnas R y RE, corresponde al número de relaciones obtenidas por el SRI sin expansión, y con expansión, respectivamente. Las columnas FE y F, corresponden a la diferencia de las relaciones recuperadas por el SRI con expansión y sin expansión, respectivamente. Las columnas P y PE, corresponde a la precisión del sistema sin expansión y con expansión respectivamente. Las columnas O y OE, corresponde a la cantidad de oraciones recuperadas por el SRI sin expansión y con expansión respectivamente y la columna DI, corresponde a la diferencia de oraciones recuperadas del sistema con expansión y el sistema sin expansión. En este caso, sólo se incrementaron el número de oraciones en las cuales existen el sinónimo correspondiente a cada concepto que forma la relación (ver Tabla 4).

Tabla 4: Resultados de la primera aproximación, para relaciones no taxonómicas.

Dominio	Ontología							SRI			
	RNT	R	F	P	RE	FE	PE	O	OE	DI	%
IA	61	61	0	1.00	61	0	1.000	108	137	29	26.85
SCORM	759	738	21	0.972	739	20	0.973	8,728	10,689	1,961	22.46

4.2.2 Segunda Triangulación

En la tabla 5, se presentan los resultados experimentales obtenidos por la triangulación propuesta y su comparación, es decir, resultados del Sistema de Recuperación de Información Booleno (SRIB) sin expansión de consultas y del Sistema de Recuperación Información Booleano (SRIB) con expansión de consultas, propuesta en este trabajo. Los resultados obtenidos por ambos algoritmos, para el caso de los conceptos, se muestran en la Tabla 5 para cada ontología revisada (Dominio). También se muestra el total de conceptos extraídos de la ontología (CO), los conceptos recuperados por el SRIB sin expansión (C), los conceptos que no obtuvieron líneas asociadas (F) y la precisión (P); los conceptos recuperados por el SRIB con expansión (CE), los conceptos que no logró recuperar el SRIB con expansión (FE) y la precisión obtenida (PE). Además, en la tabla se incluye la cantidad de oraciones obtenidas por el SRIB sin expandir (O), con expansión (OE), la diferencia del número de líneas recuperadas con expansión y sin ella (DI) y el porcentaje de incremento (%). En base a los resultados obtenidos para los conceptos, se observa que en los casos de los dominios de SCORM y Turismo principalmente, se incrementó el número de conceptos recuperados que los que se recuperan con el SRIB sin expansión. Además, la cantidad de oraciones que

Tabla 5: Resultados del SRIB con expansión y sin expansión para la segunda triangulación propuesta, para el caso de los conceptos.

Dominio	Ontología							Corpus			
	CO	C	F	P	CE	FE	PE	O	OE	DI	%
IA	276	274	2	0.992	275	1	0.996	1,994	13,275	11,281	5565.74
SCORM	1,461	1,434	27	0.991	1,448	13	0.982	23,406	418,490	395,084	1687.96
OIL	48	48	0	1.00	48	0	1.00	232,603	567,302	334699	143.89
Turismo	963	682	281	0.708	811	152	0.738	86,353	362,442	276,089	319.72

contienen los sinónimos del concepto incrementa la cantidad de líneas u oraciones asociadas a cada concepto de las ontologías, esto ocurre para cada dominio. El porcentaje de incremento de la información recuperada en número de líneas, por el SRIB con expansión es mayor al 143%, lo que indica que el concepto puede ser representado en el corpus por su sinónimo correspondiente y que esta información es adicional a la presentada por el SRIB sin expansión. En la Tabla 6 se presentan los resultados obtenidos por ambos Sistemas de Recuperación de Información con expansión y sin ella, para relaciones de tipo taxonómicas de cada ontología de dominio. La columna RT corresponde al total de relaciones taxonómicas incluidas en la ontología de dominio correspondiente. La columna RR es el total de relaciones recuperadas con información del SRI sin expansión y con expansión RRE. La columna correspondiente a F es la diferencia de las relaciones recuperadas por el SRI booleano sin expansión y con expansión (FE). La precisión del sistema sin expansión (P) y con expansión (PE). También se incluye la cantidad de oraciones recuperadas en total por el SRIB sin expansión (O) y con expansión (OE) para este tipo de relaciones, la diferencia obtenida (DI) y el porcentaje de la diferencia (%). En base a los resultados obtenidos se observa que el número de relaciones de tipo taxonómicas de las tres primeras ontologás se mantienen por los dos algoritmos diseñados, pero en el caso de la ontología de Turismo el número de relaciones se incrementa de 291 a 479 esto indica que existen relaciones en el corpus que sólo se pueden encontrar por su correspondiente sinónimo y al SRIB sin expansión no le es posible encontrarlo exactamente. También, la cantidad de oraciones asociadas a los SRIB con expansión se incrementa para las cuatro ontologías y más aún para la ontología de Turismo, reforzando nuevamente la existencia de los sinónimos de las relaciones encontradas en el corpus.

Tabla 6: Resultados de la segunda aproximación, del Sistema de Recuperación Booleano con expansión para el caso de las relaciones taxonómicas de cada ontología de dominio.

Dominio	Ontología							SRI			
	RT	RR	F	P	RRE	FE	PE	O	OE	DI	%
IA	205	205	0	1.00	205	0	1.00	782	4278	3496	447.0
SCORM	1,038	1,002	36	0.965	1,024	14	0.986	10,640	56,006	45,366	426.37
OIL	37	32	5	0.864	36	1	0.972	12,696	25,937	13,241	104.29
Turismo	1,016	291	725	0.286	479	537	0.471	5,606	37,080	31,474	561.43

En el caso de las relaciones tipo no taxonómicas, que sólo las ontologías IA y SCORM tienen, se observa que la cantidad de relaciones recuperadas es la misma para ambos sistemas. La columna RNT corresponde, al número de relaciones no taxonómicas incluídas en cada ontología de dominio correspondiente. Las columnas R y RE, corresponde al número de relaciones obtenidas por el SRI sin expansión, y con expansión, respectivamente. Las columnas FE y F, corresponden a la diferencia de las relaciones recuperadas por el SRI con expansión y sin expansión, respectivamente. Las columnas P y PE, corresponde a la precisión del sistema sin expansión y con expansión respectivamente. Las columnas O y OE, corresponde a la cantidad de oraciones recuperadas por el SRI sin expansión y con expansión respectivamente y la columna DI, corresponde a la diferencia de oraciones recuperadas del sistema con expansión y el sistema sin expansión. En este caso, sólo se incrementaron 13 relaciones recuperadas y algunas oraciones en las cuales existen el sinónimo correspondiente a cada concepto que forma la relación (ver Tabla 7).

Tabla 7: Resultados de la segunda aproximación, para relaciones no taxonómicas.

+

Dominio	Ontología							SRI			
	RNT	R	F	P	RE	FE	PE	O	OE	DI	%
IA	61	61	0	1.000	61	0	1.000	108	229	121	112.03%
SCORM	759	738	21	0.972	751	8	0.972	8,728	24,037	15,309	175.40%

5. CONCLUSIONES

Después de los experimentos realizados en esta nueva aproximación presentada con dos variantes, hemos comprobado que la presencia de más sinónimos en una consulta, hace posible que el SRIB sea capaz de recuperar mayor cantidad de información, y por ende la información solicitada por el usuario, se verá aun más enriquecida.

Esta investigación nos señala que la inserción de sinónimos, mejora notablemente los niveles de precisión del SRIB, ya que el sistema fue capaz de recuperar una mayor cantidad de información.

Como trabajo a futuro, se proponen el uso de patrones léxico-sintácticos para la extracción de relaciones de tipo sinónimia, como una forma alternativa al uso de WordNet.

6. REFERENCIAS

[1] B. G. Carranza Chávez. Diseño de un modelo de recuperación de información usando expansión de consultas basadas en ontologías en el dominio de la ciencia de la computación, 2014.

[2] G. J. Kowalski and M. T. Maybury. *Information storage and retrieval systems: theory and implementation*, volume 8. Springer Science & Business Media, 2006.

[3] C. A. C. LOZADA, E. E. MENDOZA, M. E. M. BECERRA, L. C. G. FLÓREZ, and E. L. GUZMÁN. Algoritmos de expansión de consulta basados en una nueva función discreta de relevancia. *Revista UIS Ingenierías*, 10(1), 2012.

[4] I. A. Martínez. *Análisis y optimización del recurso UMLS en la recuperación de información biomédica mediante métricas de similitud semántica*. PhD thesis, Universidad Pontificia Comillas, 2016.

[5] G. A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[6] M. F. Porter. Readings in information retrieval. chapter An Algorithm for Suffix Stripping, pages 313–316. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997.

[7] E. Quiñones Matesanz. Recuperación cruzada de información pública clínico-genómica a partir de consultas sobre repositorios clínicos privados. 2015.

[8] J. J. Ruiz Morilla. Poli representación de consultas borrosas para la extracción masiva de documentos en internet. 2016.

[9] G. Salton and M. J. McGill. Introduction to modern information retrieval. 1986.

[10] J. M. Schneider, T. Declerck, J. L. M. Fernández, and P. Martínez. Prueba de concepto de expansión de consultas basada en ontologías de dominio financiero. *Procesamiento del lenguaje natural*, 51:109–116, 2013.

[11] M. T. Vidal, D. P. Avendaño, A. M. Rendón, J. G. G. Serna, and D. V. Ayala. Evaluation of ontological relations in corpora of restricted domain. *Computación y Sistemas*, 19(1):135–149, 2015.

[12] M. T. Vidal, A. L. L. Sánchez, D. V. Ayala, B. Beltrán, and M. C. Cardona. Primera aproximación de un sistema de recuperación de información booleano con expansión semántica de consultas. *Avances en la Ingeniería del Lenguaje y del Conocimiento*, page 55.

[13] M. T. Vidal, A. L. L. Sánchez, D. Pinto, and D. V. Ayala. Aproximaciones para la expansión semántica de consultas de un sistema de recuperación de información booleano. *Avances en la Ingeniería del Lenguaje y del Conocimiento*. En proceso de prensa en revista.

[14] A. Zouaq, D. Gasevic, and M. Hatala. Linguistic patterns for information extraction in ontocmaps. In E. Blomqvist, A. Gangemi, K. Hammar, and M. del Carmen Suárez-Figueroa, editors, *WOP*, volume 929 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2012.

Algoritmos de Aprendizaje Supervisados para el Análisis de Sentimientos en el Dominio de Restaurantes y Laptops

Karen Vazquez Flores
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
krnlet@gmail.com

Mireya Tovar Vidal
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
mireyatovar@gmail.com

David Pinto
Benemerita Universidad Autonoma de Puebla
Ciudad Universitaria, 72570
Puebla, Mexico
davideduardopinto@gmail.com

RESUMEN

En el presente artículo se presenta la investigación realizada para el estudio de Minería de opiniones en dos dominios: Restaurantes y Laptos, el primero para idioma Español e Inglés y el segundo únicamente para el idioma Inglés. Este trabajo se realiza con el objetivo de detectar automáticamente la polaridad emitida por cada opinión en la colección de datos, dicha colección fue extraída de los datos compartidos por SemEval 2016 para la solución de las tareas que proponen. Para la solución se propone el uso de dos clasificadores: Máquina de Soporte Vectorial y el clasificador Naïve Bayes. Para la prueba de los datos se usa validación cruzada, se emplean dos técnicas de Aprendizaje Automático con herramientas de Python y Weka, con la finalidad de aportar también una comparación de resultados entre estas herramientas. Los resultados obtenidos muestran un mejor desempeño al lograr el 83% en el dominio de Laptops y con el uso de los clasificadores en WEKA.

Palabras clave

Análisis de Sentimientos; Aprendizaje Supervisado; Validación Cruzada

1. INTRODUCCIÓN

El creciente uso de Redes Sociales en Internet, a colaborado a la participación del estudio de la Minería de opiniones, pues son tantos los datos generados que se requiere la investigación profunda de esto. Las opiniones escritas en dichas redes, pueden tener cierto grado de polaridad de acuerdo a lo que esta opinando el usuario, en el área del Procesamiento de Lenguaje Natural se extiende el estudio de Análisis de Sentimientos o Minería de opiniones, esta área permite analizar el comportamiento de los usuarios a través de sus comentarios, reseñas u opiniones. Conocer estos datos impacta a gran variedad de ámbitos, en este artículo se muestra el ánalisis de opiniones escritas para calificar a Restaurantes y Laptops en general, las opiniones están escritas en el idioma inglés y en el idioma español para el caso del dominio de Restaurantes y en inglés para el caso del dominio de Laptops. El objetivo, es la detección de la polaridad de cada opinión, es decir, la clasificación según el grado de polaridad, que en este caso se pueden encontrar cuatro grados de polaridad: positivo, negativo, neutral y conflicto. El propósito de este estudio se logra utilizando Máquina de Soporte Vectorial como clasificador y validación cruzada para la prueba de datos, así como también el uso del clasificador Naïve Bayes. En las propuestas se emplean los mismos métodos con la programación en Python y con WEKA, posteriormente se realiza una comparación de resultados usando los mismos datos.

El artículo se distribuye de la siguiente manera: en la sección 2 se presentan los trabajos relacionados con esta investigación, en la sección 4 se presenta el algoritmo o modelo propuesto, en la sección 5.1 se presenta la información de los datos utilizados, y los resultados logrados en la sección 5. Finalmente, las conclusiones se presentan en la sección 6.

2. TRABAJOS RELACIONADOS

El análisis de sentimientos ha tenido varias participaciones de investigaciones en diferentes idiomas y con el uso de distintos métodos tanto para aprendizaje supervisado y no supervisado, cabe mencionar que la mayoría de las investigaciones se han realizado para textos escritos en el idioma inglés. A continuación se describen algunos trabajos relacionados con esta área de investigación.

En el estudio sobre Análisis de sentimientos de [14] se realiza un sistema, donde su primer objetivo es detectar la categoría de las opiniones analizadas, dichas opiniones escritas para los dominios de Restaurantes y Laptops. La clasificación de la categoría se basa en un conjunto de uno-contra-todos los clasificadores binarios, un clasificador para cada categoría que se encuentra en el conjunto de datos de entrenamiento, siendo 13 clasificadores para el dominio de Restaurantes y 79 para el de Laptops, utilizan Vowpal Wabbit tool7 para capacitar a los clasificadores binarios. En el caso del aspecto de la opinión lo modelaron como una tarea de clasificación secuencial, donde a cada palabra de la opinión se le asigna una etiqueta utilizando el esquema IOB2. El sistema logra resultados favorables en la competencia de SemEval 2015.

En el trabajo de [10] se muestra el uso de un corpus con datos provenientes de Twitter para el análisis de sentimientos, construyendo un clasificador de sentimientos que es capaz de determinar la polaridad positiva, negativa y neutra

de los datos escritos en el idioma inglés. Para el clasificador utilizaron la presencia de unigramas, bigramas y trigramas como características binarias. Los clasificadores que usan son Naïve Bayes multinomial, SVM y CRF, obteniendo mejores resultados con Naïeve Bayes. Se probaron dos clasificadores de Bayes, cada uno con características diferentes: Presencia de n-gramas e información de distribución de parte de discurso. Los resultados obtenidos son mejores con el uso de bigramas.

En [4] se propone un esquema de ponderación supervisada en base a dos factores básicos: El primero en Importancia de un Término en un Documento (por sus siglas en ingés, ITD), en este factor se exploran tres definiciones basadas en frecuencia de términos y como segundo factor esta la Importancia de un Término para expresar el Sentimiento (por sus siglas en ingés ITS), donde presentan siete funciones estadísticas para aprender la importancia sentimental de un término por su distribución estadística en los documentos positivos y negativos. Su esquema hace uso completo de información de las etiquetas disponibles para asignar pesos apropiados a un acuerdo. Para comprobar lo dicho los autores realizaron experimentos en tres conjuntos de datos de referencia. El clasificador utilizado se basa en máquinas de soporte vectorial y se utilizó la medida de exactitud como criterio de desempeño de su investigación. Cornell fue el primer conjunto de datos utilizado, el cual está constituido por reseñas de películas y contiene 1,000 datos de polaridad positiva, y 1,000 negativas. Para el segundo conjunto de datos se utiliza el corpus Movie Database con sentimiento multidominio, tomados de muchos dominios de productos de amazon.com, se utilizaron las opiniones de cuatro dominios diferentes con aproximadamente 4000 comentarios cada uno. El tercer conjunto de datos contiene 50,000 reseñas de películas altamente polares divididas entre positivas y negativas. Sus resultados experimentales muestran que el método propuesto produce mejor precisión en dos de los 3 conjuntos de datos.

Mientras que en los estudios reportados en [8] se investigó la utilidad de las características lingüísticas para detectar el sentimiento de los mensajes de Twitter, evaluando la utilidad de recursos léxicos que existen y características que capturan información sobre el lenguaje creativo usando microblogging. Para realizar los experimentos, utilizan tres corpus de mensajes de Twitter, para desarrollar y entrenar utilizan el conjunto de datos hashtagged, y el conjunto de emoticos EMOT y finalmente para la evaluación usan un conjunto de datos manuales producidos por el iSieve Corporation2 (ISIEVE). Los mejores resultados en la evaluación es con el uso de características léxicas (Palabras listadas del léxico subjetivo MPQA, son etiquetadas con su polaridad preferente: positiva, negativa o neutra). Los autores crean tres características basadas en la presencia de alguna palabra del léxico, como segundo mejor resultado es con el uso de características del microblogging en donde se crearon características binarias que capturan la presencia de emoticones y abreviaturas positivas, negativas o neutras y la presencia de intensificadores

Para el trabajo realizado y planteado en [15], los autores clasifican por tipos las expresiones de sentimientos consideradas a nivel entidad o a nivel aspecto, en el caso de entidad, una opinión es una cuádruple (e, s, h, t) donde e es una entidad, u es un objetivo de opinión, s es el sentimiento, h es el titular del sentimiento y t es el tiempo en el que se expresó el sentimiento, mientras que para el nivel de aspecto se considera una quíntuple (e, a, s, h, t) teniendo el mismo significado para los primero componentes incluyendo a, de aspecto. Primero se identifica el tipo de expresión y después se realiza un pre-procesamiento en donde se categoriza el texto mediante la agrupación o clasificación de técnicas, los documentos se representan en el espacio vectorial basado en SVM (Máquina de Soporte Vectorial). Después de esa fase se realiza la clasificación del sentimiento léxico donde la opinión puede ser positiva o negativa. La clasificación de la opinión se hace bajo diccionarios tales como WordNet o Dictionary.com[1]. Después de generar lo mencionado anteriormente los autores crean métodos de aprendizaje automático supervisados y no supervisados para categorizar el texto.

Siguiendo las investigaciones realizadas con datos provenientes de Twitter, en [7] se introduce un enfoque para la clasificación automática de sentimientos en los mensajes emitidos en twitter, la clasificación puede ser positiva o negativa, presentan los resultados de un algoritmo de aprendizaje automático usando supervición a distancia, en la que el conjunto de datos consiste en mensajes de Twitter con emoticones, en donde los emoticones sirven como etiquetas (por ejemplo: :) indicando sentimiento positivo o :(sentimiento negativo). En el enfoque propuesto se prueban con diferentes clasificadores: basado en palabras clave, Naïve Bayes, Entropia Máxima (MaxEnt) y Máquina de Soporte Vectorial (SVM) y como características usan unigramas, bigramas, unigramas y bigramas, y partes de la oración. Los resultados obtenidos muestran una precisión superior a 80% con el uso de Naïve Bayes, Entropia máxima y SVM, y con unigrama como característica.

En otro trabajo de investigacion[2] [2] proponen la clasificación de polaridad con tres posibles resultados: positivo, negativo, neutral. Para opiniones en idioma Ingles y Francés en el dominio de restaurantes. Tratan un método de sintésis basado en modelado de conjunto combinado con rasgos lingúísticos incluyendo información semántica, léxica y dependencias sintáctico-semántico para abordar la clasificación de la polaridad. Para esto los autores clasifican los términos detectados usando un representación de entidad término-centrado y luego clasifican las opiniones. Utilizan *pipeline*, asociando la probabilidad más alta de polaridad para el término o la opinión, haciendo caso omiso de los pocos casos que presentan una polaridad mixta (negativo y positivo a la vez). En las características extraídas, añaden la categoría del aspecto detectado previamente como una característica para la clasificación de la polaridad. En sus resultados muestran más satisfacción en la detección de polaridad en primer lugar para el idioma inglés y en segundo lugar para el francés.

En el caso de [6] la estrategia para detectar la polaridad de tweets en el idioma inglés es con base al clasificador Naïve Bayes. Los autores probaron dos estrategias diferentes, la primera es un *baseline* la cual es un clasificador de Naïve Bayes que aprende del corpus de entrenamiento para clasificar tres categorias (positiva, negativa y neutral) que se encuentran en ese corpus. La segunda estrategia es: Binario en el cual sólo identifican los tweets positivos y negativos con

[1]Wordnet, Dictionary.com: Diccionarios con listas de sinónimos y antónimos que pueden ser utilizados para generar automáticamente sentimientos léxicos.
[2]http://m-mitchell.com/NAACL-2016/SemEval/pdf/SemEval45.pdf

el fin de detectar los tweets sin polaridad (o neutral), siendo la segunda estrategia una mejor propuesta, alcanzando mas del 80% de precisión en el corpus. Se consideraron lemas, multipalabras, léxicos de polaridad y cambiadores de valencia (SV) como características. El sistema descrito en este trabajo lo utilizan en una empresa de tecnología de lenguaje natural aplicado a cuatro idiomas.

En este artículo se utilizan dos tipos de clasificadores que en la literatura muestran un desempeño satisfactorio en la minería de opiniones, Máquina de Soporte Vectorial (SVM) y Naïve Bayes, utilizando como característica léxica unigramas y TF-IDF para los dominios de Restaurantes y Laptops. Como contribución adicional, se realiza una comparación entre dos tipos de lenguajes como son Weka y Python.

3. APRENDIZAJE SUPERVISADO

Para la solución de la clasificación de polaridades se han aplicado dos algoritmos de clasificación de aprendizaje supervisado: SVM y Naïve Bayes, los cuales se describen brevemente a continuación

3.1 Máquina de Soporte Vectorial

Máquina de Soporte Vectorial (SVM) es una técnica para la clasificación de datos [1]. El objetivo de las máquinas de soporte vectorial es producir un modelo que se base en los datos de entrenamiento que predice los valores objetivo de los datos de prueba, dando únicamente los atributos de los datos de prueba. Las máquinas de soporte vectorial en Python cuentan con métodos de aprendizaje supervisado utilizados para la clasificación, regresión y detección de valores atípicos. La herramienta **scikit-learn** de Python es una máquina de aprendizaje en Python, que proporciona simples y eficientes herramientas para la minería de datos y análisis de sentimientos, contiene varios *vectorizers* [3] de traducción para los documentos de entrada en vectores de características.

3.2 Naïve Bayes

El clasificador Bayesiano, asigna la clase más probable dado un ejemplo descrito por su vector de características [13]. Naïve Bayes es un método sencillo e intuitivo con un funcionamiento similar a otros enfoques. Naïve Bayes combina la eficiencia (tiempo óptimo de rendimiento) con una razonable precisión [6]. Este clasificador tiene como inconveniente asumir independencia condicional entre los rasgos lingüísticos. Si las características principales son los tokens extraídos de los textos, es evidente que no pueden considerarse como independientes, ya que las palabras co-ocurren en un texto siendo de alguna manera unidas por diferentes tipos de dependencias sintácticas y semánticas. Pero incluso si Naïve Bayes produce un modelo muy simplificado, sus decisiones de clasificación son precisas. Este modelo probabilistico esta basado en la regla de Bayes, junto con un fuerte supuesto de independencia, dicho supuesto es que dada una clase (positiva o negativa) las palabras son condicionalmente independientes entre si, el supuesto no afecta tanto a la exactitud en la clasificación del texto, pero hace más rápido los algoritmos de clasificación aplicables para el problema [9]. Python permite el uso del clasificador Naïve Bayes con el kit de herramientas de lenguaje natural (NLTK)[4].

[3]Función utilizada en scikit-learn http://scikit-learn.org/
[4]//www.nltk.org/book/ch06.html

4. ALGORITMO PROPUESTO

A continuación se describe el algoritmo propuesto para la determinación de la polaridad de las opiniones proporcionadas en la Tarea 5 de SemEval 2016 [12]. Como enfoque de solución se propone el uso de aprendizaje supervisado en el cual se optó por dos sistemas de clasificación Máquina de Soporte Vectorial y Naïve Bayes, cada uno es aplicado en el lenguaje de programación Python y en la herramienta WEKA, con el fin de obtener una comparativa de resultados y saber en cual caso se comportan mejor.

4.1 Fases del modelo

El algoritmo propuesto está estructurado en tres fases. En la primera fase, pre-procesamiento de los datos, se realiza una extracción de las opiniones desde el corpus, además de la limpieza de las opiniones, ect. En la segunda fase se realiza la extracción de características 4.1.2 y en la tercera fase 4.1.3 se utilizan dos sistemas de clasificación, Máquinas de Soporte Vectorial y Naïve Bayes, aplicados independientemente al conjunto de datos.

4.1.1 Pre-procesamiento

El pre-procesado de los datos consiste de los siguientes pasos:

- **Extracción de las opiniones de los documentos en formato XML**. Filtrar únicamente opiniones del documento.
- **Limpieza de opiniones**: Eliminar palabras vacías, signos de puntuación, acentos y caracteres aislados.
- **Tokenización**: Tokenizar opiniones por palabra.
- **Stemming**. Proceso heurístico que corta el final de las palabras y con frecuencia incluye la eliminación de los afijos derivativos, por ejemplo: coches = coche, fuimos = fui, etc.
- **Filtrar Polaridades**. Filtrar opiniones de los datos de entrenamiento por polaridad como: positivo, negativo, neutral y conflicto.

4.1.2 Extracción de características

La característica léxica considerada es la *Frecuencia Inversa del Documento para un Término (TF-IDF)*. Este método indica qué tan relevante es la palabra con respecto al documento seleccionado y al corpus en general. Permite también calificar a los documentos del corpus con base en estas palabras clave, es decir, si las palabras tiene mayor peso, entonces el documento está más relacionado con ellas que uno con las mismas palabras pero con menor peso.

La biblioteca *scikit-learn* de Python, es una herramienta de aprendizaje automático [11], [3][5]. Proporciona el método *TfidfVectorizer*, el cual usa TF-IDF como esquema de ponderación para el vocabulario que forman las opiniones.

4.1.3 Sistema de Clasificación

Como sistemas de clasificación, se utilizan los clasificadores, Naïve Bayes y Máquinas de Soporte Vectorial.

En el caso de Python, para la **Máquina de Soporte Vectorial** se utiliza **SVC** [6] con un kernel líneal. En el caso

[5]http://scikit-learn.org/
[6]sklearn.svm.SVC

de **Naïve Bayes** se utiliza un extractor de características con NLTK, el cual devuelve un diccionario que indica que palabras están contenidas en la entrada.

A continuación se presentan los resultados obtenidos con la aplicación de este algoritmo.

5. RESULTADOS EXPERIMENTALES

En este apartado se muestran los datos utilizados (5.1) para el análisis de la polaridad en opiniones, las medidas utilizadas en la evaluación (5.2) y los resultados correspondientes con el algoritmo propuesto en cada dominio.

5.1 Conjunto de datos

El conjunto de datos utilizados en este trabajo fueron los proporcionados por SemEval 2016[7] para la solución de la tarea 5, subtarea 2. Se consideraron dos dominios, Restaurantes en el idioma Español e Inglés y Laptops en el idioma Inglés. En la Tabla 1 se presenta la cantidad de información de los tres dominios revisados y la cantidad de opiniones de cada conjunto de datos utilizados en las pruebas del algoritmo propuesto.

Tabla 1: Total de opiniones por conjunto de datos y dominio

Dominio	*Entrenamiento*	*Prueba*	*Gold*
Restaurantes (Español)	2,121	881	881
Restaurantes (Inglés)	1,435	404	404
Laptops (Inglés)	2,082	545	545

En la Tabla 2 se muestra el total de opiniones por tipo de polaridad en cada dominio. La etiqueta *Pos* corresponde al número de opiniones positivas, la etiqueta *Neg* es el total de opiniones negativas y la etiqueta *Con* es el total de opiniones conflicto que existen en el corpus de entrenamiento para cada dominio. En base a los datos de la tabla se logra apreciar que las opiniones con polaridad neutral y conflicto son mucho menos en comparación con la cantidad de opiniones positivas, lo que provoca que la clasificación sea más compleja al no tener los corpus balanceados.

Tabla 2: Conjunto de datos de entrenamiento por polaridad

Dominio	*Pos*	*Neg*	*Neu*	*Con*	*Total*
Restaurantes (Español)	1,519	443	101	58	2,121
Restaurantes (Inglés)	1,012	327	55	41	1,435
Laptops (Inglés)	1,210	707	123	41	2,081

5.2 Medidas de evaluación

Precisión y Recall son medidas comúnmente utilizadas en el área de recuperación de información, se basan en la comparación de un resultado esperado y el resultado efectivo del sistema evaluado [5]. Estas medidas han sido adaptadas para la evaluación de la clasificación en el análisis de sentimientos. La precisión mide la relación de correspondencias que se encuentran correctamente (tp verdaderos positivos) sobre el total de las correspondencias devueltas (tp verdaderos positivos y fp falsos positivo).

[7]International Workshop on Semantic Evaluation 2016

En la matriz de confusión, la precisión se mide con la Ecuación (1) y el *recall* con la Ecuación (2).

$$Precision = \frac{tp}{tp + fp} \tag{1}$$

$$Recall = \frac{tp}{tp + fn} \tag{2}$$

Donde:
tp: verdaderos positivos
tn: verdaderos negativos
fp: falsos positivos
fn: falsos negativos

La medida armónica entre el *recall* y la precisión es la función F_1 (ver Ecuación 3).

$$F_1 = \frac{2pr}{(p + r)} \tag{3}$$

Donde:
p: precisión
r: *recall*.

Accuracy Esta medida se refiere a la evaluación del sesgo de las predicciones, es decir, responde a la pregunta, cúal es el promedio de las predicciones correctas. En la matriz de confusión se obtiene la Exactitud de como se muestra en la formula (4)

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{4}$$

Donde:
tp: verdaderos positivos
tn: verdaderos negativos
fp: falsos positivos
fn: falsos negativos

5.3 Resultados aplicando Máquina de Soporte Vectorial

Para cumplir el objetivo, se evaluó la clasificación de Polaridad usando los tres conjuntos de datos separados de acuerdo a su dominio, en cada conjunto de datos se probó con validación cruzada tomando el 90% para entrenar al clasificador y el 10% para probar.

5.3.1 Resultados en Python

En la Tabla 3 se muestran los resultados totales de *Precisión*, *Recall* y F_1 por dominio aplicando SVM en Python.

Tabla 3: Resultados obtenidos en Python y SVM

Dominio	*Precision*	*Recall*	F_1
Restaurantes (Español)	0.74	0.81	0.77
Restaurantes (Inglés)	0.72	0.78	0.74
Laptops (Inglés)	0.83	0.87	0.84

5.3.2 Resultados en WEKA

En la Tabla 4 se muestran los totales de *Precisión*, *Recall* y F_1 por dominio aplicando SVM en Weka.

Tabla 4: Resultados obtenidos en Weka y SVM

Dominio	*Precision*	*Recall*	F_1
Restaurantes (Español)	0.786	0.821	0.799
Restaurantes (Inglés)	0.791	0.838	0.813
Laptops (Inglés)	0.762	0.816	0.786

5.3.3 *Comparación entre SVM*

Con los resultados obtenidos para Máquina de Soporte Vectorial, a continuación se muestran las figuras comparativas de resultados utilizando SVM en Python y Weka. En el dominio de Restaurantes Weka obtuvo mejores resultados en el idioma Español (ver Figura 1) e idioma Inglés (ver Figura 2). En el caso de Laptops en Inglés se puede observar (ver Figura 3) que el comportamiento de SVM es mejor en Python que en Weka.

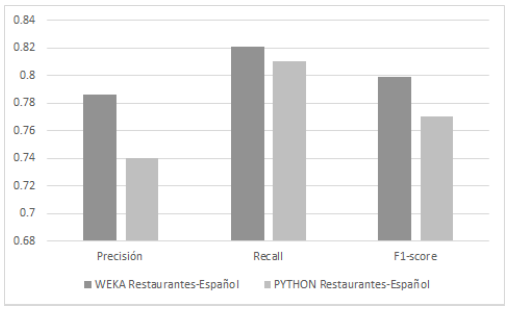


Figura 1: Dominio de restaurantes en el idioma español utilizando SVM

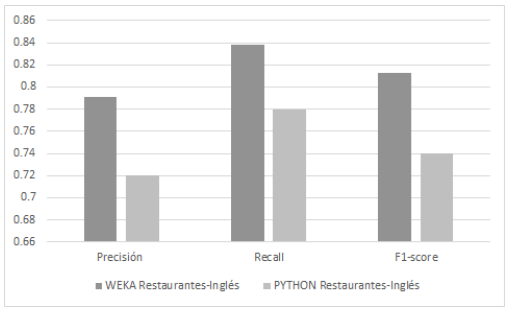


Figura 2: Dominio de restaurantes en el idioma inglés utilizando SVM

5.4 Resultados aplicando Naïve Bayes

Uitlizando Naïve Bayes como clasificador, se evaluaron los resultados con la medida *Accuracy* y los mejores resultados fueron obtenidos nuevamente con Weka.

5.4.1 *Resultados en Python*

En la Tabla 5 se muestra los resultados de la medida de *Accuracy* por Dominio con el uso del lenguaje Python.

5.4.2 *Resultados en WEKA*

En la Tabla se muestra el resultado de la medida de *Accuracy* al aplicar la herramienta de Weka.

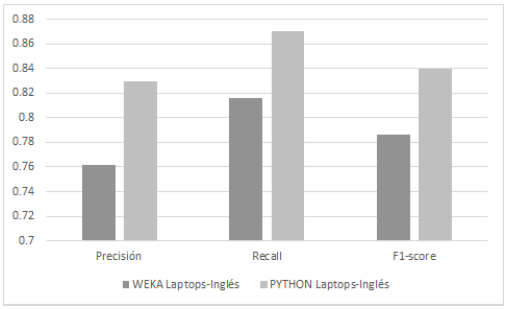


Figura 3: Dominio de laptops en el idioma inglés utilizando SVM

Tabla 5: Resultados en Python con Naïve Bayes

Dominio	*Accuracy*
Restaurantes (Español)	0.7644
Restaurantes (Inglés)	0.7615
Laptops (Inglés)	0.7474

Tabla 6: Resultados en Weka con Naïve Bayes

Dominio	*Accuracy*
Restaurantes (Español)	0.8189
Restaurantes (Inglés)	0.815
Laptops (Inglés)	0.83

5.4.3 *Comparación entre Naïve Bayes*

Con la aplicación del clasificador Naïve Bayes se obtuvieron mejores resultados al utilizar la herramienta Weka para los tres dominios. A continuación se muestran las Figuras 4, 5 y 6 para observar la comparativa del mismo clasificador pero con Python y Weka.



Figura 4: Restaurantes en español y Naïve Bayes

6. CONCLUSIONES

En el trabajo presentado se muestra el uso de dos clasificadores Máquina de Soporte Vectorial y Naïve Bayes para el análisis de sentimientos de las opiniones de los dominios de restaurantes y laptops. Por cada opinión se detecta una de cuatro posibles polaridades: positivo, negativo, neutro y conflicto. Se utilizan dos herramientas Python con nltk y Weka. En base a los resultados obtenidos se observa que en la mayoría de los casos la herramienta Weka se comportó adecuadamente al obtener resultados en presición mayores

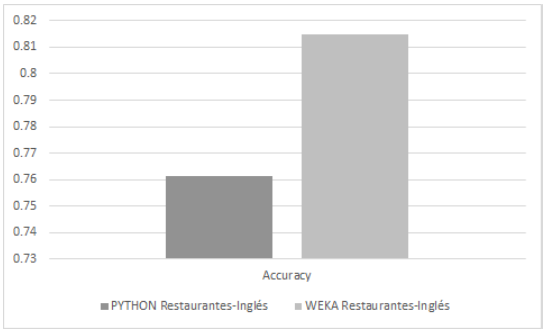


Figura 5: Restaurantes en Inglés y Naïve Bayes

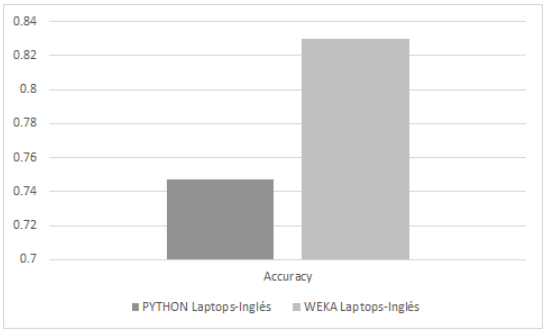


Figura 6: Laptops en Inglés y Naïve Bayes

que en Python. Estos resultados se ven afectados por la forma en la cual se crean los vectores de caraterísticas. En el caso de Python la herramienta construye los vectores y realiza la clasificaci'on, en el caso de Weka se construye el vector verificando el comportamiento de TF-IDF y se aplica el método de clasificación. Utilizando el clasificador de Máquina de Soporte Vectorial con el empleo de WEKA se logra obtener un 83% para el dominio de Restaurantes en el idioma Inglés.

Se puede concluir también que al aplicar SVM los resultados son mejores si se realiza la clasificación en Weka, mientras que al aplicar Naïve Bayes los resultados se obtiene con mayor Recall en Python. Como trabajo a futuro se considera el uso de otras características léxicas y el empleo de diccionarios, así como el uso de otros clasificadores para la tarea de análisis de sentimientos en opiniones.

7. AGRADECIMIENTOS

Esta investigación es parcialmente apoyada por el proyecto PRODEP-SEP ID 00570 (EXB-792) DSA/103.5/15/10854, por el proyecto ID 00570 VIEP-BUAP. Apoyado por el Fondo Sectorial de Investigación para la Educación, proyecto Conacyt 257357.

8. REFERENCIAS

[1] G. A. Betancourt. Las máquinas de soporte vectorial (svms). *Scientia Et Technica 27*, XI:67–72, 2005.

[2] C. Brun, J. Perez, and C. Roux. Xrce at semeval-2016 task 5: Feedbacked ensemble modeling on syntactico-semantic knowledge for aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 277–281, San Diego, California, June 2016. Association for Computational Linguistics.

[3] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[4] Z.-H. Deng, K.-H. Luo, and H.-L. Yu. A study of supervised term weighting scheme for sentiment analysis. *Expert Systems with Applications*, 41(7):3506–3513, 2014.

[5] J. Euzenat. Semantic precision and recall for ontology alignment evaluation. In *IJCAI*, pages 348–353, 2007.

[6] P. Gamallo and M. Garcia. Citius: A naive-bayes strategy for sentiment analysis on english tweets. *Proceedings of SemEval*, pages 171–175, 2014.

[7] A. Go, R. Bhayani, and L. Huang. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1:12, 2009.

[8] E. Kouloumpis, T. Wilson, and J. D. Moore. Twitter sentiment analysis: The good the bad and the omg! *Icwsm*, 11:538–541, 2011.

[9] V. Narayanan, I. Arora, and A. Bhatia. Fast and accurate sentiment classification using an enhanced naive bayes model. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 194–201. Springer, 2013.

[10] A. Pak and P. Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, pages 1320–1326, 2010.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[12] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S. M. Jiménez-Zafra, and G. Eryiğit. Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 19–30, San Diego, California, June 2016. Association for Computational Linguistics.

[13] I. Rish. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46. IBM New York, 2001.

[14] Z. Toh and J. Su. Nlangp: Supervised machine learning system for aspect category classification and opinion target extraction. 2015.

[15] P. Wlodarczak, M. Ally, and J. Soar. Opinion mining in social big data. *Available at SSRN 2565426*, 2015.



Taller:
Network Systems and Protocols

Organizadores:
J. Antonio García Macías (CICESE)
Rolando Menchaca Méndez (CIC-IPN)

An agent based framework for multi-sink wireless sensor networks

Miriam Carlos-Mancilla, Ernesto López-Mellado, Mario Siller
CINVESTAV Unidad Guadalajara
Av. Del Bosque 1145, Col El Bajío. 45019 Zapopan, Jal., México
{mcarlos, elopez, msiller }@gdl.cinvestav.mx

ABSTRACT

This paper presents a novel multi-agent based modelling framework for developing ad-hoc networking strategies in multi-sink wireless network. The framework describes the structure and activities of a wireless sensor device as a reactive agent interacting with other peers with identical capabilities. The framework is illustrated using a case study in a multi-sink wireless network.

CSS Concepts

• Computer system organization → Distributed architectures, Self-organizing autonomic computing • Networks → Networks protocols → network protocol design. Networks structure → Logical/virtual topologies. • Computing methodologies → Artificial intelligence, Distributed artificial intelligence, Multi-agent systems.

Keywords

Wireless networks; Multi-agent system; Modeling framework; Multi-Sink environments.

1. INTRODUCTION

Wireless networks have been widely studied and implemented in many applications and fields, namely surveillance, regulations and environmental protection, military applications, building automation, among others. An operational wireless network must have properties such as robustness, reliability, availability, error tolerance, security, and reduction in the energy consumption to mention some advantages. These features are achieved according to the number of devices, resources, and information available in the environment where the wireless devices are deployed.

One of the most representative characteristics of wireless networks is that the demand for resources such as computation capacity, power consumption, or bandwidth, is in many cases higher than it can provide. Thus, resource-awareness can play a key role at all operation levels of the network. Consequently, wireless networking algorithms need to be more flexible, robust and reliable.

Multi-agent systems based approaches have been widely used to deal with complex problems, which involve distributed information processing. Thus, wireless sensor networks may be naturally modelled as a Multi-Agent System (MAS) by considering each sensor as an agent. A MAS can also cope with large scale systems offering modularity, decomposing the problem, and assigning sub-problems to different (sensor) agents. As well as sensors, agents consider local information. Agents can be modelled with different roles and activities. In this proposal, agents have the same capabilities with different defined tasks according the agent role during the execution of the algorithm.

In the proposed framework, an agent manages the activities of a sensor in the environment according to a general architecture, allowing defining specific behaviours.

There are many proposals for agents and multi-agent architectures; they address specific issues and are defined according to the objective of the application. In most approaches, the agents are presented using their mental states, such as belief, desires and intentions (BDI); some of these classic approaches are presented in PRS (Procedural Reasoning System) [1], which is a framework for symbolic reactive control system in dynamic environments. It also integrates a goal-directed and event-driven behaviour and a multi-agent planning system. IRMA (Intelligent Resource-Bounded Machine Architecture) [2] is the first system that implemented BDI software model in deliberative (intentional) agents.

Later, INTERRAP[3] combines the use of patterns of behaviour with planning facilities to exploit the advantages of the reactive, behaviour-based, deliberative, and plan-based paradigms.

There are others proposals where the objective is to provide communication between the agents, guarantee a service, or optimize a specific metric. For instance, in [4] the authors present a framework where sensor nodes delegate to software agents the tasks of data collection from their neighbours. Sensors are aware of spatial characteristics of the area to specialise the handling of other services. The agents can be static or mobile; they use this strategy for determining routing and relocation procedures in a personalization method.

In this paper, we are interested in modelling a wireless sensor network as a Multi-agent system. For this purpose, a framework for developing multi-agent systems composed by networked agents with identical capabilities is proposed. This approach allows representing the activities of an agent in a simple manner by considering the goals, desires and beliefs in a system.

The remaining of this paper is organized as follows. Section II presents the agent architecture. Section III focuses on relevant researches, describing strategies for multi agent systems. Finally, in Section IV, concluding remarks are given.

2. Agent Based Framework

This section overviews the subclass of multi-agent system dealt with and presents the definition and the architecture of the interactive agents.

Every agent operates in an autonomous manner and shares information or knowledge with other neighbour agents.

2.1 Basic Concepts

In the approach held in this paper, agents are considered as *reactive*, i.e. they can respond timely to changes in the

environment, including other agents. Two types of reactive agents are proposed:

• Ordinary Reactive Agents (ORA). These agents are deployed in the environment in a random fashion; they are able to communicate with other agents in the environment and send the information to the special agents.

• Special Reactive Agents (SRA). They collect information from the environment through other ordinary agents.

The agents are defined as computational entities with some degree of autonomy; they are able to cooperate, compete, communicate, act flexibly, and manage their behaviour within the frame of their objectives. Agents are the enabling technology for a wide range of advanced applications relying on distributed and parallel processing of data, information, and knowledge relevant in domains ranging from industrial manufacturing and e-commerce to health care [5].

The ORA set, denoted by A represents the sensors deployed in the environment. The SRA set, denoted by S, represents the sink nodes that have unlimited power and no memory restrictions to process and store information from agents in A.

Formally, an agent A_i is a 3-tuple defined by $<PA_i, TA_i, Q_i>$ where:

- PA_i is the vector of perceptions of A_i; the agent i receives messages from others agents A_j; each agent manages the information perceived from other agent j in a 8-entry vector with the following information

$$PA_j = [\mathrm{ID_j}, \mathrm{RoleF_j}, \mathrm{RoleT_j}, \mathrm{E_{resj}}, \mathrm{NNeighs_j}, \mathrm{Level_j}, \mathrm{Type_j}]^{\mathrm{T}}, \quad (1)$$

  where $Level_j$, ID_j and $NNeighs_j \in \aleph$,
  $RoleF_j \in$ {Leader, Gateway, Bridge, and Member}
  $RoleT_j \in$ {Intermediate, Leaf}
  $\mathrm{Type_j} \in$ {Broadcast, Acknowledgement, DMsg, Stopmsg}
  $E_{resj} \in \Re^{>0}$, $i, j \in [1, N]$, and $m \in [1, M]$; where M, N are the total of special and ordinary agents defined in the environment respectively.

$$\text{then, } PA = \begin{bmatrix} ID_1 & RoleF_1\ RoleT_1\ E_{res1} \cdots & Level_1 \\ \vdots & \ddots & \vdots \\ ID_N & RoleF_N\ RoleT_N\ E_{resN} \cdots & Level_N \end{bmatrix}^T \quad (2)$$

- TA_i is the set of external information of a Special Agent S_m; each agent maintains the information of every special agent m in the environment; this information contains the agent's external perception ($Ts_i \subseteq Pt_i$); it is defined by the following vector:

$$TA_m = [\mathrm{ID}_m, \mathrm{RoleT}_m, \mathrm{E}_{resm}, \mathrm{NNeighs}_m]^{\mathrm{T}} \quad (3)$$

$$\text{then, } TA = \begin{bmatrix} ID_1 & RoleT_1\ E_{res1} \cdots & NNeighs_1 \\ \vdots & \ddots & \vdots \\ ID_m & RoleT_m\ E_{resm} \cdots & NNeighs_m \end{bmatrix}^T \quad (4)$$

- Q_i is the agent logical state; it contains all the basic internal variables to describe the current situation of an agent i. These variables help in the agent's decisions such as $Level_i$, ID_i, $RoleF_i$, $RoleT_i$, E_{resi}, $NNeighs_i$, $Nsink_i$, $Type_i$; where $Nsink_i \in \aleph$.

A Special agent S_m is a 2-tuple defined by (PS_m, Q_m) where:

- PS_m is the vector of perceptions; is the messages received from A agents. Each agent S_m maintains and evolves the information perceived from and agent A_i, and it is defined with the same structures defined in equations (1) and (2).

- Q_m is the state of the agent S_m; it contains all the basic internal variables to describe the S agent, these are the $Eres_m$, $Nleaders_m$, $NNeighs_m$, ID_m; it also contains the processing time, messages' requests, and so on.

2.2 A Multi-agent system

The Multi-Agent Networked System can be described as an undirected graph MAS = ($A \cup S$, E) where:

- $A = \{A_1, A_2, A_3 \ldots A_N\}$ is the set of ordinary agents deployed in the environment.
- $S = \{S_1, S_2, S_3 \ldots S_M\}$ is the set of special agents deployed in the area.
- $E \subseteq \{(A \times A) \cup (A \times S)\}$, indicates the relationship between agents which defines the topology connection of the network, as follows:
  - The pair $(A_i \times A_j) \in E$, indicates the information exchange between two any ordinary agents A_i and A_j.
  - The pair $(A_i \times S_m) \in E$ indicates the information exchange between an ordinary agent A_i and S_m a special agent.

Some relevant properties and constraints for the framework are:

- The topology can change; if an agent dies or a new is added, the agents will perform the necessary changes to guarantee the communication in the topology.
- The communication between agents is bidirectional.
- The neighbourhood of an agent includes the neighbours inside a transmission range.
- The transmission range can be defined differently for every agent.
- There is no control of one agent to any other.
- The agents are deployed randomly without any a priori information.
- Agents send their information to the agents that are connected in the topology formation.
- All agents have the same capabilities.
- Special agents receive sensory information.

2.3 Agent internal framework

The main motivation is to design and implement a new framework for networking and maintenance of a MAS to address challenges inherent to sensor constraints (energy, memory, latency, loss of messages, number of retransmissions, and so on). The proposed framework includes a local reconfiguration strategy to extend the lifetime of the network.

Agents are able to communicate, receive and process information with other agents *A and/or S according to their internal processing*. Then agents and it produce actions that may affect, change, or update the *environment*. Figure 1 shows the main components of an autonomous agent and its relationships between them. Every module performs an important activity in the behaviour of the agent.

Energy module. The energy depletion is a natural process during the execution of any process, in this architecture the energy consumption is caused by transmitting, receiving and processing information.

Mobility module. It is used to move an agent into the environment; the involved variables in this module are the speed and direction. This module also affects some functions of the process module and the reconfiguration process in the agent.

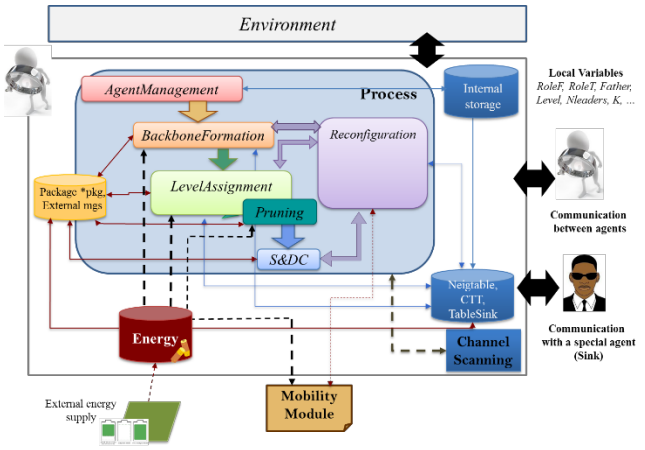


Figure 1. Agent Framework.

Channel scanning. It is used to look for an available channel to start the whole process. This module is launched at the beginning to the algorithm when it is required.

External energy supply module. This module is an additional source of energy that relax the energy constraints of an agent node. The extra energy supply is provided by a solar cell or any other energy source.

Internal storage. *Neigtable*, *CutTreeTable* (CTT), *TableSink*, and *package*, *external messages*: are variables directly related to the processing of the agent. Variables represent all the dates the agent stores in a local manner to update or change its state or stage in the application progress.

Process module. The activities of an agent are defined according to the role that an agent is able to perform. An agent plays one role at once. The activities that a node performs during the procedure of the algorithm are classified by stages. Every stage performs different activities of the complete multi-sink wireless networking algorithm.

- *AgentManagement* function. It is used to control the order and duration times of the stages in the agent. This module is the responsible to manage the incoming messages and to generate new messages in the agent and send them to other agents.
- *BackboneFormation*. It involves self-organization and clustering strategies. The backbone formation stage is the responsible to create clusters with the available agents in the environment using local information; the agents take decisions in an autonomous manner (self-organization). This stage create redundant links and assure the communication with others clusters.
- *LevelAssigment*. This module performs the formation of trees and the level assignment. The procedure is invoked from the special agents *S* towards the ordinary agents *A*. At the end of this stage, every agent *A* will have a level with respect to every *S* agent.
- *Pruning*. It performs the cutting some strategic links to assure the formation of the trees and the connectivity between the nodes. The result of this stage is the formation of multiples trees in the environment according to several situations.
- *Sensing and Data Collection (S&DC)*. Every agent collects and processes the sensed information from other agents and forwards it to its predecessor agent.
- *Reconfiguration*. This procedure can be called any time during the execution of the stages if it is required; every agent in a local way decides how to react to the reconfiguration according to the changes in the network.

The interaction between modules is performed as follows.

First, an agent can receive information from other agents (*external msg*); this information is stored using the *Neigtable*, *CTT* or *TableSink* repositories, afterwards, the agent maps the information to the *process* module for updating the internal *storage variables*; after the information processing, a new packet is generated (*package *pkg*). Then, the agent looks for an available *channel* to communicate with other agents and sends the data to the corresponding agent. During this procedure, there is an obvious energy consumption and the agent may need an external energy to keep working properly.

This procedure continues until the agent is stable in one procedure or the agent decides to finish the algorithm.

3. Related Protocols

There exists many approaches that uses multi-agent techniques to specify the behaviour of devices as agents to improve the design job and consequently, the performance of the application. These approaches have specific purposes such as providing a better routing, direct data diffusion, reducing the energy consumption, etc.

This section presents some of these related works, particularly a multi-sink wireless networking approach that provides an efficient manner to communicate agents in one environment, and allows reduced energy consumption, including a self-organization capabilities, formation of trees based on local decisions, and a data collection approaches.

3.1 Agent based approaches

A modelling framework for a multi-agent system is presented in [6]. In this work, an architecture for designing interactive agents is proposed and then the component devoted to represent and revise the agent's hybrid state (continuous and discrete) is described. This approach allows managing agent tasks that require handling variables real-valued and integer-valued.

A multi-agent framework is proposed in [7] that can assist existing routing approaches to improve their routing performance; this framework enables each sensor node to build a neighbourhood based of a set of past experiences, which are the transmitted information in a previous step. The energy consumption is considered and the better route is chosen according to defined metrics.

There are other approaches where is more important to save the energy by looking for routes and sending the information in a

direct manner; an example of this kind of works is presented in [8] where a Mobile Agent (MA) combined with a Direct Diffusion (DD) algorithm (MADD) is described. A sink node spread the data into a zone and the agent that receives the message floods the sink individually; then the sink decides the list of sources that will be visited by the MA; the nodes are visited following an importance order. The energy consumption is not considered and the cost for flooding in the list creation of important sources is too expensive; nodes are aware of its location.

A distributed framework called VISUM, for monitoring wireless networks is presented in [9]. This strategy stores the collected information at data repositories. VISUM supports the seamless integration of new devices into its monitoring framework and requires minimal maintenance by network operators. VISUM does not support large scale networks.

In [10] the use of agent paradigms for Wireless sensor networks is promoted. "Agents" and "sensors" have many common issues that can be fruitfully exploited to design efficient agent-based WSNs; such issues are: *situatedness*, which is related to the position and control of an agent in the environment, *limited resources*, *recovery of incomplete data*, *large-scale networks*, *distributed system*, and *adaptable environments*, among others.

The authors also describe the use of an agent platform (SPOTs) based on wireless body sensor networks. This paper does not present any application or strategy.

A distributed intrusion detection system for ad-hoc wireless network based on mobile agent technology is presented in [11]. It is implemented a framework that intrusion at multiple levels and takes into account decision policies. The architecture is aimed to minimize the cost of network monitoring and maintaining the system.

3.2 Agent based multi-sink wireless networking

A wireless sensor networking and operation algorithm for multi-sink environments is presented in [12]. This procedure is represented in the framework in the *Process* stage.

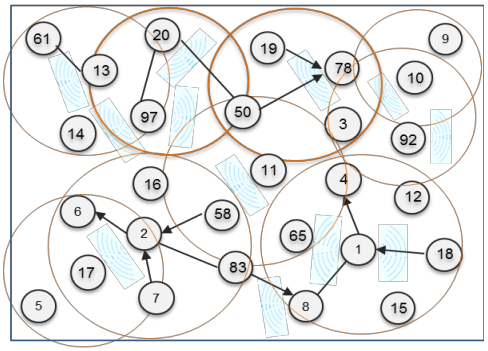
The proposed scheme, which allows low energy consumption, consists of three stages: the first one performs network formation and maintenance through distributed self-organization strategy under a policy of power consumption reduction. The second one builds the trees from the *S* agents using a load balance strategy based on local information. The third stage collects sensory data through the trees. Finally, a reconfiguration local method that modifies only the affected area considering the neighbour information is presented.

The design of an efficient clustering algorithm would be beneficial for a better performance. The agents A_i are deployed in the environment and they interact with neighbour agents A_j without any priori information.

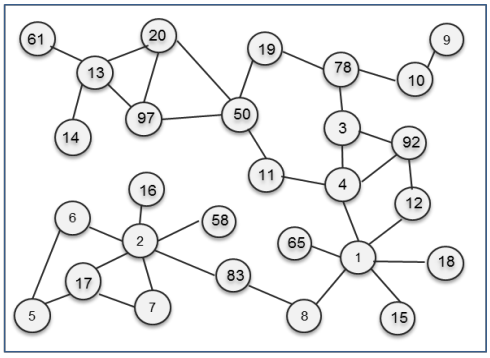
3.2.1 Backbone formation stage

For the backbone formation, a self-organization and clustering method are used. The agents A_i send a message by broadcasting method until receiving an acknowledgment message, then they register the information in a neighbour table, and then the agents decide which role to play according to the available information.

A_i agents form groups called clusters led by a Cluster Head or *Leader*, one or more *Members* connected to the Leader, one or more *Gateways* that connect others clusters and zero or more *Bridges* that connect segments of the network. The desired global behaviour is decomposed into local rules for each device. Figure 2 presents an example of the backbone formation stage, a) presents the initial state of the agents in the environment and the discovering neighbouring process; b) shows the self-organization method in which roles start choosing their role according the received information; c) presents the result of this when the stage has been stabilized. More details can be found in [12].

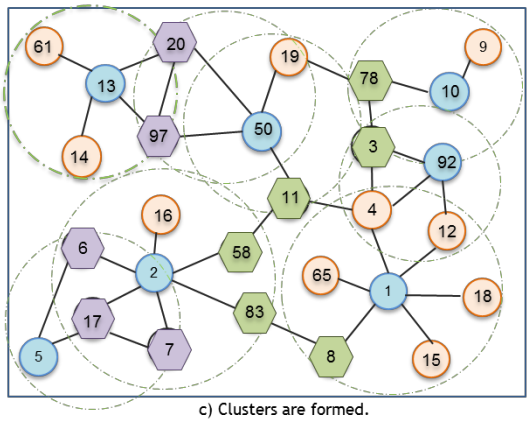


a) Initial environment; neighbor discovering, agent records the neighbors agents.



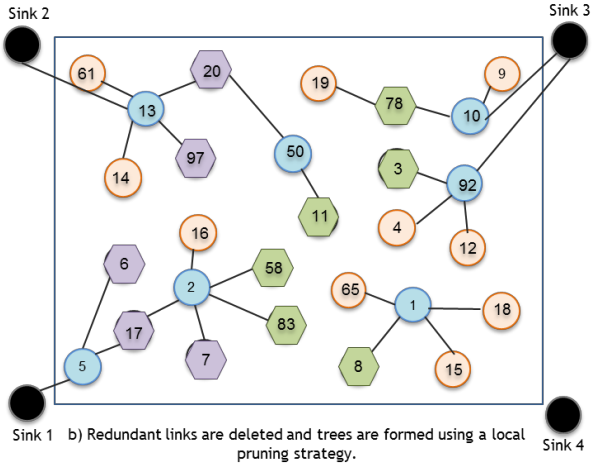
b) self-organized method is implemented for the cluster formations, agents chose in an autonomous manner their role





c) Clusters are formed.

Figure 2. Backbone Formation process.



b) Redundant links are deleted and trees are formed using a local pruning strategy.

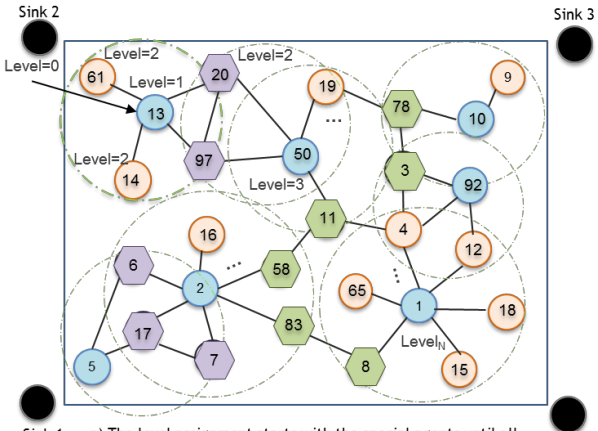
Figure 3. Tree formation procedure.

3.2.2 *Tree Formation stage*

The main benefits of creating multiple tree formations in a multi-sink environment are the balanced load between the nodes, the information redundancy, reduction of energy consumption, and reduction in the number of transmitted messages.

This stage consists of assigning levels starting from the sink nodes; *S* agents are the roots of the trees with level 0. The level assignment its spread until all agents in the environment have a level with respect to every special agent *S* defined. The relationship successor/predecessor is established and redundant links are deleted without loss of connectivity in the network topology.

In this stage, agents that play Leader, Gateway and Bridge roles become *Intermediate agent nodes* and Member roles are *Leaf agents*. For the formation of the trees, a pruning strategy is considered. The pruning strategy is decided by intermediate nodes based on a metrics, which involves the *number of links*, *residual energy*, *the number of hops*, *and the transmission range*. Figure 3 describes the tree formation procedure: Figure. 3.a) shows how a special agent send its level to its leader near agents, leader send its level to its neighbourhood and so on. This procedure continues for every *S* agent. The assignment ends when all agents have a level with every special agent. Figure. 3.b) shows the tree formation once the pruning strategy has been applied; redundant links are deleted and relationship predecessor/node is established.



a) The level assignment starts with the special agents until all nodes have a level.

3.2.3 *Sensing and Data Collection stage*

Once the pruning has finished and there are no more changes in the topology, the sensing and data collection stage is performed. Each node has a schedule to transmit/receive, sensing tasks and its idle times (Figure 4). Unexpected transmissions occur when an event is detected and the information is transmitted to its respective special agent *S*.

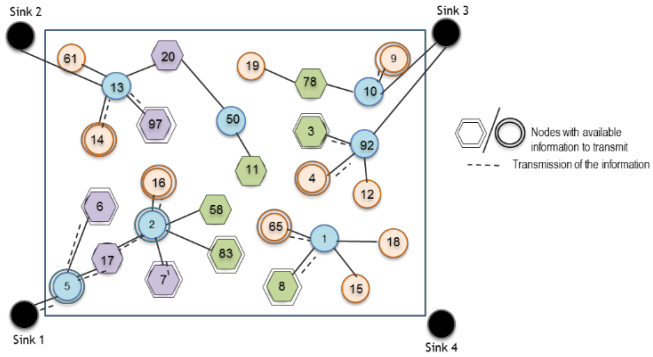


Figure 4. Sensing and Data Collection.

During any of these stages, an unexpected event can appear. When an event appears, the affected agents decide the reconfiguration level according to the local available information, without affecting the performance of the network (Figure 5); Fig. 5.a) presents the case when a leaf node dies; in this case the agents do not perform any change, the communication is not affected; Fig. 5.b) presents whether an intermediate agent dies, the agents in the same cluster has to re-establish the communication using only the local available information from their neighbours. This process does not affect the whole network.

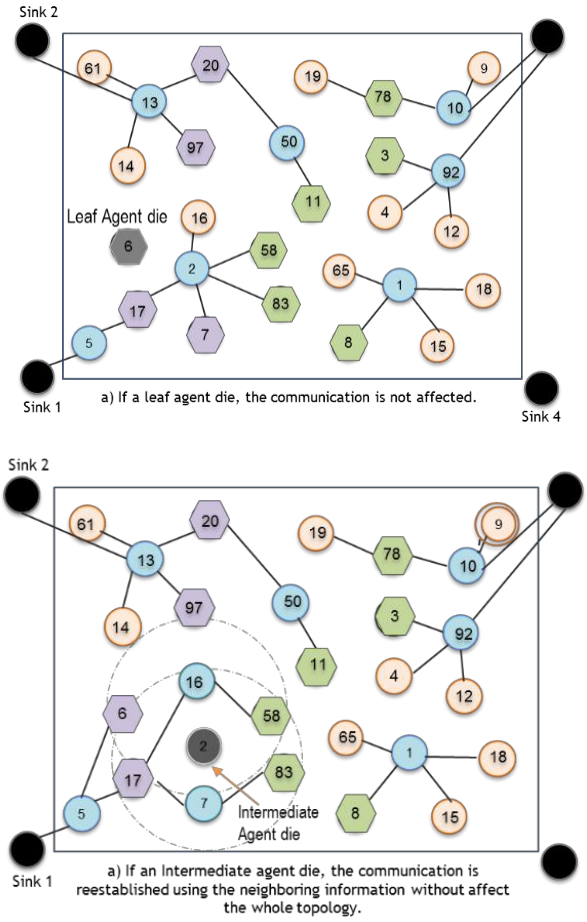


Figure 5. Reconfiguration strategy.

4. CONCLUSION

A multi-agent based framework to support the design of wireless networking strategies has been presented. This framework allows representing the complete behaviour of an agent in a wireless network. The main features of the proposed approach are the possibility to address problems regarding the energy consumption, mobility, and communication with other agents. Furthermore the components and their interactions and are clearly established. Based on these characteristics, further analysis can be started regarding some dynamical properties of the agent and the collective tasks, as well as suitable communication structures.

The strategies used in this proposal extend previous works of the authors dealing with multi-sink multi-agent based environments.

5. ACKNOWLEDGMENTS

This work has been partially supported by CONACYT, M. Carlos' Grant No. 243189.

6. REFERENCES

[1] Wooldridge, M., "An introduction to Multi-agent systems". John Wiley & Sons. England, 2002. ISBN 0 47149691X.

[2] Ingrand F., Georgeff, M., Rao, A. "An architecture for Real-Time reasoning and system control". IEEE Expert: Intelligent Systems and Their Applications, 7(6): 34-44, 1992.

[2] Bratman, M. E., Israel, D. J., and Pollack, M. E. "Plans and resource bounded practical reasoning". Computational Intelligence, 4:349-355.1988.

[4] Nabil Sahli, Nafaa Jabeur, and Mohamad Badra. "Agent-based Framework for Sensor-to-Sensor Personalization". Procedia Computer Science Volume 19, 2013, Pages 197-205, ELSEVIER editorial. The 4th International Conference on Ambient Systems, Networks and Technologies (ANT 2013).

[5] Gerhard Weiss. "Multiagent Systems", MIT Press, 2013, "nd edition. ISBN 978-0-262-01889-0.

[6] Berenice Gudiño-Mendoza, Ernesto López-Mellado. "A Modeling Framework for Developing Networked Agents Applications". 8th International Conference on Electrical Engineering, Computing Science and Automatic Control, 2011. ISBN: 978-1-4577-1013-1.

[7] Dayon Ye,Minji Zhang, and Yu Yang. "A Multi-Agent Framework for Packet Routing in Wireless Sensor Networks". Sensors (Basel). 2015 May; 15(5): 10026–10047. Published online 2015 Apr 28.

[8] Min Chen, Taekyoung Kwon, Yong Yuan, Yanghee Choi, and Victor C. M. Leung. "Mobile agent-based directed diffusion in wireless sensor networks". Published in: Journal EURASIP Journal on Applied Signal Processing archive Volume 2007 Issue 1, 1 January 2007. Pp. 219-219

[9] Camden C. Ho, Krishna N. Ramachandran, Kevin C. Almeroth, and Elizabeth M. Belding-Royer. "A scalable framework for wireless network monitoring". Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots (WMASH '04). Pp. 93-101. ACM New York, NY, USA ©2004.

[10] Giancarlo Fortino, Stefano Galzarano, Raffaele Gravina, Antonio Guerrieri. "Agent-based Development of Wireless Sensor Network Applications". Online Proceedings for Scientific Conferences and Workshops. WOA, pp.123-132

[11] O. Kachirski, and R. Guha. "Intrusion detection using mobile agents in wireless ad hoc networks". Published in: IEEE Workshop on Knowledge Media Networking, 2002. Page(s): 153 - 158, 2002.

[12] Miriam Carlos-Mancilla, Ernesto López-Mellado and Mario Siller. "A localized multi-sink multi-hop algorithm for Wireless Sensor Networking". IEEE Global Information Infrastructure and Networking Symposium (GIIS'15). Guadalajara, Jalisco. PP. 1-6. October 2015.

Cifrado Caótico Parcial para Comunicaciones de Video en Tiempo Real

Rogelio Hasimoto-Beltran, Marcos A. Calderón
Centro de Investigación en Matemáticas (CIMAT),
Jalisco S/N, Col. Valenciana,
Guanajuato, Gto. México
hasimoto@cimat.mx

Salvador Ruiz-Correa
Center for Mobile Life-CEMobili
www.cemobili.org
src@cmls.pw

RESUMEN

Las comunicaciones de datos multimedia tales como voz, imagen y video representan actualmente el tráfico dominante sobre el Internet. Antes de su transmisión, son transformados por un proceso de compresión o eliminación de redundancia, cuya salida produce códigos de longitud variable (VLC) que reducen significativamente el tamaño original de los datos, pero resultan extramadamente sensibles a errores en la comunicación. Un simple error de bit en los códigos VLC, puede propagarse y dañar la calidad de la información por periodos largos de tiempo como es el caso de audio o video (aproximadamente en el orden de milisegundos a segundos). El presente trabajo aprovecha esta vulnerabilidad a los errores, y propone una técnica novedosa de cifrado parcial como una alternative eficiente al cifrado total de datos multimedia en la cuál, solo un subconjunto del total de bits a transmitirse es encriptado. Nuestro algoritmo se basa en la introducción artificial de errores aleatorios en el flujo de bits comprimidos (para ocultar los verdaderos códigos VLC), seguido por una permutación aleatoria de segmentos de longitud variable que impiden la descompresión parcial o total de los datos transmitidos, asegurándo que el contenido sea legible solo para el destinatario. Para la generación de números aleatorios, se propone una nueva función caótica robusta y eficiente en el dominio de los enteros, cuyos parámetros deben ser conocidos por el que recibe para poder desencriptar y descomprimir la información recibida. Las principales características del algoritmo desarrollado con respecto a trabajos previos en la literatura son las siguientes: a) El nivel de seguridad es escalable; b) Es independiente del codificador de video a utilizar; c) Es el algoritmo mas rápido de cifrado parcial que se reporta en la literatura a la fecha; y finalmente d) Es robusto a ataques conocidos en la literatura como lo son el de fuerza bruta y textoplano escogido y diferencial.

CCS Concepts

• **Security and Privacy → Cryptography**

Keywords

Criptografía Caótica; Mapa caótico; Código de Longitud Variable (VLC); Código de Huffman.

1. INTRODUCCIÓN

Con el avance acelerado de tecnologías informáticas, surgieron nuevas herramientas de comunicación que han cambiado muchas actividades de la vida cotidiana, como las transmisiones de video conferencias, televisión digital, aprendizaje en línea o transferencias bancarias, etc. Se tiene hoy en día la necesidad de ofrecer seguridad a esta enorme cantidad de información que se transmite en canales inseguros como el Internet. Existen riesgos en las redes de comunicación: hay personas no autorizadas cuyo propósito es acceder a información privada. De aquí la necesidad de crear sistemas de comunicación seguros sin sacrificar o afectar (lo menos posible) la velocidad de transmisión de la información. La transmisión de datos multimedia genera demandas de seguridad que a veces son difícil garantizar; se require procesar una gran cantidad de información bajo restricciones estrictas de tiempo que requieren un mayor poder de procesamiento. Debido a esto, es imperante crear nuevos esquemas criptográficos que satisfagan dichas demandas de procesamiento aún en dispositivos móviles (cuyo poder de procesamiento es relativamente limitado, presnetando además restricciones de abastecimiento de energía).

Actualmente, existe el enfoque de algoritmos de cifrado en donde se procesa la totalidad de los datos a ocultar. Esta característica presenta algunas desventajas: a) es posible que el poder de procesamiento del dispositivo (móvil) no sea suficiente para ejecutar el algoritmo en tiempo real; y b) limitación en su escalabilidad, esto es no provee distintos niveles de seguridad de acuerdo a cada situación particular. Como una respuesta a los problemas que presentan los algoritmos de cifrado total, surgió la alternativa de *cifrado parcial*, el cual se basa en la premisa que solo es necesario ocultar una parte de la información, y esta es suficiente para proveer seguridad ante ataques de privacidad en la información. Aunado al cifrado parcial, es también deseable un cifrado no-invasivo, esto es después del proceso de compression para evitar cualquier co-dependecia entre compression y cifrado (cuando el cifrado es invasivo, el cifrado debe implentarse dentro del codificador).

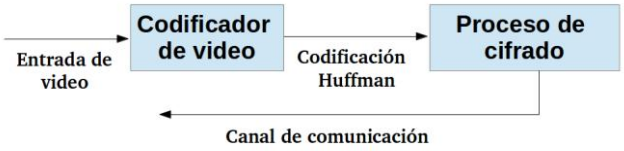


Figura 1: Proceso de empaquetado.

Diversos esquemas de cifrado parcial selectivo (solamente se protegé cierta información como son Intra frames, vectores de movimiento, etc.) han sido propuestos para las comunicaciones de multimedia (video en particular). Khanvilkar [1] et. al., propusieron un sistema cifrador selectivo para archivos mp3 donde solo se cifran campos del encabezado del formato. Meyer y Gadgast [2] diseñaron un esquema cifrador selectivo para archivos del formato MPEG-1. Tang [3] propuso un sistema cifrador que consite en utilizar permutaciones aleatorias como una alternativa al orden de zigzag al momento de mapear un bloque DCT (Discrete Cosine Transform) de 8 × 8 a un bloque de 1 × 64; de esta manera, es difícil calcular la inversa de la DCT. Nuevas propuestas de cifrado parcial selective modifican la información específica del archivo. Por ejemplo, Lian et al. [4], propusieron un sistema de cifrado que se basa en el cifrado de los macrobloques y los vectores de movimiento de un archivo en proceso de compresión. Li et al. [5] definieron un algoritmo cifrador que se enfoca en el cifrado de los vectores de movimiento y los coeficientes de la transformada de coseno de un archivo a comprimir, la operación de cifrado de este trabajo está basada en la operación lógica XOR, inversiones de signo y mezclado aleatorio. El algoritmo de Jiang et al. [6] modifica los coeficientes de la transformada de coseno, los vectores de movimiento y los componentes de predicción, usando un enfoque caótico. Spinsante et al. [7] propone cifrar los parámetros de cuantificación del algoritmo de compresión, así como los componentes de predicción, en este trabajo, se utiliza la operación XOR y el enfoque Linear Feedback Shift Register (LFSR) como modo de cifrado. En el trabajo de Lian et al. [8], se cifran los vectores de movimiento, y los macrobloques, se utiliza un paradigma de cifrado por flujo. Wang et al. [9] proponen cifrar los componentes de predicción y los vectores de movimiento, algo importante a notar en este algoritmo, es que el volumen de datos cifrados es muy bajo. Un cifrador caótico fue propuesto por Yuan et al. [10] que se enfoca en el cifrado del bit de signo de los coeficientes de la transformada de coseno, una vez más, el volumen de datos cifrados es muy bajo en este algoritmo.

Los trabajos de cifrado parcial hasta ahora mencionados, dependen fuertemente del formato utilizado para la compresión de datos y afectan la razón de compression (esquemas invasivos). Por lo tanto, su aplicación se restringe a un tipo específico de información comprimida o formato. En este trabajo, se propone un sistema cifrador de tipo *parcial no-selectivo* como una respuesta a las desventajas de los algoritmos selectivos actuales ya mencionados, y que a su vez sea amigable con las tecnologías móviles con capacidad de procesamiento limitado. El propósito del sistema es sencillo: cifrar una parte del archivo a transmitir de una manera independeinte del codificador (cifrado después de la compresión), que sea escalable (defina grados de seguridad internamente) y que sea lo suficientemente rápido para aplicarse a transmisiones de video en tiempo real.

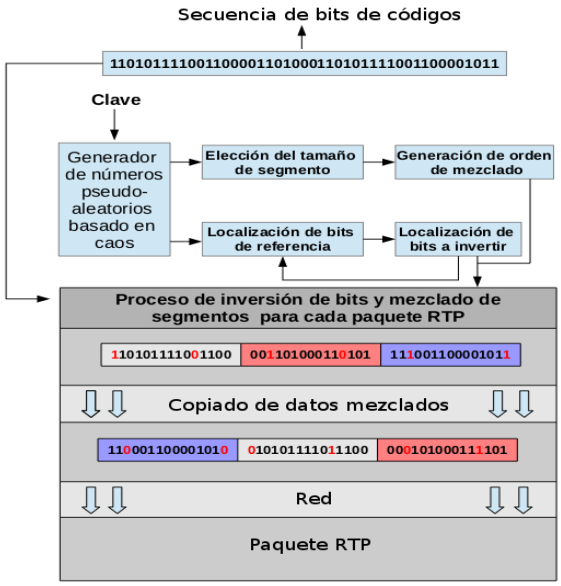


Figura 2: Esquema cifrador propuesto. Fuente: Almasalha-Hasimoto [11].

2. ESQUEMA PROPUESTO

El algoritmo de cifrado que se propone en este trabajo se ejecuta después del proceso de compression de video y empaquetado (Figura 1). El tamaño del fragmento de información o paquete a cifrar puede variar de los 300 bytes a los 1400 bytes, dependiento de las propiedades de transmisión del medio. El esquema consiste en cifrar los códigos VLC (de Variable Length Code) de un paquete por medio de las siguientes operaciones: *inversión de bits* y *permutación aleatoria de segmentos*, donde un segmento es el resultado de dividir el paquete en fragmentos del mismo tamaño. La posición de los bits a ser invertidos y la permutación de los segmentos dependen de un novedoso generador de números pseudoaleatorios seguro. La Figura 2, muestra un diagrama general del esquema de cifrado propuesto, formado por tres componentes principales: a) Generador seguro de números pseudoaleatorios basado en caos (PRNG); b) Proceso de inversión de bits; y c) permutación aleatorio de segmentos.

2.1 Generador de Números Aleatorios

La seguridad del esquema de cifrado propuesto ante diferentes ataques depende principalmente de la robustez del PRNG utilizado (resistencia a diversos ataques de Criptoanálisis, valores generados con alto grado de aleatoriedad, conjuntos grandes para la elección de parámetros y valores iniciales). Debido a que los actuales generadores de numerous aleatorios son vulnerables a ataques debido a su dependencia de la semilla de longitude finita, en este trabajo proponemos un PRNG basado en teoría del caos. La razón principal es que los Sistemas Caóticos Discretos (DCS) presentan propiedades que son necesarias en criptografía: sensibilidad a las condiciones iniciales, dependencia de los parámetros elegidos y trayectorias (numerous aleatorios) impredecibles. Las primeras dos propiedades están relacionadas con el fenómeno de difusión, y la última con el fenómeno de confusión. Estas propiedades son la base para desarrollar sistemas de cifrado seguros. A

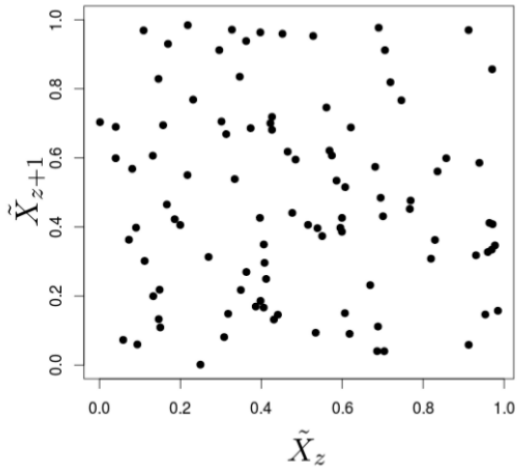


Figura 3: Malla tipo Rényi para el conjunto de vectores: $(\tilde{X}_i, \tilde{X}_{i+1})$. Donde $j = 5$, $b = 513$, $X_0 = 0x87345377$.

continuación, se explica el proceso de generación de números pseudoaleatorios por medio del mapa caótico Rényi

2.1.1 Generador de Números Pseudoaleatorios

Para el proceso aleatorio de inversión de bits y permutado de segmentos, se propone la utilización de mapas caóticos acoplados para la generación de números pseudoaleatorios. El acoplamiento entre mapas ofrece diversas ventajas, entre ellas se encuentra la longitud extendida del ciclo del PRNG; esto es el ciclo de repetición de los números aleatorios generados se vuelve muy grande lo que incrementa la seguridad del PRNG. Otra de las ventajas es que N mapas caóticos acoplados estan representados por 2N variables, mismas que un atacante debe encontrar antes de tomar control del PRNG. El PRNG propuesto se compone de N mapas caóticos Rényi ($1 \leq i \leq N$) y se puede escribir como:

$$X_{i,j} = f(X_{i,j-1}) + \varepsilon \& H(X_{i,j-1},\ldots,X_{N,j-1}) \quad (1)$$

$$H(X_{i,j-1}, \ldots, X_{N,j-1}) = \bigoplus_{i=1}^{N} (X_{i,j-1}).$$

Donde **ε** representa un número aleatorio por cada iteración, & es el operador lógico AND, ⊕ es el operador lógico XOR, y H es la function de acoplamiento definida como la operación XOR de todas las trayectorias anteriores del mapa acoplado. De esta manera, cada número aleatorio X, representa la influencia de cada uno de los N mapa caóticos Rényi $f(X) = \lfloor \beta X \rfloor \ mod\ 2^n$, donde 1<**β** representa el parámetro caótico. El valor de salida de la i-ésima iteración de $X_{i+1}=f(X_i)$, se usa como entrada en la siguiente iteración $i+1$. En la Figura 3 se muestra el efecto no-lineal y la aleatoriedad de la Eq.1, cuyos valores se distribuyen en todo el dominio. Adicionalmente se evaluó la aleatoriedad del PRNG propuesto (Eq.1) usando el conjunto de pruebas propuesto por el National Institue of Standard and Technologu (NIST), aprobando cada una de ellas (para mayor informaci´øn de las pruebas de aleatoriedad propuestas por NIST, ver referencia [12]). Los valores de las N variables iniciales ($X_{i,0}$) de los N parámetros **β**ᵢ deben ser distintos para cada mapa, lográndo así una mayor aleatoriedad en el sistema.

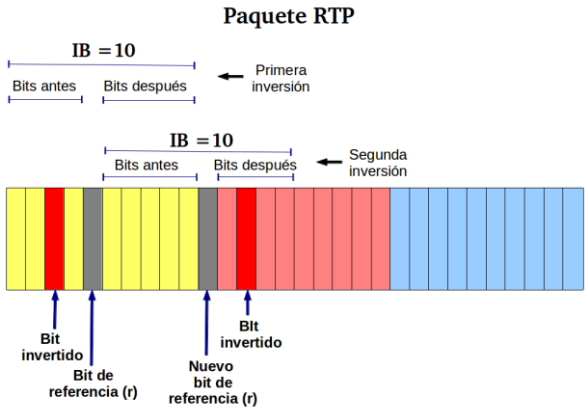


Figura 4: Esquema de inversión propuesto.

3. PROCESO DE CIFRADO

A continuación se explica de manera detallada los dos componentes principales del algoritmo de cifrado: inversión de bits y permutado de segmentos.

3.1 Inversión de Bits

La información a cifrar se divide en paquetes, el sistema de cifrado recibe cada uno de ellos y ejecuta uno de los pasos que conforman el algoritmo propuesto: destruir el significado de la secuencia de código por medio de la inversión de bits; de esta manera, resulta una tarea muy complicada el predecir la secuencia de original. Para lograr lo anterior, se propone invertir un bit cada IB bits, donde

$$IB = f \times PROM \times MEPL \quad (2)$$

PROM es el tamaño promedio de los códigos de Huffman (valor conocido para cada codificador de video) [12], MEPL es la longitud promedio de propagación de un error causado por la inversión de un bit dado en número de códigos, f representa un factor de seguridad que define que tan frecuente se modificarán los bits en el flujo de bits y cuyo valor debe estar en el rango $1/3 \leq f \leq 1$. El concepto puede ser explicado mediante el siguiente ejemplo: Supongamos que el conjunto A = {a1, a2, a3, a4, a5} se puede representar con los códigos {01, 1, 000, 0010, 0011} respectivamente, y que una secuencia de la forma: a2a1a3a2a1a2… es codificada como {1, 01, 000, 1, 01, 1, ...}. Ahora, supongamos que por un error, el valor del primer bit es invertido a 0, entonces la secuencia 0010001011· · · decodificada representa ahora la siguiente cadena a4a4a2a2 · · ·. El significado correcto ocurre después de cinco códigos (en la secuencia original), por lo tanto, se puede concluir que la longitud del error de propagación (MEPL) es de 5. En este trabajo, se utiliza el teorema propuesto por Zhou y Zhang [13] para calcular el valor de MEPL. En la Eq.2, PROM×MEPL es el promedio de la longitud de propagación de error en bits, por lo que (PROM×MEPL/3) ≤ IB ≤ (PROM×MEPL). Las características anteriores hacen que el esquema de cifrado propuesto sea un sistema de seguridad escalable, si f toma valores cercanos a su límite inferior, una mayor cantidad de bits son invertidos en cada paquete, incrementándose la seguridad del sistema. El valor de f define también el rendimiento o velocidad del sistema, este será mejor al sacrificar la seguridad.

El proceso de selección de bits a ser invertidos es el siguiente (ver Figura 4): **1)** Identificar un bit de referencia r en un grupo G

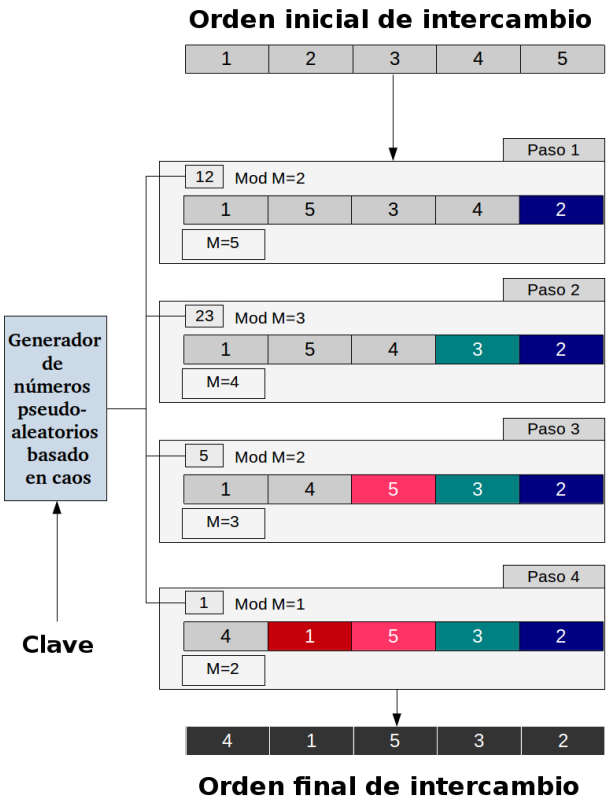


Figura 5: Esquema de mezclado aleatorio de segmentos. Fuente: Almasalha-Hasimoto [11].

de bits de tamaño IB, el bit r se deberá elegir de manera aleatoria, pero con la siguiente restricción: si consideramos que los IB bits de G son enumerados de la siguiente manera: $b_1 b_2 \cdots b_{IB-1} b_{IB}$, solo se podrá elegir cualquiera de los siguientes bits internos excepto b_1 y b_{IB}; **2)** Una vez que se ha elegido al bit de referencia r, se localiza ahora el bit a invertir. A partir del bit r, es posible cifrar uno de los bi bits del lado izquierdo o uno de los bd bits del lado derecho (incluyendo a r), esto es posible porque en el paso anterior, se aplicó la restricción de que el primer y último bit del conjunto G, no pueden ser bits de referencia r, por lo tanto, siempre hay bits a ambos lados del bit r. La decisión de invertir un bit en el lado izquierdo o derecho del bit de referencia, depende de la aleatoriedad de los números generados por los esquemas acoplados de mapas caóticos Rényi (ver sección anterior); si el número generado es par, se invertirá un bit en el lado izquierdo del bit de referencia r, si el número generado es impar, se invertirá un bit en el lado derecho del bit r; **3)** Al igual que en el paso anterior, el bit a invertir depende de los números pseudoaleatorios generados por el sistema de mapas caóticos acoplados. Al número pseudoaleatorio se le aplica la operación módulo (mod bi ó mod bd), que se ajusta al número posible de bits que se pueden invertir en el lado derecho (bd) o izquierdo (bi) del bit de referencia r; **4)** Cuando ya se ha invertido el bit seleccionado, elegimos un nuevo punto de inicio para realizar la siguiente inversión: será un bit posterior al bit actual de referencia r. A partir del nuevo bit de inicio, se busca un nuevo bit de referencia r de la misma manera que en el paso 1 y se repite el ciclo.

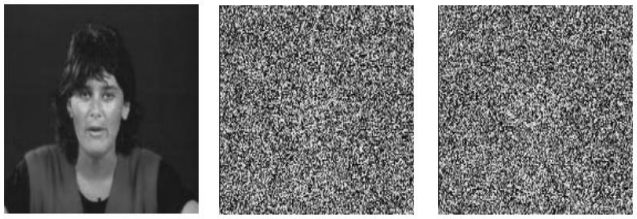


Figura 6. Sensibilidad a cambios mínimos en el algoritmo de descifrado para la imagen "miss". Izquierda: Imagen original. Centro: Imagen descifrada. Derecha: Imagen descifrada con la inversión de un bit de la función de acoplamiento H.

3.2 Proceso de Permutación de Segmentos

Los pasos para el mezclado de segmentos del paquete es el siguiente: **1)** Elección del número de segmentos 20<N<60 para cada paquete. Si N es grande, (N = 60), la complejidad del proceso de descifrado aumenta, ya que se necesitan realizar 60 permutaciones (pero aumenta la seguridad); **2)** El tamaño en bytes de cada segmento se define mediante S=P/N, donde P es el Tamño del paquete en bytes; **3)** Una vez que se ha definido el arreglo de segmentos, comienza un proceso iterativo para lograr la permutación de forma aleatoria: Se recorren N − 1 segmentos en el siguiente orden: N, N −1, ... , 3, 2. Supongamos que en una iteracion dada, se procesa el i-ésimo segmento, donde $2 \leq i \leq N$, este segmento se intercambiará con un segmento j, donde $1 \leq j \leq i-1$, la elección de j debe ser aleatoria, lo anterior se logra mediante la expresión j=T mod I; donde T es un valor pseudoaleatorio generado por el conjunto de mapas caóticos en Eq.1. La Figura 5, muestra la permutación de segmentos para cada uno de los paquetes de entrada.

3.3 Proceso de Descifrado

Para obtener el paquete original a partir de su correspondiente cifrrado, es necesario que el receptor posea la misma llave con la que se cifró dicha infromación, se debe enfatizar que el algoritmo propuesto es un sistema cifrador simétrico. Esto es, tanto el cifrado como el descifrado deben generar los mismos números pseudoaleatorios, lo que significa que los mapas caóticos en ambos lados son inicializados de manera idéntica. Por lo tanto, es posible conocer cuales fueron los bits invertidos y la manera en cómo se mezclaron los segmentos de cada paquet. Una ventaja del algoritmo propuesto, es que los números pseudoaleatorios utilizados durante la permutación de segmentos son independientes de los números pseudoaleatorios que se necesitan en la etapa de inversión de bits.

A diferencia del algoritmo de cifrado, el primer paso en el descifrado es la reordenación de los segmentos del paquete. Posteriormente se aplica el proceso inverso de inversión de bits para obtener el archivo original.

4. RESULTADOS

A continuación, se presentan los resultados obtenidos al aplicar el algoritmo cifrador propuesto a datos multimedia. Se analiza en primer lugar el comportamiento del PRNG propuesto así como las etapas de inversión de bits y permutación de segmentos por medio de la revisión del grado de aleatoriedad de las secuencias generadas. Se presenta un estudio analítico de la seguridad del algoritmo, donde se analiza la complejidad de obtener la información original una vez cifrada. En los experimentos se utilizó una llave de 256 bits (su tamaño puede dependiendo del número de mapas caóticos a utilizar. La llave

IB = 6	IB = 8	IB = 10	IB = 12	IB = 18	Archivos de prueba
126.24	175.21	228.56	275.12	430.29	1 music.flv
126.59	176.04	229.32	277.43	431.87	2 dolphins.3gp
127.21	177.14	230.3	277.89	432.71	3 helicopter.divx
126.75	176.37	228.97	275.62	429.36	4 tutorial.avi
126.6	176.11	229.87	277.54	431.25	5 wildlife.wmv
127.07	175.13	228.41	277.45	430.48	6 psycho.mpg

Figura 7: Velocidad de cifrado en Megabits/segundo para distintos valores de IB.

se utiliza para inicializar los mapas caóticos en Eq.1 (parámetros y valores inciales), donde N=4 mapas; 2 para el proceso de inversión de bits y 2 para la etapa de mezclado aleatorio de segmentos. La llave se puede crear con cualquier generador de números pseudoaleatorios y se comparte entre los nodos involucrados. El tamaño de los paquetes se fijó en 400 bytes con 20 segmentos, se evaluó IB para los siguientes valores {6, 8, 10, 12, 18}, donde MEPL y f se fijan de acuerdo a los valores de IB.

En Figura 6, se observa una imagen del video original de "miss" (izquierda) y su correspondiente imagen cifrada (centro) que representa como se vería el video en el recibidor si no se aplicara el proceso de descifrado. No se revela absoultamente nada de de la información original. Además de este comportamiento, un buen sistema de cifrado debe también ser sensible a los cambios en la llave en ambos procesos, cifrado y descifrado. Esto significa que si la misma información la ciframos con una llave distinta, los cifrados deben de ser completamente diferentes e ilegibles. Igualmente, si al descifrar se utiliza una llave distinta, no se podrá tampoco recuperar la información original. El algoritmo propuesto responde satisfactoriamente a cambios en la llave de un solo bit, independientemente de su significancia. En Figura 6 (derecha), la llave a descifrar fué modificada en el bit menos significativo obteniéndose una imagen totalmente ruidosa y estadísticamente independiente (correlación zero entre las imágenes) de cualquier descifrado con llave distinta a la original. Nuestro algoritmo es robusto a ataques conocidos en la literature como lo son texto plano conocido y elegido. Si el atacante decide atacar los mapas caóticos para conocer los valores iniciales y parámetros, necesitaría del orden de 2^{256} operaciones, esto es el tamaño de la llave (la cual es variable y se puede incrementar). Por otro lado, si el ataque es en los procesos de inversión de bits y permutado de segmentso, se require $O(N) = R!\,IB^{J/IB}$ operaciones, donde J es el tamño del paquete en bits (en este caso 400*8=3200), IB=6 (mínimo), y R! representa el número de permutaciones de R segmentos (R=20). El sistema es prácticamente invulnerable.

Adicionalmente, el algoritmo propuesto es el único en la literature que ofrece una seguridad escalable, debido al valor $1/3 \leq f \leq 1$ en Eq.2. El nivel máximo de seguridad se alcanza cuando f=1/3 (IB = 6); sin embargo, se sacrifica la velocidad del algoritmo; con f = 1 (IB=18), la velocidad del algoritmo aumenta considerablemente, pero la seguridad disminuye. Con base en las consideraciones anteriores, se puede medir el volumen de datos cifrados, que es el porcentaje de información de un archivo que se ha modificado al momento de cifrarlo. Para $6 \leq IB \leq 18$, el porcentaje de inversión de bits oscila entre 5.56 % y 16.67 %. Para determinar si el algoritmo propuesto es suficiente para llevar a cabo cifrado de video en tiempo real, se midió su rendimiento o velocidad de ejecución en un equipo de cómputo de mediano poder con las siguientes características: Memoria RAM: 7 GB, Procesador: AMD A10-5800B APU (introducido en el 2012), 3.8 GHz, 4 núcleos. Los resultados se muestran en la Figura 7, para distintos valores de IB. Dependiendo del nivel de seguridad, el sistema presenta una velocidad entre 126-430 Mbs (Mega Bits por segundo), pudiéndo cifrar video en tiempo real para uso de Internet hasta ultra HD de 25Mbs en líneas de television por cable.

5. CONCLUSIONES

En este trabajo, hemos presentado un esquema de cifrado parcial y escalable basado en mapas caóticos para comunicación segura de video en tiempo real. Una de las principales ventajas del sistema de cifrado propuesto es su flexibilidad con respecto a la enorme cantidad de estándares de compresión que actualmente existen; no es necesario conocer a los detalles de los codificadores de video para lograr un cifrado eficaz, seguro y eficiente. Los mapas caóticos Rényi presentan robustez a los valores iniciales, esto significa que las inicializaciones de los mapas puede realizarse de manera aleatoria; en general, las trayectorias muestran un comportamiento caótico de acuerdo a las pruebas NIST. El sistema desarrollado logra velocidades de hasta 430Mbs, suficiente para cifrar television digital en Utra HD.

Las pruebas de seguridad realizadas muestran que el algoritmo de cifrado parcial propuesto, produce histogramas del texto plano y su texto cifrado independientes y extremadamente sensibles a los cambios de bit en la llave inicial. Se concluye que algoritmo cumple con los estándares de seguridad definidos en el área de criptografía caótica.

6. AGRADECIMIENTOS

Este trabajo fué parcialmente financiado por el Consejo Nacional de Ciencia y Tecnología de México-CONACYT, bajo la convocatoria de Apoyo a Proyectos Nacionales con solicitud #247802.

7. REFERENCIAS

[1] S. Khanvilkar and A. Khokhar. NIS07-1: Efficient Transmission of MP3 Streams over VPNs. In Global Telecommunications Conference, 2006. GLOBECOM-06. IEEE, pages 1–5, nov 2006.

[2] J. Meyer and F. Gadegast. Security Mechanisms for Multimedia Data with the Example MPEG-1 Video, 1995.

[3] Lei Tang. Methods for Encrypting and Decrypting MPEG Video Data Efficiently. In Proceedings of the 4th ACM International Multimedia Conference, pages 219–230, Boston, Massachusetts, USA, 1996.

[4] Shiguo Lian, Jinsheng Sun, Guangjie Liu, and Zhiquan Wang. Efficient video encryption scheme based on advanced video coding. Multimedia Tools and Applications, 38(1):75–89, 2008.

[5] Y. Li, L. Liang, Z. Su, and J. Jiang. A new video encryption algorithm for H.264. IEEE 5th International Conference on Information, Communications, and Signal Processing, pages 1121–1124, 2005.

[6] J. Jiang, Y. Li, and L. Liang. Research and improvement of the video encryption algorithm for H.264. Acta Electronica Sinica 9, (9):1725–1726, 2007.

[7] S. Spinsante, F. Chiaraluce, and E. Gambi. Masking video information by partial encryption of H.264/AVC coding parameters. 2005.

[8] S. Lian, Z. Liu, Z. Ren, and Z. Wang. Selective Video Encryption Based on Advanced Video Coding. Advances in Multimedia Information Processing - PCM 2005, 3768:281–290, 2005.

[9] Xiaofeng Wang, Nanning Zheng, and Lihua Tian. Hash key-based video encryption scheme for H.264/AVC. Signal Processing: Image Communication, 25:427–437, 2010.

[10] C. Yuan, Y. Zhong, and Y. He. Selective video stream encryption algorithm based on chaos.Chinese Journal of Computers, 27(2):257–263, 2004.

[11] F. Almasalha, R. Hasimoto-Beltran, A. A. Khokhar. Partial Encryption of Entropy-Coded Video Compression Using Coupled Chaotic Maps. Entropy 2014, 16, 5575-5600

[12] National Institute of Standards and Technology (NIST). FIPS-197: Advanced Encryption Standard (AES).

[13] G. Zhou and Z. Zhang. Synchronization Recovery of Variable-Length Codes. IEEE Transactions on Information Theory, 48(1):219–227, jan 2002.

Implementación de Canal Seguro para la Protección de Imágenes de Actas de Escrutinio

Oscar Ruiz H.
Departamento de Eléctrica y Computación
Instituto de Ingeniería y Tecnología, UACJ
Ciudad Juárez, Chih, México
al150668@alumnos.uacj.mx

Víctor Morales R.
Departamento de Eléctrica y Computación
Instituto de Ingeniería y Tecnología, UACJ
Ciudad Juárez, Chih, México
victor.morales@uacj.mx

RESUMEN

En este artículo se describe la implementación de un conjunto de mecanismos que permiten establecer un canal seguro de comunicación entre aplicación y servidor. La necesidad de la implementación de un canal seguro de comunicación es una parte primordial del proyecto ERRS. Proyecto desarrollado por las empresas Scytl y Tecnologías Rethos en conjunto con la Universidad Autónoma de Ciudad Juárez. Debido a que ERRS pretende manejar imágenes de actas de escrutinio surge la necesidad de asegurar la integridad de éstas y proteger el medio por el cual serán transmitidas. Para establecer un canal de comunicación cifrado y autenticar a los participantes se implementó MTLS. La integridad de la información enviada se protege y verifica mediante la implementación de JWT. Para la autenticación de usuarios y distribución de certificados y claves privadas se implementó el protocolo *key roaming*.

Palabras clave

Canal Seguro, Auditoría electoral, JWT, MTLS.

1. INTRODUCCIÓN

La transparencia en los procesos electorales aporta credibilidad a las instituciones electorales y legitimidad a estos procesos. Al realizar auditorías del proceso de votación se puede garantizar la integridad del proceso electoral.

Election Results Reliable Sampling (ERRS) es un proyecto de investigación y desarrollo tecnológico. Desarrollado por las empresas Scytl y Tecnologías Rethos en conjunto con la Universidad autónoma de ciudad Juárez. ERRS se propone como un innovador sistema de verificación del resultado de elecciones mediante el desarrollo de tecnologías para la validación del proceso de escrutinio electoral. El proceso de validación propuesto se llevará a cabo a partir de la captura de imágenes, mediante dispositivos móviles, de las actas de resultados de una muestra representativa de colegios electorales. Dichas actas serán protegidas mediante avanzadas tecnologías criptográficas en los mismos dispositivos que las han capturado y durante su transmisión, para garantizar su integridad antes y después de ser enviadas al servidor central. Las actas capturadas se procesarán mediante tecnologías de procesamiento de imágenes en el mismo dispositivo o en el servidor, para su posterior comparación con los resultados oficiales.

Debido a la sensibilidad de la información que pretende manejar ERRS, surge la necesidad de implementar tecnologías avanzadas de seguridad que permitan garantizar la integridad, privacidad y autenticación de toda la plataforma, especialmente de las imágenes de actas de escrutinio, usuarios intervinientes y canales de comunicación.

En este trabajo se describe la implementación de mecanismos que permiten establecer un canal seguro de extremo a extremo. Para lograr esto, en cada conexión entre la aplicación y el servidor se utiliza un canal de comunicación cifrado y se autentican ambas partes. Además se realiza la autenticación de usuarios y distribución de certificados y claves privadas, mediante el protocolo *key roaming*. Este protocolo fue desarrollado e implementado por Scytl, empresa que ofrece soluciones de voto electrónico y también participa en el desarrollo del proyecto ERRS. Por último la aplicación protege la integridad de la información que será transmitida hacia el servidor.

2. MECANISMOS DE SEGURIDAD

El Instituto Nacional de Estándares y Tecnología (NIST por sus siglas en inglés, *National Institute of Standards and Technology*) define en [1] a la seguridad de la información como la protección de la información y de los sistemas de información de acceso, uso, divulgación, alteración, modificación o destrucción no autorizados con el fin de garantizar:

- Confidencialidad: Sólo se le permite el acceso a la información a usuarios autorizados.
- Integridad: Protección de la información contra modificación o destrucción no autorizadas.
- Disponibilidad: Garantizar el acceso oportuno y confiable para el uso de la información por parte de usuarios válidos.

En la mayoría de las aplicaciones que utilizan una comunicación entre al menos dos partes (dispositivos, usuarios o procesos) se desea cumplir con uno o más de los servicios de seguridad previamente descritos. Ejemplos de estas aplicaciones son los sistemas de comercio electrónico, correo electrónico, voz sobre IP, etc. La dificultad para lograr una comunicación segura en un entorno de red abierta, por ejemplo en Internet o alguna red inalámbrica, consiste principalmente en que el canal de comunicación es público, por lo tanto es relativamente trivial para un intruso interceptar una comunicación y capturar la información transmitida, e incluso eliminarla o sustituirla. Debido a esta necesidad se han desarrollo una variedad de protocolos y/o mecanismos que permiten establecer un canal seguro de extremo a extremo. Tal es el caso de los protocolos SSH [2], IPSec [3], SSL/TLS [4] y HTTPS [5], así como de mecanismos que se apoyan en dichos protocolos, como es el caso de la implementación de redes privadas virtuales. Un canal seguro puede establecer mecanismos para la autenticación de una o ambas partes de la comunicación, de acuerdo a la necesidad, y mecanismos que permitan proteger la confidencialidad de la información transmitida, generalmente a través de técnicas de

cifrado, así como para proteger la integridad de la misma (o al menos poder detectar cuando la información ha sido alterada).

3. IMPLEMENTACION DE CANAL SEGURO DE COMUNICACIÓN

Con el sistema ERRS se pretende que los usuarios del sistema capturen la imagen del acta de escrutinio mediante un dispositivo móvil para posteriormente ser enviada al servidor para su procesamiento y almacenamiento.

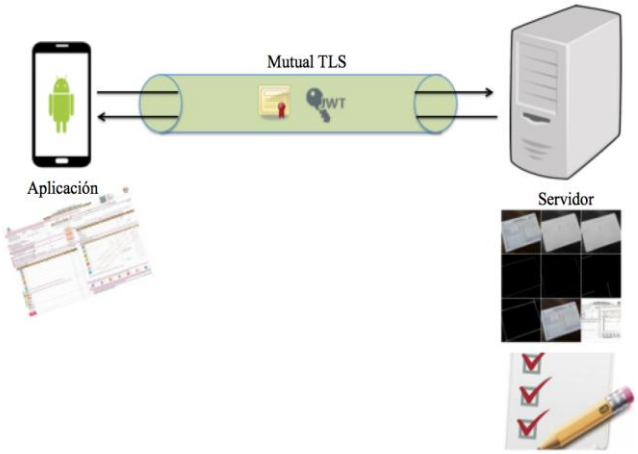


Figura 1. ERRS

Cada vez que la aplicación establezca comunicación con el servidor se debe de realizar mediante un canal seguro de comunicación. Para lograr esto se realizó la implementación de *Mutual Transport Layer Security* (MTLS) [6], que ofrece una comunicación cifrada y la autenticación de ambas partes. En MTLS el servidor autentica al cliente y el cliente autentica al servidor, esto es posible gracias al uso de certificados digitales. La aplicación tiene instalado el certificado del servidor y su propio certificado. El servidor de igual manera contiene su propio certificado y el certificado de la aplicación. De esta manera, al intentar establecer comunicación, los certificados que se intercambian son comparados con los que previamente se instalaron en ambas partes. Así se asegura que sólo las partes autorizadas pueden establecer comunicación. Durante la autenticación de los participantes también se negocian los parámetros que serán utilizados para establecer una canal cifrado sí la autenticación es exitosa. Una descripción general del funcionamiento de MTLS se puede observar en la figura 2.

Una vez que la aplicación establece un canal seguro de comunicación se puede enviar información al servidor siempre y cuando el usuario esté autenticado en la aplicación. Para la autenticación de usuarios se realizó la implementación del protocolo *Key Roaming*. Los dos objetivos principales del protocolo son: realizar una autenticación de usuarios robusta, de manera transparente para el usuario, y la distribución de certificados.

La robustez de *key roaming* radica en que para cada usuario existe un *keystore* que contiene su clave privada y su certificado. Este *keystore* está protegido con una clave derivada de las credenciales del usuario. De este modo se asegura que solamente él pueda abrirlo y hacer uso de su contenido.

Para la implementación de *key roaming* se requiere un proceso de creación de credenciales que consiste en:

1. Definir el número de usuarios N. Para definir N también se deben de considerar usuarios extras, ya que *key roaming* no permite recuperar cuentas. Para esto se debe de contar con credenciales extras y cuando un usuario olvide o extravíe sus credenciales se le asignan unas nuevas.

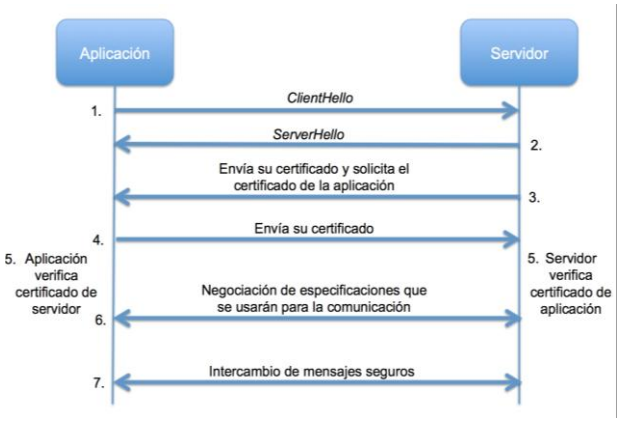


Figura 2. MTLS.

2. Se crean N nombres de usuarios con sus respectivas contraseñas. Los nombres de usuario y contraseña son almacenados en un archivo para posteriormente ser distribuidos a los usuarios. Cabe mencionar que no existe la necesidad de almacenar estas credenciales en el servidor porque nunca hará uso de ellas, por lo tanto una vez distribuidas se deberá eliminar el archivo que las contiene.

3. Se genera el *keystore* de cada usuario. Para cada keystore se genera una clave privada y con esa clave se genera un certificado digital que contiene la clave pública. Este certificado es firmado por una *Autoridad Certificadora* (CA) para futuras validaciones. Para nombrar y cerrar el keystore se calcula un *Password-Based Key Derivation Function 2* (PBKDF2) en base a las credenciales de cada usuario. Esta función sirve para generar una clave criptográficamente segura basándose en una contraseña.

Para nombrar el archivo se usa una PBKDF2, que a partir de una concatenación de nombre de usuario + contraseña y una semilla *salt1* (Constante) dá como resultado el ID, que será usado para identificar el *keystore*. Como ya se mencionó, el *keystore* deberá asegurarse con una clave, ésta se genera casi de manera similar al ID, con la única variación que la constante, es decir *salt2*, usada como base es diferente a la que se usa para generar el ID. Hacer uso de dos constantes diferentes nos permite generar dos claves completamente diferentes y sin relación.

Cada *keystore* será almacenado en un archivo PKCS12 [7], formato especial para almacenar objetos criptográficos en un solo y protegerlos con una clave. Estos archivos deberán quedar almacenados en el servidor.

Con *key roaming* aparentemente la autenticación se realiza de la forma tradicional, el usuario tiene un nombre de usuario y una contraseña que las introduce en la aplicación y son enviadas a un servidor para validarlas. La única diferencia es que el servidor no compara lo recibido contra un registro de usuarios válidos.

La aplicación envía el nombre del *keystore* que le corresponde al usuario y entonces el servidor responde con el *keystore* de ese usuario si lo encuentra y en caso contrario le responde solo con un error. Entonces la aplicación al recibir el *keystore* calcula la contraseña para abrirlo y poder hacer uso de su clave privada y

certificado, es decir se distribuyen la clave privada y certificado para posteriormente enviar información de forma segura hacia el servidor. En la figura 3 se describe el funcionamiento de *key roaming*.

El uso de *key roaming* agrega algunas ventajas y también comparte algunas con el método de autenticación tradicional. En la tabla 1 se muestra una comparación de ambos métodos. Ambos métodos son transparentes para el usuario y no requieren de equipo especializado para su implementación. Una desventaja de *key roaming* es que no permite recuperar cuentas, esto debido a que no se cuenta con un registro de las credenciales de usuarios. Sin embargo, esta desventaja también se puede tomar como un punto a favor ya que al no existir un registro, desaparece la posibilidad de un ataque para la obtención de estas credenciales. Otra ventaja de *key roaming* es que al tener la posibilidad de crear las credenciales, se crean contraseñas fuertes. Por último la ventajas mas importante es que permite la distribución de certificados digitales para su posterior uso, en este caso asegurar y validar la integridad de la información enviada.

Tabla 1. Tabla comparativa

Características	**Autenticación con Usuario y Contraseña**	**Autenticación con *key roaming***
Uso de base de datos para usuarios válidos	Necesario	No necesario
Recuperación de cuentas	Disponible	No disponible (deseado en algunos escenarios)
Uso de contraseñas fuertes	No garantizado	Garantizado
Susceptible a Ataque de fuerza bruta	Si	Se vuelve más complejo
Robustez	Se limita a la protección de la contraseña	Basado en algoritmos criptográficos
Requiere equipo especializado para su implementación	No	No

Debido a la sensibilidad de la información que la aplicación enviará al servidor, existe la necesidad de asegurar su integridad y autenticidad. Para cumplir con lo antes mencionado se utilizaron implementaciones de criptografía asimétrica como se describe a continuación. La aplicación firma digitalmente con la clave privada del usuario la información que se desea enviar y adjunta el certificado del usuario en el mensaje.

Cuando el servidor recibe el mensaje, extrae el certificado del usuario, después comprueba la validez del certificado y por último verifica la integridad de la información recibida con la clave pública contenida en el certificado del usuario.

Para lograr lo anterior se realizó la implementación de JSON Web Token (JWT) [8], que es un estándar abierto (RFC 7519) que define una forma compacta y auto contenida para la transmisión segura de información entre quienes se comunican, esto mediante un objeto JSON. En la figura 4 se explica el proceso de envío y verificación de información entre el servidor y la aplicación.

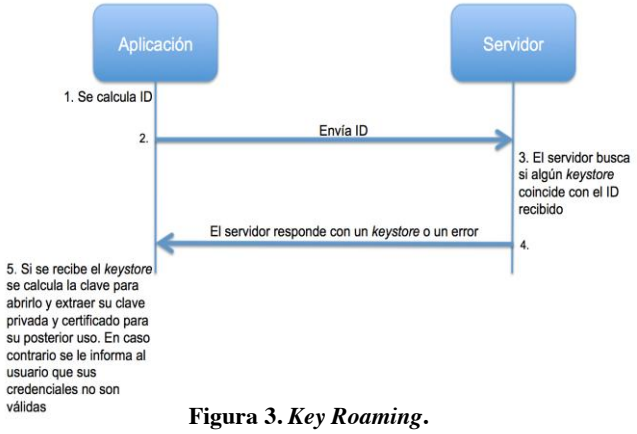


Figura 3. *Key Roaming*.

Se desarrolló un prototipo de aplicación móvil para sistema operativo Android desde donde el usuario se autentica y envía un mensaje, imagen o ambos. Todo esto a través de un canal de comunicación seguro. Desde la aplicación el usuario introduce su nombre de usuario y contraseña y presiona iniciar sesión. Si las credenciales son correctas, el servidor envía el *keystore* del usuario y se abre una nueva ventana.

Del *keystore* recibido se extrae el certificado y la clave privada del usuario. En la nueva ventana ya se puede enviar información al servidor de forma segura. Cuando el usuario envía información se firma para proteger su integridad durante el trayecto hacia el servidor. Una vez recibida por el servidor, éste verifica la integridad de la información, si no se encuentran incongruencias se almacena y se informa al usuario que el servidor ha recibido la información correctamente desplegándole un aviso en la aplicación.

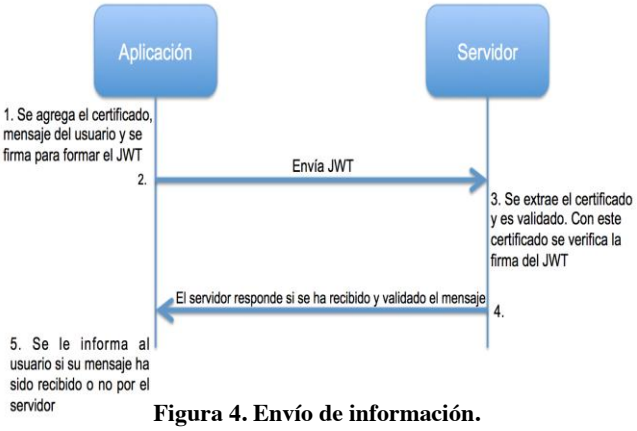


Figura 4. Envío de información.

4. CONCLUSIONES

Al realizar la implementación de JWT para el envío seguro de información se puede garantizar que de realizarse alguna modificación no autorizada durante el trayecto hacia el servidor, ésta puede ser detectada. JWT protege y ofrece la posibilidad de verificar la integridad. Esto gracias a que se firma con la clave privada del usuario y se verifica con la clave pública. Para la fácil distribución de esta claves se realizó la implementación de *key*

roaming. Este protocolo también ofrece una autenticación robusta de usuarios aparentando ser un método de autenticación tradicional. Con esto se ayuda a que el usuario no tenga que realizar pasos extras para una autenticación robusta. Por ejemplo, cuando se utiliza el método de autenticación con token y credenciales del usuario se requiere del uso de un dispositivo adicional, esto lo convierte más complejo de utilizar comparado con solamente utilizar credenciales del usuario.

Al implementar MTLS se establece un canal de comunicación cifrado para garantizar que la información no pueda ser visible en caso de ser interceptada por alguien no autorizado. Además de ofrecer el cifrado también se asegura que quien intenta establecer comunicación con el servidor es una persona autorizada y viceversa. Al implementar este conjunto de mecanismos se establece un canal seguro de extremo a extremo entre la aplicación y el servidor.

Debido que el proyecto ERRS se encuentra en desarrollo, como trabajo futuro se realizarán pruebas al prototipo desarrollado para medir el nivel de seguridad que ofrece esta implementación y también realizar mejoras. También se trabajará en implementar y probar mecanismos de seguridad que garanticen la integridad, autenticidad y confidencialidad de todos los módulos restantes del sistema para dar como resultado un esquema de seguridad.

5. AGRADECIMIENTOS

El proyecto ERRS se lleva a cabo a través del fondo bilateral CONACYT-CDTI bajo el convenio 2013-218454.

6. REFERENCIAS

[1] Glossary of Key Information Security Terms. (2013, Mayo) NIST. [Online]. http://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7298r2.pdf

[2] Daniel J. Barrett and Richard Silverman, *SSH, The Secure Shell: The Definitive Guide*. Sebastopol, California, USA: O'Reilly, 2001. [Online]. ftp://ftp.bupt.edu.cn/pub/Documents/Networking.TCP_IP/O'Reilly%20-%20SSH.The%20Secure%20Shell%20-%20The%20Definitive%20Guide%20Ebook%20.pdf

[3] S. Frankel and S. Krishnan. (2011, Febrero) Internet Engineering Task Force. [Online]. https://tools.ietf.org/pdf/rfc6071.pdf

[4] Holly Lynne McKinley. (2003) InfoSec Reading Room. [Online]. https://www.sans.org/reading-room/whitepapers/protocols/ssl-tls-beginners-guide-1029

[5] E. Rescorla. (2000, Mayo) Internet Engineering Task Force. [Online]. https://tools.ietf.org/pdf/rfc2818.pdf

[6] Microsoft. [Online]. https://technet.microsoft.com/en-us/library/bb964088(v=office.12).aspx

[7] K. Moriarty, M. Nystrom, S. Parkinson, A. Rusch, and M. Scott. (2014, Julio) Internet Engineering Task Force. [Online]. https://tools.ietf.org/pdf/rfc7292.pdf

[8] M. Jones, J. Bradley, and N. Sakimura. (2013, Noviembre) Internet Engineering Task Force. [Online]. https://tools.ietf.org/pdf/rfc2818.pdf

A Li-Fi Based Application with an Android Device

Sergio Sandoval-Reyes
CIC Instituto Politecnico Nacional
Av. Juan de Dios Batiz s/n, Col. Nueva
Industrial Vallejo, Mexico City 07738
+52 55 5729-6000 xt 56502
sersand@cic.ipn.mx

Victor Hugo Herver-Segura
CIC Instituto Politecnico Nacional
Av. Juan de Dios Batiz s/n, Col. Nueva
Industrial Vallejo, Mexico City 07738
+52 55 5729-6000 xt 56502
vicherver@gmail.com

ABSTRACT
Li-Fi is a technology that uses visible light from a Light Emitter Diode (LED) to transmit high speed data to a photo detector, which is connected to a device. Location awareness refers to devices that can passively or actively determine their location with respect to a well-known location wireless communications device. Here we describe a Li-Fi based application, in which a tablet with a light sensor and within the range of a Li-Fi lamplight, displays a picture, access a Web page, or plays a video.

CCS Concepts
• C.2.1 [**Computer-Communication Networks**]: Data Communications—*Wireless Communication*

Keywords
Li-Fi; Location awareness; Android.

1. INTRODUCTION
Li-Fi is a term that was coined by Harald Haas[1, 2] and is a form of visible light communication (VLC) and a subset of optical wireless communications (OWC). This OWC technology uses light from light-emitting diodes (LEDs) as a medium to deliver networked, mobile, high-speed communication in a similar manner to Wi-Fi [3]. Visible light communications (VLC) works by switching the current to the LEDs off and on at a very high rate, too quick to be noticed by the human eye. While Wi-Fi is close to full capacity, Li-Fi has almost no limitations on capacity. The visible light spectrum is 10,000 times larger than the entire radio frequency spectrum [4]. Researchers have reached rates of over 10 Gbit/s, which is much faster than typical fast broadband [5].

Location awareness refers to devices that can passively or actively determine their location with respect to a well-known location wireless communications device. Location awareness enables new applications for ubiquitous computing systems and mobile phones or tablets. This paper describes a Li-Fi based application, in which a smart phone or tablet with a light sensor and within the range of a Li-Fi lamplight, immediately trigger events like displaying a picture, accessing a Web page, or playing a video.

2. RELATED WORK

2.1 LEDs--leads-you-right-to-a-discount

In [7, 8] ByteLight an English Company focused on using LEDs to provide high-speed data communications—the technology referred to as Li-Fi, felt their technology was better suited to helping people find their way around large indoor spaces. The idea was the following: suppose you are in a department store that has replaced a number of its traditional lightbulbs with ByteLights. The lights, flickering faster than the eye can see, would emit a signal to passing smart phones. Your phone would read the signal through its camera, which would direct the smart phone (Figure 1), to pull up a deal offering a discount on a product on a nearby rack.



Figure 1. A ByteLight's LED bulbs system. Taken from http://www.acuitybrands.com/solutions/services/bytelight-services-indoor-positioning

2.2 Philips shopping assistant with LEDs

In [9] Phillips Lighting showed a LED-based indoor location detection technology. The system is similar to ByteLight, in terms of location-determination technology that relies on one-way communication between networked LED-based luminaires and customers' smartphones. In both cases the customer requires however, to download an app to utilize the technology. The communication link from luminaires to the smartphone would deliver the location data and offers. Figure 2. It also sells kits to experiment with the Li-Fi technology. See Figure 3.

2. A Location Aware Li-Fi Application
In the following we develop a location-aware Li-Fi application in which when an Android phone or tablet with a light sensor, is within the range of a Li-Fi lamplight, it will immediately display a picture, access a Web page, or play a video, depending of the ID code of the lamplight. In order to do that, we will use the Oledcomm GEOLiFi XS Kit [11]. This kit contains: a) 3 GEOLiFi LED lamps, b) 1 GEOLiFi Android tablet with an embedded sensor light; c) 1 extra GEOLiFi dongle to be used with a smartphone; and d) 1 GEOLiFi SDK Library. See Figure 4.

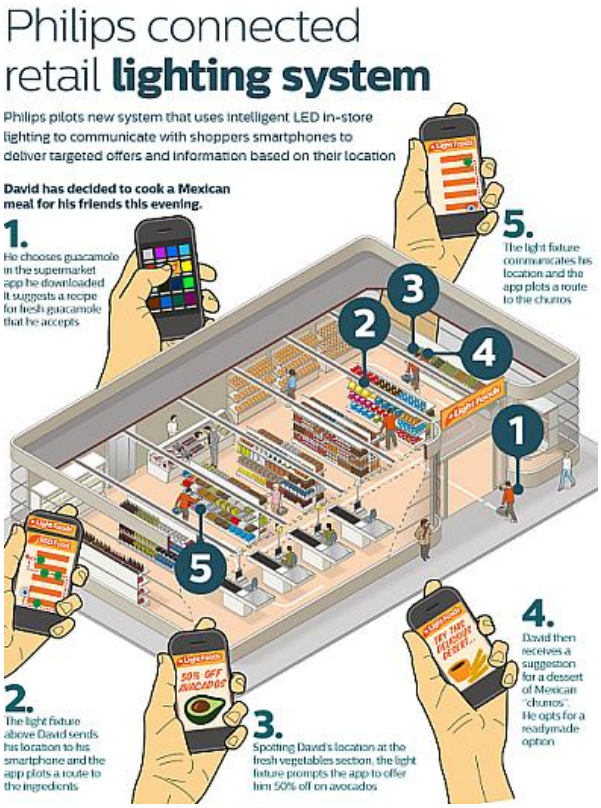


Figure 2. Philips Lighting LED-based indoor location detection technology. Taken from: http://spectrum.ieee.org/tech-talk/computing/networks/philips-creates-store-shopping-assistant-with-leds-and-smart-phone



Figure 3. Oledcomm Li-Fi. Taken from: http://www.oledcomm.com/home/lifi



Figure 4. The Oledcomm GEOLiFi XS Kit. Taken from http://www.oledcomm.com/product/kit-lifi-xs

2.1 Displaying a Picture with LiFi

The GEOLiFi SDK Library has a small tutorial of how to use the Oledcomm GEOLiFi API. We begin loading the last version of Java SDK and Android Studio. After that we create a new Android application specifying the project and packages names with (in our case), mínimum SDK 8, target SDK (API 19) and compiling with SDK 20. The most important part is in the MainActivity class in the *LiFiLocation ()* method, Figure 5.

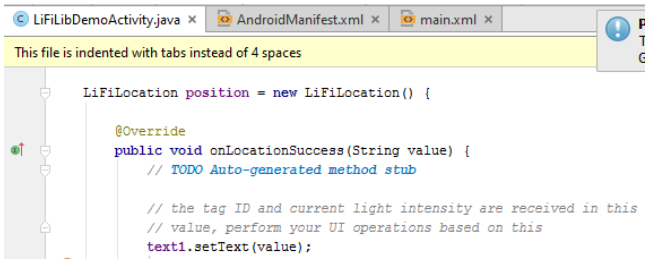


Figure 5. The LiFiLocation () method.

This method is used for retrieving geolocation of a phone or tablet receiving data from a Li-Fi lamp. When this happens the corresponding Id lamp will be available. In our case, the Id of every one of the three lamps were:

a) Lamp 1: 0xd5d5
b) Lamp 2: 0x6390
c) Lamp 3: 0x6e71

Having known the Ids of the three lamps, we simply use the first Id (0xd5d5) for displaying a picture in the tablet with the follwing java code, Figure 6:

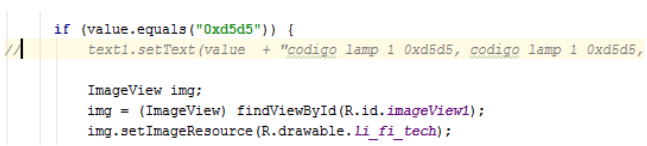


Figure 6 Code for displaying a picture when the tablet receives the ligth code from lamp1.

Of course, the *li_fi_tech.png* picture must be already available in the *res --> drawable directory*.

2.2 Displaying a Web page with Li-Fi

The following java code allows access to a Web page when the tablet receives the light code (0x6390) from lamp 2. Figure 7.

```
} else if (value.equals("0x6390")) {
            text1.setText(value + "código lamp 2 0x6390, codigo lamp 2 0x6390,

            WebView webView;
            String home="http://www.oledcomm.com";
            webView = (WebView) findViewById(R.id.webView);
            webView.loadUrl(home);
```

Figure 7. Accessing a Web page when the tablet receives the light code from lamp 2.

2.3 Displaying a video with Li-Fi

To display a video using Li-Fi, we use the light code *0x6e71* and the following java code, Figure 8. The *video.mp4* must be already stored in this case, in the res→raw directory.

```
} else if (value.equals("0x6e71")) {
            text1.setText(value + "código lamp 3 0x6e71, codigo lamp 3 0x6e71, codigo lamp 3 0x6e71");

            getWindow().setFormat(PixelFormat.UNKNOWN);
            VideoView mVideoView;
            mVideoView = (VideoView) findViewById(R.id.surface_view);
            String uriPath ="android.resource://com.example.lifilibrarydemo/"+ R.raw.video;
            Uri uri2=Uri.parse(uriPath);
            mVideoView.setVideoURI(uri2);
            mVideoView.setVideoPath("android.resource://com.example.lifilibrarydemo/"+ R.raw.video);
            mVideoView.setVideoPath("http://www.ebookfrenzy.com/android_book/movie.mp4");
            mVideoView.start();
            mVideoView.requestFocus();
} else {

      text1.setText("no se pudo leer codigo lamp 3, no se pudo leer codigo lamp 3  " + value);
   }
```

Figure 8. Displaying a video with Li-Fi.

3. EXPERIMENTS AND RESULTS

For the experiments we use as was mentioned an Oledcomm GEOLiFi XS Kit with an Android OS 4.2.4 (Figure 9) tablet, thus as another Toshiba AT100 tablet with Android OS 3.1, but with a light sensor connected into the airphone jack (Figure 10).



Figure 9. GeoLiFi Kit Android tablet 4.2.4 and three Li-Fi lamps.

Figures 10 and 11 shown the corresponding picture, Web page (overlapping the picture) and the video. In figure 11 it can also be noticed the embedded Li-Fi sensor (down to the right from the camera lens).

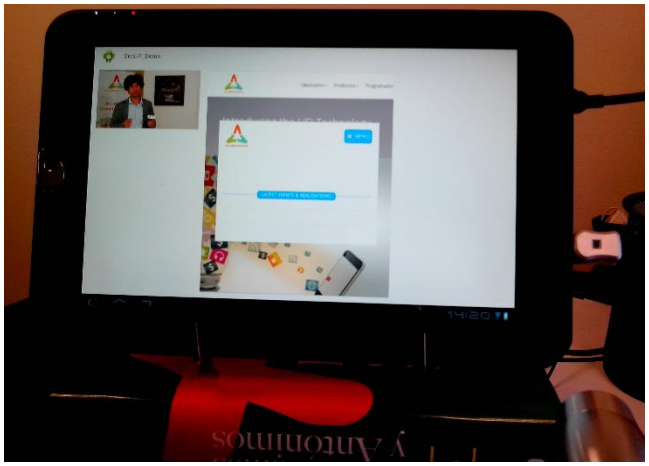


Figure 10. Toshiba tablet with attached Li-Fi sensor into the airphone jack.

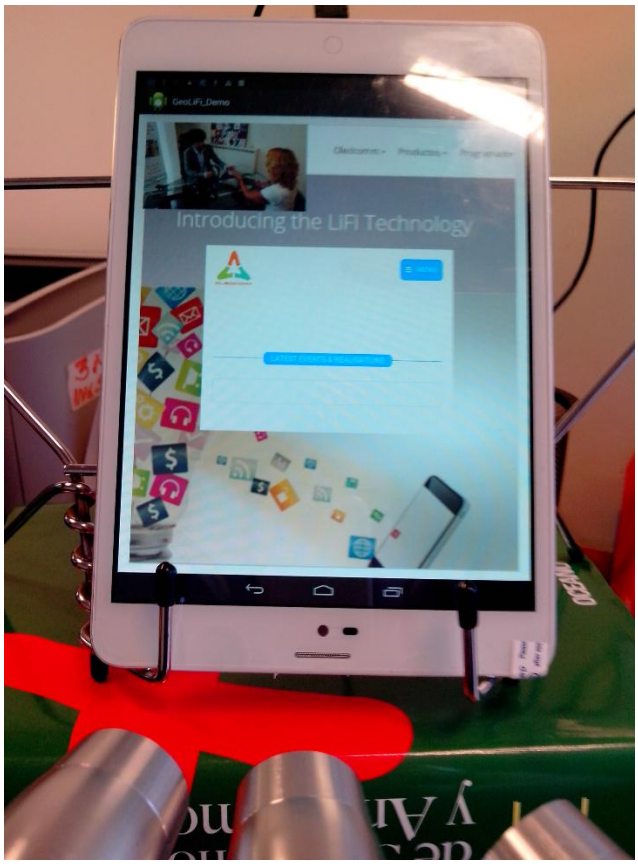


Figure 11. Tablet with Li-Fi sensor embedded.

4.1 Discussion of Results

Both tablets worked fine with the Li-Fi lamps. However the tablet with the attached Li-Fi sensor was less sensible to the Li-Fi light than the Oledcomm tablet.

4. CONCLUSIONS

A location-aware Li-Fi application was developed using three Li-Fi lamps and Android tablets with a Li-Fi sensor. When the tablet was within the range of a Li-Fi lamplight, the light sensor read the light code of the lamp under test and triggered events like

displaying a picture, accessing a Web page, or playing a video, according to the respective light code.

5. ACKNOWLEDGMENTS

This work was supported by the *Instituto Politecnico Nacional*, SIP Project number 20161439.

6. REFERENCES

[1] TED Ideas worth spreading. "Harald Haas: Wireless data from every light bulb". 2011. http://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb .

[2] Tsonev, D., Videv, S. and Haas, H. December 18, 2013. "Light fidelity (Li-Fi): towards all-optical networking". Proc. SPIE (Broadband Access Communication Technologies VIII) 9007 (2). DOI: http://dx.doi.org/10.1117/12.2044649 .

[3] Sherman, J. 2013. "How LED Light Bulbs could replace Wi-Fi", Digital Trends. http://www.digitaltrends.com/mobile/light-bulb-li-fi-wireless-internet/

[4] Haas, H. 2013. "High-speed wireless networking using visible light". SPIE Newsroom. DOI: http://dx.doi.org/10.1117/2.1201304.004773

[5] Vincent, J. 2013. "Li-Fi revolution: internet connections using light bulbs are 250 times faster than broadband. http://www.independent.co.uk/news/science/li-fi-revolution-internet-connections-using-light-bulbs-are-250-times-faster-than-broadband-8909320.html

[6] Wikipedia. 2011. "Location awareness", https://en.wikipedia.org/wiki/Location_awareness

[7] MIT Technology Review. 2012. "LEDs Could Lead You Right to a Discount". https://www.technologyreview.com/s/426869/leds-could-lead-you-right-to-a-discount/

[8] AcuityBrands. 2016. "ByteLight Services: Indoor positioning". http://www.acuitybrands.com/solutions/services/bytelight-services-indoor-positioning

[9] IEEE Spectrum. 2014. "Philips creates shopping assistant with LEDs and Smart Phone". http://spectrum.ieee.org/tech-talk/computing/networks/philips-creates-store-shopping-assistant-with-leds-and-smart-phone

[10] Oledcomm Li-Fi. 2016. http://www.oledcomm.com/home/lifi

[11] Oledcomm, GEOLiFI XS Kit. 2016. http://www.oledcomm.com/product/kit-lifi-xs