SOCIEDAD MEXICANA DE CIENCIAS DE LA COMPUTACIÓN

SMCC

CONACYT
Consejo Nacional de Ciencia y Tecnología

APLICACIONES DE LA COMPUTACIÓN

INTELIGENCIA ARTIFICIAL: MÉTODOS, ALGORITMOS Y APLICACIONES

APLICACIONES DE CÓMPUTO SUAVE Y ANALÍTICA DE DATOS

APLICACIONES EN EL SECTOR ELÉCTRICO

CIBERSEGURIDAD Y CRIPTOGRAFÍA

SISTEMAS INTELIGENTES EN INTERNET DE LAS COSAS

INFORMÁTICA MÉDICA PARA LA SALUD Y EL BIENESTAR

UNIVERSIDAD AUTÓNOMA DE COAHUILA

CIMA

CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS APLICADAS

APICACIONES DE LA COMPUTACIÓN

Coordinadores:
María Lucia Barrón Estrada
José Antonio Camarena Ibarrola
Karina Mariela Figueroa Mora
Valeria Soto Mendoza

Aplicaciones de la Computación
@ María Lucia Barrón Estrada
@ José Antonio Camarena Ibarrola
@ Karina Mariela Figueroa Mora
@ Valeria Soto Mendoza

@ Universidad Autónoma de Coahuila
Boulevard Venustiano Carranza sin número
Colonia República Oriente, C.P. 25280
Saltillo, Coahuila, México

ISBN: pendiente

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

La totalidad de los capítulos que integran este libro fueron arbitrados y sometidos a evaluación externa, mediante un dictamente de doble ciego, donde participaron especialistas en el área. Las versiones finales se enriquecieron con las aportaciones de los pares académicos. Asímismo, se empleó un software especializado para análisis de similitudes con respecto a otras obras. La Dirección de Investigación y Posgrado de la UAdeC mantiene en resguardo los resultados de este proceso.

Esta obra fue financiada con recursos del Consejo Nacional de Ciencia y Tecnología.

Primera Edición: Agosto, 2020
Diseño de portada: Laura Gómez Cruz

Libro Electrónico, disponible en línea con acceso directo en: www.smcc.mx
Sociedad Mexicana de Ciencias de Computación

Con apoyo y colaboración de:

El Centro de Investigación en Matemáticas Aplicadas y la Facultad de Sistemas de la Universidad Autónoma de Coahuila (UAC)
Consejo Nacional de Ciencia y Tecnología
Tecnológico Nacional de México – Instituto Tecnológico de Culiacán (TecNM-IT Culiacán)
Universidad Michoacana de San Nicolás de Hidalgo (UMSNH)
Instituto Nacional de Electricidad y Energías Limpias (INEEL)
Asociación Mexicana de Interacción Humano-Computadora (AMEXIHC)

Presentación

A nivel mundial, con el uso de computadoras se han revolucionado las actividades del quehacer humano. Los avances tecnológicos y científicos producidos en el área de las ciencias de la computación han modificado radicalmente muchas actividades al grado de dejar obsoletos objetos/procesos que en su momento fueron indispensables.

Desde la llegada de la primera computadora a México, hace ya más de 60 años, en nuestro país se han producido innumerables cambios positivos en todas las áreas; además, la incorporación de internet impulsó un crecimiento acelerado en ramas como la salud, la comunicación, la seguridad, el entretenimiento, las actividades bancarias, entre muchas más no menos importantes, generando significativos desarrollos tecnológicos que están disponibles de forma natural para toda la sociedad.

Cada día se van descubriendo nuevas áreas de oportunidad en la que el trabajo de la comunidad científica del área de las ciencias de la computación se vuelve el núcleo de todos estos desarrollos tecnológicos con los que convivimos día a día.

Este libro presenta una pequeña muestra del trabajo de investigación más reciente que se realiza en diversas instituciones a nivel nacional y exhibe los avances alcanzados en distintas áreas donde, a través de las ciencias de la computación se logra mejorar algún proceso, solucionar algún problema, facilitar la toma de decisiones u optimizar alguna actividad.

Este libro de *Aplicaciones de la Computación* contiene 32 capítulos que corresponden al mismo número de trabajos de investigación desarrollados en diferentes instituciones y donde participan tanto estudiantes como investigadores. Los trabajos de las primeras cinco secciones se agruparon de acuerdo con el eje temático donde realizan su mayor aportación. Estas secciones son: Inteligencia artificial: métodos, modelos y aplicaciones; Aplicaciones de Ciberseguridad, Criptografía e Informática Médica; Aplicaciones de Cómputo Suave y Analítica de Datos; Sistemas Inteligentes en Internet de las Cosas; y, Aplicaciones en el Sector Eléctrico. La sexta sección del libro, llamada Trabajos en desarrollo, presenta siete trabajos que han alcanzado un avance significativo, pero se encuentran aún en etapa de desarrollo principalmente por estudiantes de posgrado.

La Sociedad Mexicana de Ciencias de la Computación se complace en presentar el libro *Aplicaciones de la Computación* que es producto de un arduo trabajo de revisión en un proceso doble ciego, donde participó un comité técnico integrado por 62 expertos de diversas instituciones quienes después de evaluar minuciosamente cada trabajo realizaron una dictaminación y selección para conformar este ejemplar.

Agradecemos infinitamente a todos los participantes su colaboración para hacer posible este trabajo.

Maria Lucia Barrón Estrada
José Antonio Camarena Ibarrola
Karina Mariela Figueroa Mora
Valeria Soto Mendoza

Saltillo, Coahuila, agosto 2020.

Instituciones

AMexIHC - Asociación Mexicana de Interacción Humano Computadora
BUAP - Benemérita Universidad Autónoma de Puebla
CETYS Universidad - Centro de Enseñanza Técnica y Superior
CICESE - Centro de Investigación Científica y de Educación Superior de Ensenada
CentroGeo - Centro de Investigación en Ciencias de Información Geoespacial
CIC- Centro de Investigación en Computación
CIIA - Centro de Investigación en Inteligencia Artificial
CIMAT - Centro de Investigación en Matemáticas
CIMA - Centro de Investigación en Matemáticas Aplicadas
CINVESTAV - Centro de Investigación y de Estudios Avanzados
CITEDI - Centro de Investigación y Desarrollo de Tecnología Digital
CENIDET - Centro Nacional de Investigación y Desarrollo Tecnológico
CONACYT - Consejo Nacional de Ciencia y Tecnología
IIASA - International Institute for Applied Systems Analysis.
INAOE - Instituto Nacional de Astrofísica, óptica y Electrónica
INEEL - Instituto Nacional de Electricidad y Energías Limpias

IPN - Instituto Politécnico Nacional
IT Cd. Guzmán - Instituto Tecnológico de Cd. Guzmán
IT Cd. Madero - Instituto Tecnológico de Cd. Madero
IT Culiacán - Instituto Tecnológico de Culiacán
IT Matamoros - Instituto Tecnológico de Matamoros
IT Saltillo - Instituto Tecnológico de Saltillo
ITSON - Instituto Tecnológico de Sonora
IT Zitácuaro - Instituto Tecnológico de Zitácuaro
ITESM - Instituto Tecnológico y de Estudios Superiores de Monterrey
IIASA - International Institute for Applied Systems Analysis
TecNM - Tecnológico Nacional de México
UIowa - The University of Iowa
UABC - Universidad Autónoma de Baja California
UACH - Universidad Autónoma de Chihuahua
UAdeC - Universidad Autónoma de Coahuila
UASLP - Universidad Autónoma de San Luis Potosí
UAEM - Universidad Autónoma del Estado de México
UCOL - Universidad de Colima
UdeG - Universidad de Guadalajara
UDLAP - Universidad de las Américas Puebla
UNISON - Universidad de Sonora
UMSNH - Universidad Michoacana de San Nicolás de Hidalgo
UNAM - Universidad Nacional Autónoma de México
UPP - Universidad Politécnica de Pachuca
UIowa - University of Iowa.
UV - Universidad Veracruzana

Comité técnico

Abigail María Elena Ramírez Mendoza, CINVESTAV-IPN
Abraham Efraim Rodríguez Mata, CONACYT/TecNM-IT Culiacán
Alejandra Guadalupe Silva-Trujillo, UASLP
Alicia Martínez Rebollar, TecNM-CENIDET
Ana I. Martínez-García, CICESE
Ángel G. Andrade, UABC
Araceli Celina Justo López, UABC
Brenda L. Flores-Ríos, UABC
Carlos A. Reyes-García, INAOE
Cuauhtémoc Rivera Loaiza, UMSNH
Daniel Fajardo-Delgado, TecNM-IT Cd. Guzmán
Efraín Ruiz, TecNM-IT Saltillo
Emmanuelle Ruelas Gómez, UABC
Félix F. González-Navarro, UABC
Fernando Fausto, UdeG
Fernando Orduña Cabrera, IIASA
Francisco Domínguez-Mota, UMSNH
Gabriel López-Morteo, UABC

Gerardo Salvador Romo Cárdenas, UABC
Gerardo Tinoco-Guerrero, UMSNH
Gilberto Ochoa-Ruiz, ITESM-Guadalajara
Gina Gallegos-García, CIC, IPN
Héctor Rodríguez Rangel, TecNM-IT Culiacán
Hugo Estrada Esquivel, TecNM-CENIDET
Javier Ortiz Hernández, TecNM-CENIDET
Joibel Tadea Giménez Mogollón, UABC
Jorge Eduardo Ibarra-Esquer, UABC
José ángel González Fraga, UABC
José Antonio Ortega, CITEDI
Juan José Flores, UMSNH
Juan Pablo García Vázquez, UABC
Juan Tapia, CITEDI
L. Leticia Ramírez Ramírez, CIMAT
Laura Cruz-Reyes, TecNM-IT Cd. Madero
Lil Rodríguez, INAOE
Lorena Castro-García, UIowa
Luis A. Castro, ITSON
Luis Morales, CONACYT-UMSNH
Manuel Mejía-Lavalle, TecNM-CENIDET
Marcela Quiroz Castellanos, CIIA-UV
Miguel ángel León Chávez, BUAP
Miguel González-Mendoza, ITESM
Nareli Cruz Cortés, CIC-IPN
Nataly Medina, CETYS Universidad
Noel Enrique Rodríguez Maya, TecNM-IT Zitácuaro
Norma Candolfi Arballo, UABC
Oscar Montiel, CITEDI
Pedro César Santana Mancilla, UCol
Rafael Morales Gamboa, UdeG
Ramón Imad Verdes, CITEDI
Raymundo Cornejo García, UACH
René Cruz Flores, UAEM
René Navarro, UNISON
Rocío Aldeco-Pérez, UNAM
Rodrigo López Farias, CONACYT -CentroGeo

Valeria Soto-Mendoza, CIMA-UAC
Víctor Estrada-Manzo, UPP
Víctor-Alejandro González-Huitrón, CONACYT/TecNM-IT Culiacán
Wendy Aracely Sánchez Gómez, TecNM-IT Matamoros
Wendy Flores-Fuentes, UABC
Yasmín Hernández, INEEL
Zobeida Jezabel Guzmán Zavaleta, UDLAP

Agradecimientos

Dra. Alicia Martínez Rebollar, TecNM-CENIDET

Dra. Ana Isabel Martínez García, CICESE

Dra. Anabel Martínez Vargas, UPP

Dr. Félix Fernando González Navarro, UABC

Dr. Héctor Efraín Ruiz y Ruiz, TecNM-IT Saltillo

Dr. Héctor Rodríguez Rangel, TecNM-IT Culiacán

Dr. René Francisco Navarro Hernández, UNISON

Dra. Rocío Aldeco Pérez, UNAM

Índice general

Agradecimientos **XII**

I Inteligencia Artificial: Métodos, Algoritmos y Aplicaciones **2**

1. ***Human Activity Recognition from Wrist Wearable Device Sensors using Data Mining Techniques*** **4**
Marcial Sandoval Gastelum, Felix F. González Navarro and Brenda L. Flores Rios

2. ***Software para Emisión de Alertas Basadas en Detección de Movimiento por Señales WiFi*** **13**
Emmanuel López Hernández, Felix F. González Navarro y Brenda L. Flores Rios

3. ***A meshfree formulation for solving PDEs in applied engineering problems*** **22**
Felix R. Saucedo Zendejo and Edgar O. Resendiz Flores

4. ***Aplicación de Algoritmos de Optimización en una Red Neuronal de Función de Base Radial*** **30**
Wendy Elizabeth Corpus Rojas, Rolando J. Praga Alejo David Salvador González González e Isidro Guzmán Flores

5. ***DinoBlock: Una Aplicación Móvil para el Pensamiento Computacional*** **39**
Ramón Zatarain Cabada, María Lucia Barrón Estrada y Ulises Uriel Cabrera Rubio

6. ***Una nueva herramienta basada en tecnología de videojuegos para enseñar teoría musical básica*** **48**
Baruc de Los Santos Izquierdo, Flor Radilla López, Eduardo Vazquez Santacruz y Elvia Elvira Morales Turrubiates

7. ***Educational Data Mining for finding patterns in First Semester Engineering Students Outcomes to identify potential Success or Failure*** **57**
Jorge Eduardo Ibarra Esquer, Felix F. González Navarro, Daniel Hernández Balbuena and Brenda L. Flores Rios

II Aplicaciones de Ciberseguridad, Criptografía e Informática Médica **66**

8. ***A Single-Chain System Implementation in C++*** **68**
Benjamin Oribe Mendieta and Luis Lizama

9. ***Studying the association of users' depressed moods and posting on Instagram*** **77**

Ivan Alejandro Encinas Monroy, Luis Héctor Sánchez Gámez, Jessica Beltran and Luis A. Castro

10. ***Buscando vulnerabilidades en aplicaciones de Android relacionadas al COVID 19*** **82**
Diana Serrano Delgado, Fernando Velázquez López, Rafael Llamas Contreras y Hugo Gonzalez

III Aplicaciones de Cómputo Suave y Analítica de Datos 91

11. ***Una Propuesta Metodológica para el Proceso de Toma de Decisiones aplicada en Centros de Soporte a la Toma de Decisiones*** **93**
Alejandro Flores Benitez, José Martín Molina Espinosa y Manuel Alvarez Madrigal

12. ***Spatial network analysis of Tijuana's metropolitan transport network*** **102**
Belgica Valencia Garcia and Adan Hirales Carbajal

13. ***Generación automática de controladores difusos*** **111**
Amaury Sanchez Negrin, Selene Lilette Cardenas Maciel, Jorge Antonio Lopez Renteria, Armando Martinez Graciliano y Nohe Ramon Cazarez Castro

14. ***Optimization of the GTAW manual welding process, considering random covariates*** **120**
Yaneth Guadalupe Robles de León, David Salvador González González and Argelia Faviola Miranda Pérez

15. ***Turisteando Ando: una App con Kernel Evolutivo*** **129**
Misael Lopez Sanchez, Anabel Martinez Vargas, Ma. Cosío León y Juan Pablo García Vázquez

16. ***PSO in Ridge Regression to eliminate multicollinearity in a PTA welding process*** **138**
Jorge Saúl Rodríguez Vázquez, Rolando Javier Praga Alejo and David Salvador González González

17. ***Redesign of a back-propagation neural network in a manufacturing process*** **147**
Rocio Natali Carranza Santoy, Rolando Javier Praga Alejo and David Salvador González González

IV Sistemas Inteligentes en Internet de las Cosas 156

18. ***Variante Algorítmica tendiente a dar Eficiencia a Búsquedas en el entorno de Internet*** **158**
Alejandro Aranda Díaz y Manuel Mejía Lavalle

19. ***Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío (MAIShOP)*** **167**
Miguel Angel García Morales, Héctor Joaquín Fraire Huacuja, Laura Cruz Reyes, Claudia Guadalupe Gómez Santillán, Nelson Rangel Valdez y María Lucila Morales Rodríguez

20. ***Optimización del Problema Multiobjetivo de la Selección de Proyectos con Parámetros Difusos*** **175**
Daniela López García, Héctor Fraire Huacuja, Laura Cruz Reyes, Nelson Rangel Valdez, Claudia Guadalupe Gómez Santillán, María Lucila Morales Rodríguez y Fausto Antonio Balderas Jaramillo

21. ***Optimizing the selection of mechanical exoskeletons in the industry: An approach from TOPSIS and MOORA*** **182**
Grisel Alvarez, Samuel Sandoval and Alberto Ochoa

V Aplicaciones en el Sector Eléctrico 191

22. ***Predicción del Recurso Solar mediante Diferencias*** **193**
Juan Jose Flores, Baldwin Cortes, Jose Cedeño, Antony Morales, Hector Rodriguez y Felix Calderon

23. ***Metodología para la estimación del diámetro en líneas de transmisión mediante reconstrucción 3D*** **202**
Victor Alejandro Gonzalez Huitron, Abraham Efraim Rodriguez Mata y Modesto Medina Melendrez

24. ***Arquitectura de un Lago de Datos para Sistemas de Medición Inteligente*** **209**
Juan Carlos Olivares Rojas, Luis Enrique Ruiz Martinez, Enrique Reyes Archundia, José Antonio Gutiérrez Gnecchi y Ismael Molina Moreno

25. ***Diseño e implementación de hardware propio de un interruptor de energía eléctrica con comunicación WiFi para aplicaciones en Industria 4.0*** **218**
Roberto Morales, Y. Guadalupe Lazaro, Adrian Reyes, J. Federico Casco y H. Patricia Martínez

VI Trabajos en Desarrollo 227

26. ***Generación de un grafo probabilístico en base a bitácoras de micro servicios RESTful*** **229**
Javier Iracheta y Sara Garza

27. ***Equipos de Laboratorio a Escala Piloto con Realidad Aumentada a partir de modelados 3D, como Apoyo en el Proceso de Aprendizaje en el Nivel Superior*** **237**
Elena Ortiz, Esbeidy Gomez, Julio César Sandria y Hugo Daniel González

28. ***Seguimiento de Objetos en la cadena de suministros de la industria de manufactura*** **246**
Ana Karen Hernandez, Alexis Jair López y Perfecto Malaquías Quintero Flores

29. ***Detección de la distancia de objetos mediante el uso de disparidad estéreo en un vehículo autónomo*** **255**
Josue Ezequiel Juarez Manjarrez y Victor Alejandro Gonzalez Huitron

30. ***Valoración de la salud mental de estudiantes universitarios a partir de datos de uso de teléfonos inteligentes*** **260**
Regina González Gómez, Valeria Soto Mendoza y Jesús Alejandro Navarro Acosta

31. ***Una propuesta de sello digital para los CFDI utilizando criptografía de curvas elípticas*** **269**
Diego Abraham Olvera Mendoza, Fidel González Gutiérrez y Arturo González Gutiérrez

32. ***Horizontal Fragmentation of Data Warehouses Using Decision Trees*** **276**
Nidia Rodríguez Mazahua, Lisbeth Rodríguez Mazahua, Asdrúbal López Chau and Giner Alor Hernández

Índice de Autores 281

Parte I

Inteligencia Artificial: Métodos, Algoritmos y Aplicaciones

Introducción

Actualmente la Inteligencia Artificial es el campo de las ciencias computacionales que mayor crecimiento e impacto ha tenido en los últimos años, en todos los campos de las ciencias y de la ingeniería.

El aprendizaje de máquina, el aprendizaje profundo, el lenguaje natural y la visión por computadora son las áreas que actualmente están siendo empujadas con mayor fuerza por grandes corporaciones y universidades, como las grandes apuestas del futuro tecnológico y desarrollo mismo de la humanidad. En este sentido, su avance en México ha sido sólido, con la participación de muchos investigadores y alumnos en gran cantidad de Instituciones de Educación Superior.

Así, en un esfuerzo por integrar artículos de alto impacto; la sección de Inteligencia Artificial: Métodos, Algoritmos y Aplicaciones presenta siete artículos relacionados con las áreas antes mencionadas que son: El reconocimiento inteligente aplicado a la detección de movimientos humanos usando tecnología WiFi y dispositivos inteligentes; en la Tecnología Educativa Inteligente donde se ofrecen aplicaciones móviles para la enseñanza y fomento de la creatividad; en la búsqueda de patrones o conocimiento en grandes bases de datos; en el desarrollo del importante campo de las redes neuronales con novedosos métodos aplicados a procesos industriales; y en el campo de las matemáticas aplicadas en las áreas de ingeniería.

Félix Fernando González Navarro

1

Human Activity Recognition from Wrist Wearable Device Sensors using Data Mining Techniques

MARCIAL SANDOVAL GASTELUM, FELIX F. GONZÁLEZ NAVARRO AND BRENDA L. FLORES RIOS

Human Activity Recognition from Wrist Wearable Device Sensors using Data Mining Techniques

Marcial Sandoval-Gastelum[1][0000−0001−6320−8368], Felix F. González-Navarro[1][0000−0002−9627−676X], and Brenda L. Flores-Rios[1][0000−0002−2502−6595]

Instituto de Ingeniería
Universidad Autónoma de Baja California
Mexicali, México[1]
{marcial.sandoval,fernando.gonzalez,brenda.flores}@uabc.edu.mx

Abstract. The quantity of in-home workers has increased in the recent years and remote professional roles have emerged. For this reason, gesture and daily human activity recognition technology development are important to this growing sector of the workforce in society. In this paper, we propose a detailed human activity recognition system based on in machine learning that works on data provided by an accelerometer and gyroscope. The purpose of this study is to find a model that classifies four human movements (walking upstairs, walking downstairs, standing up from a chair and sitting on a chair) that can be used in any wrist wearable device that has accelerometer and gyroscope. Four machine learning techniques provided the best accuracy results: Support Vector Machines, Bagged Trees Ensamble, K Nearest Neighbors and a six hidden layer neural network, reaching a maximum overall accuracy of 80.8%.

Keywords: Human Activity Recognition · Accelerometer · Wrist band · Neural Network · Support Vector Machine · Ensemble Learning

1 Introduction

The quantity of In-home workers has increase in the recent years and remote professional roles have emerged. According to [1], independent contractors, freelancers, on-call workers, temporary-help agency workers or people who are contracted out in their main job, represent about 16% of the US workforce and it was about 10% a decade ago. Among this population, the elder people that live alone are the most vulnerable, since any accidental fall can be mortal [2]. Therefore, daily human activity recognition technology development is getting a lot of attention in the research community, not only to provide tools that help this sector of society to perform better, more comfortably and efficient job, but also for security reasons, that is the reason why domestic healthcare of the elder people is one of the main applications of this studies.

According to [3], Human activity recognition or HAR, allows machines to analyze and comprehend several human activities from input raw data sources and it has been applied in surveillance systems, behavior analysis, gesture recognition, patient monitoring systems, ambient assisted living and a variety of healthcare systems that involve direct interaction or indirect interaction between human and smart devices. In terms of the type of sensor that is used for activity monitoring, activity recognition can be generally classified into two categories: vision-based and sensor-based activity recognition. The first is based on the use of visual sensing facilities, such as video cameras, to detect human behavior and environmental changes.The second category is based on the use of emerging sensor network technologies for activity monitoring [4], e.g. wearable inertial sensors.

Due to the rapid cost drop, its portable size and availability of inertial sensors, they have become one of the most utilized source of the raw data needed for HAR detection [5].

In this paper we propose a detailed human activity recognition system based in machine learning that works on data provided by an accelerometer and gyroscope, embedded wearable inertial sensors that can be found on many commercial smartwatch devices. Two of the most common smartwatches were considered as a benchmark to this research: the iWatch from Apple and the Fit Bit. According to their SDK, the sampling rate of the accelerometers of both devices are hardware dependant, constrained by the smarthphones bandwidth on which they are linked to, working at least at 100 hz. The Microsofts' Band 2 wrist band is the chosen wrist wearable devices due to the vast amount of sensors available and flexibility of its SDK. It has three different sampling rate that to work with: 8 hz, 32 hz and 62 hz, the last value was selected to be used in the experiments of the present research, which has no conflict with the other devices on the benchmark since they sample at a higher rate and a down sampling could be performed in order to use the final model on their system.

This work is organized as follows: Experiment settings, software and hardware used to collect and analyze the data are detailed in Section 3. Section 4 summarizes experimental results. Finally, conclusions and directions for future work are presented in Section 5.

2 Related Work

There is a vast variety of possible movements that have been studied previously using inertial sensors (e.g. accelerometers, gyroscopes, magnetometers) due to its portable size and low cost, from general activities such as going upstairs, going downstairs, standing, sitting, walking or running [6] [7] [8] to office related activities such as writing, phoning, drinking or working on the computer [9].

The quantity of sensors may vary, in [7], a single tri-axial accelerometer was placed on the testers wrist in contrast with [6], on which two tri-axial accelerometers were used, one on the waist and another on the wrist.

It is found that most experiments on the literature have been on a specific experiment environment and the sensors used are not found easily on commercial wearable devices. The purpose of this study is to find a model that classifies four different human movements: walking up stairs (UP), walking downstairs (DWN), standing up from a chair (STD) and sitting on a chair (SIT), that can be used in any wrist wearable device that have accelerometer and gyroscope, but also that can be used on the wearers free living environment.

3 Experimental Settings

Three male students of ages from 28 to 29, weight from 60 kgs to 80 kgs and height from 1.68 m. to 1.78 m. provided the experiments input data by performing each one of the four movements 30 times, creating a 360 human activity sample database called *movement* data set.

Figure 1 illustrates the architecture of the proposed experiments.

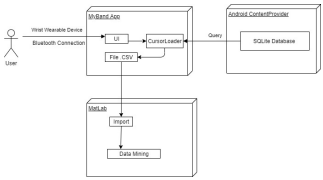


Fig. 1: Proposed architecture for human activity recognition

3.1 Hardware

Microsofts' Band 2 wrist band was the wearable device selected to provide the input data, due to its vast quantity of sensors: heart rate monitor, accelerometer, gyroscope, GPS, ambient light sensor, skin temperature sensor, UV exposure sensor, capacitive sensor, Galvanic Skin Response sensor and barometer [10].

Accelerometer and gyroscope sensors were chosen to provide the datasets' input raw data, since they work at the same sample rate (62Hz), which is high enough to achieve acceptable recognition rate and convenient for data mining processes, but also because they can be found on other commercial wrist wearables and it facilitates further recreation of the experiments.

3.2 Software

An Android app called *MyBand* for data collection was developed, using the Microsofts' Band SDK. A BandAccelerometerEventListener and a BandGyroscopeEventListener were implemented to collect each sensors raw data and store

it locally on a smartphone. Once a movement sample is completed, a csv file with the samples data is exported.

Subsequently, Mathworks numerical computing environment *Matlab* [11] was used as data analysis software for data mining.

3.3 Data Mining Techniques

The data mining process phases carried out in this work are presented below:

Preprocessing The selected technique for data imputation to handle null values on the input data set was K Nearest Neighbors Algorithm with k = 3. The data was normalized to have 0 mean and standard deviation of 1 with respect to each feature, meaning that each feature has been normalized independently.

Classification Algorithms The following classification techniques provided the best accuracy results:

Support Vector Machines Two parameters have been tuned to find the best SVM model: Box constraint 'C' and the kernel 'k' used to map the inercial sensors data to the feature space.

The tested values for 'C' are defined on 1:

$$C = \{1 <= c <= 16\} \tag{1}$$

The Kernel function 'k' used to compute the feature space values in this experiments are presented as follows:

Suppose G(xj,xk) is element (j,k) of the feature space matrix, where xj and xk are p-dimensional vectors representing observations j and k in movements data set X.

1. Gaussian or Radial Basis Function (RBF) kernel:

$$G(x_j, x_k) = e^{-||x_j - x_k||^2} \tag{2}$$

2. Linear kernel:

$$G(x_j, x_k) = x_j{}' x_k \tag{3}$$

3. 3rd grade Polynomial kernel:

$$G(x_j, x_k) = (1 + x_j{}' x_k)^3 \tag{4}$$

A grid search has been performed with these two parameters on each axis in order to find the model with best accuracy.

Tree Ensamble The data was trained on an ensemble of decision trees using bagging predictor method.

K Nearest Neighbors The k-nearest neighbor algorithm assumes that all the instances correspond to points in a space of n dimensions [14]. In this experiment, the nearest neighbors of an instance within the movements data set are retrieved using the Euclidean distance given on 5:

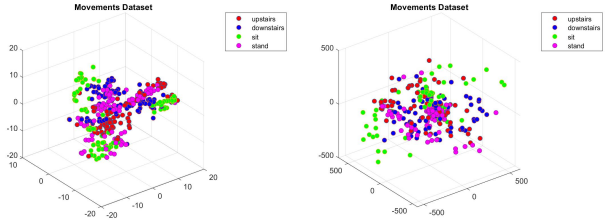
$$D(X, Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} \tag{5}$$

Neural Network Several hyperparameters have been tuned to reach better results, including network hidden layers, hidden neurons per layer, learning rate and epoch iterations.

Validation Methods All model metrics have been validated using the following technique.

10 Fold Cross Validation The dataset has been randomly split into ten non-overlapping groups and generate data "folds", having one designated for testing and the other for training [12]. Since the input data set included 360 samples, in this case, each fold contains 36 samples.

Visualization Methods In this study, two visualization methods have been used: PCA and t-SNE. The 60 feature sample movements input data set has been visualized in Figure 2 using both techniques.



(a) tSNE (b) PCA

Fig. 2: Movements Dataset

4 Results

On the SVM experiment, the highest accuracy of 78.61% was reached using a Gaussian kernel function and a box constraint level of 2 following 'One vs All' multi class method. An ensemble of 25 decision trees was trained using the 'Bagging' Method, the maximum number of splits within the ensemble was set to 330 and reached a 80.8% accuracy. Using the Euclidean distance metric and 1 nearest neighbor, KNN technique reached an overall 75% true positive accuracy.The final neural network architecture is shown in Figure 3. The values of the weights and biases in training were evaluated using a gradient descend with variable learning rate algorithm. A 80% of accuracy was obtained by setting the initial learning rate at 5e-06, that increased 0.05% per iteration, since the learning rate was very small, the training process time needed to be long and a maximum number of epochs was set to 64000, no iteration reached that number.

The true positives rates are presented in Figure 4.

Fig. 3: Neural Network Architecture

		Actual			
	Class	1	2	3	4
Predicted	1	72%	22%	1%	4%
	2	8%	89%	2%	1%
	3	1%	12%	84%	2%
	4	10%	7%	11%	72%

(a) SVM

		Actual			
	Class	1	2	3	4
Predicted	1	80%	11.11%	1.11%	7.77%
	2	7.77%	83.33%	5.55%	3.33%
	3	3.33%	11.11%	82.22%	3.33%
	4	5.55%	1.11%	7.77%	85.55%

(b) Bagged Tree Ensemble

		Actual			
	Class	1	2	3	4
Predicted	1	76%	9%	0%	16%
	2	17%	67%	4%	12%
	3	8%	11%	78%	3%
	4	6%	1%	13%	80%

(c) KNN

		Actual			
	Class	1	2	3	4
Predicted	1	88.37%	10.47%	1.16%	0%
	2	15.29%	76.47%	7.06%	1.18%
	3	0.92%	12.84%	69.92%	16.51%
	4	0%	2.5%	8.75%	88.75%

(d) Neural Network

Fig. 4: Output Confusion Matrices

5 Conclusions

Among the reasons that influenced strongly on the experiments results are the inconsistency of human movements in different individuals performing the same

activity, e.g. the rate of arm movements, one of the testers almost did not move the arm while performing UP and DWN movements, while another did a lot, in the case of the SIT and STD movements, the participants differ on the movement performing time. For that reason, it was found that the maximum and minimum absolute values within each testers data set was very different.

The low quantity of samples impacted dramatically on the Neural Network training performance in compared with the rest of the data mining techniques used. More sample collection from different participants can be seen as future work to get more improvements in the presented results.

Deep Learning techniques can also be used to reach better accuracy. DeepInsight methodology published by Sharma et al. in [13] is being tested currently in a new set of experiments, no better results have been found yet but it has shown very promising outputs, it also can be seen as a good opportunity for future work.

Acknowledgments

The authors would like to thank the Consejo Nacional de Ciencia y Tecnología (CONACYT) and Instituto de Ingeniería, Universidad Autónoma de Baja California, Mexicali, Baja California, México for their financial support given. The authors also are grateful to the participants in the data collection process.

References

1. Kwok, R. Flexible working: Science in the gig economy. Nature 550, 419–421 (2017)
2. González-Cañete, F. J., Casilari, E.: Consumption Analysis of Smartphone based Fall Detection Systems with Multiple External Wireless Sensors. Sensors (Basel, Switzerland), 20(3) (2020).
3. Dang,L. M. et al.: Sensor-based and vision-based human activity recognition: A comprehensive survey. Pattern Recognition. Volume 108 (2020)
4. L. Chen, J. Hoey, C. D. Nugent, D. J. Cook and Z. Yu: Sensor-Based Activity Recognition, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 6, 790-808, (2012)
5. Wang, Y., Cang, S., Yu, H.: A survey on wearable sensor modality centred human activity recognition in health care. Expert Systems with Applications, Volume 137. Pages 167-190. (2019)
6. Inês P. Machado, A. Luísa Gomes, Hugo Gamboa, Vítor Paixão, Rui M. Costa:Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization. Information Processing and Management Volume 51, Issue 2, Pages 204-214. (2015)
7. M. Panwar et al.: CNN based approach for activity recognition using a wrist-worn accelerometer, 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) , pp. 2438-2441 , (2017)
8. Zhu, R., Xiao, Z., Cheng, M., Zhou, L., Yan, B., Lin, S. Wen, HongKai: Deep Ensemble Learning for Human Activity Recognition Using Smartphone. 1-5. (2018)

9. Shuangquan Wang, Jie Yang, Ningjiang Chen, Xin Chen and Qinfeng Zhang: Human activity recognition with user-free accelerometers in the sensor networks. 2005 International Conference on Neural Networks and Brain, Beijing, pp. 1212-1217. (2005)
10. Microsoft Support Homepage, https://support.microsoft.com. Last accessed 9 March 2020
11. Mathworks Matlab Homepage, https://www.mathworks.com/products/matlab.html. Last accessed 9 March 2020
12. Seni, G., Elder, J.: Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions. Synthesis Lectures on Data Mining and Knowledge Discovery **2**, (2010)
13. Sharma, A., Vans, E., Shigemizu, D. et al. DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. Sci Rep 9, 11399 (2019)
14. Real-Moreno, O., Castro-Toscano, M. J., Rodrízuez-Ouiñonez, J. C., Hernández-Balbuena, D., Flores-Fuentes, W., Rivas-Lopez, M: Implementing k-nearest neighbor algorithm on scanning aperture for accuracy improvement. In IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society (pp. 3182-3186). IEEE. (2018)
15. Rodríguez-Quiñonez, J. C., Sergiyenko, O., Hernandez-Balbuena, D., Rivas-Lopez, M., Flores-Fuentes, W., Basaca-Preciado, L. C.: Improve 3D laser scanner measurements accuracy using a FFBP neural network with Widrow-Hoff weight/bias learning function. Opto-Electronics Review, 22(4), 224-235. (2014)
16. Flores-Fuentes, W., Rodríguez-Quiñonez, J. C., Hernandez-Balbuena, D., Rivas-López, M., Sergiyenko, O., Gonzalez-Navarro, F. F., Rivera-Castillo, J.: Machine vision supported by artificial intelligence. In 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE) (pp. 1949-1954). IEEE. (2014)
17. Flores-Fuentes, W., Rivas-Lopez, M., Sergiyenko, O., Gonzalez-Navarro, F. F., Rivera-Castillo, J., Hernandez-Balbuena, D., Rodríguez-Quiñonez, J. C.: Combined application of power spectrum centroid and support vector machines for measurement improvement in optical scanning systems. Signal Processing, 98, 37-51. (2014)

2

Software para Emisión de Alertas Basadas en Detección de Movimiento por Señales WiFi

EMMANUEL LÓPEZ HERNÁNDEZ, FELIX F. GONZÁLEZ NAVARRO Y BRENDA L. FLORES RIOS

Software para Emisión de Alertas Basadas en Detección de Movimiento por Señales WiFi*

Emmanuel López Hernández[0000−0002−9521−2967], Felix F. González-Navarro[0000−0002−9627−676X] y Brenda L. Flores-Rios[0000−0002−2502−6595]

Instituto de Ingeniería
Universidad Autónoma de Baja California
Mexicali, México
{elopez26,fernando.gonzalez,brenda.flores}@uabc.edu.mx

Resumen: Se presenta una investigación centrada en los problemas de seguridad en casas habitación ante riesgos delincuenciales, y el cuidado de la salud en personas con capacidades disminuidas. Se aborda el desarrollo de una aplicación capaz de generar mensajes de alerta, como respuesta al movimiento de una persona en una habitación, mediante el uso de señales inalámbricas, y algoritmos de Machine Learning.
Como resultado se obtuvieron siete Datasets, relacionados a siete pruebas realizadas. Cada uno incluye información de timestamp, Amplitud y Fase, de los subcarriers de la señal CSI recolectada por las tres antenas del dispositivo RX.

Palabras clave: Channel State Information, CSI, RSS, WiFi, Seguridad, Salud

1 Introducción

De acuerdo con [1], desde sus inicios en 1997, las tecnologías Wi-Fi (o WiFi) han pasado a través de muchas generaciones. Wi-Fi, acrónimo de wireless fidelity, es una marca de tecnología inalámbrica cuyo propietario es la Wi-Fi Alliance, y está enfocado en mejorar la interoperabilidad de los productos de redes inalámbricas de área local, basados en los estándares de la IEEE 802.11. Dicha tecnología es soportada por múltiples sistemas operativos, y resulta ideal para la conectividad del hogar u oficina, y es normalmente usada en aplicaciones como internet, telefonía IP, juegos, entre otras. Cada uno de sus estándares soporta diferentes tasas de transferencia de datos y frecuencias de banda.

- 802.11a. Hasta 54 Mbps, y frecuencia de 5 GHz.
- 802.11b. Hasta 11 Mbps, y frecuencia de 2.4 GHz.
- 802.11g. Hasta 54 Mbps, y frecuencia de 2.4 GHz.

* Se agradece al Conacyt y a la UABC por el apoyo financiero para realizar la investigación.

- 802.11n. Hasta 600 Mbps y frecuencias de 2.4 GHz y 5 GHz [2].
- 802.11ac. De 500 Mbps en una sola antena, hasta más de 1 Gbps mediante técnicas MIMO multi-usuario, y frecuencia de 5 GHz [3].
- 802.11ax. Hasta 4.8 Gbps y frecuencias de 2.4 GHz y 5 GHz [4].

Haciendo uso de la tecnología inalámbrica basada en el estándar 802.11n, se plantea desarrollar una aplicación para la detección de movimiento, mediante la recolección de señales CSI, con la finalidad de enviar mensajes que alerten a los usuarios, tanto para la seguridad de casas habitación, como del cuidado de la salud.

El resto del artículo se estructura de la siguiente manera: En la Sección 2 se menciona el trabajo relacionado donde se implementan soluciones basadas en WiFi y CSI; en la Sección 3 se especifican las herramientas de software, hardware y Machine Learning a utilizar; la Sección 4 trata acerca de los resultados de las pruebas realizadas; por último, en la Sección 5 se presentan las conclusiones sobre la presente investigación.

2 Trabajo Relacionado

Existen varias referencias en la literatura sobre el uso de CSI en la detección de patrones de información. [5] propone el uso de métodos de aprendizaje profundo basados en MLP y 1-D CNN, para predecir la ubicación de una persona en un entorno interior, mediante la recolección de señales RSS y CSI, entre un router con dos antenas emisoras, y una tarjeta NIC Intel 5300 con tres antenas receptoras. Se construyeron cuatro redes neuronales de aprendizaje profundo, demostrando una mejor precisión por parte de CSI.

En [6] se utiliza el método de fingerprinting basado en CSI para la identificación del posicionamiento en el interior. En sus pruebas se utiliza un router TP Link y una laptop Lenovo con una tarjeta NIC 5300 y un sistema operativo Ubuntu; se extraen la asimetría y curtosis de la señal CSI; y se emplea una red neuronal de aprendizaje profundo de tres capas para la fase de entrenamiento offline. Se concluye que este método supera a otros, como: Deepfi, ML y LAD; logrando un promedio de error inferior, pudiéndose adaptar a diversos entornos.

El método desarrollado en [7] se conoce como FDF-PIHD (Frequency Domain Fingerprint-based Passive Indoor Human Detection), y tiene como objetivo la detección de la actividad humana pasiva. Aplicable, al cuidado de adultos mayores y seguridad contra intrusos. Hace énfasis en 3 escenarios: sin personas, en reposo, y en movimiento. Con esta técnica, se logran extraer ciertos datos de las señales CSI que permiten generar “fingerprints”, que identifican de manera única cada Dataset. Para las pruebas se usó un Access Point, un router, y una computadora con una tarjeta de red Intel 5300 de 3 antenas. Se logró una tasa de detección superior al 90%, para cualquiera de los escenarios propuestos.

[8] presenta a ”SleepGuardian”, un sistema de Radio Frecuencias que utiliza dispositivos WiFi de bajo costo, y está enfocado en el cuidado de la salud de las personas mientras duermen; se basa en el procesamiento de señales, así como

técnicas de "Edge computing" y "Machine Learning". Entre los resultados de pruebas realizadas en entornos reales, se ha determinado que este sistema es confiable y eficiente, y que se puede integrar sin problema en instalaciones que ya cuenten con infraestructura de señales WiFi.

[9] ofrece un sistema de detección humana en situaciones de peligro, pero en el sanitario. Utiliza una cantidad mínima de datos de entrenamiento, por tratarse de datos enfocados a condiciones de peligro. En los resultados se demostró que el Sistema logró un alto desempeño al detectar la postura de las personas en situaciones de peligro y en un entorno "Sin Linea de Visión", en este caso el baño.

[10] muestra un sistema de detección de rutinas de ejercicios llamado WiFit, que usa señales WiFi para identificar el tipo y el número de repeticiones de cierto ejercicio, que una persona realiza. Se utilizaron un par de dispositivos WiFi, un transmisor y un receptor, ambos equipados con las herramienta "CSI Tool3". Éste aprovecha el "Efecto Doppler" que se presenta cuando el movimiento humano modifica la frecuencia Doppler debido al cambio de trayectoria. Se obtuvo una precisión del 95.8% en el reconocimiento de ejercicios y 99% al contar las repeticiones.

En [11] se propone el rastreo de signos vitales humanos, tanto frecuencia respiratoria como cardiaca mientras la persona duerme, mediante el uso de la red WiFi y el CSI. Se consideran variables en el dominio de tiempo y frecuencia, para lograr la estimación de la respiración y los latidos del corazón. Como resultado se demuestra la captura precisa de los signos vitales durante la etapa de sueño, así como un rendimiento comparable o mejor al de las técnicas tradicionales utilizadas, sin elementos invasivos en el cuerpo, y a bajo costo.

3 Configuraciones del Experimento

Se utiliza SCRUM como Metodología Ágil, con el propósito de planificar los objetivos a corto, mediano y largo plazo para el Desarrollo del Software de Emisión de Alertas, primero con el desarrollo de un prototipo funcional que servirá para la realización de las pruebas; y posteriormente un software más robusto, a miras de implementarse para la seguridad de casas habitación, monitoreo de adultos mayores, entre otras aplicaciones.

Las herramientas de software utilizadas, son:

- Sistema Operativo: Ubuntu 14.04 LTS Desktop con ambiente gráfico.
- Lenguaje de desarrollo: Python 3 y MATLAB R2019b.
- Librerías para la gestión de señales WiFi de los autores Daniel Halperin [12] y Hirokazu Narui [13].

De manera predeterminada, se trabaja en un entorno de Ubuntu en ambos dispositivos NUC (Tx y Rx), donde ya se encuentran instaladas las librerías de Halperin. Además, se utilizan como apoyo, los scripts del autor Hirokazu Narui, para la conversión de los Dataset generados, de .dat a .csv; así como también, de la gestión de su contenido.

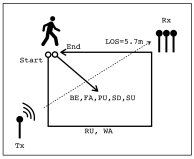


Fig. 1. Mapa de distribución del laboratorio.

Fig. 2. Izquierda: Colocación de Tx o Transmisor con 1 antena. Derecha: Colocación de Rx o Receptor con tres antenas.

En cuanto al espacio de las pruebas controladas, se ha desarrollado un mapa de distribución física (ver Fig. 1).

Los dispositivos hardware que se emplean son 2 Intel NUC Kit D54250WYKH [14]: un transmisor (Tx) con una antena y un receptor (Rx) con tres antenas; además de un monitor, un teclado y un mouse (ver Fig. 2).

Desde un punto de vista general, son cuatro las grandes etapas que deben seguirse para la puesta en marcha del modelo de clasificación, tal como se observa en la Fig. 3, esto para asegurar el análisis de datos de forma sistemática y repetible: Obtención de datos y generación del Dataset, Preprocesamiento, Extracción de Características, y Clasificación o Reconocimiento. Para el escenario de desarrollo experimental del software, se toma en cuenta un paso previo a la Clasificación, que es la Selección y Validación del Modelo de Minado de Datos.

De esta forma, se explorarán las siguientes técnicas dentro de estas etapas:

- Preprocesamiento: Normalización, que traduce los datos muy grandes a una menor escala con el fin de manejarlos de una manera más controlada; Imputación de datos, para dar una aproximación de los datos faltantes; Eliminación de ruido, para reducir o eliminar el ruido de las señales CSI.
- Extracción de Características: Se explorarán dos tipos de características, PCA (Principal Component Analysis), el recurrentemente usado en la literatura científica; y se extraerán variables en el dominio del tiempo y frecuencia de las señales.

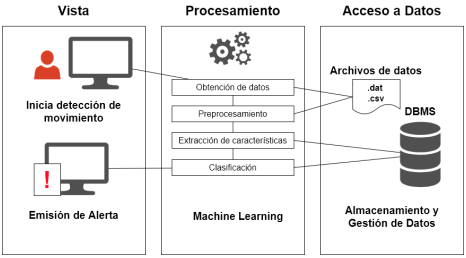


Fig. 3. Arquitectura del Software.

- Algoritmos de Clasificación: Se utilizarán algoritmos bayesianos como son Naive Bayes, LDA (Linear Discriminant Classifier) y QDA (Quadratic Discriminant Classifier); Vecinos Cercanos; Máquinas de Vectores de Soporte con varios tipos de kernel; Redes Neuronales Artificiales.
- Selección y Validación de Modelos: Para la selección de parámetros de los algoritmos de clasificación se utilizará 10-fold crossvalidation. Y para la validación final se creará un Dataset de prueba.

4 Resultados

Se definieron siete casos de pruebas controladas en el espacio especificado en un lapso de 20 segundos por actividad: Go to Bed, Fall, Pick Up, Run, Stand Up, Sit Down, Walk. Por cada actividad se genera un Dataset, tal como se muestra en la Fig. 4.

Con el uso de algunas librerías en Matlab, de Halperin y Narui, así como la adaptación de un script de Narui, cada Dataset se convirtió a una matriz que contiene 181 variables (representadas en columnas): La primera corresponde al timestamp; las siguientes noventa a la Amplitud = 30 subcarriers x 3 antenas de Rx (A, B y C); y las últimas 90 a la Fase = 30 subcarriers x 3 antenas de Rx (A, B y C).

Para el Preprocesamiento se implementó un script de autoría propia, donde se llevan a cabo Operaciones de Imputación, Normalización y Reducción del Ruido sobre los datos, dando como resultado un Dataset Preprocesado; y para la Extracción de Características, se adaptó un script de Narui para la generación de Gráficos de Actividad por cada Antena Rx, así como Gráficos de PCA, donde se extrajeron tres Principales Componentes.

En la Fig. 5 se muestra el resultado de la Actividad Run, donde se grafican los datos de la Matriz obtenida de CSI por las tres antenas Rx a lo largo del timestamp de 20 segundos; además, para el mismo tipo de actividad, en la Fig. 6,

Nombre	Fecha de modificación	Tipo	Tamaño
BE_200311_14_15.dat	11/03/2020 02:24 p. m.	Archivo DAT	4,196 KB
FA_200311_14_19.dat	11/03/2020 02:26 p. m.	Archivo DAT	4,196 KB
PI_200311_14_13.dat	11/03/2020 02:20 p. m.	Archivo DAT	4,198 KB
RU_200311_14_06.dat	11/03/2020 02:15 p. m.	Archivo DAT	4,198 KB
SD_200311_14_11.dat	11/03/2020 02:18 p. m.	Archivo DAT	4,197 KB
SU_200311_14_12.dat	11/03/2020 02:19 p. m.	Archivo DAT	4,193 KB
WA_200311_13_52.dat	11/03/2020 02:03 p. m.	Archivo DAT	4,197 KB

Fig. 4. Datasets generados por actividad.

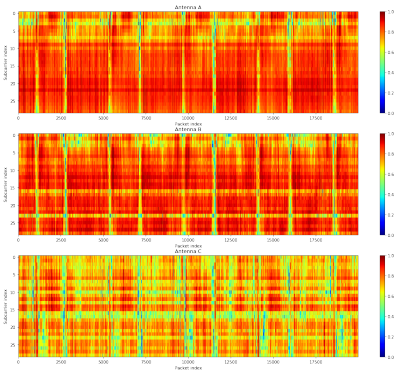


Fig. 5. Amplitud de la Actividad RUN por cada Antena de Rx

se muestran tres Principales Componentes de la señal de Amplitud que representan tres registros del resultado de implementar PCA, a lo largo del timestamp. Con esta información se observa que cuando el sujeto de pruebas pasa entre los dispositivos, la señal sufre una perturbación, en un intervalo aproximadamente de 2 a 2.5 segundos.

Con los resultados obtenidos en las pruebas, se continuará con la Construcción de un Modelo a entrenar, que al ser puesto en marcha en tiempo real, envíe alertas a los usuarios, sobre el movimiento registrado en el interior de una habitación.

5 Conclusiones

Se percibe el impacto actual que tienen las tecnologías de detección de movimiento, como en el caso de las señales CSI, que se pueden propagar con tan solo utilizar dispositivos de red comerciales asequibles para relativamente cualquier hogar que

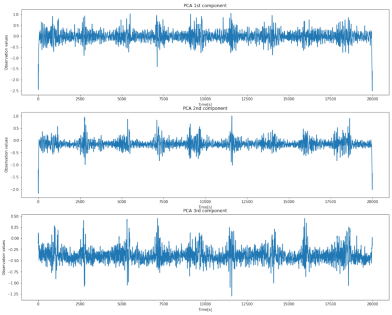


Fig. 6. Gráfico de PCA para 3 Componentes Principales de la Actividad RUN.

cuente con Internet. El empleo de estas tecnologías resulta multidisciplinario, porque abarca tanto Tecnologías de Red y Desarrollo de software, como Data Mining y Machine Learning.

En base a los resultados se puede concluir, que en la etapa en la que se encuentra el software, este permite Preprocesar correctamente los datos generados, así como la Extracción de Características y su Representación Gráfica, y de manera visual se percibe que ocurren perturbaciones en la señal, al momento que el sujeto de pruebas se cruza en el camino de los dos dispositivos NUC. Esta información resulta de gran utilidad porque visualmente se aprecian las diferencias entre los distintos tipos de Actividades realizadas.

Se planea que en un futuro, una vez terminado el software, se presente como una solución al problema de asalto a casas habitación, al ser capaz de detectar movimiento en el interior, y advertir a los propietarios, o encargados de la vivienda, edificio, oficina, etc., sobre la presencia de un intruso y poder tomar las acciones pertinentes. El alcance del envío de los mensajes no se limitará a una ventana de alerta en una computadora, como es el caso de lo que se trabaja en el experimento, sino que se considera puede abarcar desde un mensaje de correo electrónico, hasta una alerta personalizada mediante una app en el teléfono móvil.

En el área de la salud, para el caso de personas con capacidades disminuidas, el software permitirá alertar a los familiares que no estén presentes en la habitación, si la persona ha sufrido una caída, o está realizando alguna actividad que represente un riesgo para su salud y así poder intervenir oportunamente.

Sin embargo, los alcances del software no se limitan a los antes mencionados, pudiendo tener una mayor y amplia gama de posibilidades en el mercado.

Referencias

1. Hammoshi, M., Alani, R.: A Proposed Model to Implement Load and Throughput of W LAN implemented as Wi-Fi System. INC2010: 6th International Conference on Networked Computing, 1-7 (2010)
2. Deniau, V., Gransart, C., Romero, G., Simon, E., Farah J.: IEEE 802.11n Communications in the Presence of Frequency-Sweeping Interference Signals. IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY **Vol. 59, No. 5**, 1625-1633 (2017)
3. Jang, S., Bahk, S.: A Channel Allocation Algorithm for Reducing the Channel Sensing/Reserving Asymmetry in 802.11ac Networks. IEEE TRANSACTIONS ON MOBILE COMPUTING **Vol. 14, No. 3**, 458-472 (2015)
4. Cisco, https://www.cisco.com/c/en/us/products/collateral/wireless/white-paper-c11-740788.html. Último acceso 21 Jul 2020
5. Hsieh, C., Chen, J., Nien, B.: Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information. IEEE Access **Vol. 7**, 33256-33267 (2019)
6. Li, X., Shi, J., Zhao, J.: DeFe: indoor localization based on channel state information feature using deep learning. IOP Publishing, 1-9 (2019)
7. Han, C., Tan, Q., Sun, L., Zhu, H., Guo, J.: CSI Frequency Domain Fingerprint-Based Passive Indoor Human Detection. Information, 1-14 (2018)
8. Gu, Y., Wang, Y., Liu, Z., Liu, J., Li, J.: SleepGuardian: An RF-based Healthcare System Guarding Your Sleep from Afar. IEEE NETWORK MAGAZINE, 1-18 (2019)
9. Zhang, Z., Ishida, S., Tagashira, S., Fukuda, A.: Danger-Pose Detection System Using Commodity Wi-Fi for Bathroom Monitoring. Sensors, 1-16 (2019)
10. Li, S., Li, X., Lv, Q., Zhang, D.: Demo: WiFit - A Bodyweight Exercise Monitoring System with Commodity Wi-Fi. ACM UbuComp, 396-399 (2018)
11. Liu, J., Chen, Y., Wang, Y., Chen, X., Cheng, J., Yang, J.: Monitoring Vital Signs and Postures During Sleep Using WiFi Signals. IEEE INTERNET OF THINGS JOURNAL **Vol. 5, No. 3**, 2071-2084 (2018)
12. Halperin, D., https://github.com/dhalperi/linux-80211n-csitool. Último acceso 22 Mar 2020
13. Narui, H., https://github.com/ermongroup/Wifi_Activity_Recognition. Último acceso 22 Mar 2020
14. Intel, https://ark.intel.com/products/81164/Intel-NUC-Kit-D54250WYKH. Último acceso 22 Mar 2020

3

A meshfree formulation for solving PDEs in applied engineering problems

FELIX R. SAUCEDO ZENDEJO AND EDGAR O. RESENDIZ FLORES

A meshfree formulation for solving PDEs in applied engineering problems

Felix R. Saucedo-Zendejo[1] and Edgar O. Reséndiz-Flores[2]

[1] Centro de Investigación en Matemáticas Aplicadas, Universidad Autónoma de Coahuila, Camporredondo S/N, Saltillo, Coahuila, CP 25115, Saltillo, Coahuila, México
fesaucedoz@uadec.edu.mx
http://www.cima.uadec.mx/

[2] Tecnológico Nacional de México / Instituto Tecnológico de Saltillo, División de Estudios de Posgrado e Investigación, Blvd. V. Carranza 2400 Col. Tecnológico CP 25280, Saltillo, Coahuila, México
eresendiz@itsaltillo.edu.mx
http://www.its.mx/

Abstract. This paper reports the development and application of a Generalized Finite Differences Method (GFDM) formulation for solving partial differential equations in applied engineering problems. Its meshfree nature provides the advantage of capturing domain fragmentation or deformation without the need of adaptive remeshing algorithms which steams some of the drawbacks in traditional mesh-based approaches. The main features of this meshfree method and details of its implementation are presented and finally, numerical solutions of some test problems solved with this approach are reported which show that this formulation is feasible and suitable for solving the partial differential equations arising in several engineering problems.

Keywords: WKS7 · Finite Pointset Method · FPM · Solid Deformation · Free Surface Flows.

1 Introduction

Many industrial applications require the production of components with precise shapes, dimensions, and surface fineness, which could be very complex in general since it could involve complex free-surface flows, cooling, solidification and a stress and strain evolution during the production process. The end features in the manufactured components and the number of defects are the result of the coupling of the physical phenomena involved in the whole process. Therefore, in order to get homogeneous components with the desired characteristics and an acceptable low amount of defects, it is necessary to know and control all the physical phenomena involved and to conduct a proper design procedure [9].

Unfortunately, it is very difficult to improve and optimize the quality in such components with experimental techniques since it is impossible to exactly

measure velocities, temperatures, pressures and stresses without interfering with the process. Numerical simulation is commonly used to improve such processes since it provides a large amount of information that cannot be obtained through other methods [5].

Numerical mesh-based methods such as the Finite Element Method (FEM), Finite Difference Method (FDM), Finite Volume Method (FVM), and more recently, meshless methods as Smoothed-particle hydrodynamics (SPH) have already been used to analyze this kind of processes [8,6,1,3]. The advantages of meshless methods over mesh-based methods are that they use a set of nodes to discretize the problem domain and its boundaries without requiring any information about the relationship between nodes such that they do not form an element mesh which lets to model deformations and discontinuities in the domain without the drawbacks of mesh-based methods. Moreover, they have the flexibility to add or remove nodes wherever and whenever needed and it lets to easily develop adaptive schemes [2].

A truly meshfree GFDM is the so called Finite Pointset Method (FPM) developed by J. Kuhnert [7]. It has proven to be far superior to traditional mesh-based and some other meshfree methods to treat fluid dynamics problems with rapidly changing do-mains, free-surface or multiphase flows, and heat transfer problems [13,14,15,4,11]. This is a Lagrangian strong-form method which uses the weighted least-squares (WLSM) interpolation scheme to approximate the spatial derivatives [15]. It has many advantages over other methods since it is able to naturally and easily incorporate any kind of boundary conditions without requiring any special treatment or stabilization and it is really simple to implement. Therefore, as an attempt to develop a robust simulation tool, in this paper we propose the application of a novel FPM formulation [11] for solving partial differential equations in applied engineering problems. In order to get some insight on its performance we compute the numerical solution of a benchmark test. The structure of the paper is as follows: Section 2 shortly describes the governing equations and the numerical procedure used to solve them. Section 3 describes the basic ideas behind FPM and the special formulation for the mold filling problem followed by the numerical test presented in Section 4 with its corresponding results. Finally, some conclusions are given in the last section.

2 Mathematical model

Many physical phenomena such as heat transfer or linear elasticity are governed by an elliptical partial differential equation whose general form reads

$$A_1\psi + \mathbf{B}\cdot\nabla\psi + C_1\Delta\psi - F = 0 \tag{1}$$

with prescribed Dirichlet, Neumann ($\partial\psi/\partial\mathbf{n} = \varphi$) or Robin ($A_2\psi + C_2\partial\psi/\partial\mathbf{n} = \varphi$) boundary conditions, where $A_i, \mathbf{B}$, C_i, F and φ are known.

In order to solve this equation with their corresponding boundary and initial conditions in a natural and simple way, we follow the corresponding FPM

discretization [12] under this setting considering the general case with Robin boundary conditions for completeness.

3 The Finite Pointset Method (FPM)

In this section we describe the main ideas of the FPM method proposed by [7]. The FPM is a member of the family of the GFDM and it is based on the WLSM.

Let Ω be a given domain with boundary $\partial\Omega$ and suppose that the set of points $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$ are distributed with corresponding function values $\psi(\mathbf{r}_1)$, $\psi(\mathbf{r}_2)$, $\ldots$, $\psi(\mathbf{r}_N)$. The problem is to find an approximate value of ψ at some arbitrary location $\psi(\mathbf{r})$ using its discrete values at particles positions inside a neighbourhood of $\mathbf{r}$. To define the set of nodes and the neighbourhood of $\mathbf{r}$, a weight function $w(\mathbf{r}-\mathbf{r}_i)$ is introduced

$$w_i = w(\mathbf{r}-\mathbf{r}_i) = \begin{cases} e^{-\gamma\|\mathbf{r}-\mathbf{r}_i\|^2/h^2}, \text{ if } \frac{\|\mathbf{r}-\mathbf{r}_i\|}{h} \leq 1 \\ 0 \qquad\qquad else \end{cases} \tag{2}$$

where h is the smoothing length, γ is a positive constant whose value is considered to be 6.5, and $\mathbf{r}_i$ is the position of the $i-$th point inside the neighbourhood. A Taylor's series expansion of $\psi(\mathbf{r}_i)$ around $\mathbf{r}$ reads

$$\psi(\mathbf{r}_i) = \psi(\mathbf{r}) + \sum_{k=1}^{3} \psi_k (r_{k_i} - r_k) + \frac{1}{2}\sum_{k,l=1}^{3} \psi_{kl} (r_{k_i} - r_k)(r_{l_i} - r_l) + \epsilon_i, \tag{3}$$

where ϵ_i is the truncation error of the Taylor's series expansion, r_{ki} and r_k represent the $k-$th components of the position vectors $\mathbf{r}_i$ and $\mathbf{r}$, respectively. ψ_k and ψ_{kl} ($\psi_{kl} = \psi_{kl}$) represent the set of first and second spatial derivatives at node position $\mathbf{r}$. The values of ψ_k and ψ_{kl} can be computed minimizing the error ϵ_i for the n_p Taylor's series expansion of $\psi(\mathbf{r}_i)$ corresponding to the n_p nodes inside the neighbourhood of $\mathbf{r}$. This system of equations can be written in matrix form as $\mathbf{e} = M\mathbf{a} - \mathbf{b}$, where $\mathbf{e} = [e_1, e_2, e_3, \cdots, e_{n_p}]^t$, $\mathbf{a} = [\psi, \psi_1, \psi_2, \psi_3, \psi_{11}, \psi_{12}, \psi_{13}, \psi_{22}, \psi_{23}, \psi_{33}]^t$, $\mathbf{b} = [\psi(\mathbf{r}_1), \psi(\mathbf{r}_2), \cdots, \psi(\mathbf{r}_{n_p})]^t$, $M = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_{n_p}]t$, $\mathbf{s}_i = [1,\ \Delta r_{1_i},\ \Delta r_{2_i},\ \Delta r_{3_i},\ \Delta r_{11_i},\ \Delta r_{12_i},\ \Delta r_{13_i},\ \Delta r_{22_i},\ \Delta r_{23_i},\ \Delta r_{33_i}]^t$, $\Delta r_{k_i} = r_{k_i} - r_k$, $\Delta r_{kl_i} = (r_{k_i} - r_k)(r_{l_i} - r_l)$ and $\Delta r_{kk_i} = 0.5(r_{k_i} - r_k)(r_{k_i} - r_k)$, for $k, l = 1, 2, 3$ and $k \neq l$. The unknown vector $\mathbf{a}$ is obtained through WLSM by minimizing the quadratic form

$$J = \sum_{i=1}^{n_p} w_i \epsilon_i^2, \tag{4}$$

which reads $(M^t W M)\mathbf{a} = (M^t W)\mathbf{b}$, where $W = \text{diag}(w_1, w_2, \cdots, w_{n_p})$. Therefore,

$$\mathbf{a} = (M^t W M)^{-1}(M^t W)\mathbf{b}. \tag{5}$$

In this way we automatically get the values of ψ and its derivatives at points $\mathbf{r}$.

3.1 FPM formulation for general elliptic partial differential equations

Consider now the numerical solution of the elliptic partial differential equation (1). In the FPM representation of the above problem, equation (1) must be taken together with the system of n_p Taylor's series expansion of $\psi(\mathbf{r}_i)$ around $\mathbf{r}$. In this case, the matrices we need to compute by each particle in Ω take the following form: $\mathbf{b} = [\psi(\mathbf{r}_1), \psi(\mathbf{r}_2), \cdots, \psi(\mathbf{r}_{n_p}), F]^t$, $M = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_{n_p}, \mathbf{s}_E]t$ and $W = \text{diag}(w_1, w_2, \cdots, w_{n_p}, 1)$, where $\mathbf{s}_E = [A_1, B_1, B_2, B_3, C_1, 0, 0, C_1, 0, C_1]^t$ and $\mathbf{B} = [B1, B2, B3]t$. If $\mathbf{r}_i \in \partial\Omega$, additionally we have to add the general boundary condition.

If we consider a general Robin boundary condition for example, in this case the matrices we need to compute by each particle in $\partial\Omega$ take the following form: $\mathbf{b} = [\psi(\mathbf{r}_1), \psi(\mathbf{r}_2), \cdots, \psi(\mathbf{r}_{n_p}), F, \varphi]^t$, $M = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_{n_p}, \mathbf{s}_E, \mathbf{s}_B]t$ and $W = \text{diag}(w_1, w_2, \cdots, w_{n_p}, 1, 1)$, where $\mathbf{s}_B = [A_2, C_2 n_1, C_2 n_2, C_2 n_3, 0, 0, 0, 0, 0, 0]^t$.

If we define $\mathbf{q} = [q_1, q_2, \cdots, q_{10}]^t$ as the first row of $(M^t W M)^{-1}$ and the terms in the moving least squares solution $\mathbf{a} = (M^t W M)^{-1}(M^t W)\mathbf{b}$ are worked out, we can see that the following linear system of equations arises

$$\begin{aligned}\psi(\mathbf{r}_j) - \sum_{i=1}^{n(j)} w_i \left(q_1 + q_2 \Delta r_{1_i} + q_3 \Delta r_{2_i} + q_4 \Delta r_{3_i} + q_5 \Delta r_{11_i} + q_6 \Delta r_{12_i}\right. \\ + q_7 \Delta r_{13_i} + q_8 \Delta r_{22_i} + q_9 \Delta r_{23_i} + q_{10} \Delta r_{33_i}) \psi(\mathbf{r}_i) = [A q_1 + B_1 q_2 \\ + B_2 q_3 + B_3 q_4 + (q_5 + q_8 + q_{10}) C] D + (n_1 q_2 + n_2 q_3 + n_3 q_4) E,\end{aligned} \tag{6}$$

where $\psi(\mathbf{r}_j)$ denotes the unknown function value at particle j and $n(j)$ the number of $j-$th particle neighbours. Since equation (6) is valid for $j = 1, 2, \cdots, N$, this can be arranged in a full sparse system of linear equations $L\mathbf{F} = \mathbf{P}$ which can be solved by iterative methods. In the same way, coupled vector boundary value problems are treated similarly. For further information on this kind of problems, we refer to [11]. Therefore, all kind of problems governed by elliptic partial differential equations can be solved with this formulation, just adding appropriate entries in the corresponding systems of equations [12,11].

Fig. 1: Geometrical description of the two-dimensional cantilever beam problem.

4 Numerical Example

In this section the suitability and feasibility of this FPM formulation in order to solve an elliptic partial differential equation arising in an applied problem will be evaluated. It corresponds to the standard benchmark problem, described in Figure 1, where a cantilever beam of height D and length L is fixed on the left hand side and a tangential load is imposed on the right hand side. This problem is governed by the Navier - Cauchy equations and the specific values are set as: $L = 12\,m$, $D = 2\,m$, $E = 1\,kPa$, $\nu = 0.33$ and $P = 1\,N/m$. The exact solution for the stresses and the corresponding displacements was derived in [16] and reads

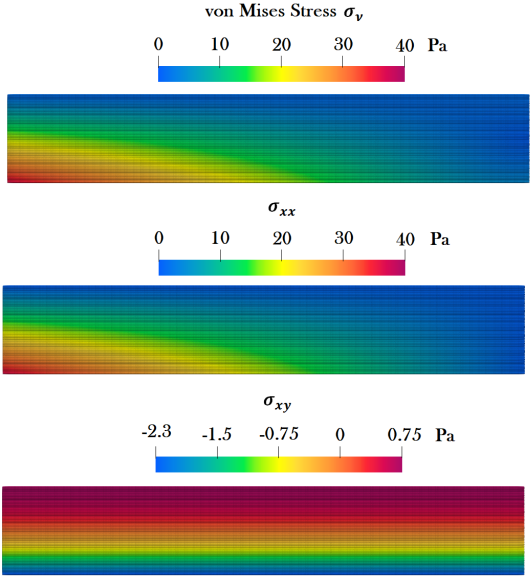


Fig. 2: Stresses distribution.

$$\sigma_{xx} = \frac{-P}{I}(L-x)\left(y-\frac{1}{2}\right),\ \sigma_{yy}=0,\ \sigma_{xy}=\frac{-Py}{2I}(y-D) \tag{7}$$

and

$$u = -\frac{P\left(y-\frac{D}{2}\right)\left[(6L-3x)\,x+(2+\nu)+\left(y^2-Dy\right)\right]}{6EI} \tag{8}$$

$$v = -\frac{P\left(y-\frac{D}{2}\right)^2(L-x)+\frac{1}{4}(4+5\nu)\,D^2x+(3L-x)\,x^2}{6EI} \tag{9}$$

$$\tag{10}$$

where$E$ is the Young's modulus, P is the total load force, $I=\frac{1}{12}D^3$ is the moment of inertia around the horizontal axis and ν is the Poisson's ratio. For the numerical solution this displacements are applied at all boundary points as boundary conditions.

In this example, the problem domain was discretized with approximately 20211 points. The corresponding numerical results for the stresses σ_{xx}, σ_{xy} and σ_ν are shown in Figure 2, where the the von Misses stress is defined as

$$\sigma_\nu = \sqrt{\sigma_{xx}^2-\sigma_{xx}\sigma_{yy}+\sigma_{yy}^2+3\sigma_{xy}^2} \tag{11}$$

This figure shows smooth and stable stress patterns. The σ_{xx} and σ_ν are compatible with the physical configuration. These pictures indicate the robustness of this FPM formulation for the solution of these partial differential equations since its numerical prediction matches very well with the exact solution, with a more numerically simple management of the boundary conditions. For further numerical examples we refer to [10,11].

5 Conclusions

Based on the numerical performance shown in the numerical example we can conclude that the current approach is suitable and feasible for solving partial differential equations in applied engineering problems. It is stable and has enough accuracy in order to capture the physical behavior observed in the governed processes. Since this formulation is a truly meshfree method it could be used for the study and analysis of complex problems involving high deformations and domain fragmentations with a great computational advantage since it does not need to compute any numerical quadrature and it does not need remeshing approaches. Further, it is able to naturally and easily handle any kind of boundary conditions without requiring any special treatment or stabilization and it is really simple to implement. Therefore, it could be a promising numerical tool for the simulation of industrial processes involving complex flows and other phenomena described by elliptic partial differential equations as heat transfer and elastic deformation. Consequently, it depicts a rich source of research opportunities.

References

1. Bašić, H., Demirdžić, I., Muzaferija, S.: Finite volume method for simulation of extrusion processes. Int. J. for Numer. Meth. Engng. **62**(4), 475–494 (2005)
2. Belytschko, T., Krongauz, Y., Organ, D., Fleming, M., Krysl, P.: Meshless methods: An overview and recent developments. Comput. Methods Appl. Mech. Engrg. **139**(1-4), 3–47 (1996)
3. Cleary, P.W., Savage, G., Ha, J., Prakash, M.: Flow analysis and validation of numerical modelling for a thin walled high pressure die casting using SPH. Comput. Part. Mech. **1**(3), 229–243 (2014)
4. Jefferies, A., Kuhnert, J., Aschenbrenner, L., Giffhorn, U.: Finite Pointset Method for the Simulation of a Vehicle Travelling Through a Body of Water. In: Griebel, M., Schweitzer, M.A. (eds.) Meshfree Methods for Partial Differential Equations VII, Lecture Notes in Computational Science and Engineering, vol. 100, pp. 205–221. Springer Berlin Heidelberg (2015)
5. Kermanpur, A., Mahmoudi, S., Hajipour, A.: Numerical simulation of metal flow and solidification in the multi-cavity casting moulds of automotive components. J. Mater. Process. Technol. **206**(1-3), 62–68 (2008)
6. Kimatsuka, A., Ohnaka, I., Zhu, J.D., Ohmichi, T.: Mold filling simulation with consideration of gas escape through sand mold. Int. J. Cast Metal Res. **15**(3), 149–152 (2003)
7. Kuhnert, J.: General Smoothed Particle Hydrodynamics. Ph.D. thesis, Technische Universität Kaiserslautern (1999)
8. Lewis, R.W., Ravindran, K.: Finite element simulation of metal casting. Int. J. Numer. Meth. Engng. **47**(1-3), 29–59 (2000)
9. Rao, T.V.R.: Metal casting: Principles and practice. New Age International, 1 edn. (2007)
10. Reséndiz-Flores, E.O., Saucedo-Zendejo, F.R.: Numerical simulation of coupled fluid flow and heat transfer with phase change using the finite pointset method. International Journal of Thermal Sciences **133**, 13–21 (2018)
11. Saucedo-Zendejo, F.R., Reséndiz-Flores, E.O.: A new approach for the numerical simulation of free surface incompressible flows using a meshfree method. Comput. Method. Appl. M. **324**, 619–639 (2017)
12. Tiwari, S., Kuhnert, J.: Grid free method for solving the poisson equation. Berichte des Fraunhofer ITWM **25** (2001)
13. Tiwari, S., Kuhnert, J.: A meshfree method for incompressible fluid flows with incorporated surface tension. Revue Europeenne des Elements **11**(7-8), 965–987 (2002)
14. Tiwari, S., Kuhnert, J.: Particle Method for Simulation of Free Surface Flows. In: Hou, Y., Tadmor, E. (eds.) Hyperbolic Problems: Theory, Numerics, Applications, pp. 889–898. Springer Berlin Heidelberg (2003)
15. Tiwari, S., Kuhnert, J.: Modeling of two-phase flows with surface tension by finite pointset method (FPM). J. Comput. Appl. Math. **203**(2), 376–386 (2007)
16. Zhang, X., Song, K.Z., Lu, M.W., Liu, X.: Meshless methods based on collocation with radial basis functions. Computational Mechanics **26**, 333–343 (2000)

4

Aplicación de Algoritmos de Optimización en una Red Neuronal de Función de Base Radial

WENDY ELIZABETH CORPUS ROJAS, ROLANDO J. PRAGA ALEJO DAVID SALVADOR GONZÁLEZ GONZÁLEZ E ISIDRO GUZMÁN FLORES

Aplicación de Algoritmos de Optimización en una Red Neuronal de Función de Base Radial.

Wendy E. Corpus Rojas[1] Rolando J. Praga Alejo[1-2] David S. González González[1-2] Isidro Guzmán Flores [1]

[1] facultad de Sistemas, Universidad Autónoma de Coahuila, Ciudad universitaria Carretera a México km. 13, Arteaga, Coahuila, México.
[2] corporación Mexicana de Investigación en Materiales S.A. de C. V. (COMIMSA) Calle Ciencia y Tecnología 790, Frac. Saltillo 400, Saltillo, Coahuila, México.

Abstract. Los procesos de soldadura son de los más empleados en la industria manufacturera. Aunado esto a los recientes avances tecnológicos de fabricación se tiene la necesidad de mejorar los procesos de manufactura y así, ofrecer productos de alta calidad. Por lo tanto, es útil la aplicación de modelos que permitan la planificación, diseño y simulación de los procesos, y así, eliminar el uso de la técnica de "prueba y error". Una herramienta adecuada para poder cumplir con dichas tareas son las Redes Neuronales de Función de Base Radial. Para que la red describa el comportamiento del proceso adecuadamente, es necesario tener los mejores centros de la Función de Base Radial. Al tener los centros óptimos, la Red tendrá un mejor ajuste y una manera de optimizar tales centros es mediante Algoritmos evolutivos de optimización implementados en la Red Neuronal. Se propone un modelo que prediga el comportamiento de un proceso de soldadura PTA implementando un Algoritmo Genético, Algoritmo de Recocido Simulado y un Algoritmo PSO, siendo este último en conjunto con la Función de Base Radial Gaussiana el que permitió obtener el mejor ajuste de la Red, según los criterios AIC, BIC, CAIC y HQIC.

Keywords: Función de base radial, Algoritmos Evolutivos, Criterios de Información.

1 Introducción.

Las Redes Neuronales de Función de Base Radial (RNFBR) son herramientas que permiten a la industria cumplir con los objetivos de mantener un control de sus procesos y asegurar la calidad de sus productos. Modelan y predicen el comportamiento de un proceso, eliminando el uso de la técnica de prueba y error. Este articulo muestra cómo se logró que una RNFBR represente adecuadamente la realidad de un proceso de soldadura PTA (Plasma transferido por Arco) empleando ciertos algoritmos evolutivos de optimización para obtener ciertos parámetros de red óptimos (centros de la función de base radial)

Se propone analizar e implantar dichos algoritmos, para encontrar una matriz de centros óptima mediante los algoritmos: Algoritmo Genético (AG) [1], Algoritmo de Recocido Simulado (ARS) [2] y Algoritmo PSO [3].

1.1 Estado del Arte.

El proceso de soldadura de plasma transferido por arco (PTA), cuenta con características únicas, dichas características la hacen una técnica de soldadura. El PTA es un proceso de arco eléctrico que utiliza un electrodo no consumible de tungsteno o de aleación de tungsteno con protección de gas inerte y alta densidad enérgica [4].

Las RNFBR son usadas para el modelado y predicción de procesos de soldadura en [5]. Generalmente, las características físicas más importantes de la soldadura son la geometría del cordón, así como la altura, ancho y penetración. [6] desarrolla una RNFBR para lograr obtener los valores aceptables de los parámetros de geometría del cordón. Algunos autores han utilizado algoritmos evolutivos para mejorar el comportamiento de la Red y así obtener un buen ajuste al momento de predecir el comportamiento del proceso de fabricación, algunos ejemplos se encuentran en [7] y [8] Los cuales usan un Algoritmo PSO y un Algoritmo de Recocido Simulado en conjunto con un Algoritmo Genético respectivamente. En este documento, se plantean los Criterios de Información de Akaike (AIC). Es un predictor que se basa en la teoría de la información de Kullback-Leiber. Las aplicaciones de este criterio se pueden encontrar en [9] quien utiliza AIC para seleccionar el número de variables óptimas para realizar un modelo de cointegración, y en [10] quienes analizan la predicción de series de tiempo utilizando Redes Neuronales para encontrar la arquitectura óptima de la red.

2 Métodos

2.1 Red Neuronal de Función de Base Radial.

Una red neuronal es un método que está diseñado para modelar la forma en que el cerebro realiza alguna una tarea. Según [11] la red involucra tres capas diferentes. La relación multivariada de entrada-salida de la red viene dada por:

$$y_{ij}(x) = \sum_{j=1}^{K} \sum_{i=1}^{N} w_{mj} G(\|X_i - C_m\|) + b + e_{ij} \quad (1)$$

Donde y_{ij} es la *ij-esimo* elemento de la matriz de salida, Y_{NK}, *N* es el número de observaciones (datos), *K* es el número de salidas en la Red Neuronal de Función de Base Radial (respuestas). Se denota *M* como el número de neuronas ocultas (o los centros de la Función de Base Radial), w_{mj} es el *mj-esimo* elemento de la matriz de pesos $W_{M+1,K}$, m=1,2,…,M, M+1 es el número de neuronas ocultas, se agrega el bias (b) , X_{NM} son los patrones de entrada, C_m es la matriz cuadrada de centroides $M \times M$, E_{NK} es la matriz de residuales de e_{ij}.

Un estimador imparcial de los pesos *W*, *es* dado por mínimos cuadrados ordinarios multivariados:

$$W = (G'G)^{-1}G' * Y \quad (2)$$

G es dado por la distancia $X_i - C_m$ aplicando la función de base radial. Hay varias funciones de base radial. La más comúnmente usada es la función Gaussiana (3), la

multicuadratica inversa (4), la función multicuadratica (5), la función spline (6) y la función cubica (7).

$$G(X_N - C_M) = exp(-\|X_N - C_M\|)^2 \tag{3}$$

$$G(X_N - C_M) = \frac{1}{\sqrt{1+ (\|X_N - C_M\|)^2}} \tag{4}$$

$$G(X_N - C_M) = \sqrt{1 + (\|X_N - C_M\|)^2} \tag{5}$$

$$G(X_N - C_M) = (\|X_N - C_M\|)^2 ln(\|X_N - C_M\|) \tag{6}$$

$$G(X_N - C_M) = (\|X_N - C_M\|)^3 \tag{7}$$

G se define por la forma:

$$\mathrm{G} = \begin{bmatrix} \mathrm{G}(X_1 - C_1) & \mathrm{G}(X_1 - C_2) & \cdots & \mathrm{G}(X_1 - C_\mathrm{m}) \\ \mathrm{G}(X_2 - C_1) & \mathrm{G}(X_2 - C_2) & \cdots & \mathrm{G}(X_2 - C_\mathrm{m}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{G}(X_\mathrm{N} - C_1) & \mathrm{G}(X_\mathrm{N} - C_2) & \cdots & \mathrm{G}(X_\mathrm{N} - C_\mathrm{m}) \end{bmatrix} \tag{8}$$

La distancia $X_1 - C_m$ puede ser reemplazada por la distancia de Mahalanobis *r* dada por la ecuación (9):

$$r = \sqrt{(X - C_m)' \sum^{-1} (X_i - C_m)} \tag{9}$$

2.2 Algoritmo Genético.

La salida de la función de base radial depende de la distancia entre los patrones de entrada y los centros de la red. Existen diferentes métodos de determinar dichos centros: ya sea por agrupamiento o por métodos de optimización. Uno de ellos es el algoritmo genético, el cual es un método de optimización basado en el proceso biológico de la evolución [1].

El proceso consiste en una selección al azar de individuos de una población actual. Dichos individuos serán los padres de la siguiente generación, los cuales evolucionarán hacia una solución óptima.

2.3 Algoritmo PSO.

El algoritmo de optimización por enjambre de partículas (particle swarm optimization, PSO) es un método de búsqueda iterativo y colectivo, con énfasis en la cooperación, inspirado en el movimiento coordinado de los animales que viven en grupos. El objetivo común de todos los miembros de la población consiste en encontrar la ubicación más favorable dentro de un determinado espacio de búsqueda.

La población o grupo también se denomina enjambre (*swarm*), y la trayectoria de cada partícula en el espacio de búsqueda es ajustada dinámicamente mediante la alteración de su velocidad. Estas partículas vuelan a través del espacio de búsqueda, teniendo dos capacidades de razonamiento esenciales: La memoria de la mejor posición individual alcanzada y el conocimiento de la mejor posición global [12].

2.4 Algoritmo Recocido Simulado.

La simplicidad y buenos resultados de este algoritmo en diversos problemas, la ha convertido en una herramienta muy popular, con cientos de aplicaciones en los más variados campos [13]. El recocido simulado es una variante de la búsqueda local que permite movimientos ascendentes para evitar quedar atrapado prematuramente en un óptimo local. Su nombre viene de la idea en que está basado, es un algoritmo diseñado en los años 50 para simular el enfriamiento de un material (un proceso denominado "recocido"). El fundamento de este algoritmo se basa en el campo de la termodinámica estadística.

2.5 Selección de la Función de Base Radial.

Esta es una fase que se basa en evaluar ciertos criterios; en este caso, los criterios de Información de Akaike, esta métrica denominada criterio teórico de información mínima [14]. Los algoritmos consideran una función de evaluación. En este caso, las extensiones multivariadas de los Criterios de Información de Akaike se presentan en la ecuación (10):

$$\mathrm{AIC_q} = \mathrm{nlog}\left|\widehat{\sum_q}\right| + 2\mathrm{qp} + \mathrm{p(p+1)} \tag{10}$$

Donde q y p son el número de parámetros de entradas y salidas en la matriz $\widehat{\sum_q}$.
Dichos criterios son empleados como una función de aptitud dentro de cada uno de los algoritmos mencionados anteriormente. También es posible utilizar una extensión AIC multivariada para procesos con más de una respuesta.
La ecuación de extensiones multivariadas es la corrección $\mathrm{AIC_q}$ ($\mathrm{CAIC_q}$) la cual se describe en (11):

$$\mathrm{AIC_{cq}} = \mathrm{(n-q-p-1)log}\left|\widehat{\sum_q}\right| + \mathrm{(n+q)p} \tag{11}$$

El criterio de información bayesiano BIC (Bayesian Information Crierion) está basado en la función de probabilidad y se relaciona con el AIC en los modelos de ajuste. El BIC multivariante se define por la ecuación (12):

$$\mathrm{BIC_q} = \mathrm{nlog}\left|\widehat{\sum_q}\right| + \mathrm{qplog(n)} \tag{12}$$

El Criterio de Información de Hannan y Quinn HQIC (Hannan and Quinn Information Criterion) es un criterio para la selección de modelos. Es una alternativa al AIC y al BIC. El HQIC, como el BIC, pero a diferencia del AIC no es un estimador de una divergencia Kullback-Leiber. La ecuación 13 muestra el HQIC multivariante:

$$HQIC_q = nlog\left|\widehat{\sum_q}\right| + 2qploglog(n) \quad (13)$$

También se puede emplear el criterio extendido R_q^2, que es una expresión de matriz análoga del coeficiente de determinación R^2. R_q^2 se describe en la ecuación (14):

$$R_q^2 = (Y'Y - n\acute{y}\acute{y}')^{-1}(\widehat{B_q}'G_q'Y - n\acute{y}\acute{y}') \quad (14)$$

Para convertir (14) a su forma escalar es posible usar $tr(R_q^2)/p$, donde p es el número de y's, entonces $0 \leq tr(R_q^2)/p \leq 1$.

3 Resultados experimentales.

Tabla 1. Datos del Proceso. Fuente: Elaboración propia.

	Variables de entrada			Variables de Salida		
Run	Powder Feed rate (g/min)	Process Speed (cm/min)	Welding Current (Amp)	Area HAZ (mm^2)	Length HAZ (mm)	Weld Bead Width (mm)
1	26	80	60	4.908	5.74	3.667
2	21	100	60	5.184	5.336	2.798
3	18	80	80	4.859	6.122	3.707
4	21	100	80	3.693	4.623	2.847
5	21	100	70	3.623	4.343	2.786
6	26	100	70	3.504	4.433	2.995
7	26	120	80	3.711	4.511	2.268
8	21	100	70	3.737	4.593	3.083
9	26	120	60	3.371	4.21	2.611
10	18	100	70	3.345	4.23	2.922
11	21	100	70	3.541	4.291	2.813
12	21	100	70	4.092	4.872	2.889
13	18	80	60	4.674	5.732	3.855
14	21	100	70	4.023	4.325	2.964
15	21	100	70	3.47	4.062	2.724
16	21	100	70	3.478	4.23	2.726
17	18	120	80	3.202	3.915	2.326
18	18	120	60	3.081	3.584	2.3
19	21	100	70	3.447	4.301	2.651
20	21	100	70	4.016	4.87	2.812
21	21	120	70	2.719	3.351	2.245
22	21	100	70	3.526	4.36	2.602
23	26	80	80	4.584	5.767	3.77
24	21	80	70	4.662	5.661	3.793
25	21	100	50	3.498	4.372	2.778
26	21	80	50	4.449	5.526	3.544
27	21	120	50	3.197	3.931	2.167
28	18	120	50	2.985	3.384	2.19
29	21	120	50	3.028	3.673	2.214
30	26	120	50	2.099	3.521	2.25
31	26	160	50	1.813	2.246	0.996
32	26	140	50	2.406	3.054	1.619
33	26	180	50	1.945	2.474	1.233

El modelo de Red Neuronal de Función de Base Radial fue aplicado en un proceso de soldadura PTA, generando tres variables de entrada: Velocidad de alimentación del

polvo (g/min), Velocidad del proceso (cm/min) y Corriente de soldadura (Amp). Este proceso de soldadura cuenta con tres características de calidad: Área HAZ (mm^2), longitud de HAZ (mm) y ancho del cordón de soldadura (mm).

La red neuronal de función de base radial cuenta con una matriz de 33 entradas, siendo estas el número de observaciones realizadas por las 6 variables de interés mencionadas anteriormente. La matriz de las variables de entrada (X) fue aplicada en los algoritmos de optimización para generar una matriz de centros de 3 x 3 con 3 subgrupos de las tres variables, representando todo el grupo de observaciones, de forma que cada punto de la matriz es el centro de cada subgrupo.

Los centros de la Función de Base Radial se calculan y optimizan por los algoritmos con el fin de encontrar una matriz que, según la función de base radial maximice el valor de la función fitness, (en el caso del coeficiente de determinación R_q^2) y para el caso de los Criterios de Información de Akaike (AIC) definidos de (10) a (13), los minimice. A continuación, se muestran en la Tabla 2 los resultados obtenidos del coeficiente de determinación R_q^2, y los Criterios de Información con cada algoritmo para las tres variables de salida. Para obtener dichos resultados cada algoritmo realizó 300 iteraciones.

Tabla 2. Resultados de Criterios de Información Para Algoritmo Genético, PSO y Recocido Simulado. Fuente: Elaboración propia.

AG	***Gauss***	***Multi C.***	***Multi C.I.***	***Cubica***	***Spline***
R_q^2	**0.9116**	0.8914	0.8252	0.6564	0.6564
AIC_q	**247.52**	302.25	-13.67	1459.79	1695.97
AIC_{cq}	**109.76**	150.35	-95.13	1056.6	1254.19
BIC_q	**244.92**	296.44	-11.49	1456.76	1703.83
$HQIC_q$	**269.73**	305.607	-8.09	1475.23	1709.31
PSO	***Gauss***	***Multi C.***	***Multi C.I.***	***Cubica***	***Spline***
R_q^2	0.9048	0.8164	0.7340	0.6588	0.6584
AIC_q	795.54	656.84	36.51	-21189.89	-15640.93
AIC_{cq}	419.13	414.16	-49.703	-16831.28	-11095.89
BIC_q	671.22	607.65	117.51	-22199.95	-14182.00
$HQIC_q$	680.70	626.67	147.549	-21915.48	-15562.93
ARS	***Gauss***	***Multi C.***	***Multi C.I.***	***Cubica***	***Spline***
R_q^2	0.6736	0.6766	0.6833	0.6564	0.6564
AIC_q	2528.39	1058.58	52.305	686.1	1095.22
AIC_{cq}	1928.17	743.59	-35.97	460.33	772.66
BIC_q	2548.29	1045.79	47.82	680.64	1066.38
$HQIC_q$	2584.7	1058.22	67.93	698.83	1075.8

En la Figura 1 se observa la relación en y_1: Área de HAZ entre la respuesta de la RNFBR y la respuesta real como objetivo. La Figura 2 muestra la relación entre el

objetivo y la predicción obtenida para la variable y_2: Longitud de HAZ. Del mismo modo, la Figura 3 muestra de manera gráfica la relación entre el objetivo y la predicción dada por la RNFBR para la variable y_3: Ancho del cordón de soldadura. Estos resultados se obtuvieron al emplear el algoritmo Genético y la función Gaussiana ya que fue la combinación de centros dados por este algoritmo y función de activación que permitió que la red tuviera un buen ajuste, al obtener los valores más cercanos a cero de los criterios de información, y un valor alto en el coeficiente de determinación R_q^2, el cual fue de 0.9116 para cada variable de salida (Y1, Y2 y Y3).

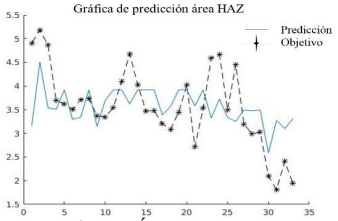


Figura 1. Gráfica predicción para Área HAZ. Fuente: Elaboración propia.

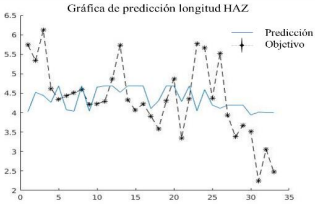


Figura 2. Gráfica predicción para Longitud HAZ. Fuente: Elaboración propia.

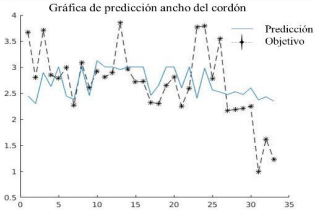


Figura 3. Gráfica predicción para Ancho del cordón. Fuente: Elaboración propia.

Conclusiones.

La RNFBR fue empleada en conjunto con algoritmos evolutivos para predecir las respuestas de un proceso de soldadura PTA con tres variables de respuesta. Se utilizaron los Criterios de Información multivariados AIC_q como función de adecuación en cada uno de los algoritmos, lo cual permitió la comparación de las diferentes Funciones de Base Radial y la obtención de los centros óptimos para la elección del mejor modelo que describió adecuadamente el comportamiento del proceso. Dicha función, en este caso fue la función Gaussiana empleada en el algoritmo Genético, debido a los valores más cercanos a cero de AIC_q y a su vez el estadístico R_q^2 muestra que el modelo de RNFBR representó adecuadamente el proceso de soldadura PTA, ya que es muy cercano a 1. Se puede concluir que para este caso la función con mejores resultados será la Gaussiana, y que el algoritmo que permitirá obtener mejores resultados al momento de optimizar los centros, será el algoritmo Genético.

Referencias.

1. Holland, J., (1975). Adaptation in natural and artificial systems. University of Michigan Press: Ann Arbor.
2. Kirkpatrick, S., Gelatt CD, Vecchi MP (1983). Optimization by simulated annealing.
3. Kennedy, J., & Eberhart, R. C. (1995). Particle Swarm Optimization. In Proceedings of the IEEE conference on neural networks IV, Piscataway, NJ.
4. Hiromassa T. E. (2008). Evaluación de los parámetros del procesamiento sobre recargue PTA de la Universidad Federal de Paraná, la División de Tecnología, Departamento de Ingeniería Mecánica, Curitiba, Paraná, Brasil.
5. Acherjee, B. (2019). FEM-ANN Sequential Modelling of Laser Transmission Welding for Prediction of Weld Pool Dimensions.
6. Ali N. Ahmed, C. W.-S. (2016) RBF-NN-based model for prediction of weld bead geometry. Neural Comput & Applic.
7. Dhaliwal S., & Pattnaik S., (2017). Development of PSO-ANN Ensemble Hybrid Algorithm and Its Application in Compact Crown Circular Fractal Patch Antenna Design.
8. Saravanan, R., Asokan P., & Vijayakumar K., (2003). Machining Parameters Optimisation for Turning Cylindrical Stock into aContinuous Finished Profile Using Genetic Algorithm and Simullated Annealing.
9. Naser H., 2017. Oil Price Fluctuation, Gold Returns and Inflationary Pressure: An Empirical Analysis Using Cointegration Approach.
10. Hou, M., Yang, Y., Liu, T., & Peng, W. (2018). Forecasting time series with optimal neural networks using multi-objective optimization algorithm based on AICc. Frontiers of Computer Science, 12(6), 1261-1263.
11. Broomhead D., Lowe D. S., 1988. Multi-variable functional interpolation and adaptive networks. Complex Systems 2.
12. Yu, X. and Gen, M.: Introduction to Evolutionary Algorithms, Springer (2010).
13. Dowsland K., & Díaz B., 2003. Diseño de Heurística y Fundamentos del Recocido Simulado. Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial.
14. Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19716-723.

5

DinoBlock: Una Aplicación Móvil para el Pensamiento Computacional

RAMÓN ZATARAIN CABADA, MARÍA LUCIA BARRÓN ESTRADA Y ULISES URIEL CABRERA RUBIO

DinoBlock: Una Aplicación Móvil para el Pensamiento Computacional

[1]Ramón Zatarain-Cabada, [1]María Lucia Barrón-Estrada, [1]Ulises Uriel Cabrera-Rubio

Tecnológico Nacional de México/IT de Culiacán, Juan de Dios Bátiz 310 Pte., Col. Guadalupe, Culiacán Sinaloa, CP 80220, México

{rzatarain, lbarron, ulises.cabrera}@itculiacan.edu.mx

Resumen. En este artículo presentamos una aplicación móvil enfocada en apoyar el aprendizaje del pensamiento computacional, además de apoyarnos de las técnicas de bloques de Google. El objetivo fue: combinar técnicas de gamificación, Google Blockly, lógica difusa y monitoreo del estado afectivo del estudiante. La aplicación cuenta con dos tipos diferentes de juegos: el primero es tipo puzzle y el segundo es de recolección de objetos. El sistema al final de cada ejercicio por medio de lógica difusa evalúa las siguientes variables: emoción del estudiante, tiempo, errores, ayudas y categoría actual, para asignar que ejercicios podrá realizar el estudiante. Las emociones del estudiante son: aburrido, comprometido, emocionado, enfocado, interesado y relajado. Estas se obtienen a través de tecnologías de inteligencia artificial y apoyándonos de la librería OpenCV para el preprocesado de imagen. La aplicación fue probada con estudiantes de maestría y se presentan los resultados de las pruebas que se obtuvieron.

Palabras clave: Pensamiento computacional, lógica difusa, computación afectiva, aplicación Android.

DinoBlock: A Mobile Application for Computational Thinking

Abstract. In this article presents a mobile application focused on supporting computational thinking learning, addition to supporting Google's block techniques. The objective was: to combine gamification techniques, Google Blockly, fuzzy logic and monitoring of the student's affective state. The application has two different types of games: the first is a puzzle type and the second is an object collection game. The system at the end of each exercise, using fuzzy logic, evaluates the following variables: student emotion, time, errors, help and current category, to assign which exercises the student will be able to do. The student's emotions are bored, engaged, excited, focused, interested, and relaxed. These are obtained through artificial intelligence technologies and using the OpenCV library for image preprocessing. The application was tested with master's students and the results of the tests that were obtained are presented.

Keywords: Computational thinking, fuzzy logic, affective computing, Android application.

1 Introducción

Hoy en día el pensamiento computacional ha tomado gran relevancia ya que es fundamental para la resolución de problemas y la comprensión del comportamiento humano [1]. Muchos países buscan la forma de desarrollar el pensamiento computacional a temprana edad de los estudiantes. Según Wing el pensamiento computacional "Es un proceso de pensamiento en el que las soluciones están representadas en una forma que un agente de procesamiento de información puede llevar a cabo de manera efectiva" [2]. Es por eso por lo que aprender los conceptos del pensamiento computacional representa todo un desafío para un estudiante joven sin conocimientos elementales de computación y programación. Por otra parte, la computación afectiva desempeña un papel muy importante en el aprendizaje de las personas. Actualmente estos sistemas deben de ser capaces de capturar y reconocer los estados afectivos del usuario a través de mediciones en el rostro, voz o cuerpo, procesar esta información clasificándola, y generar una respuesta. A través de los años se ha desarrollado sistemas capaces de reconocer las emociones mediante diferentes medios: expresiones faciales, señales de habla, señales del cerebro, entre otras, con la finalidad de mejorar el interés en el aprendizaje. Ekman & Rosenberg realizaron un estudio de emociones donde utilizan las siguientes categorías: Sorpresa, Alegría, disgusto, tristeza, miedo e ira [3].

El objetivo de esta investigación consiste en desarrollar una aplicación móvil para apoyar el aprendizaje del pensamiento computacional, apoyándonos de diferentes tecnologías como lo son: lógica difusa con la cual personalizamos el aprendizaje de los estudiantes; aprendizaje máquina para el análisis de emociones del estudiante; computación afectiva para usar la emoción del estudiante en su proceso de aprendizaje; Toda la aplicación móvil implementada en un entorno de programación visual con Blockly. El desarrollo de esta aplicación permite que los estudiantes comprendan mejor y de una forma más sencilla los conceptos del pensamiento computacional.

El artículo se estructuró de la siguiente manera: en la sección 2 se describe algunos de los trabajos más importantes en el área del pensamiento computacional en la educación. En la Sección 3 describimos la arquitectura del sistema, el reconocedor de emociones usando expresiones faciales, el sistema de gamificación y el sistema de lógica difusa. En la Sección 4 describimos las pruebas que se realizaron y los resultados obtenidos de dichas pruebas. Por último, la Sección 5 presenta conclusiones y el trabajo a futuro.

2 Trabajos Relacionados

En esta sección se describe algunos de los trabajos relacionados en el área de la computación afectiva y del pensamiento computacional en la educación, además de incluir una descripción de las contribuciones del proyecto en el área.

Uno de los proyectos mas importantes es la organización Code.org, el cual esta basado en web y está dedicado a expandir el acceso a la informática en las escuelas y aumentar la participación de mujeres y jóvenes. Ofrecen diferentes áreas para aprender, como lo son Ciencias de la Computación [4].

El proyecto desarrollado por José Manuel [5] cuenta con dos dispositivos: Raspberry Pi y Picoboard los cuales permiten utilizar diferentes sensores para utilizarlo de guitarra programable con Scratch. El proyecto esta enfocado en: entender los principios fundamentales de la ciencia de la computación, analizar y resolver problemas mediante la aplicación de técnicas computacionales y escribir un programa relativo a estos problemas. Se realizaron pruebas con 93 estudiantes.

El reconocimiento facial Alizadeh y Fazel utiliza 7 categorías de emoción: enojado, asco, miedo, feliz, triste, sorpresa y neutral. El corpus consta de unas 37,000 imágenes en escala de grises de 48×48 píxeles. Con el propósito de aumentar los datos, utilizaron imágenes reflejadas volteando imágenes en el conjunto de entrenamiento horizontalmente. El sistema utiliza una red neuronal CNN, el cual alcanza una precisión de: enojado (53%), asco (70%), miedo (46%), feliz (80.5%), triste (63%), sorpresa (62.5%), neutral (51.5%) [6].

El proyecto propuesto por Patwardhan y Knapp [7] es un sistema que utiliza múltiples modalidades: cara, cabeza, mano, cuerpo y habla para el reconocimiento afectivo, para ello utilizan un sistema que les permita la detección de color y profundidad (Kinect). Este dispositivo les permite la extracción de características faciales, así como el movimiento de las articulaciones. En la parte de la clasificación utilizan aprendizaje supervisado (Maquina de Vectores de Soporte). En los resultados solo se basaron en 3 emociones: felicidad (84%), enojo (77%) y tristeza (71%).

La principal contribución de este trabajo de investigación es la integración de un reconocedor de emociones centrado en el aprendizaje con un sistema de lógica difusa para la asignación del nivel y puntaje del usuario, además de un sistema de gamificación con la finalidad de disminuir los abandonos o la falta de motivación en las actividades de aprendizaje, todo ello bajo la plataforma móvil de Android.

3 DinoBlock

En esta sección se describe las partes de DinoBlock: arquitectura seleccionada, el sistema de reconocedor de emociones centradas en el aprendizaje, el sistema de lógica difusa y el sistema de gamificación. La metodología que se uso para el desarrollo de esta aplicación fue el desarrollo iterativo. Esta permite obtener resultados importantes desde las primeras iteraciones, es flexible a los cambios y podemos conocer el progreso real del proyecto [8].

3.1 Arquitectura

La arquitectura se seleccionó pensando en los objetivos. La arquitectura de DinoBlock seleccionada fue de 3 capas. Esta arquitectura proporciona una buena distribución de las responsabilidades jerárquicamente, además de reducir el acoplamiento. Las

capas son: presentación, lógica de negocios y de datos. En la figura 1 se puede visualizar la arquitectura con sus 3 capas. La capa de presentación es la encargada de todo lo referente a la interfaz de usuario del sistema y esta dividida en: Actividades, Diálogos y Fragmentos. La capa de lógica de negocios está representada por: la autentificación del usuario, los niveles de cada juego, reconocedor de emociones, los tutoriales, manejo de LibGDX, manejo de Blockly, gamificación y la lógica difusa. La capa de datos es la encargada del acceso de la información, por lo cual se utilizó Firebase ya que proporciona una base de datos de manera local y tanto como de manera remota.

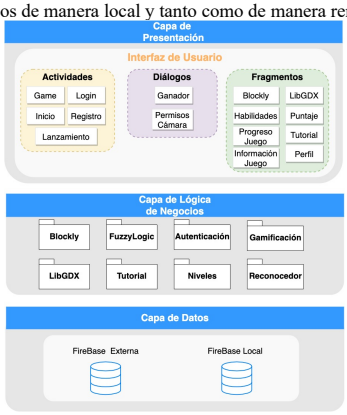


Figura 1: *Arquitectura del sistema. Fuente: Elaboración propia.*

3.2 Sistema Reconocedor de Emociones

El sistema reconocedor de emociones es el encargado de obtener una imagen del rostro del estudiante, procesarla y proporcionar una emoción, dicha emoción se utiliza en el sistema de lógica difusa. El sistema de reconocedor de emociones se compone en 4 componentes fundamentales: el componente de Cámara, Acelerómetro, OpenCV y Tensorflow Lite. Dado que trabajamos en un dispositivo móvil se requiere optimizar los recursos lo mas posible. OpenCV lo utilizamos en gran medida para el preprocesamiento de la imagen.

La clasificación consiste en reconocer el rostro mediante una imagen de 150 x 150, la cual, por medio de aprendizaje profundo, entrega 6 posibles valores: aburrido, comprometido, emocionado, enfocado, interesado y relajado. Para la clasificación se usó una red neuronal convolucional (CNN por sus siglas en ingles). El corpus se creó utilizando una diadema EMOTIV Insight con 5 canales de electroencefalograma por medio de un programa computacional [9].

3.3 Sistema de Lógica Difusa

El sistema de lógica difusa se encarga de evaluar al estudiante en cada nivel, utilizando para ello los datos que se generan a través del ejercicio. Una vez que el estudiante completa dicho ejercicio se encarga de utilizar alguno de los datos y posteriormente procesar las variables. El sistema proporciona una variable de salida en un rango de 0 – 100, tomamos ese valor para asignar la puntuación y la cantidad de estrellas. La cantidad de estrellas es un medio de gamificación que estamos utilizando para motivar a los estudiantes. Usamos la librería jFuzzyLogic, la cual utiliza el lenguaje de control difuso (FCL por sus siglas en ingles), además que permite trabajar en la plataforma java [10].

En las variables de entrada se definieron 5 posibles valores: errores, tiempo, ayudas, categoría y emoción. Se definió una variable de salida la cual es puntaje. En la figura 2 se muestra como se definió las Variables lingüísticas.

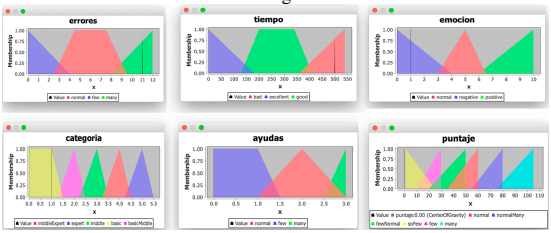


Figura 2: *Variables lingüísticas. Fuente: Elaboración propia.*

De acuerdo con la experiencia de tres profesores, se optó por darle una mayor importancia a las variables errores y tiempo, y darle una menor importancia a las variables ayudas y emoción. Las reglas difusas representan el conocimiento difuso. Estas reglas están representadas mediante IF-THEN y son utilizadas para determinar el valor de la variable de salida "Score". Realizando todas las combinaciones posibles el sistema de lógica difusa cuenta con 119 reglas difusas.

3.4 Sistema de Gamificación

El sistema de gamificación se utiliza con la finalidad de disminuir los abandonos o la falta de motivación en las actividades de aprendizaje. Se utilizan diferentes mecánicas para el sistema entre ellas son: acumulación de puntos, escalado de niveles y las clasificaciones. El sistema proporciona los puntos del estudiante en cada nivel y las estrellas con un rango de 0 a 6. Dichos puntos los proporciona el sistema de lógica difusa por medio de sus reglas.

Se tiene una tabla de posiciones por juego donde se muestran los 10 jugadores con mayor puntaje. A los 3 primeros jugadores se les asigna una medalla. Cada nivel cuenta con dos estados: nivel bloqueado y nivel completado. El nivel bloqueado es aquel donde el estudiante no puede acceder, ya que no a completado uno de los niveles anteriores. El nivel completado es aquel donde el estudiante termino dicho nivel y este tiene asignado una cantidad de estrellas.

4 Pruebas y Resultados

Esta sección se divide en dos partes: la primera se enfoca en la ejecución de un usuario, la segunda en los resultados obtenidos en las pruebas realizadas.

4.1 Sesión de Trabajo

En la figura 3 se muestra un ejemplo de una sesión de trabajo con la aplicación. La aplicación inicia con una animación la cuál, una vez finalizada, muestra la pantalla de inicio de sesión. Para ello, el estudiante tendrá 2 opciones para iniciar sesión: inicio de sesión con google o inicio de sesión con correo y contraseña. Una vez iniciada la sesión se mostrará la pantalla principal, la cual proporciona información de los 2 tipos de juego, aquí podremos desplazarnos en los menús: Juego 1, Juego 2 y Perfil. El estudiante podrá escoger entre los diferentes niveles (dependiendo el nivel que se encuentre desbloqueado). Una vez cargado el nivel, se mostrará: un contador de tiempo, errores, ayudas, y se podrá visualizar las instrucciones del nivel (Figura 3A). El sistema de reconocedor de emociones capturara una imagen cada 30 segundos y almacenara la información, además de mostrar un pequeño icono de una cámara para que el usuario pueda voltear a ver la cámara (Figura 3B). El objetivó es que el estudiante utilice bloques existentes para resolver el nivel (Figura 3C). En caso de que el estudiante resuelva el nivel, el sistema de lógica difusa calculara los puntos obtenidos del nivel dependiendo de las variables (errores, tiempo, ayudas, emoción). Posteriormente se mostrará con un dialogo con el puntaje obtenido y la cantidad de estrellas obtenidas (Figura 3D). El estudiante tendrá dos opciones: ir al siguiente nivel o ir a la pantalla de niveles.

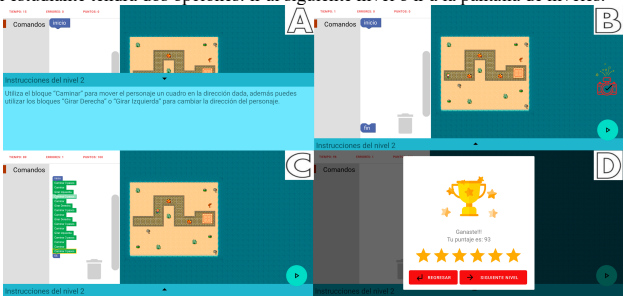


Figura 3: *Sesión de trabajo. Fuente: Elaboración propia.*

4.2 Resultados

Con el objetivo de evaluar el funcionamiento correcto y las interfaces de la herramienta y recibir valiosa retroalimentación, se realizaron pruebas en un aula con 10 estudiantes de maestría de Ciencias de la Computación del Instituto Tecnológico de Culiacán.

En la evaluación se les asignó un tiempo de 60 minutos para la realización de los ejercicios. Los alumnos compartieron comentarios para realizar algunas mejoras en la

aplicación. En la tabla 1 se muestran algunos de los comentarios realizados por los estudiantes.

Tabla 1. *Comentarios de retroalimentación.*

No	Comentario
1	Colocar imagen de puzzle armado en el cuadro de instrucciones en juego 1
2	Poner una guía para contar el numero de espacios a caminar en juego 2
3	Situar un puntaje cuando se termina un nivel
4	Emplear estrellas en niveles para visualizar que este completo
5	Crear un bloque Caminar X3 en juego 2

Fuente: Elaboración propia.

La emoción que mas se presento fue: emocionado, seguida de comprometido. Esto indica que en la prueba los usuarios disfrutaron usando la aplicación.

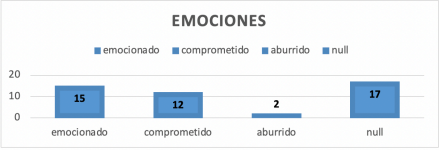


Figura 4: *Gráfico de emociones frecuentes. Fuente: Elaboración propia.*

Como se puede ver en la figura 4 muestra que el sistema tiene margen de mejora, ya que por ser un dispositivo móvil afecta mucho el grado de inclinación con el que es sujetado y es por eso por lo que aparece la etiqueta "null". Por otra parte, conseguimos que los usuarios no se aburrieran usando la aplicación ya que las emociones que

5 Conclusiones y Trabajo Futuro

Desarrollar una aplicación para apoyar el aprendizaje del pensamiento computacional no es una tarea fácil. El objetivo de esta investigación fue el desarrollar una aplicación móvil para el aprendizaje del pensamiento computacional con la ayuda de la computación afectiva. Se realizo una prueba con estudiantes de maestría donde se demostró que la aplicación es amigable con los usuarios y tiene un impacto positivo en ellos.

DinoBlock utiliza: un sistema multijuegos, el cual se utiliza para ayudar al estudiante a aumentar el interés en el aprendizaje; un sistema de bloques, que permite crear algoritmos de manera mas fácil (en vez de código). Todo esto ayuda al estudiante a aprender los conceptos de el pensamiento computacional de forma divertida y fácil.

La aportación de esta investigación consistió en: incluir un sistema reconocedor de emociones, sistema de lógica difusa y gamificación en la plataforma móvil Android. Las herramientas utilizadas fueron: Android Studio, Java, Kotlin, Tiled, jFuzzyLogic, OpenCV, Tensorflow Lite, LibGDX, FireBase y Blockly. El sistema reconocedor de emociones centradas en el aprendizaje inicialmente se desarrolló para Windows (PC y laptops), por lo cual se implementó para móviles.

Existen varias limitaciones con la investigación existente que necesitan trabajarse en un futuro. Una limitación es que las pruebas se realizaron con un pequeño grupo de estudiantes de maestría del Instituto Tecnológico de Culiacán. Otra limitante es en la parte del reconocedor de emociones, ya que esta implementado en un dispositivo móvil se necesita estar monitoreando el grado de inclinación, ya que este impacta negativamente en la clasificación.

Como trabajo futuro se planea realizar experimentos con dos grupos de estudiantes de nivel secundaria, para analizar mejor el impacto de la herramienta. La primera prueba es TAM (Modelo de Aceptación Tecnológica). La cual permite comprender y evaluar la aceptación de los usuarios hacia nuevas tecnologías. La segunda es una prueba cognitiva, esta se realizará mediante: un pretest (Se obtiene una referencia base sobre los conocimientos que tienen los estudiantes); intervención (donde consiste en emplear la aplicación DinoBlock); y un posttest (Permite realizar una comparación entre método tradicional, contra nuestra aplicación DinoBlock). Además, que dada las pruebas que se realizaron se llego a la conclusión que se modificara el sistema de reconocedor de emociones, ya que al ser un dispositivo móvil el ángulo del dispositivo lo afecta mucho.

Referencias

1. Yadav, A., Gretter, S., Good, J., & McLean, T. (2017). Computational thinking in teacher education. In Emerging research, practice, and policy on computational thinking (pp. 205-220). Springer, Cham.
2. Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. Educational Research Review, 22, 142-158.
3. Bosquez, V., Sanz, C., Baldassarri, S., Ribadeneira, E., Valencia, G., Barragan, R., ... & Camacho-Castillo, L. A. (2018). LA COMPUTACIÓN AFECTIVA: EMOCIONES, TECNOLOGÍAS Y SU RELACIÓN CON LA EDUCACIÓN VIRTUAL. Revista de Investigación Talentos, 5(1), 94-103.
4. Bučková, H., & Dostál, J. (2017). Modern approach to computing teaching based on Code.org. In 10th International Conference of Education, Research and Innovation (pp. 5091-5096).
5. López, J. M. S., & Gutiérrez, R. C. (2017). Pensamiento computacional y programación visual por bloques en el aula de Primaria. Educar, 53(1), 129-146.
6. Rashid, T. A. (2016, September). Convolutional neural networks based method for improving facial expression recognition. In The International Symposium on Intelligent Systems Technologies and Applications (pp. 73-84). Springer, Cham.
7. Patwardhan, A., & Knapp, G. (2016). Multimodal affect recognition using Kinect. arXiv preprint arXiv:1607.02652.
8. Jalote, P. (2012). An integrated approach to software engineering. Springer Science & Business Media.
9. González-Hernández, F., Zatarain-Cabada, R., Barrón-Estrada, M. L., & Rodríguez-Rangel, H. (2018). Recognition of learning-centered emotions using a convolutional neural network. Journal of Intelligent & Fuzzy Systems, 34(5), 3325-3336.
10. Cingolani, P., & Alcalá-Fdez, J. (2013). jFuzzyLogic: a java library to design fuzzy logic controllers according to the standard for fuzzy control programming. International Journal of Computational Intelligence Systems, 6(sup1), 61-75.

6

Una nueva herramienta basada en tecnología de videojuegos para enseñar teoría musical básica

Baruc de Los Santos Izquierdo, Flor Radilla López, Eduardo Vazquez Santacruz y Elvia Elvira Morales Turrubiates

Una nueva herramienta basada en tecnología de videojuegos para enseñar teoría musical básica

Baruc de los Santos Izquierdo,-y, Flor Radilla-López, Eduardo Vázquez-Santacruz, Elvia Morales Turrubiates

Universidad Autónoma del Carmen. Ciudad del Carmen Campeche, México
161447@mail.unacar.mx, fradilla@pampano.unacar.mx, evazquez@pampano.unacar.mx, emorales@delfin.unacar.mx

Abstract. One of the main problems for people who want to learn to play an instrument is the auditory part since it is something very difficult to develop for those who do not have that ability naturally. The present work aims to create a support tool through a video game for the musical learning of people mainly children. It is a tool in arcade-type video game format developed with the support of the Unity software, in which the character will go on a race through a staff that will reveal the musical notes of the song that has previously been selected from the list and based on the difficulty level also chosen. This software tool will be a support for children to help them memorize musical pieces in addition to develop their musical ear. The video game creation process was carried out using the cascade software development methodology (Requirements Analysis, Design, Implementation and Testing / Adjustments) since the user requirements were clearly known.

Keywords: Videojuego, Unity, arcade, play, instrument musical notes, cascade software development.

1 Introducción

Existen diferentes métodos en la actualidad que sirven de apoyo en el aprendizaje de distintas disciplinas tales como la música. Esta disciplina plantea diversos desafíos, que van desde el aprender el nombre de las notas, reconocerlas en el pentagrama, la ejecución de algún instrumento o incluso la composición musical. Estos desafíos se tornan más difíciles para aquellos alumnos que muestran un proceso de aprendizaje más tardado o que requieren de un reforzamiento constante de los temas. Con frecuencia se requiere de estos alumnos un esfuerzo mayor que muchas veces termina por frustrar la ruta de aprendizaje esperada [4]. Debido a esto la industria del software para la educación ha estado en constante crecimiento desde hace varios años desarrollando aplicaciones que apoyan y refuerzan el aprendizaje; gracias a esta labor, hemos podido experimentar mejoras en el proceso de aprendizaje de disciplinas como matemáticas, gramática, geografía, por mencionar algunas. El uso de software para la educación

propicia un aprendizaje autodidacta, simple y con mayor grado de interés [3], poniendo en práctica los conocimientos y aprendizajes de los alumnos a través de desafíos interactivos. A su vez, permite un mayor número de usuarios, mayor diversidad y una mejor inclusión ya que con una correcta planificación e implementación el software puede adaptarse a las necesidades individuales de cada usuario. El software para la educación musical tiene una gran importancia ya que ayuda al alumno en la concentración, comprensión y memorización de la teoría musical. También refuerza la habilidad y el conocimiento musical sin importar el estilo de aprendizaje (auditivo, visual o quinestésico) que predomine en el alumno ya que se aprovechan al máximo los diferentes canales de percepción y son enfocados específicamente al desarrollo musical. Teniendo en cuenta lo anterior se ha desarrollado un prototipo de software educativo que permite a través de un videojuego que el usuario se relacione con elementos como las notas, el pentagrama, las escalas que existen dentro de una pieza y la composición musical. El desarrollo de esta herramienta se basa esencialmente en lograr que las personas desarrollen la habilidad y conocimiento musical sin importar el tipo de aprendizaje predominante de cada usuario ya que se aprovechan al máximo los diferentes canales de percepción que son enfocados específicamente al desarrollo musical. Esta herramienta se creó considerando una metodología de desarrollo de software tipo cascada dado que se conocían con claridad los requerimientos del usuario [1,2].

2 Trabajo relacionado

Existen diversos esfuerzos con este enfoque de aplicación. Tenemos por ejemplo la herramienta "atrapanotas" que se puede acceder a través del navegador y tiene como propósito practicar la lectura de las notas y el reconocimiento auditivo de cada una. Su modo de juego implica generar un clic en el nombre de la nota que aparece en la pantalla antes de que desaparezca como lo indica la Fig 1 [5]. Por otro lado, se tiene la aplicación *Guitar Hero* de la empresa Activision que es compatible en las plataformas Xbox, PS3, PS4, Wii, Xbox 360 y iOS.

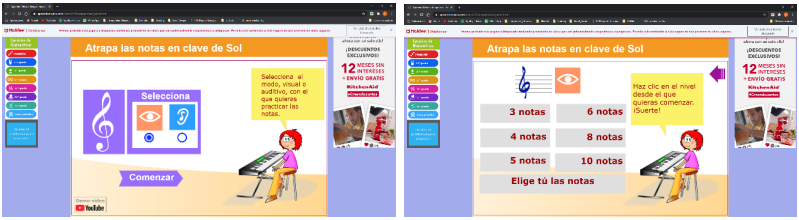


Fig. 1. Modo de juego de la herramienta atrapanotas [5]

Su propósito fundamental es el entretenimiento y su modo de juego implica que a través del controlador se deberán ir tocando de forma rítmica las "notas" que vayan apareciendo en la pantalla como lo muestra la Fig. 2 [6].

Fig. 2. Modo de juego de la herramienta *Guitar Hero* [6]

Nuestro trabajo pretende ofrecer una herramienta de fácil acceso principalmente para los niños que permita disfrutar de la música y al mismo tiempo aprender aspectos básicos de ella.

3 Objetivo

Desarrollar una herramienta de software educativo por medio de un videojuego con contenido interactivo para el aprendizaje musical dirigida principalmente al público infantil para desarrollar la concentración, comprensión y memorización de piezas musicales y difusión de música clásica, y permita además proporcionar apoyo a los docentes en la tarea de evaluación en los cursos de música. La herramienta permitirá mejorar la habilidad de los estudiantes en la identificación de elementos tales como las notas y su valor de tiempo y fomentará la apreciación por la música clásica en las nuevas generaciones ya que las piezas incluidas pertenecen a este género musical.

4 Metodología

La metodología de nuestro trabajo comprende el desarrollo de las siguientes fases (Fig. 3) : 1) Capturar y analizar los requerimientos considerando las necesidades del docente y estudiantes en el proceso de enseñanza-aprendizaje sobre la teoría musical, 2) Diseñar con base en los requerimientos definidos los componentes visuales utilizando herramientas de diseño y las funciones de programación a desarrollar, 3) Implementar las funciones diseñadas para proporcionar comportamiento y dinamismo a los componentes visuales, 4) Ejecutar pruebas para verificar el correcto funcionamiento y en caso de ser necesario realizar ajustes.

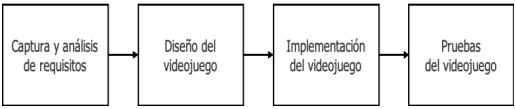


Fig. 3. Proceso de desarrollo de software. Fuente: elaboración propia

Herramientas utilizadas

1. Para la creación y animación del personaje se utilizó Mixamo de Adobe
2. Para la integración y desarrollo se utilizó Unity 2018.3.5f1
3. Para el diseño de interfaz se utilizó Adobe Illustrator.
4. También se utilizaron diferentes paquetes de assets, descargados de Unity Asset Store, para los elementos complementarios del escenario de juego.
5. El lenguaje de programación empleado es C#.

5 Avances

La primera tarea que se realizó fue la captura de los requisitos funcionales que integran el videojuego, esta etapa es muy importante debido a que se definen las funcionalidades que integrarán el videojuego [1,2]. Como se puede apreciar en la Tabla 1 la definición de algunos requisitos está compuesta por un identificador de requisito, descripción y criterio de aceptación. En esta tabla se aprecia la caracterización de la herramienta a partir de las necesidades del usuario.

Número de requisito	Descripción	Criterio de Aceptación
RF_RUNSIC_01	Menú principal para selección de niveles	Identificar el menú de inicio en el juego, al abrirlo.
RF_RUNSIC_02	Teclas de control “a”, “d”, “flecha derecha” “flecha izquierda” y “barra espaciadora”.	El ítem personaje podrá manipularse con las teclas mencionadas.
RF_RUNSIC_07	Cantidad inicial de vidas (Oportunidades)	Cada nivel inicia con 5 “vidas”
RF_RUNSIC_08	Pérdida de una “vida”	Al no pasar por la nota correspondiente se restará una “vida”

Tabla 1. Requisitos funcionales del videojuego desarrollado. Fuente: elaboración propia

Además de establecer los requisitos funcionales fue necesario definir un nombre del videojuego y considerando que va enfocado hacia el aprendizaje musical, se determinó nombrarlo “RUNSIC” por las palabras en inglés RUN (en inglés correr) y SIC (en inglés terminación de la palabra MUSIC). El nombre surge de la tarea que realiza el personaje dentro del videojuego: correr en un trayecto a través de las notas correctas, se ha implementado esto en la designación del nombre tomando el verbo correr en inglés “RUN” y añadiéndole un prefijo de la segunda vertiente del videojuego, que es la música, en inglés, “MUSIC”. El género del juego será de tipo arcade para computadora personal. Posteriormente como parte de las tareas de diseño, se especificó el tema de jugabilidad el cual consiste de un diagrama musical con un personaje principal llamado “Runbot” como se muestra en la Fig. 4. Éste seguirá el sonido de las notas Do, Re, Mi, Fa, Sol, La, Si, Do para lograr completar la canción con el movimiento del personaje manipulado con los controles del teclado (flecha derecha, flecha izquierda, barra espaciadora). El juego tendrá tres niveles. Respecto a las herramientas para diseño e implementación se utilizó el motor de videojuegos Unity versión 2018.3.5f1 y la herramienta Maximo para la creación del personaje principal.

Fig. 4. Personaje principal Runbot. Fuente: elaboración propia

Para el diseño de interfaces se consideró una paleta de colores en tonos fríos (azul, violeta, turquesa). La interfaz principal de inicio de juego se muestra en la Fig. 5.

Fig. 5. Menu de inicio de juego. Fuente: elaboración propia

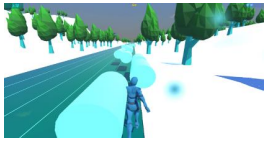


Fig. 6. Pantalla de inicio de juego. Fuente: elaboración propia

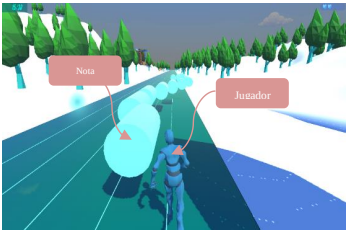


Fig. 7. Pantalla de progreso de juego. Fuente: elaboración propia

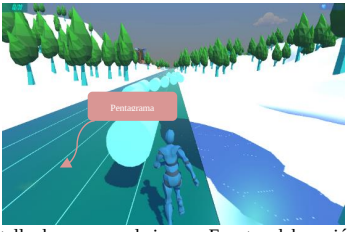


Fig. 8. Pantalla de progreso de juego. Fuente: elaboración propia

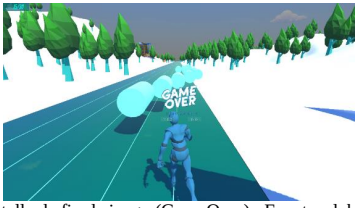


Fig. 9. Pantalla de fin de juego (GameOver). Fuente: elaboración propia

Las figuras Fig. 6, Fig 7, Fig 8, Fig 9 muestran algunos aspectos de funcionalidad del videojuego. El usuario debe mover al personaje hacia los lados, de acuerdo con el carril en donde se ubique la próxima nota (Do, Re, Mi, Fa, Sol, La, Si), dentro de la pista (pentagrama) y permanecer dentro de la nota de acuerdo con la duración de ésta ya que representa el valor de tiempo que la nota tendría en un pentagrama real. En caso de saltarse alguna nota, el jugador perderá una "vida" de las cinco totales con las que cuenta al inicio del juego. El movimiento del personaje será mediante de las flechas del teclado o en su defecto con las letras A-

Izquierda, D-Derecha. A diferencia de los trabajos relacionados antes mencionados, la propuesta de RUNSIC no es abordar el manejo de un instrumento especifico, sino facilitar el aprendizaje de la teoría musical básica, por lo que se elige una pista en forma de pentagrama, como elemento principal de la escena, y notas con el valor de tiempo igual al de un pentagrama en la vida real, para que el alumno pueda asimilar estos conceptos. Durante el juego, el personaje hace un recorrido interactuando con un pentagrama a través del cuál el jugador recibe información propiamente del conocimiento básico del mundo de la música. Como se puede apreciar, el jugador va acumulando puntos conforme se vayan eligiendo las notas correctas con respecto a la melodía en curso. El jugador tendrá cinco oportunidades ("vidas") a lo largo del nivel del juego. El juego considera tres niveles de complejidad que deberán ser superados para considerar que el usuario es capaz de reproducir una melodía de manera correcta conforme avanza el juego. Se maneja un diseño de estilo *lowpoly*, esto con fines prácticos en las pruebas de jugabilidad. Ya que esta primera etapa se enfoca las funciones e interacciones del usuario con los elementos que componen la escena. Para el menú se eligieron colores que evocaran lo electrónico y que fuese llamativo el contraste entre ellos. Dentro del espacio de juego, se eligieron colores más claros, que fueran identificables y llamativos. Elegimos tonos azules ya que transmiten tranquilidad y propician la comprensión. Mientras que los colores verdes aportan al usuario concentración y mejora la eficiencia.

6 Resultados esperados

Actualmente nuestra herramienta se encuentra en un nivel de madurez tecnológica que se puede considerar como prototipo funcional. Es un desarrollo que contiene todos los elementos básicos de un videojuego educativo sin embargo podemos incrementar niveles de complejidad y mejorar los detalles gráficos y estéticos de la herramienta. Pretendemos aplicar esta herramienta en áreas infantiles como guarderías, preescolares y primarias para fomentar el gusto por la música desde la temprana edad. En este sentido, la herramienta tendrá que ser analizada y mejorada considerando los comentarios de los usuarios finales como son profesores, padres de familia e incluso los mismos infantes. En general, nuestra herramienta tiene las funciones básicas necesarias pero su mejora dependerá mucho de la retroalimentación de los actores del ambiente productivo en el que se puede implantar. Se ha elaborado únicamente el primer

nivel del juego, en este el personaje atraviesa la nota próxima, marcada en un carril correspondiente. Para el nivel dos existirán saltos que podrán darse con la barra espaciadora, estos saltos representarán los silencios dentro de las piezas musicales. En el nivel tres, el usuario tendrá que elegir la nota correcta, ya que aparecerán dos notas en diferente carril. Esto permitirá evaluar y reforzar los conocimientos en cuanto a composición musical y escalas. Las piezas musicales serán elegibles y estarán disponibles en los tres niveles, adaptadas a la dificultad de cada uno. Los elementos dentro de la escena del videojuego serán ampliados y rediseñados, para permitir al usuario elegir entre diferentes escenarios y personajes. La herramienta deberá ser analizada por los usuarios finales, que incluirá a los profesores, en la facilidad para la evaluación del progreso académico de los alumnos; y a los alumnos, en cuanto a la jugabilidad. Se prevé también un proceso de implementación de las mejoras posibles, identificadas durante el análisis.

7 Conclusiones

Esta herramienta es producto de diversos esfuerzos que permiten hacer evidente la jugabilidad. El juego como tal implica diferentes conocimientos previos para la realización del videojuego en general como la programación, el diseño, la interactividad, animación, creatividad, el entorno del videojuego, la jugabilidad [1,2,3,4]. Este producto es una herramienta importante en muchos contextos dado que sirve como base para el desarrollo de herramientas digitales para la educación. Además, permitirá realizar mejoras para implantarlo en ambientes productivos educativos, mencionados previamente.

Referencias

1. Lidon Mañas, Marc. Unity 3D. S.A. Marcombo. ISBN: 9788426726827. 2018
2. González Jiménez, Daniel. Diseño de Videojuegos Da Forma a Tus Sueños. Ra-Ma S.A. Editorial y Publicaciones (2015) 176 páginas. ISBN: 8499645488 ISBN-13: 9788499645483
3. Espigares Pinazo, M. J., García Pérez, R., Tejada Gimenéz, J., & Rebollo Catalán, M. Á. (2014). El discurso del profesorado de Educación Musical en la innovación educativa con TIC: posicionamientos en la evaluación del software Tactus. Revista Electrónica Complutense de Investigación en Educación Musical, 11, 14.
4. Reyes Rivero, L., Céspedes Gómez, G., & Molina Cedeño, J. (2017). Tipos de aprendizaje y tendencia según modelo VAK. Tecnología Investigación Y Academia, 5(2), 237-242.
5. Aprendo Música, https://aprendomusica.com/n-03atrapanotas.html, last accessed 2019/12/9.
6. GuitarHero, https://www.guitarhero.com/mx/es/game, last accessed 2019/12/9

7

Educational Data Mining for finding patterns in First Semester Engineering Students Outcomes to identify potential Success or Failure

JORGE EDUARDO IBARRA ESQUER, FELIX F. GONZÁLEZ NAVARRO, DANIEL HERNÁNDEZ BALBUENA AND BRENDA L. FLORES RIOS

Educational Data Mining for finding patterns in First Semester Engineering Students' Outcomes to identify potential Success or Failure

Jorge E. Ibarra-Esquer[1[0000-0003-2636-5051]], Félix F. González-Navarro[2[0000-0002-9627-676X]], Daniel Hernández-Balbuena[1[0000-0002-0055-4797]], and Brenda L. Flores-Rios[2[0000-0002-2502-6595]]

[1] Universidad Autónoma de Baja California, Facultad de Ingeniería. Mexicali, B.C., 21280, México
[2] Universidad Autónoma de Baja California, Instituto de Ingeniería. Mexicali, B.C., 21280, México
{jorge.ibarra, fernando.gonzalez, dhernan, brenda.flores}@uabc.edu.mx

Abstract. In this paper we present the first experimental results on an ongoing project to apply educational data mining in the identification of early patterns that lead to success or failure in engineering students. Data for students of the College of Engineering were obtained from institutional repositories at Universidad Autónoma de Baja California and have been used to track student behavior from an academic perspective, leading to the implementation of mainly corrective strategies based on past and current outcomes. However, there is a need to move a step further and try to predict the chance for failure since early semesters, helping establish preemptive strategies that decrease failure and dropout rates. As a first approach, under the assumption that academic factors can be used to find patterns that allow this type of prediction, the academic results for the first semester of a cohort of engineering students was chosen to find whether and which of the available variables contribute the most to this prediction. Under a data mining perspective, new data was inferred, and an unsupervised learning method was applied to this data in order to create clusters that were compared to recent academic data for that same group of students. While results are inconclusive for some of the categories, there is some insight on variables that allow to identify success and failure from the outcomes of their first semester.

Keywords: Educational data mining, clustering, dropout.

1 Introduction

Educational Data Mining (EDM) is a discipline that uses educational systems' data to create useful information by applying techniques from fields such as machine learning, statistics, and data mining [1]. Among those, clustering techniques are found to be one of the most important pre-processing strategies for finding patterns and categorizations that help understand learning styles, behavior, and performance of students.

As reported in [2], current efforts in using data mining in the higher education sector are oriented to provide further dimensions associated with academic activities, evaluation of learning materials and curriculum improvement, identifying factors associated with student success, failure, and dropout intention, institutional planning and strategies, as well as understanding the instructors' support and administrative decision making. An interesting approach on using EDM on reducing dropout rates of engineering students is described in [3], using decision tree classifiers to predict dropout probability

In this paper we briefly describe a process for using the academic records of first-semester engineering students to identify if these data contain enough information to identify students that might withdraw or dismiss from their majors. Section 2 presents the methodology that was followed for the study. Section 3 describes how every step of the methodology was carried and the results obtained. Conclusions are presented in section 4.

2 Methodology

This work is performed as part of an institutional effort to track student performance and create data and statistics that describe behavior and outcomes at institutional and specific major levels, grouping them under cohorts for most of the required metrics. Data gathering and analysis that constitute the current approach started in early 2018, expanding to include every student at the College of Engineering (Facultad de Ingeniería Mexicali - FIM) of Universidad Autónoma de Baja California since the academic year of 2009, when the current curricula was adopted. The designed approach opted to track not only data at a major or college level, but also on an individual basis, which allows us to move a step further and not only create statistics and data based on past data, but also predict student outcomes.

On an overall basis, the CRISP-DM methodology has been used for the different stages of the student performance tracking project. This methodology was proposed by [4] and defines a process model that provides a framework for carrying out data mining projects that is independent of both the industry sector and the technology used. According to CRISP-DM, a data mining project consists of a cyclical process that can be break down in six phases: Business understanding, Data understanding, Data preparation, Modeling, Evaluation, and Deployment. These phases can be iterated several times in the data mining project, in addition to internal iterations between them. In our current iteration, grades for the first semester courses obtained by students of the 2015-2 cohort were extracted and analyzed to create new variables and identify whether and which of them can be used to predict success or failure in their engineering studies. This cohort was selected as being one of the most recent that provides a set of students that can be categorized under all the possible outcomes listed in Table 1.

These 5 categories allow to better specify and track students according to their global outcomes, as well as giving a deeper level of detail to the generated statistics. For the purposes of the current experiment, categories 1 and 2 are considered as "success", while categories 3-5 are considered as "failure". Student data from the institutional repositories allow to identify two additional categories, as students that got accepted into

the university but never registered to their first semester, and students that transferred to another college within the university. Both cases are treated as outliers as there is no way to track individual data corresponding either to their first semester, for students that transferred previous to the semester the current student tracking approach started, or their current status through their majors, as we only have permissions to access data from students enrolled in the FIM.

Table 1. Categories for student outcomes.

Category	Description
1 - Active	Students that are currently enrolled at the university, with credits pending through their major.
2 - Graduate	When a student achieved all the requirements set by the university and specific major and is no longer enrolled at the university.
3 - Withdrew	Students that have pending requirements through their major but are not currently enrolled at the university. There is no academic restriction for these students to continue their major.
4 - Change of major	Students that at some point through their studies decided to switch majors. Statistically these students are treated as dropouts for their original major.
5 - Dismissed	Students that can't continue their studies, mainly due to poor academic achievement.

As stated earlier, several iterations through the CRISP-DM phases have allowed us to obtain an understanding of data that can now be used to try to infer or predict future outcomes from their early performance. The following section describes an experiment carried out with first semester data for students in the 2015-2 cohort, in terms of the first five stages in the process model of CRISP-DM.

3 Finding success and failure indicators in students' academic data

3.1 Business understanding

The engineering programs at FIM make students course a two-semester common core before opting for a specific major. There are 12 compulsory courses, six per semester, that must be passed before proceeding with their major. About 650 students are welcomed to the college each semester and most of them register for the six courses. Students must achieve a minimum grade of 60 to pass a course, and they are given the choice to present an supplementary exam in case they not reach the passing grade; if they fail in both opportunities or opt not to present the extraordinary exam, then the course has to be taken for a second time the following semester.

An academic year consists of two semesters, identified by year and period as YYYY-P. We use those same identifiers to refer to student's cohorts. Period 2 is the first semester of the academic year, while period 1 of the following year is the second. There

are significant differences in terms of failure in specific courses and dropout rates between students of those cohorts, so we find advisable to analyze data by cohort instead of year.

3.2 Data understanding

Data is obtained from institutional repositories as comma separated files including a record for each grade every student has obtained. This data also contains references to the major, credits achieved and pending, as well as the current status of the student, being this one of the categories described in Table 1. Each row in the file has 54 variables and may contain from 50 to 60 rows per student in the cases of students that have finished their studies. Not all of these variables were required for the current study.

Data for each major are obtained in a separate file, containing students' records dating back to the early 1990s. When a student finishes the common core and enters any of the programs, all of the records referring to this student are transferred to the major and are no longer found in the common core historic file. In other words, as our interest is to work with first-semester records, they must be gathered from all the available files to be processed.

3.3 Data preparation

Variables were arranged in a single row for each of the students in cohort 2015-2. Each row contains student identification, final grade for each of the required courses, grade for the extraordinary exam of each course, if presented, total credits obtained in their first semester, current status of the student, current advance in credits through the major, the total number of semesters the student spent or has spent since entering the FIM, and the total number of semesters it took to finish the common core. From this data, additional variables were created, as grade average in the first semester and percentage of first semester credits earned in that semester.

The initial dataset contained 686 students . However, several were discarded as outliers due to the following reasons. 5 students that never enrolled for their first semester courses and 6 students that transferred to another college; these are the cases that were discussed in section 2. 39 students that transferred from another college after finishing their common core courses; these students were discarded as they did not course their first semester in the FIM. Finally, 41 students that withdrew just after or during their first semester. The final dataset consists of 595 students that started their major at the CE, enrolled for at least 2 semesters, and we were able to obtain current data and status as either success (n_S=387) or failure (n_F=208).

3.4 Modeling

Our main interest is to contribute to the creation of tools, information, and knowledge aimed to decrease dropout and failure rates for engineering students. These would complement and support strategies based on tutoring and mentoring students. Student performance tracking at FIM shows that most students withdraw before finishing their common core courses, so our efforts are on identifying and predicting trends from the

first academic data we have available, being these the outcomes of their first semester. Based on these assumptions, and being aware that college dropout is due to a combination of many variables besides academics, we selected different subsets of variables to conduct an unsupervised learning experiment using the well-known K-means clustering algorithm and try to find some features that could provide an early warning on students that might abandon their studies.

Data corresponding to first semester includes grades for the six courses, grades for the supplementary exams in case they were presented, and credits earned in the semester. Additional data were constructed from the previous data, as grade average and percentage of credits earned.

In a first set of experiments all variables were scaled to the range of 0 to 1, i.e., grades were divided by 100 and credits earned by the total of achievable credits in the first semester. Different subsets and combinations of variables were fed into the kmeans() function in Matlab, using squared Euclidean distance and setting parameters for 5 clusters, as the categories in Table 1, and 2 clusters, for success and failure. A first observation was that data from supplementary exams did not actually contribute to the results, so these grades were discarded from the sets. Not all of the students presented this type of exams, so data was incomplete, and even through imputation (0 for failed courses where the exam was not presented, and 1 when the student passed the course) they did not alter the composition of the clusters.

Grades for three of the courses were also removed from the data in these experiments. These are mostly soft skills developing courses which had high approval rates (above 95% for each course) and high average grades (85 to 87 in a scale of 100). The remaining three courses, Differential Calculus, Linear Algebra, and Chemistry continuously present some of the higher failure rates among common core courses (25, 14, and 18% respectively when considering only students from the 2015-2 cohort) and had lower average grades (74, 86, and 82).

In a second experiment, grades from all six compulsory courses were converted to binary data where 0 indicates a grade below average and 1 a grade at or above average. A new variable was also included where 0 represented a student that failed at least one course and 1 a student that approved all the courses. This provided a better distribution for data, as for most variables, about 80% of the original values were distributed in the top 20%.

3.5 Evaluation

Several experiments were conducted using different subsets of variables. The first set of experiments, as described in the previous section, offered hints of proper clustering but not for all the categories. Setting k= 5 for the K-means algorithm usually formed two large clusters where most of the graduated and active students (Success categories) were properly placed (above 90% accuracy for most of the subsets of variables used). However, students in the other 3 categories (Failure categories) were scattered between the five clusters with no clear indication of a larger cluster. Clustering was not significantly improved when using other measures for distance (cosine and correlation) or the

K-medoids algorithm. When setting k=2 most of the students were usually grouped into a large cluster regardless of the category.

An additional transformation of the data was performed, normalizing the grades values. This slightly improved the true-positive rates (Success categories correctly clustered) but had no effect on the clustering of the failure categories.

A last variation of this experiment was executed. In this case, the percentage of credits earned was removed, keeping only the grades of the compulsory courses. The effect on the clustering was that the true-negative rate (Failure categories correctly clustered) improved to nearly 79%, with higher rates obtained when using only the normalized values for the three courses with higher failure rates, but it also caused the true-positive rate to drop just below 60%.

Even though none of the sets of variables resulted in clusters that correctly grouped both success and failure categories, some interesting insights were gained. First of all, first-semester grades seem to provide a better indicator of risk of failure than credits earned. Additionally, students changing majors, that current tracking records classify as dropouts, are more similar to active students than to students withdrawing from their studies.

The most promising results were obtained from the second experiment, using the seven binary variables with the K-means clustering algorithm. Hamming distance was used and k was set as 2, trying to form clusters for success and failure, as described previously. Table 2 shows the average confusion matrix from 10 executions of the algorithm initialized with random centroids. Matthews Correlation Coefficient (MCC) for this matrix is 0.45 which is an indicator of a good performance.

Fig. 1 shows the clustering results using the grade average for the six compulsory courses in the horizontal axis and the first semester pass rate in the vertical axis to create a representation of the seven variables in two dimensions. For better visualization, the figure is separated in two graphs, the first one shows true-positives and negatives and the second false-positives and negatives; in addition, a logarithmic scale is used for the horizontal axis to separate points along the axis. Even though clusters appear to have a clear distinction, putting all the points in a single graph would show a huge overlap, which makes it difficult to properly separate them.

Table 2. Confusion matrix for the K-means algorithm (k=2).

Current status	Clustering results		Rates	
	Success	Failure	Success	Failure
Success (387)	322	65	83% ± 3%	17% ± 3%
Failure (208)	81	127	39% ± 2%	61% ± 2%

If we extract the subsets of students for the categories Graduate (n_G=42) and Dismissed (n_D=39), as described in Table 1, the clustering true rates improve significantly. The confusion matrix in Table 3 shows the results for these categories.

Table 3. Confusion matrix for the K-means algorithm (k=2) for Graduate and Dismissed students.

Current status	Clustering results		Rates	
	Success	Failure	Success	Failure
Graduate (42)	40	2	95% ± 3%	5% ± 3%
Dismissed (39)	8	31	21% ± 2%	79% ± 2%

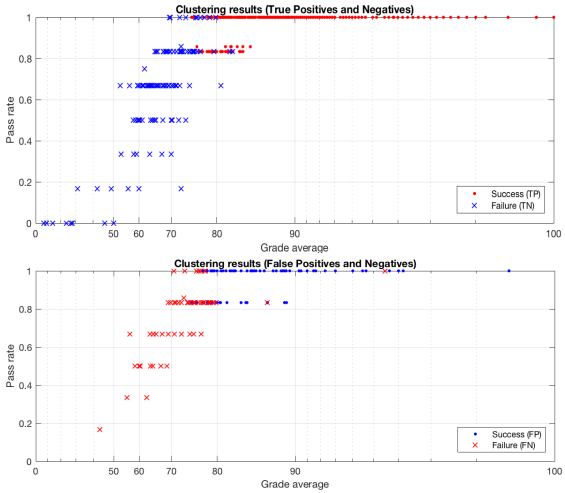


Fig. 1. Clustering results for the second experiment.

As part of tutoring activities at FIM, we have usually tracked active student status using a formula that calculates risk of dropout by dividing the credits earned by the number of semesters that the student has coursed, and then using this value to infer how many semesters each student would need to complete the required credits. Students are then tagged as normal, in risk, or critical depending on how much more time they would need, and this information is provided to tutors to aid them in their advising activities. Taking as a reference the recommendations from accreditation organizations and the internal regulations of the university, a student should finish studies in at most 1.5 times the semesters defined in the curricula. At the FIM, we have set this same factor for the common core, expecting students to finish this stage in at most three semesters. Using our current formula, a student that earns less than two thirds of first semester credits is tagged as critical at the beginning of the second semester, regardless of the courses

passed or failed. Using the same clustering results but tagging students according to whether they finished the common core in maximum 3 semesters (n_N=404) or did not finish the common core or did it in 4 or more semesters (n_C=191), the clustering results were as shown in Table 4. MCC for this matrix is 0.67, which suggests the algorithm performs a better clustering for a short-term outcome than when used for categorizing students by their long-term status. Applying the formula based on credits earned, 71 students would have been properly labeled as critical after their first semester, instead of the 148 that the clustering method correctly identifies as potential risk of failure. The importance of identifying students with risk of failure since early semesters and taking actions tending to improve their status is evidenced by the fact that 146 of the n_c students (76%) has a current status in one of the failure categories.

Table 4. Confusion matrix for the K-means algorithm (k=2) for students finishing the common core in the expected time.

Tracking status	Clustering results		Rates	
	Success	Failure	Success	Failure
Normal (404)	360	44	89% ± 2%	11% ± 2%
Critical (191)	43	148	23% ± 0.5%	77% ± 0.5%

4 Conclusions

A data mining experience with first-semester data from a cohort of engineering students was described. Using an unsupervised learning algorithm, it was possible to identify variables from their first semester outcomes that may lead to infer the possibility of further success or failure in their studies. Clustering results suggest the method that has recently been used at FIM to identify students with risk of failure can be improved. We acknowledge that dropout is influenced by external factors but being able to identify a trend since early semesters can support institutional strategies for decreasing dropout and failure rates. Further analysis is needed on the results, as well as repeating the experiment with students from different cohorts to confirm our findings and proceed with the development of prediction algorithms and tools based on the identified variables.

References

1. Dutt, A., Ismail, M., Heerawan, T.: A systematic review on educational data mining. IEEE Access 5, 15991–16005 (2017).
2. Aldowah, H., Al-Samarraie, H., Fauzy, W.: Educational data mining and learning analytics for 21st century higher education: A review and synthesis. Telematics and Informatics 37, 13–49 (2019).
3. Pal, S.: Mining educational data to reduce dropout rates of engineering students. I.J. Information Engineering and Electronic Business 2, 1–7 (2012).
4. Wirth, R., Hipp, J.: CRISP-DM: Towards a standard process model for data mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, pp. 29-39. Springer-Verlag (2000).

Parte II

Aplicaciones de Ciberseguridad, Criptografía e Informática Médica

Introducción

A medida que la digitalización de los servicios y el consecuente uso de internet ha ido creciendo hasta convertirse en una parte integral de nuestras vidas, la ciberseguridad y el desarrollo de la criptografía se han colocado como requisitos esenciales que garantizan la seguridad de los servicios. Al mismo tiempo, el gran éxito de esta digitalización en la mayoría de los sectores industriales y gubernamentales ha convertido a estos sectores en un blanco propicio para el abuso por partes hostiles, que se aprovecharán de los fallos técnicos y humanos para comprometer los sistemas y robar información privada y financiera.

Por otra parte, la informática médica para la salud y el bienestar social trata con el uso de recursos informáticos, dispositivos, y métodos que se requieren para optimizar la adquisición, almacenamiento, recuperación y uso de la información en el área de la salud, para propiciar que los pacientes tengan una mejor calidad de vida. La permanente demanda de aplicaciones prácticas de las tecnologías computacionales para la atención integral del paciente, el apoyo al quehacer del médico, de servicios clínicos, y la emergencia de nuevos retos de salud y tecnologías para apoyar la atención y cuidados de pacientes vía remota, convoca la participación de grupos de trabajo interdisciplinarios.

En esta sección se presentan tres trabajos, de los cuales dos se encuentran en el área de Ciberseguridad y Criptografía en el ámbito de investigación, innovación y desarrollo tecnológico; el tercer trabajo comparte una experiencia de investigación y desarrollo tecnológico en el ámbito de la Informática médica

Rocío Aldeco
Ana Isabel Marínez-García

8

A Single-Chain System Implementation in C++

BENJAMIN ORIBE MENDIETA AND LUIS LIZAMA

A Single-Chain System Implementation in C++

Benjamin Oribe-Mendieta[1][0000−0001−6656−5811] and
Luis A. Lizama-Perez[1][0000−0001−5109−2927]

Universidad Politécnica de Pachuca, Ex-Hacienda de Santa Bárbara, 43830 Pachuca, México boribem@micorreo.upp.edu.mx
http://www.upp.edu.mx/posgrado

Abstract. We present an implementation of the single-chain algorithm, which constitutes a new alternative to blockchain methods. The single-chain system does not need mining nodes because each user registers it's own transactions that can be distributed among users over the network. The results shown that our implementation reduces the time required to insert a transaction's register to the single chain. The single chain method has been designed for the quantum era since it does not rely on cryptography based on modular arithmetic.

Keywords: hashchain · digital signature · chain · c++ · cryptography · blockchain.

1 Introduction

The blockchain technology [1] has brought new opportunities as a choice to the actual centralized computer systems. This is due to the implementation of cryptographic techniques and a P2P network approach. The first implementations of this technology have shown the potential and limitations of blockchain technology for smart contracts in economy.

Many systems have been based on blockchain. Most of them try to improve the bitcoin functionality. Among these platform, some of the most popular are Ethereum [2] and Monero [3].

Some of drawbacks of blockchain implementations are the dependency of the mining nodes in terms of their computing power and storage capacity in order to create and store new blocks for the blockchain and keep them stored securely.

To solve the mining problem some solutions have been implemented to improve time or resources needed to add blocks to he blockhain such as trying to make the mining parallel[4], but many bet to change the proof-of-work system in the blockchain for methods like proof-of-stake[5], or proof-of-activity[6]. IOTA implements general nodes that can approve and issue transactions.

The cryptographic digital signing systems are considered vulnerable to attacks by quantum methods [7], blockchain is no inmune to this, so methods for non deterministic wallets [8] have been developed, as systems based on post-quantum cryptography. In 2019 it was introduced a new concept for block chain called single-chain [9].

The proposed system states that client nodes of the system can create blocks by themselves in a personal chain that stores transactions, not blocks of transactions, those chains can be managed by the owners in a decentralized ledge with the same level of security provided by regular blockchain systems.

In this paper we implement prototype of single-chain to analyze if all concepts described in the single-chain article are functional, try to improve on some of the theoretical concepts of single-chain as well as compare how this early prototype compares with mainstream blockchain systems.

2 The Single-Chain System

We show an implementation of the single-chain based on Lizama's work [9]. Let us explain briefly such system: having a new message called m, the sender (Alice) has to compute a secret pseudo random secret seed (s). The hash function f is applied $2^n - 1$ times over s the starting point of the chain. Here, $2^n - 1$ is a fixed length for the chain and n is an integer value from 128 to a value in order to keep a desired security level. The head of the chain (public hash) will be expressed as $f^{2^n-1}(s)$. Using the hash (h) of the original message a middle point between s and $f^{2^n-1}(s)$ is defined as $f^{2^n-1-h}(s)$.

With the public hash and the middle point (signing hash) sent publicly to Bob (along the message m), he computes h and apply f^h to the middle point to validate $f^{h_0}(f^{2^n-1-h_0}(s_0)) = f^{2^n-1}(s_0)$ thus verifying the signature as it can be seen in Fig. 3.

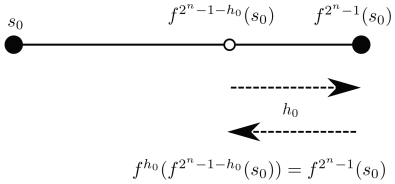


Fig. 1. The public head-chain $f^{2^n-1}(s_0)$ is obtained whenever the hash function is applied h_0 times to the signing hash $f^{2^n-1-h_0}$. It must be noticed that in absence of the signing hash is infeasible to return from the public head-chain to the signing point, because it implies reversing the hash function.

But, this system is not completely functional when trying to apply 2^n times a hash function when $n >= 128$. This is due to the computational time it requires

on a regular computer. So a new approach was to reduce the length of the hash-chain. Thus, the hash of the message is divided in k blocks as shown in Fig. 2. As an example, if the hash of the message is 128 bit long, we can divide it into 4 blocks of 32-bit: h_{00}, h_{01}, h_{10} and h_{11}. Now, the length of the hash chain is reduced according to the relation $f^{2^{n/k}-1}(s)$.

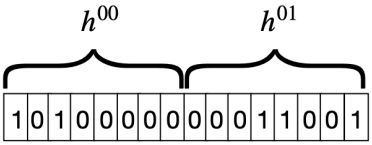


Fig. 2. Is illustrated the dividing process of a 16-bit hash into two 8-bit blocks based on bit significance, thus generating two blocks h^{00} and h^{01}.

2.1 The Key-File

To compute the chains it will be needed to create a random seed (s) value for every block the message is divided in. Then the list of secret seeds $s_1 ... s_k$ is defined. Because that information is private and used to validate the user identity it cannot be kept in the database, so is stored in a file in the user system that will be called key-file.

When the values $n = 256$ and $k = 16$ the key-file should store $16 \cdot 256$ bits for each message to be signed. This bring issues about the size of the key-file after the user sends more messages. The original work generates only the first two secret keys randomly, then creates the third seed from the concatenation of the two previous seeds and compute the hash to create the next s value, as explained in Equation 1.

$$s_0, s_1, s_2 = f(s_0 \| s_1), s_3 = f(s_1 \| s_2), ... \tag{1}$$

With this method, no matter the value of k, the number of seeds stored in the key-file will be no longer than two, so if $n = 256$, the key file should store 512 bits per message, and the rest of the values of s can be computed by concatenating the values of previous seeds.

For this project we designed a solution that looked for a static file size. In the new solution the complete set of k seeds of the first two messages should be generated pseudo-randomly and stored in a section of the key-file called `alpha`. When generating the third message the seeds will be created by the concatenation from the corresponding seeds of the previous messages and hashing them into one new 256 bits number as shown in equation 2.

$$s_{j_0} = H(s_{j-1_0} \| s_{j-2_0}), s_{j_1} = H(s_{j-1_1} \| s_{j-2_1}), ... \quad (2)$$

To finish the key-file structure there is a section called `omega`. In this section the values of the secret seeds from the last two messages are stored, so there is no need to compute the next group of seeds from the ground. `alpha` will only store the sets of seeds from the first two messages, in case a complete validation of the user chain is needed.

When new seeds are added, the older ones are removed, they will keep the key-file in a constant file size of approximately 7KB for a $k = 16, n = 256$ configuration.

2.2 The Security of Single Chain

To prevent that a third user could choose a signing hash between $f^{2^n-1-h}(s)$ and $f^{2^n-1}(s)$, a security hash is calculated from two consecutive messages. The security point is generated as $f^{2^n-1-h-\delta x_{i+1}}(s_i)$ where δx, or *chain ratio*, is defined in equation 3.

$$\delta x_i = h_i - \frac{h_{i-1}h_i}{2^n - 1} \quad (3)$$

The signature verification is done validating $f^{\delta x_{i+1}}(f^{2^n-1-h-\delta x_{i+1}}(s_i) = f^{2^n-1-h}(s_i)$ instead of $f^h(f^{2^n-1-h}(s_i)) = f^{2^n-1}(s_i)$, cutting down the computational effort needed for the signature verification process.

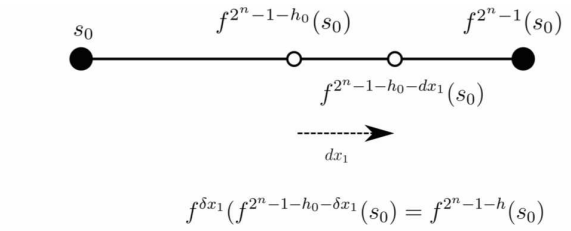


Fig. 3. The message can be validated applying δx_1 hashes to $f^{2^n-1-h_0}(S_0)$ that is shorter than applying .

2.3 The Web Service

To implement the signature process, a cloud storage web service would be used, as shown on Fig. 4 where the head hash, signing hash, the security hash and

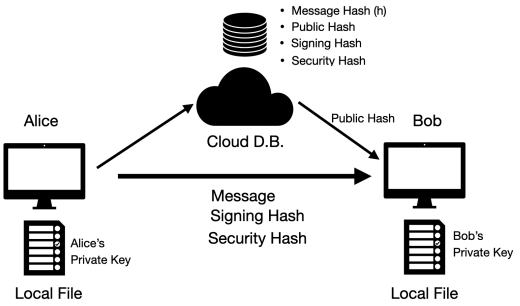


Fig. 4. The public information is stored in a cloud data base system while the private keys are generated from the secret seeds (s) which are saved in a local key-file. To sign message 1 Alice sends to Bob: Message 1, Sign1 and Sec0. Bob retrieves from the database the Public Key Public1. In the database will be stored HashMsg1, Sign1, Sec0 and Public2.

the hash of the message are kept in the web service while the seed values s are stored in a key-file at the user's computer side.

Suppose that only a hash chain is used to sign a message (actually 16 chains are required). The database structure to store the sensitive data in the web service is shown in Table 1. After the message 1 is signed, it is stored Sign1, the hash of the message HashMsg1, the security hash Sec0 and the Public Key Public1. Then the public hash for the next message Public2 is announced in the database. To sign a new message is computed Sign2 and the security hash Sec1, both will be stored in the database along the hash HashMsg2 and the Public Key Public3 and so on (see Table 1).

Table 1. Illustrative representation of the data storage in a theoretical database where only a hash chain is needed to sign a message.

ID	Public Key	Message Hash	Signature	Security
S0	Public0	HashMsg0		
S1	Public1	HashMsg1	Sign1	Sec0
S2	Public2	HashMsg2	Sign2	Sec1
S3	Public3	HashMsg3	Sign3	Sec2

3 System Development

The prototype was built in C++, with the Crypto++ library for its multiple hash functions and arbitrary-precision arithmetic. The JSON(Javascript Object Notation) was chosen to implement the key-file format using the "JSON for Modern C++" library. All tests were done on a Macbook Pro Mid 2015 with an Intel Corei5 processor, 8GB RAM memory and MacOS 10.15 OS. The software was compiled on a GCC C++ compiler and the Crypto++ 8.2 library.

3.1 Database Design

A database was designed to storage the chains information. The complete list of tables and relationship of the basic database can be seen in the Entity-Relationship Diagram (ERD) in figure 5.

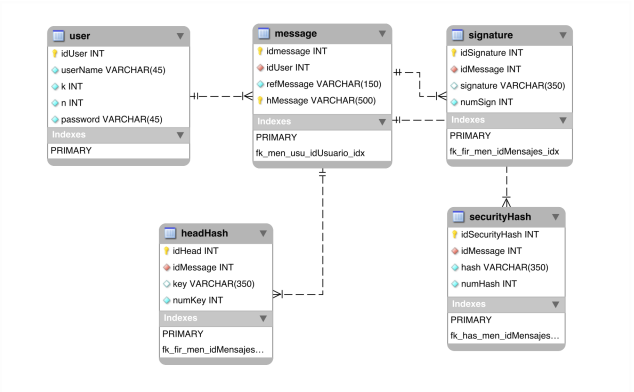


Fig. 5. Prototype of Database Base Entity Relationship diagram. The design allows the dynamic use of k per user separating the head-chain, security hash and signature hash in separated tables.

The database was designed with five tables that keep all public info implied in the single-chain management process. The **users** table keeps the users information and the n and k info. The **messages** table keeps the reference to the original message and it's hash. The **signature**, **security hash** and **publichash** tables keep the hashes generated for each purpose, storing k hashes per message.

The hash values are kept in the database as strings and then casted as arbitrary-precision integer (`BigInt`) types once recovered by the application for their further use. Keeping the n and k and the hashes separated in separated tables gives the system prototype dynamic security per user.

4 Results

One of the main challenges of this prototype was the time to compute $2^n - 1$ hashes in order to produce the public-head hashes provided $n = 256$ and $k = 16$, so 65536 hashes have to be computed to get the public Key. Four hash libraries were tested to improve the hash chaining time, the test results can be seen on table 2.

Table 2. Time to compute a hash-chain with length $2^{16} - 1$,$2^{20} - 1$ and $2^{22} - 1$.

Library	Time (ms)		
	2^{16}	2^{20}	2^{22}
DigestPP	591	8,936.61	155,627.10
PicoSHA	635	9,446.60	150,953.55
CryptoPP	610	10,019.22	153,996.90

The test reveals that little modifications in the variable that performs the conversion process using Crypto++ produce similar results to those obtained with dedicated libraries for hashing. The final version of the hash-chaining method was reduced 14 seconds in average.

We looked at many platforms to analyze the size of their blockchain in order to compare the storage requirement of three of the most popular ones [10,11,12,13]. The results of this comparison can be seen in Table 3.

Table 3. Size of common blockchain platforms and the transaction size average which is calculated from the total value of the blockchain size requirement at each platform.

Platform	Transactions	Blockchain Size (Gb)	Avg. Trans. Size(Kb)
Bitcoin	540,942,076	330.43	0.64
Ethereum	739,670,000	336.46	0.47
Monero	7,807,181	63.75	8.56

To get a comparison between the requirement of database storage we made an experiment inserting 1,000 messages in our single-chain implementation system and checked out the size growth of the database tables. Each hash tables grew in 2.7Mb in total, so the average storage for the 33 rows needed to store the message and it's signature is 8.57Kb.

Compared with the data shown in table 3, a single-chain transaction is 0.01Kb bigger than Monero, the crypto currency with the biggest transaction storage average. So further improvements for this field are still in order for future development.

5 Conclusions

The single-chain system relies on hash-chains, so it does not require the use of proof-of-work. Here, we introduced a modified design for the key-file system that reduces the storage requirement because it consists in a basic database structure. Our implementation proves that the single-chain system is completely functional over a regular computational environment reducing the time required to insert a transaction's register to the single chain.

The system can still be improved so that it would achieve a distributed database. The next development will focus in the use of threads to perform the signing process and possibly the use of hash functions written in assembly language for embedded and mobile phones.

References

1. S. Nakamoto *et al.*, "Bitcoin: A peer-to-peer electronic cash system," 2008.
2. V. Buterin *et al.*, "A next-generation smart contract and decentralized application platform," *white paper*, vol. 3, no. 37, 2014.
3. N. Van Saberhagen, "Cryptonote v 2.0," 2013.
4. S. S. Hazari and Q. H. Mahmoud, "A parallel proof of work to improve transaction speed and scalability in blockchain systems," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0916–0921, IEEE, 2019.
5. A. Kiayias, A. Russell, B. David, and R. Oliynykov, "Ouroboros: A provably secure proof-of-stake blockchain protocol," in *Annual International Cryptology Conference*, pp. 357–388, Springer, 2017.
6. I. Bentov, C. Lee, A. Mizrahi, and M. Rosenfeld, "Proof of activity: Extending bitcoin's proof of work via proof of stake [extended abstract] y," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 3, pp. 34–37, 2014.
7. J. J. Xu, "Are blockchains immune to all malicious attacks?," *Financial Innovation*, vol. 2, no. 1, p. 25, 2016.
8. C. M. Sciabarra, *Ayn Rand: The Russian Radical*. Penn State Press, 2013.
9. L. A. Lizama-Perez, "Digital signatures over hash-entangled chains," *SN Applied Sciences*, vol. 1, p. 1568, Nov 2019.
10. "Etherscan." https://etherscan.io. Accessed: 2020-06-21.
11. "Blockchair - monero." https://blockchair.com/monero. Accessed: 2020-06-21.
12. "Bitinfocharts - ethereum." https://bitinfocharts.com/ethereum/. Accessed: 2020-06-21.
13. "Blockchain.com." https://www.blockchain.com/charts/blocks-size. Accessed: 2020-06-21.

9

Studying the association of users' depressed moods and posting on Instagram

IVAN ALEJANDRO ENCINAS MONROY, LUIS HÉCTOR SÁNCHEZ GÁMEZ, JESSICA BELTRAN AND LUIS A. CASTRO

Studying the association of users' depressed moods and posting on Instagram

Iván Alejandro Encinas Monroy[1], Luis Héctor Sánchez Gámez[2], Jessica Beltran[2], Luis A. Castro[1]

[1] Sonora Institute of Technology (ITSON), Ciudad Obregon, Mexico
[2] CITEDI-IPN, Tijuana, Mexico
ivanalexx@msn.com, lh.sg26@gmail.com, jbeltranm@ipn.mx, luis.castro@acm.org

Abstract. An increasing number of young adults are reporting feelings of loneliness or depressed moods. Mental health professionals typically use psychometric inventories, self-report and interviews with patients for diagnosis and follow-up. In many cases, however, patients can hide, overlook, or minimize feelings or occurrences that happen off the clinic, which can nonetheless be of interest for professionals. It is highly likely that patients can show these types of feelings on online social networks through images or words. In this work, we study the association of feelings related to users' mental health and their posts on Instagram. We recruited 50 young adults and analyzed their posts over the course of 32 days. We asked our participants to answer the Beck's depression inventory (BDI-II) and the Positive and Negative Affect Schedule, extended form (PANAS-X). Preliminary results suggest that keeping track of the mental state at the moment of posting on Instagram can be useful for inferring depressed moods in users. The results of this work can be used by mental health professionals to learn about the patient's well-being, which could otherwise be difficult to obtain through conventional face-to-face encounters with patients.

Keywords: Instagram, content analysis, mental health.

1 Introduction

The prevalence of major depressive disorders has increased in the last few years. All around the world, 300 million people suffer from depression, making it the leading cause of disability, costing about 1 trillion dollars each year in the U.S. [1]. Although seeking help has become increasingly common [2], there is still a need for accessible early mental health assessment. In this regard, online social networking sites are typically used for creating and maintaining interpersonal relationships, managing the person's own presentation, or allowing users to express their own vision of the world [3]. In particular, Instagram is mainly about promoting yourself and your daily activities, which makes it interesting to study mental health patterns of their users [4].

Typically, mental health professionals use psychometric inventories, self-report and interviews with patients for diagnosis and follow-up. In many cases, patients can

hide, overlook, or minimize feelings or occurrences that happen off the clinic, which can nonetheless be of interest for professionals. It is highly likely that patients show these types of feelings on online social networks through images or words [5]. Research shows promising new ways to do this through analysis of data shared in social media. For example, the posts' own content may present information for pattern detection. Tags (or *hashtags*) provide additional information about the topic of the post and can be helpful to quickly identify posts related to depression [6]. In other cases, the mere use of language and hashtags show relationships with the inner motivations and ideas of an user [7]. Alternate methods for detecting depressive mood have been proposed. For instance, based on certain patterns such as rating images with lower levels of valence can identify when a user had been depressed before [8].

There is recent work on prediction methods of different mental disorders through information shared online [9]. In this regard, these works use historical information labeled by the researchers themselves according to predefined emotion categories to train computational models. This approach is used because it allows a quantitative approach to the problem. However, you may run into the problem where the information is not appropriately labeled, or the perceived sentiment of the post cannot be entirely trusted. No case has been found in the reviewed literature where data are labeled by the user, and, at the same time, users' data are available for a quantitative analysis. In this work, we study the relationship of users' reported feelings and their mental health, with the purpose of improving early screening and detection in adults.

2 Methods

We recruited 50 subjects using convenience sampling method in northwest Mexico, from which 35 subjects (13 male) completed 32 days of data collection. Participants' average age was 23.51 years (SD = 3.36), ranging from 19 to 40 years old. We recruited participants through social networks, and other electronic media (e.g., text messages). Our participants were healthy without prior record of a mental illness. All participants signed an informed consent. There was no monetary incentive for them.

Those who agreed to participate answered a demographic questionnaire. We used the Beck Depression Inventory (BDI-II) in Spanish [10] as ground truth. The BDI-II is a questionnaire about people's depressive symptoms. The BD-II can categorize subjects in four classes: minimal range = 0–13, mild depression = 14–19, moderate depression = 20–28, and severe depression = 29–63. According to our participants' answers they were categorized for analysis: minimal range = 17 (9 male), age=23.88 (SD=1.32); mild = 5 (2 male), age=26.6 (SD=7.54); moderate = 7 (1 male), age=22.28 (SD=2.36); severe = 6 (1 male), age=21.6 (SD=1.21).

We also used the Expanded Positive and Negative Affect Schedule (PANAS-X) to measure a person's positive and negative affect to assess our participants' feelings after each post. General Positive/Negative Affect scores represent the degree of positive (e.g., enthusiastic, excited, inspired) and negative emotions (e.g., afraid, scared, nervous, jittery), each of which are composed by ten emotion items in the PANAS-X.

Our participants were instructed to use their Instagram account as they would normally do. We asked them to answer the BDI-II at the beginning of the study, and the PANAS-X each time they posted an entry on Instagram through an app that worked through the Instagram API.

3 Preliminary Results

We obtained a total of 325 entries. In average, each user made 9.28 posts (SD: 4.94) through the 32 days of the study. Taking into account reported affect scores, preliminary results show that there are differences in the way users with depressive moods feel when posting in comparison with the rest of the users (see Figure 1).

For the General Positive Affect (GPA), users with depressive moods (i.e., mild, moderate and severe) had lower scores (GPA=20.98, SD=8.84) when compared with users with minimal depressive moods (GPA=25.00, SD=9.22), which was statistically significant (t=4.01, df=312, P<0.0001). As for the General Negative Affect (GNA), our participants with depressive moods had a significantly higher score (GNA=17.74, SD=6.51) when compared with users with minimal depressive moods (GNA=14.76, SD=8.09), which was statistically significant (t=-3.67, df= 321, P<0.0001).

Also, participants with severe depression had a mean GPA of 18.64 (SD=7.30), while the rest of the classes had an average GPA score of 23.72 (SD=9.35) (t=-4.49, df= 97, P<0.0001). On the other hand, participants with severe depression also reported higher scores of the GNA=21.62 (SD=8.41) when compared with the average GNA=15.26 (SD=6.87) of the rest of participants (t=5.3, df=71, P<0.0001). These findings concurs with previous findings that indicate diminished positive emotions for the severe depressed [8]. However, our finding differs in the case of the GNA score, which were also lower than expected. Our preliminary results indicate that the self-labeling at the time of posting is related with their depressive moods. These labeling can be useful to further analyze the content of the post (e.g., text, image).

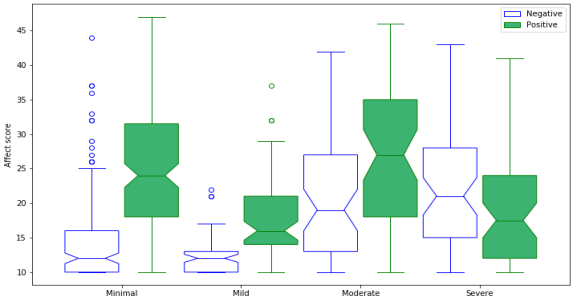


Fig. 1. Positive and negative affect scores across our participants' BD-II score

4 Conclusions

We found a significant difference between what people with depressive moods feel at the time of posting when compared to participants with minimal depressive moods. This suggests that computational resources could be devoted to early identifying people with depression, or patients' progress during treatment. As future work, we will analyze the content of posts, i.e., text, tags, images, and the posting behavior (e.g., posting frequency and postings times). We will use computer vision algorithms to identify objects in the images that can be related with depression, such as the number of faces and potentially extract emotions from images. Also, we will preprocess the data from image and text analysis to feed machine learning algorithms aimed to predict scores of the BDI-II and the self-reported emotion through PANAS-X.

References

1. W. H. Organization, "World Mental Health Day 2017," 2017. [Online]. Available: https://www.who.int/mental_health/world-mental-health-day/2017/en/. [Accessed: 23-Jun-2020].
2. J. Hunt and D. Eisenberg, "Mental Health Problems and Help-Seeking Behavior Among College Students," *J. Adolesc. Heal.*, vol. 46, no. 1, pp. 3–10, 2010, doi: 10.1016/j.jadohealth.2009.08.008.
3. N. Van House and M. Davis, "The social life of cameraphone images," *Image Capture Shar. New Soc.*, pp. 1–10, 2005.
4. L. Manikonda, V. Vamsikrishna, and M. S. Kambhampati, "Tweeting the mind and Instagramming the heart: Exploring differentiated content sharing on social media," *Proc. 10th Int. Conf. Web Soc. Media, ICWSM 2016*, no. Icwsm, pp. 639–642, 2016.
5. M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar, "Discovering shifts to suicidal ideation from mental health content in social media," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2098–2110, 2016, doi: 10.1145/2858036.2858207.
6. N. Andalibi, P. Ozturk, and A. Forte, "Depression-related imagery on Instagram," *Proc. ACM Conf. Comput. Support. Coop. Work. CSCW*, vol. 2015-Janua, pp. 231–234, 2015, doi: 10.1145/2685553.2699014.
7. C. Lee and D. Chau, "Language as pride, love, and hate: Archiving emotions through multilingual Instagram hashtags," *Discourse, Context Media*, vol. 22, no. June, pp. 21–29, 2018, doi: 10.1016/j.dcm.2017.06.002.
8. D. M. Sloan, M. E. Strauss, and K. L. Wisner, "Diminished response to pleasant stimuli by depressed women," *J. Abnorm. Psychol.*, vol. 110, no. 3, pp. 488–493, 2001, doi: 10.1037/0021-843X.110.3.488.
9. A. G. Reece and C. M. Danforth, "Instagram photos reveal predictive markers of depression," *EPJ Data Sci.*, vol. 6, no. 1, 2017, doi: 10.1140/epjds/s13688-017-0110-z.
10. J. Sanz, A. L. Perdigón, and C. Vázquez, "Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propiedades psicomé tricas en población general. The spanish adaptation of Beck ' s Depression Inventory-II (BDI-II):," *Clin. y Salud*, vol. 14, no. 3, pp. 249–280, 2003.

10

Buscando vulnerabilidades en aplicaciones de Android relacionadas al COVID 19

Diana Serrano Delgado, Fernando Velázquez López, Rafael Llamas Contreras y Hugo Gonzalez

Buscando vulnerabilidades en aplicaciones de Android relacionadas al COVID-19

Diana Serrano-Delgado[1], Fernando Velázquez-López[1], Hugo Gonzalez[1][0000−0002−7322−4019], and Rafael Llamas-Contreras[1]

Laboratorio de Innovación e Investigación en Ciberseguridad e Inteligencia Artificial, Universidad Politécnica de San Luis Potosí. San Luis Potosí, México
{160087,160174,hugo.gonzalez,rafael.llamas}@upslp.edu.mx

Resumen Derivado de la pandemia originada por el brote del virus SARS-CoV2, coloquialmente llamado Coronavirus o COVID-19, surgieron diversas aplicaciones móviles con la finalidad de pronosticar y brindar apoyo a la ciudadanía mediante formularios orientados a la detección de la enfermedad. En México, las aplicaciones de este tipo son desarrolladas bajo el liderato de instituciones científicas y médicas. Todas las aplicaciones analizadas como parte experimental de este trabajo, cuentan con el apoyo de instituciones médicas. De estas aplicaciones, sólo una de ellas fue desarrollada en estados unidos, el resto fueron desarrolladas en México. El común denominador de estas aplicaciones es la vulnerabilidad a ataques en red donde es posible interceptar credenciales de usuario, incluyendo contraseñas mediante el uso de herramientas como Burp Suite, asimismo cada una de las respuestas que el usuario señala en el formulario son visibles durante su envío a los servidores de cada aplicación. Es imporante dar a conocer estas vulnerabilidades encontradas para mitigar el impacto de estos fallos.

Keywords: COVID-19 · Ciberseguridad · Aplicaciones Móviles · Android

1 Introducción

Ante la emergencia sanitaria derivada del COVID-19, la cual impactó a cientos de países, surgieron múltiples herramientas digitales que permitieron generar estadísticas a partir de formularios de datos. Dichas herramientas fueron distribuidas como aplicaciones individuales y otras tantas como complementos de sitios mundialmente famosos. Particularmente en México, muchas de las aplicaciones distribuidas a lo largo del país sirvieron como fuente de cifras oficiales para medir el índice de contagios y decesos. Derivado del uso frecuente de estas herramientas, se plantearon incógnitas referentes a la seguridad de la aplicación para transmitir o almacenar los datos del usuario al interactuar con estas, por ello se decidió profundizar en la seguridad de algunas de las aplicaciones distribuidas en México. La característica principal de estas aplicaciones es que todas se encargan de pronosticar el padecimiento del COVID-19 tomando como fuente de

información las bases de datos de organismos de salud y la introducción de datos por parte del usuario en los formularios.

Para la realización de esta investigación se hace el uso de herramientas de análisis de tráfico en dispositivos Android. La herramienta seleccionada es Burp Suite y para realizar la simulación de un dispositivo móvil se usará la herramienta de emulación NOX. El entorno operativo donde se llevarán a cabo las pruebas será en Microsoft Windows 10. Se configura Burp Suite para que funcione en conjunto con el emulador NoxPlayer, y se pueda analizar el tráfico de red de las peticiones que se hagan en el navegador web del teléfono y las aplicaciones seleccionadas.

2 Trabajo relacionado

El análisis de la seguridad de aplicaciones para Android es un tema que se ha investigado anteriormente, usualmente enfocandose en un tipo específico de aplicaciones. Las aplicaciones relacionadas con banca móvil han sido de las más analizadas en busca de debilidades en el transporte de la información, por ejemplo los trabajos [14, 2, 4]. También relacionado con banca móvil, pero enfocado en el análisis de las carteras electrónicas esta el trabajo [7]. Otros trabajos se enfocan en los problemas de seguridad de las comunicaciones de las aplicaciones para Android [9, 17].

Por otro lado, recientemente se han desarrollado trabajos relacionados con las posibles brechas de privacidad relacionadas con las aplicaciones que rastrean los contactos de las personas diagnosticadas con coronavirus. Con el objetivo principal de protección de la privacidad se desarrollo DP3T, rastreo de la proximidad con preservación de la privacidad descentralizado [12], también se han desarrollado aplicaciones en Estados Unidos pensadas en la privacidad de los usuarios, por ejemplo [16], y también [10] que incluye no solo la aplicación sino también un protocolo abierto. La coalición TCN busca promover las mejores prácticas para este tipo de aplicaciones [15].

En este trabajo nos enfocamos en la seguridad de las aplicaciones, dejando fuera del alcance del mismo las afecciones a la privacidad.

3 Objetivo de la investigación

Esta investigación busca exponer vulnerabilidades de transmisión de datos en aplicaciones móviles para Android, que estén relacionadas con el COVID-19. Al dar a conocer estas vulnerabilidades, tratamos de evitar que el problema escale a un nivel mayor.

3.1 Modelado de amenazas

Para este trabajo la amenaza principal es la intercepción y el robo de información. Esto se puede generar mientras la información está en reposo o mientras está en tránsito. Durante los experimentos limitamos nuestras interacciones

con el robo de información en tránsito para no afectar los servidores donde es almacenada.

4 Metodología

La metodología usada se puede resumir como sigue:

Diseño El diseño de esta investigación es experimental, ya que se harán pruebas con aplicaciones móviles reales.

Enfoque El enfoque usado en esta investigación es un modelo mixto (cualitativo y cuantitativo), ya que se esperan resultados con variantes y puntos de vistas únicos.

Técnicas La técnica elegida será la observación en un entorno simulado y controlado.

4.1 Procedimientos

1. Se identificaron las principales variables de las que se quería obtener información. Para esto se hizo un análisis de las aplicaciones disponibles para dispositivos con el sistema operativo Android y se seleccionaron 4 de ellas. Según funcionalidad, número de descargas, permisos que requieran y procedencia de desarrolladores.
2. Se eligió una herramienta para simular el entorno de un teléfono móvil de Android (el emulador NoxPlayer [8]).
3. Se eligió una herramienta para analizar el tráfico en dispositivos móviles Android (la herramienta de Burp Suite.[13])
4. Se configuró el proxy del dispositivo simulado para que se pudiera analizar el tráfico que generan las aplicaciones, adaptando el dispositivo con la configuración de Burp Suite.
5. Se realizó un análisis de paquetes, con el objetivo de encontrar información que el usuario ingresaba en los formularios.

4.2 Muestras

El universo de la investigación se conformó con cuatro aplicaciones móviles desarrolladas para Android. Todas estas enfocadas en el área de la salud y en especial con el COVID-19.

1. CDS (Centinela de la salud)
2. Preven App
3. CovApp
4. Ada Health

Se decidió escoger las anteriores aplicaciones por algunos aspectos en común; CovApp muestra una gran similitud a la aplicación de CDS, ya que comparten la misma funcionalidad de encuesta y se realizaron de la mano con instituciones

importantes del estado de San Luis Potosí y Querétaro respectivamente. Preven APP muestra una funcionalidad distinta pero que se asocia directamente con herramientas de apoyo a la ciudadanía en caso de presentarse síntomas. Esta aplicación solicita un inicio de sesión personal. Por último, ADA Health cuenta con un mayor número de descargas, además de que requiere permisos poco comúnes de uso en el dispositivo, tiene más funcionalidades y es notorio que se ha trabajado por más tiempo en su desarrollo.

5 Hallazgos

Las versiones más actualizadas de Android, posteriores a 6.1, incluyen protección avanzada contra la instalación de certificados personalizados, los cuales son necesarios para ciertos ataques de hombre en medio a las aplicaciones. Las aplicaciones analizadas son compatibles con versiones anteriores, permitiendonos instalarlas en un dispositivo emulado con versión Android 6.0, la cual todavía permite instalar sin problema el certificado personalizado de Burp Suite para realizar el ataque.

Una vez comprendido el funcionamiento en conjunto del emulador con Burp Suite, la intercepción del tráfico es sencilla. Cabe mencionar que es necesario saber interpretar los paquetes que se generan cuando se intercepta el tráfico, por lo que una persona sin conocimientos técnicos de informática, no podrá hacer un uso malintencionado de las aplicaciones.

Los permisos más comunes que exigían las aplicaciones fueron de ubicación, acceso completo a la red, cámara, teléfono y almacenamiento. Estos permisos tienen sentido por el tipo de aplicaciones, que mapean geográficamente a las personas para obtener la situación actual de salud por persona y por región donde se utilizan.

6 Resultados

Es importante señalar las diferencias entre las aplicaciones de una manera más precisa, como se aprecia en la Tabla 1. Los resultados obtenidos por aplicación se exponen a continuación, destacando los problemas principales detectados.

CDS En esta aplicación, desarrollada por la Universidad Autónoma de San Luis Potosí en conjunto con el IPICYT [5], se puede descargar en la tienda de aplicaciones de Google [6]. En su versión analizada presenta una filtración importante de credenciales, las cuales se cree que corresponden a una cuenta de correo electrónico creada para uso específico de la aplicación. Si bien, puede que esta cuenta de correo no tenga algúna repercusión sobre la filtración de otras credenciales de la organización, es una mala práctica de desarrollo el dejar visibles cuentas de correo y contraseñas pues podría repercutir en los datos que el usuario proporciona al hacer uso de esta App. Esta aplicación no utiliza transmisión segura (https) ni alguna protección extra al transmitir información por la red, por lo que interceptar la información personal es viable sin configuraciones extra.

Tabla 1. Tabla comparativa de aplicaciones. Fuente: Elaboración propia.

Nombre	Actualización	Descargas	Desarrollador	Permisos	Funcionalidad	Versión Android	Vulnerabilidades
CDS 1.2.2	2 May. 2020	Más de 500	Servicios de Salud SLP, IPICYT FM-UASLP	Ubicación, almacenamiento, acceso completo a la red, conexiones de red, conexiones de WIFI	Brinda una autoevaluación inmediata sobre la enfermedad COVID-19	4.4	Comunicación en texto plano, credenciales en el código, información interceptable
COVAPP 1.2	7 Abr. 2020	Más de 5,000	CQ CIAS Querétaro	Acceso completo a la red	Aplicación para el autodiagnóstico e COVID -19	6.0	Información interceptable
PrevenApp 1.0	2 May. 2020	Más de 1,000	Partnergrammer Chihuahua	Cámara, ubicación, teléfono,almacenamiento, acceso completo a la red, evitar que el teléfono entre en modo de inactividad, API de referencia de Play, Ejecución al inicio, recibir datos de internet	Analiza las ubicaciones en las que el usuario a estado, y las compara con las de otros, si encuentra que estuviste expuesto a el virus manda avisos.	4.2	Credenciales interceptables
ADA 3.5.0	12 May. 2020	Más de 5,000,000	Ada Healt	Ubicación, almacenamiento, acceso completo a la red, API de referencia de Play, Utilizar huellas digitales	Revisar síntomas y descubrir la causa de enfermedad	4.3	Credenciales interceptables.

Esta aplicación genera un identificador único por instalación, que es la forma de dar seguimiento a los usuarios geolocalizados.

Preven App Una aplicación desarrollada para uso especial en el estado de Chihuahua, se puede descargar de la tienda oficial de Google [11]. En esta aplicación el inicio de sesión presenta problemas, al crear una cuenta de usuario e iniciar sesión las credenciales viajan al servidor en forma de texto plano y no cuenta con ningún método de cifrado para proteger la integridad de las contraseñas. En el escenario supuesto donde el usuario utilice una misma contraseña para todas sus cuentas, este problema puede considerarse grave, ya que expone a los usuarios a perder información o ser víctimas de suplantación de identidad. Aunque el problema de reuso de credenciales no es directo de la aplicación, puede convertirse en un vector de ataque.

CovAPP Aplicación desarrollada para el estado de Querétaro, cuya funcionalidad es reportar la situación de contagios de Covid-19. Se puede descargar de la tienda oficial de Google [3]. Esta aplicación no expone credenciales ya que no tienen la posibilidad de inicio de sesión, pero tampoco tiene credenciales para acceder al servicio web donde se almacena y procesa la información. Una debilidad, sin embargo, es que los datos del cuestionario contestados por el usuario son visibles mediante la intercepción de tráfico.Esta aplicación guarda mucha similitud con la aplicación CDS, pues no guarda la información personal identificable del usuario, y realiza un diagnóstico a través de un cuestionario para que el mismo usuario pueda seguir las indicaciones adecuadas en caso de resultar con la enfermedad. Esta aplicación genera un identificar único por dispositivo para dar seguimiento.

Ada Health Esta aplicación no solo se centra en la detección del coronavirus, sino que también detecta otras enfermedades basado en los síntomas del usuario, además que proporciona consejos útiles dependiendo del padecimiento del usuario. Se puede descargar desde la tienda oficial de Google [1]. Las debilidades encontradas son que expone los datos del usuario al inicio de sesión, de igual manera que Preven App. Lo relevante de esta App es que lleva años en el mercado, siendo descargada por más de 5 millones de usuarios, exponiendolos al filtrado de sus credenciales.

7 Conclusiones y recomendaciones

Analizando las aplicaciones mencionadas en este trabajo, se puede apreciar que las vulnerabilidades en los dispositivos móviles por medio de aplicaciones son comunes, por lo que el campo de investigación en esta área es bastante extenso. Algunas veces, los desarrolladores no toman en cuenta las mejores prácticas en seguridad y privacidad bajo temas emergentes, como el covid-19. En México no está difundida ampliamente la concientización en temas de seguridad informática para usuarios ni para desarrolladores, resultando en riesgos para los datos y en aplicaciones vulnerables.

Con la pandemia actual, se han empezado a desarrollar algunas aplicaciones que funcionan como aplicaciones de rastreo de contacto, cuyas funcionalidades

pretenden no ser lo suficientemente confiables para los propósitos con los que se vende la app. El objetivo de estas aplicaciones en el contexto de esta pandemia es rastrear personas que sean positivas al Covid-19 para que se tenga un registro de gente que podría estar infectada y que la gente llegue a tomar precauciones o prevenir.

Lo anterior resulta ser una idea interesante, pero a veces por reaccionar rápido no se toman en cuenta medidas de privacidad y seguridad, sobre todo si los desarrolladores están enfocados en la parte médica, sin tener asesores en materia de ciberseguridad. Por esto, las vulnerabilidades que representa dichas aplicaciones para dispositivos móviles pueden ser variadas. Por ejemplo, la aplicación de PrevenAPP tiene como funcionalidad principal el rastreo, y los permisos que ésta exige son mucho más extensos que las aplicaciones de encuesta o autodiagnóstico. Cabe señalar que las especificaciones de las aplicaciones sí hacen énfasis en que los datos no serán difundidos, pero puede que esten expuestos al robo de información. Aun así, cuando se realizó el análisis de tráfico no se detectaron datos de geolocalización en esta aplicación, como en la aplicación de CDS, pero se logro tener acceso a algunas credenciales.

Con el análisis de las demás aplicaciones se concluye que es comprensible que de momento no se tenga el suficiente cuidado al desarrollarlas, por la falta de un asesor en ciberseguridad, y que la filtración de las credenciales y otros datos introducidos por el usuario son debido a que las aplicaciones se desarrollaron de manera expedita. Sin embargo, las aplicaciones fueron desarrollados por organismos gubernamentales y académicos, por lo que estas fallas no deberían estar justificadas. Es necesario que los organismos encargados del desarrollo de estas aplicaciones tomen medidas para evitar que los datos de los usuarios se expongan y que no se pase por alto el tema de la seguridad en el desarrollo de las mismas.

7.1 Recomendaciones

Se proponen una serie básica de recomendaciones generales para mejorar la seguridad e impactar en la privacidad de estas aplicaciones.

- Seguir las mejores prácticas y recomendaciones para desarrollo de aplicaciones móviles.
- Integrar a un experto en ciberseguridad en el proyecto.
- Actualizar las aplicaciones con los marcos de trabajo y recomendaciones específicas para este tipo de aplicaciones.

También se realizan recomendaciones técnicas específicas:

- Aunque limita el público objetivo, aumentar la versión compatible de Android que ya integra protección contra certificados personales, o implementar un método de cifrado para transmitir la información.
- Utilizar métodos de transmisión seguros (https).
- Utilizar métodos para fortalecer el uso de certificados, como el "cert pinning".
- Proteger el código contra análisis, utilizando métodos de obfuscación de código.

Finalmente, existen personas que poseen herramientas y el suficiente conocimiento para realizar el análisis de posibles vulnerabilidades, y a través una campaña de detección de errores se pueden coordinar los reportes de vulnerabilidades y solucionarlas para las siguientes versiones.

Referencias

1. Ada Health: Ada – your health companion (2020), https://play.google.com/store/apps/details?id=com.ada.app
2. Chothia, T., Garcia, F.D., Heppel, C., Stone, C.M.: Why banker bob (still) can't get tls right: A security analysis of tls in leading uk banking apps. In: International Conference on Financial Cryptography and Data Security. pp. 579–597. Springer (2017)
3. CQ CIAS Querétaro: Aplicación para el autodiagnóstico de covid-19, geq (2020), https://play.google.com/store/apps/details?id=mx.cqcias.covapp
4. Darvish, H., Husain, M.: Security analysis of mobile money applications on android. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 3072–3078. IEEE (2018)
5. IPICYT: Aplicación indica tu nivel de riesgo por covid-19 (2020),
6. IPICYT: Servicios de salud slp, ipicyt, fm-uaslp (2020), https://play.google.com/store/apps/developer?id=Servicios+de+Salud+SLP,+IPICYT,+FM-UASLP
7. Kaur, R., Li, Y., Iqbal, J., Gonzalez, H., Stakhanova, N.: A security assessment of hce-nfc enabled e-wallet banking android apps. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). vol. 2, pp. 492–497. IEEE (2018)
8. Nox (HongKong) Limited: Noxplayer, the simpler the better (2020), https://www.bignox.com/
9. Onwuzurike, L., De Cristofaro, E.: Danger is my middle name: experimenting with ssl vulnerabilities in android apps. In: Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks. pp. 1–6 (2015)
10. PACT Team: Pact: Private automated contact tracing (2020), https://pact.mit.edu/
11. Partnergrammer: Prevenapp (chihuahua) (2020), https://play.google.com/store/apps/details?id=com.partnergrammer.prevenapp
12. Payer, M., Barman, L.: Decentralized privacy-preserving proximity tracing (2020), https://github.com/DP-3T
13. PortSwigger: The burp suite family (2020), https://portswigger.net/burp
14. Reaves, B., Bowers, J., Scaife, N., Bates, A., Bhartiya, A., Traynor, P., Butler, K.R.: Mo (bile) money, mo (bile) problems: Analysis of branchless banking applications. ACM Transactions on Privacy and Security (TOPS) **20**(3), 1–31 (2017)
15. TCN Coalition members: Tcn coalition (2020), https://tcn-coalition.org/
16. University of Washington: Together we will help (2020), https://covidsafe.cs.washington.edu/
17. Wang, Y., Xu, G., Liu, X., Mao, W., Si, C., Pedrycz, W., Wang, W.: Identifying vulnerabilities of ssl/tls certificate verification in android apps with static and dynamic analysis. Journal of Systems and Software **167**, 110609 (2020). https://doi.org/https://doi.org/10.1016/j.jss.2020.110609, http://www.sciencedirect.com/science/article/pii/S016412122030087X

Parte III

Aplicaciones de Cómputo Suave y Analítica de Datos

Introducción

Los avances en algoritmos de Cómputo Suave y técnicas de Analítica de Datos han permitido que se desarrollen sistemas computacionales capaces de mejorar la toma de decisiones y optimizar la información. Esta característica de los sistemas permite que las decisiones sean más estratégicas y de menor riesgo. Esta sección está dedicada a introducir nuevas tendencias, enfoques, métodos, así como aplicaciones innovadoras de cómputo suave y análisis de datos.

En el primer trabajo, "Una Propuesta Metodológica para el Proceso de Toma de Decisiones aplicada en Centros de Soporte a la Toma de Decisiones", los autores diseñan una metodología para hacer más eficiente la toma de decisiones.

En el segundo, "Spatial network analysis of Tijuana's metropolitan transport network", se modela, mediante un grafo acíclico dirigido (DAG), el sistema metropolitano de transporte de la ciudad de Tijuana.

En, "Generación automática de controladores difusos", los autores proponen una herramienta computacional que apoya en el diseño de Controladores Lógicos Difusos.

En el cuarto, "Statistical comparison of models: Fuzzy Inference System vs Fuzzy Least Square", los autores presentan un análisis comparativo entre el Sistema de inferencia difusa y Sistema Difuso de mínimos cuadrados para resolver un problema de máquinas.

En, "Optimization of the GTAW manual welding process, considering random covariates", los autores aplican la técnica NSGA-II para resolver un problema de optimización durante el proceso de soldadura GTAW.

En el sexto trabajo, "Turisteando Ando: una App con Kernel Evolutivo", utilizan un algoritmo genético estacionario para diseñar itinerarios turísticos personalizados para una aplicación móvil.

Y, por último, en "PSO in Ridge Regression to eliminate multicollinearity", los autores aplican la técnica PSO para reducir los efectos negativos de la correlación en problemas multivariable.

Anabel Martínez Vargas

11

Una Propuesta Metodológica para el Proceso de Toma de Decisiones aplicada en Centros de Soporte a la Toma de Decisiones

ALEJANDRO FLORES BENITEZ, JOSÉ MARTÍN MOLINA ESPINOSA Y MANUEL ALVAREZ MADRIGAL

Una Propuesta Metodológica para el Proceso de Toma de Decisiones aplicada en Centros de Soporte a la Toma de Decisiones

Alejandro Flores Benitez[1], José Martín Molina Espinosa[2], and Manuel Alvarez Madrigal[3]

[1] Instituto Tecnológico y de Estudios Superiores de Monterrey, CDMX, México.
a01655736@itesm.mx
[2] Instituto Tecnológico y de Estudios Superiores de Monterrey, CDMX, México.
jose.molina@tec.mx
[3] Instituto Tecnológico y de Estudios Superiores de Monterrey, CDMX, México.
mmadrigal@tec.mx

Resumen La complejidad de los sistemas hace que la toma de decisiones se vuelva un proceso duro, largo y costoso tanto en tiempo como en recursos. El uso estructurado de técnicas que involucren múltiples criterios de decisiones puede mejorar al *proceso de toma de decisiones* (PTD). La constante evolución de los problemas, incremento en el número de variables, información ambigua o nula, complica el PTD. Para poder simplificar dicho proceso se recurre al uso de un espacio físico que facilite y optimice dicho procedimiento, tal espacio incrementa en gran medida el porcentaje de éxito, la creación y uso de *"Centros de Soporte a la Toma de Decisiones"* (CSTD) como una herramienta de apoyo crea la necesidad de adaptar el esquema clásico PTD hacia uno que explote todas las capacidades y funcionalidades que ofrece un CSTD.

Keywords: Centros de Soporte a la Toma de Decisiones · Proceso de Toma de Decisiones.

1. Introducción

Herbert A. Simon, nos dice "la mayoría de las personas son solo parcialmente racionales, y de hecho son emocionales o irracionales en la parte restante de sus acciones" [3]. La teoría clásica y neoclásica nos dice que el objetivo principal de la toma de decisiones es ser racional al recopilar primero toda la información relevante sobre los temas que se investigan. El siguiente paso es generar todas las alternativas posibles, examinar las consecuencias de esas alternativas y finalmente elegir la alternativa óptima [6]. Jean-Louis Le Moige define **decidir** como identificar y resolver los problemas que se le presentan a toda organización [8] y George P. Huber nos habla de que existirá un **problema** cuando hay diferencia entre la situación real y la situación deseada [5]. Entonces, una decisión es la respuesta adecuada de un ser ante una situación que requiere una acción [10]. Otros autores como Jay Forrester entienden la toma de decisiones como "el

proceso de transformación de la información en acción", en la cual tomar una decisión supone escoger la mejor alternativa de entre las posibles y para ello se necesita información sobre cada una de estas alternativas y sus consecuencias respecto al objetivo [4], cada decisión es el resultado de un proceso dinámico.

El uso de herramientas para el apoyo en el PTD no es nuevo, desde hace varios años se conocen sistemas computacionales como: *Decision Support System* (DSS), *Group Decision Support Systems* (GDSS), *Executive Information System* (EIS), *Expert Decision Support Systems* (EDSS). Por otro lado, este tipo de sistemas de software son muchas veces implementados en espacios físicos, bajo este esquema encontramos las Salas de Control, los *War Room* y los Centros de Soporte a la Toma de Decisiones, los cuales comparten características similares, cabe destacar que estos espacios proporcionan una alta capacidad de despliegue visual que permite la representación de una gran cantidad de información de un solo golpe de vista, así mismo permite que los asistentes puedan tener interacción con los datos y así poder apoyar hacia una toma de decisiones en común.

El uso de este tipo de espacios toma sentido en el PTD ya que controlar las condiciones que rodean a dicho proceso es una parte crucial para el aumento de probabilidades de éxito, *"los tomadores de decisiones no pueden ser racionales a menos que el tomador de decisiones tenga un control perfecto sobre los factores ambientales y sus capacidades mentales."* [11].

2. Trabajo Relacionado

El concepto principal detrás de los CSTD nace a partir del proyecto conocido como *"Decision Theater"* (DT) [12], que fue utilizado por primera vez para designar un nuevo enfoque de enseñanza en la toma de decisiones de marketing. A partir de entonces varios CSTD han sido creados en universidades y centros de investigación, entre los que destacan, los de Arizona State University [26], McCain Institute [14], University of Alaska Fairbanks [25], Newcastle University [16], Global Climate Forum [15], Project: SMARTCITY Decision Theater [23], World Resources SIM Center [17], Tecnológico Nacional de México Campus Laguna, Laboratorio Nacional de Políticas Públicas [21], Tecnológico de Monterrey, en los campus: Monterrey [22], Ciudad de México [24] y sede Mixcoac [20], por mencionar algunos.

El proyecto de investigación del Laboratorio binacional para la gestión inteligente de la sostenibilidad energética y la educación tecnológica [18], define un CSTD como:

> *"Un Centro para la Toma de Decisiones es una herramienta tecnológica de visualización de datos para dar soporte a la Toma de Decisiones a partir de la ejecución y visualización en tiempo real de modelos de procesos complejos como los sociales, económicos, políticos y de desarrollo sostenible"* [24].

Un CSTD está compuesto por un conjunto de elementos de hardware y de software que operan en sinergia otorgando un espacio con un esquema de visualización inmersiva a un conjunto de tomadores de decisiones, para que sea capaz

de adquirir y analizar información relevante acerca de un problema complejo, recibiendo apoyo en el PTD con base en evidencias. Dicha sala aloja un despliegue de gran formato, a través de pantallas o proyecciones, las cuales se encuentran distribuidas en un semi-círculo, formando un ambiente de visualización donde es posible presentar tableros de control gráficos para la toma de decisiones. En la Fig.(1), podemos observar un CSTD.

Figura 1. Centro de Soporte para la Toma de Decisiones, ITESM, Campus Ciudad de México. Fuente: Elaboración propia

El Centro de Toma de Decisiones del ITESM, Campus Ciudad de México, cuenta con 7 pantallas de 65 pulgadas con resolución UHD (3,840×2,160), en una disposición de semi-círculo la cual otorga una experiencia inmersiva a los asistentes presenciales, cabe mencionar que el centro cuenta con equipo multimedia para videoconferencia, de esta forma el número de asistentes a las sesiones crece a medida de las necesidades del problema a abordar. Con respecto al software utilizado en este centro para realizar el proceso de toma de decisiones en sus diferentes etapas podemos mencionar programas como: VENSIM, ARENA, Microsoft Office, PowerShow, TeamViewer, Tableau, R/RStudio, MATLAB/SIMULINK, iThink/Stella.

En un CSTD se ejecuta un PTD gracias a la plataforma de hardware y software que lo componen. Pero además, el PTD debe de seguir un conjunto estructurado de pasos que aseguren la obtención de un consenso o una toma de decisiones que dé respuesta al problema en cuestión. Si bien los CSTD antes mencionados vislumbran algunos indicios de la metodología utilizada durante su PTD, esta no está completamente documentada. Podemos mencionar, a modo de ejemplo, el trabajo de Boukherroub, D'amours y Rönnqvist [1] los cuales indagan sobre la estructura de un CSTD pero no muestran de manera completa una metodología a seguir. El CSTD del ITESM Campus Ciudad de México [24], propone 4 etapas como parte del PTD, sin embargo, dicha metodología no está completamente documentada.

3. Metodología

La investigación aquí presentada fue realizada mediante la aplicación de un método analítico, con el fin de realizar un estudio que minimizara tiempo y recursos [9]. Utilizando como base las 4 etapas implementadas en el PTD del ITESM [24], se propone una metodología que adopte dichas etapas de una manera estructurada.

4. Propuesta de Metodología para el PTD en un CSTD

Se propone una metodología para el PTD basada en la Metodología de Sistemas Suaves (SSM) [2] y Metodología de Análisis Multicriterio (MCDA) [13], aplicado a las características de un CSTD, esta propuesta lleva el nombre de **Metodología de Decisiones en Centros Multipantalla (MDMC)** por sus siglas en inglés.

La SSM se puede describir como una técnica cualitativa la cual se ocupa de problemas situacionales en los cuales hay una actividad con un alto componente social, político y humano, esto distingue a la SSM de otras metodologías que se ocupan de los problemas DUROS los cuales están a menudo más orientados a la tecnología, los pasos que describen está metodología son:

1. Situación problema no estructurada
2. Expresión de la situación problema
3. Definiciones raíz de sistemas relevantes
4. Los modelos conceptuales
   4a. Concepto de sistema formal
   4b. Otros pensamientos de sistemas
5. Comparación de paso 4 con paso 2
6. Cambios deseables y factibles
7. Acción para mejorar la situación problema

La MCDA constituye una herramienta de apoyo en el proceso de toma de decisiones [7], la cual se enfoca en evaluar alternativas según criterios individuales y combinarlos en una evaluación general, que a su vez se utiliza para facilitar el camino hacia la decisión de manera que intervengan distintos puntos de vista, aunque sean contradictorios, los pasos que describen está metodología son:

1. Delimitar el contexto de evaluación
2. Identificar los criterios de evaluación
3. Construir las escalas de evaluación
4. Construir las funciones de valor
5. Ponderar las escalas de evaluación
6. Evaluar las opciones
7. Hacer análisis de sensibilidad y de robustez

Dadas las características de estas dos metodologías se decidió trabajarlas en conjunto para enriquecer el proceso empírico que se identificó en los CSTD estudiados. El CSTD del ITESM Campus Ciudad de México propone 4 etapas generales para realizar el PTD, las cuales se muestran en la Fig. (2):

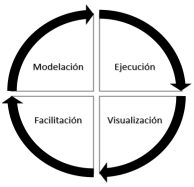


Figura 2. Proceso para la Toma de Decisiones aplicado por el ITESM, Campus Ciudad de México. Fuente: Elaboración propia

Es bien conocido que la toma de decisiones es extremadamente intuitiva cuando se consideran problemas de criterio único, ya que solo necesitamos elegir la alternativa con la calificación de preferencia más alta. Sin embargo, cuando la toma de decisiones necesita evaluar múltiples criterios, diferentes ponderaciones en criterios de decisión, aumento de variables y los diversos puntos de vista de las áreas involucradas, este proceso aumenta considerablemente su complejidad. Sumado a estas problemáticas tenemos el componente humano que se vuelve de gran relevancia en un CSTD ya que las decisiones en este ambiente se toman a través de un consenso, el cual puede tornarse largo y desgastante para todas las partes involucradas.

Por todo lo antes mencionado proponemos una metodología para el PTD, la cual se implementa en CSTD, que se compone de 7 pasos y están esquematizados en la Fig. (3)

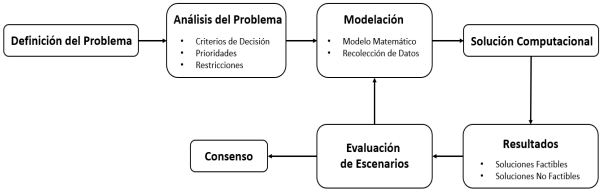


Figura 3. Metodología de Decisiones en Centros Multipantalla (MDMC). Fuente: Elaboración propia

1. **Definición del Problema** Traducción de un problema ambiguo a uno bien definido, mediante la transformación de la información de entrada a una oración estructurada que describa tanto condiciones iniciales como condiciones deseadas.
2. **Análisis del Problema** Proceso de descomposición del problema para localizar las acciones cuya modificación concluya con la solución del mismo, durante este proceso analizaremos los siguientes aspectos:
   - **Criterios de decisión** Conjunto de normas o criterios mediante los cuales se analizan las diversas alternativas.
   - **Prioridades** Conjunto de variables con mayor peso a considerar.
   - **Restricciones** Conjunto de limitantes las cuales impiden una modificación sobre una acción determinada.
3. **Modelación** Formalización del problema en su forma idealizada a una representación real mediante diversos métodos.
   - **Modelo matemático** Conjunto de métodos clásicos aplicados al problema para la obtención de alternativas de solución.
   - **Recolección de datos** Obtención, análisis y limpieza de datos para alimentar el modelo matemático.
4. **Solución Computacional** Implementación del modelo en un software especializado ya existente o creado para el problema.
5. **Resultados** Clasificación del conjunto de alternativas obtenidas mediante el paso anterior.
   - **Soluciones factibles** Elementos cuyas características satisfacen todas las prioridades y restricciones previamente planteadas.
   - **Soluciones No factibles** Elementos fuera de los rangos establecidos por las prioridades y restricciones.
6. **Evaluación de Escenarios** Sesión presencial donde las alternativas seleccionadas son presentadas en forma de escenario, el cual debe presentar visualización en gran formato, es interactivo y de fácil comprensión para los tomadores de decisiones. Durante este paso es posible regresar al paso 4 para retroalimentar la modelación, esto y cuando sea pertinente.
7. **Consenso** Sesión de trabajo presencial donde los tomadores de decisión acuerdan una alternativa como solución aceptable.

5. Trabajo Futuro

Nuestra propuesta metodológica está siendo evaluada en el marco de un proyecto de investigación conjunta con el grupo de Energía, Economía y Dinámica de Sistemas de la Universidad de Valladolid [19]. El proyecto en cuestión aborda como problema central la creación de un modelo del sector energético Mexicano, para la creación de escenarios prospectivos. Dicho trabajo se está implementando usando dinámica de sistemas como núcleo del modelo, y se están siguiendo las etapas definidas en el presente articulo.

6. Conclusiones

Para explotar al máximo las funcionalidades que ofrecen los Centros de Soporte a la Toma de Decisiones (CSTD), es necesario que se realice un Proceso de Toma de Decisiones (PTD) estructurado que asegure una propuesta de solución. A pesar de que existen varios CSTD en diversas instituciones a nivel mundial, los casos de éxito que presentan no exhiben de manera directa la metodología de PTD que utiliza cada uno de ellos. En este trabajo se realiza una propuesta metodológica en el PTD, basada en la Metodología de Sistemas Suaves (SSM) y Metodología de Análisis Multicriterio (MCDA), para ser aplicada en un CSTD. Nuestra propuesta lleva el nombre de **Metodología de Decisiones en Centros Multipantalla (MDMC)**, y está compuesta por 7 pasos, que reúnen los principales aspectos en el PTD, la cual abarca desde el proceso de definición del problema, hasta la decisión conjunta que tomarían los participantes en una sesión presencial en un CSTD.

Reconocimientos

Este trabajo fue desarrollado en el marco del "Fondo Sectorial CONACyT-Secretaría de Energía Sustentabilidad Energética" en el proyecto 266632 "Laboratorio binacional para la gestión inteligente de la sustentabilidad energética y la educación tecnológica"; con el apoyo de CONACyT en el programa de becas nacionales 860459, así como por el Tecnológico de Monterrey, Campus Ciudad de México. Los autores agradecen la colaboración de investigación sostenida con el equipo del Teatro de Decisiones® de la Universidad Arizona State University.

Referencias

1. Boukherroub, T., D'amours, S., Rönnqvist, M.: Sustainable forest management using decision theaters: Rethinking participatory planning. Journal of Cleaner Production **179**, 567–580 (2018)
2. Checkland, P.: Pensamiento de sistemas, práctica de sistemas. Limusa (2000)
3. Estrada Gallego, F.: Herbert a. simon y la economía organizacional. Cuadernos de Economía **26**(46), 169–199 (2007)
4. Forrester, J.W., Pereiro de Manzanal, M., et al.: Dinámica industrial. Biblioteca de ciencias económicas. (1972)
5. Huber, G.P.: Toma de decisiones en la gerencia. No. D10 H877, Trillas (1984)
6. Kalantari, B.: Herbert a. simon on making decisions: enduring insights and bounded rationality. Journal of management History **16**(4), 509–520 (2010)
7. Karatas, M., Sulukan, E., Karacan, I.: Assessment of Turkey's energy management performance via a hybrid multi-criteria decision-making methodology. Energy **153**, 890–912 (2018)
8. Le Moigne, J.L.: Les systèmes de décision dans les organisations. Presses universitaires de France (1974)
9. Lopera Echavarría, J.D., Ramírez Gómez, C.A., Zuluaga Aristizábal, M.U., Ortiz Vanegas, J.: El método analítico como método natural (2010)

10. Ríos, R.G., Sánchez, C.G., Gómez, E.S.: Toma de decisiones empresariales: un enfoque multicriterio multiexperto. Ingeniería Industrial **28**(1), 8 (2007)
11. Simon, H.A.: Models of bounded rationality: Empirically grounded economic reason, vol. 3. MIT press (1997)
12. Tolle, L.J.: A decision theatre designed for the laboratory instruction and observation of marketing decision-making. In: Combined Proceedings Spring and Fall Conferences of the American Marketing Association, Series. vol. 33, pp. 79–83 (1971)
13. Tzeng, G.H., Huang, J.J.: Multiple attribute decision making: methods and applications. CRC press (2011)
14. VVAA: About the decision theater, `https://www.mccaininstitute.org/decision-theater-mission/`, [Web; accedido el 04-11-2019]
15. VVAA: Decision theater zur nachhaltigen mobilität, `https://www.kopernikus-projekte.de/aktuelles?news=Decision_Theater_zur_nachhaltigen_Mobilitaet`, [Web; accedido el 01-12-2020]
16. VVAA: Decision theatre for collaborative decision-making, `https://www.ncl.ac.uk/ceser/research/decision-support/decision-theatre/`, [Web; accedido el 01-20-2020]
17. VVAA: Decision theatre for collaborative decision-making, `http://www.wrsc.org`, [Web; accedido el 01-12-2020]
18. VVAA: Energia lab, `https://energialab.tec.mx/`, [Web; accedido el 01-15-2020]
19. VVAA: Grupo de energía, economía y dinámica de sistemas, `https://geeds.es/`, [Web; accedido el 03-17-2019]
20. VVAA: La escuela de gobierno y transformación pública, `https://tec.mx/es/noticias/santa-fe/institucion/conoce-la-sede-de-la-escuela-de-gobierno-del-tec-en-cdmx-fotogaleria`, [Web; accedido el 02-15-2019]
21. VVAA: Laboratorio nacional de políticas públicas, `https://www.lnpp.mx/galeria/teatro-de-decisiones-2/`, [Web; accedido el 01-18-2019]
22. VVAA: Núcleo estratégico de decisiones, `http://www.centrodelagua.org/ned.aspx`, [Web; accedido el 01-15-2019]
23. VVAA: Project: Smartcity decision theater, `http://www.liophant.org/projects/dt/`, [Web; accedido el 02-02-2020]
24. VVAA: Teatro de toma de decisiones, `https://energialab.tec.mx/es/teatro-de-toma-de-decisiones`, [Web; accedido el 01-10-2019]
25. VVAA: Uaf's decision theater north uses tech to connect, `https://news.uaf.edu/uafs-decision-theater-uses-tech-to-connect/`, [Web; accedido el 04-12-2019]
26. VVAA: What is decision theater?, `https://dt.asu.edu/`, [Web; accedido el 01-03-2019]

12

Spatial network analysis of Tijuana's metropolitan transport network

BELGICA VALENCIA GARCIA AND ADAN HIRALES CARBAJAL

Spatial network analysis of Tijuana's metropolitan transport network

Valencia-García, B.[1][0000−0003−2088−6390] and Hirales-Carbajal, A.[1][0000−0002−5067−7250]

CETYS Universidad, Calz Cetys 813, Lago Sur, Tijuana, B.C. México
belgica.valencia@cetys.edu.mx, adan.hirales@cetys.mx

Abstract. Tijuana is the 6th largest city in Mexico. Due to its social-geographic characteristics, major avenues and arteries are often congested. Since 2006 and early 2018 onward, Tijuana's public transport fleets have undergone major renovation, by replacing small capacity vehicles, vans, and taxis, with larger capacity ones (busses). However, the city transport network properties have not been analyzed. In order to understand its structural properties and identify the locations of important nodes, we model and evaluate the transport network as weighted directed graph and analyzed its overall structural properties. We locate important nodes by analyzing each node in function of three criterion's: degree, distance and rank based attributes. We use non-dominated sorting to find nodes in the Pareto 1 front that maximize all criterion's. Among major findings, we find that for a commuter traveling along the network diameter or radius requires in average 16 and 10 route transfers correspondingly. Four, out of nine, delegations concentrate the most important nodes. Showing that the network needs restructuring.

Keywords: Tijuana transport network · structural properties · spatial analysis.

1 Introduction

Tijuana, being part of the San Diego – Tijuana megalopolis agglomeration, has a complex transport network that adapts to the immediate needs of its population. Its sociodemographic characteristics are very dynamic, it experiences: high volume of domestic, transborder, and transcounty traffic; high international exchange results in long lines at US-Mexico truck ports of entry; high migratory flows, among other characteristics [10]. Immigration phenomenon have accelerated the city growth, with more than 50% of its settlements of irregular origin, mostly in periphery [1, 10]. Thus the topological configuration of Tijuana is of regular and irregular human settlements.

To our knowledge the Tijuana *Mass Transit Network* (MTN) has not been analyzed. Previous studies have focus on understanding socio-cultural aspects and public policy, i.e., sustainable urban mobility [2], quality of service [4, 11], network supply [8], relationship between urban form and the environment [15],

and public policy and governance of the transit network [4, 11]. In this study, our focus is on analyzing the network structural and node properties.

Three classes of routes operate atop Tijuana's transit network, namely: MTN, which operate vehicles with an average capacity of 30 passengers; taxis, with an average capacity of 12 passengers; and Bus Rapid Transit (BRT), with capacities of 50 and 80 passengers. Of each class, there are 129, 157, and 22 routes accordingly. In this study, we analyze the MTN that result from merging 104 independent routes corresponding to the mass transport class. Route data was provided by the offices of the Municipal Public Transport Directorate.

2 The network model

The transport network is composed by a set of routes that aggregate geodesic coordinates and roads that routes traverse. Two kinds of roads are distinguished: bidirectional, which allow traffic in two directions, and unidirectional. Road capacity is not considered. Routes are preset and run in circular paths. A geodesic coordinate model an intersection, a location within a route, or a location within a curved road. From this point onward, the term geodetic coordinate is referred to as node and road as edge.

Formally, we model the transport network as a weighted directed graph $G = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V}$ the set of nodes and the $\mathcal{E}$ the set of edges. The number of elements in $\mathcal{V}$ and $\mathcal{E}$ are denoted by V and E. Node $n_i \in \mathcal{V}$ is defined by the tuple (φ_i, λ_i), with φ_i and λ_i the geodetic coordinates (latitude and longitude). A matricial representation of node adjacency is given by A a NxN square matrix with entries a_{ij} equaling 1 if the edge $e_{ij} \in \mathcal{E}$, otherwise zero. The edge weight w_{ij} represents the length of the road segment between nodes i and j. Such quantity is measured in meters. The Haversine formula [5] is used to compute the great-circle length between nodes i and j.

The *degree* k_i of node i is the number of incident edges with the node defined as $k_i = \sum_{j \in \mathcal{V}} a_{i,j}$. For directed graphs, the node in-degree is denoted as $k_i^{in} = \sum_{j \in \mathcal{V}} a_{j,i}$ and the out-degree as $k_i^{out} = \sum_{j \in \mathcal{V}} a_{i,j}$. The set of incident nodes of i is denoted as $\mathcal{N}_i$. The immediate predecessor and successor node sets are defined as $\mathcal{N}_i^{in}$ and $\mathcal{N}_i^{out}$ correspondingly.

Let d_{ij} be the shortest path (or geodesic) between nodes i and j, such that $i \neq j$. Further, let $L = \frac{1}{V(V-1)} \sum_{i,j \in \mathcal{V}, i \neq j} d_{ij}$ the average shortest path length defined as the mean geodesic lengths among all pair of nodes. Let n_{jk} be the number of all shortest paths between nodes j and k. Further, let $n_{jk}(i)$ be the number of shortest paths between nodes j and k that traverse through node i.

Lastly, the number of triangles in the graph is defined as $\delta(G)$ and triplets (possible triangles, chains of nodes) of two directed ties between three nodes are defined as $\tau(G)$. See [3] for an exhaustive summary of network metrics.

3 The data processing pipeline

Inadequate planning of urban transport systems leads to congestion, highly redundant routes, and from the perspective of route ownership, an unequal distribution of wealth. Thorough planning requires assessing whether applied transport policies result in a sustainable, resilient, and fair mobility infrastructure. A quantitative approach towards such end entails modeling the transport network, evaluating its structural and dynamic properties, and assessment of observed results.

At the highest level of granularity, network global metrics analyze complex relations among data with the aim of characterizing and finding similarities between discovered features and known graph models. Finer grain metrics, reveal node and linking properties of the DAG i.e. node clustering degree, betweenness, among others. In this work, we focus on characterizing the topology at global and node level. We evaluate the network quality by applying metrics summarized in Table 1. Due to space constrains we refer the reader to [3,7,12,13,16] for their description and interpretation.

Table 1. Network and node level metrics. Source: Own elaboration.

	Metric	**Expression**
Centrality	Degree	$k_i = \sum_{j \in \mathcal{V}} a_{i,j}$
	Knn	$k_{nn,i} = 1/k_i \sum_{j \in \mathcal{N}_i} k_j$
	Pagerank	pr_i, see [9,13]
	Betweenness	$b_i = \sum_{j,k \in \mathcal{V}, j \neq k} \frac{n_{jk}(i)}{n_{jk}}$
	Closeness	$c_i = \frac{1}{\sum_{j \in \mathcal{V}; i \neq j} d_{ij}}$
Network	Assort. Coeff.	r, see [12]
	Diameter	$d = \max_{j \in \mathcal{V}} d_{ij}$
	Transitivity	$t = \frac{3\delta(G)}{\tau(G)}$

In order to facilitate reproducible and verifiable results the data set, metrics, data processing pipeline, and the project anaconda development environment are made available via the project site [1].

The transport network was built by concatenating 104 out of the 129 independent routes. The remaining 25 routes that were not included correspond to incomplete routes. Thus, node importance may decrease throughout nodes these routes traverse in our results. The size of $\mathcal{V}$ is of 20,411 nodes. Routes share up to 80% of nodes. Nodes that share a location or are close to each other, i.e. withing the street width, were merged. The size of $\mathcal{V}$ reduced to 4461 non-redundant nodes. The transport network was digitally edited with JOSM Ver. 15238, an extensible editor for Open Street Map. The resulting set of routes was stored in geojson format and is available for download in the project site.

[1] See *https://github.com/ahiralesc/tijuana_vehicle_network_analysis.git*

The data processing pipeline consists of the following phases: *data preparation phase*, which load and merges the transport routes. The greater circle distance between nodes is also computed during this phase. Nodes and edge data are modeled as dictionaries; *graph preparation phase*, the dictionaries are transformed into a directed graph; *metric application phase*, user selected global and node level metrics are applied; and *statistical analysis and data visualization*. Statistical analysis was done by using simple descriptive statistics. All processed data is stored in CSV (Comma-Separated Values) and pickle format.

Node importance is evaluated by categorizing node metric into three classes, namely: *degree* based (deg); based on *distance* and neighborhood based (dist); and *rank* base (rank). Metrics in the first class are $k_i, k_i^{in}, k_i^{out}, k_{nn}$, in the second b_i and c_i, and in the third pr_i. Metrics in each class are equally weighted and linearly combined. The non-dominated sorting technique is used to extract points in each Pareto front [6].

4 Results

4.1 Network properties

Network level properties are summarize in Table 2. A diameter, or maximum eccentricity of 190 was obtained. In average a commuter traveling along the diameter requires 16 route transfers. Such quantity is four times larger than that in Rio de Janeiro's bus transfer network [14]. The number of transfers can be minimized to 10 however it increases the length of the route. The minimum eccentricity (or radius) contains 99 edges. A total of 4461 pairs of nodes length equals the radius. In average 10 bus transfers are needed to traverse along a path with length equaling the radius. Once again, the number transfers can be minimized to 6 at the expense of increasing the length of the path. The Tijuana 2017 municipal development plan report that in average 4 bus transfers are made by 46% of the MTN commuters [8]. No statement is made regarding the length of the trip. An additional two transfers would be required if the length of those trips equals the radios. Note the above analysis assumes bus transfers occur locations where routes intersect.

The MTN assortativity is of 0.37. Such value suggest that the network is not centralized in a few large hubs and that these are interconnected. Hub nodes often had a degree of 7-9 and constituted approximately 1.2% of the network nodes. The network transitivity t is low (0.01). In general, a network with poor transitivity have large path lengths [18]. These results confirm that the diameter is large.

Table 2. Global network properties. Source: Own elaboration.

r	d	t
0.37	190	0.01

4.2 Node importance

Fig 1 displays the weighed node importance scores according to the applied linearization. Nodes with less importance cluster near the origin, whereas those with highest importance tend to the unit origin. Results show that most of the nodes cluster near the origin with a range of [0,0.4]. A cloud like pattern is formed withing this region. However, as the nodes span out they begin to lean towards the degree coordinate axis. Suggesting that node importance is skewed by the node degree. We found a total of 20 non-dominated nodes in the first Pareto front. These correspond to the most important nodes in the transport network (according to the proposed technique).

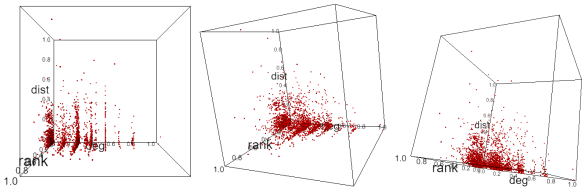


Figure 1. Linearization of metrics into 3D space. Source: Own elaboration based on [6].

The 10 most important nodes are depicted with blue markers. See Fig. 2. Nodes color in red correspond to the later 10 nodes with second most importance. These two sets of nodes are in the first Pareto front. The first set of nodes are concentrated in 4 of the 9 delegations the city is composed of, namely: La Mesa, La Presa Abelardo. L. Rodriguez, La Presa Este, and San Antonio de los Buenos. Areas containing these nodes are magnified for better visualization. Intuitively, one would expect large intersections to concentrate many routes, i.e. areas 7 and 3. Visual inspection of the remaining areas show that large route concentration occurs in main pathways in regions with lattice like topology.

Table 3 lists the range of mean measures corresponding to the first 10 Pareto fronts. Quantitatively, nodes in the first four Pareto fronts are very similar. Thus, intuitively, they share a similar level of importance and cumulatively increase the number of nodes in this class to 164 nodes.

The mean path length from any reachable node to any node in the Pareto 1 front is 54. Such length is smaller than the radius by 45 edges. Nodes in the front 1 may be good candidates for bus stops and attract a greater share of trips increasing the centrality of the network [17]. Nevertheless, their feasibility needs to be assess as these locations currently model locations with high route intersection.

Mean betweenness centrality is low. Approximately 6% of all shortest paths go through a node in Pareto 1 and 2 nodes sets. A maximum betweenness cen-

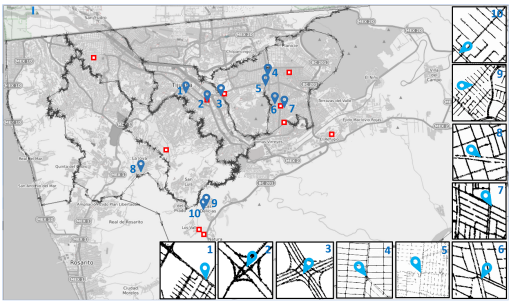


Figure 2. First Pareto front nodes. Source: Own elaboration based on [6].

trality of 0.21 was found. However, this occurs in only one location. Even if we extend the range from 0.1 to 0.21 only 1% of the nodes categorized as important have this degree of centrality.

Table 3. Mean metrics values. Source: Own elaboration.

Front	k_i	b_i	c_i	$k_{nn,i}$	pr_i
1	5.350	0.065	0.017	2.950	0.001
2	4.974	0.060	0.017	2.893	0.000
3	4.905	0.053	0.017	2.734	0.000
4	4.943	0.047	0.017	2.421	0.000
5	4.721	0.040	0.016	2.197	0.000
6	4.375	0.036	0.016	2.270	0.000
7	4.264	0.032	0.016	2.267	0.000
8	4.151	0.030	0.016	2.181	0.000
9	4.020	0.026	0.016	2.104	0.000
10	4.010	0.023	0.016	2.021	0.000

A mean closeness is not significantly high, 0.017 and 0.016, this might be explained because: routes tend to run in parallel along main avenues; they transit through roundabouts; or follow circuits in suburbs.

5 Discussion

In this work, we analyzed the Tijuana Mass Transit Network (MTN) structural and node importance properties by modeling it as a weighted DAG. In our model, edge weights quantify the great-circle length between adjacent nodes,

while nodes route geodetic coordinates. A four phase data processing pipeline is proposed. Node importance is evaluated by categorizing node metrics into three classes, namely: degree; distance and neighborhood; and rank based. A weighted linear combination of metrics in each class is applied. We apply non-dominated sorting technique to find nodes in each Pareto front.

Findings show that a commuter traveling along the network diameter or radius requires in average 16 and 10 route transfers correspondingly. The network is assortative, approximately 1.2% of its nodes had a high degree of route concentration. According to the proposed node selection criterion, important nodes are concentrated in 4 out of the 9 delegations that the city is composed of. Nodes in the Pareto front 1 can reduce the radius by 45 edges.

To our knowledge, this is the first study to address the analysis of the MTN properties. In subsequent studies we transform the network model from a weighed DAG to a multi-graph, since we lose knowledge of independent routes with the current representation. We also plan to evaluate the robustness with the aim of identifying the locations and events that may render the network unable to operate.

References

1. Tito Alegria and Gerardo Ordóñez Barba. *Legalizando la ciudad: asentamientos informales y procesos de regularización en Tijuana.* El Colegio de la Frontera Norte, March 2005.
2. Miguel Angel Avelar-Dueñas. Sistema de transporte de tijuana: Análisis de su proceso de modernización e implicaciones en la movilidad urbana. Un acercamiento desde la sustentabilidad. mathesis, El Colegio de la Frontera Norte, Tijuana, B.C., México, 2014.
3. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, February 2006.
4. Perla Rocío Castro-Cebreros. Procesos de concesión del transporte público en Tijuana, y efectos en la calidad del servicio. mathesis, El Colegio de la Frontera Norte, Tijuana, B.C., Mexico, 2012.
5. Nitin R. Chopde and Mangesh K. Nichat. Landmark based shortest path detection by using A* and haversine formula. *International Journal of Innovative Research in Computer and Communication Engineering*, 1(2):298–302, April 2013.
6. Carlos A. Coello Coello, Gary B. Lamont, and David A. Van Veldhuizen. *Evolutionary algorithms for solving multi-objective problems (Genetic and Evolutionary Computation).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. Evolutionary.
7. J. G. Foster, D. V. Foster, P. Grassberger, and M. Paczuski. Edge direction and the structure of networks. *Proceedings of the National Academy of Sciences*, 107(24):10815–10820, May 2010.
8. Instituto Metropolitano de Planeación IMPLAN. Plan municipal de desarrollo 2017-2019. Tijuana, B.C.: México. p.129, 2017.
9. Amy N. Langville and Carl D. Meyer. A survey of eigenvector methods for web information retrieval. *SIAM Review*, 47(1):135–161, January 2005.
10. Judith Ley García. Movilidad y transporte público: El arreglo urbano en Tijuana. *Revista Universitaria UABC*, 4:8–13, January 2006.

11. Lilia Cristina Millán-Velázquez. ¿Transporte público o mi automóvil? Factores socioculturales determinantes en la elección modal de desplazamiento en Tijuana, B.C. mathesis, El Colegio de la Frontera Norte, Tijuana, B.C., México, 2018.
12. M. E. J. Newman. Mixing patterns in networks. *Physical Review E*, 67(2), February 2003.
13. Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999.
14. Louise Pumar, Rafael Barbastefano, and Diego Carvalho. Bus transport network analysis in Rio de Janeiro based on topological models using social networks. *CoRR*, abs/1808.05692, 2018.
15. Dalila Ramos-Ortiz. Relación entre forma urbana, movilidad a través de transporte motorizado y emisiones de Gei en Tijuana, B. C. mathesis, El Colegio de la Frontera Norte, Tijuana, B.C., México, 2014.
16. Thomas Schank and Dorothea Wagner. Approximating clustering coefficient and transitivity. *J. Graph Algorithms Appl.*, 9(2):265–275, 2005.
17. J. Scheurer, Carey Curtis, and Sergio Porta. Spatial network analysis of public transport systems : developing a strategic planning tool to assess the congruence of movement and urban structure in Australian cities. *Australasian Transport Research Forum*, 2007.
18. Tanuja Shanmukhappa, Ivan Wang-Hei Ho, and Chi Kong Tse. Spatial analysis of bus transport networks using network theory. *Physica A: Statistical Mechanics and its Applications*, 502:295–314, jul 2018.

13

Generación automática de controladores difusos

AMAURY SANCHEZ NEGRIN, SELENE LILETTE CARDENAS MACIEL, JORGE ANTONIO LOPEZ RENTERIA, ARMANDO MARTINEZ GRACILIANO Y NOHE RAMON CAZAREZ CASTRO

Generación automática de controladores difusos

A. Sanchez–Negrin, S. L. Cardenas–Maciel, J. A. Lopez–Renteria, A. Martinez–Graciliano, and N. R. Cazarez–Castro[0000−0001−6787−3947]

Tecnológico Nacional de México-Instituto Tecnológico de Tijuana
{amaury.sanchez18,jorge.lopez,aamg_a}@tectijuana.edu.mx,
{lilettecardenas,nohe}@ieee.org

Resumen En la construcción de sistemas de inferencia difusa, una tarea compleja es establecer correctamente la relación de las variables de entrada y salida a través de las reglas y el ajuste de las funciones de membresía. Se propone un software que apoye en la generación de familia de sistemas difusos aplicando una metodología que en forma automatizada modifica los parámetros de funciones de membresía. Se verifica la oportunidad que tiene el software como una herramienta de apoyo al diseño al ser aplicada en un problema de regulación en un motor, el software crea una familia de controladores difusos que logran el objetivo de control mostrando al usuario variantes en el desempeño entre ellos.

Keywords: sistema difuso, sistema de control, variación de parámetros.

1. Introducción

La lógica difusa fue introducida por Zadeh en 1965 [17] y una de las áreas más explotadas mediante es la teoría de control [11]. Además, cuenta con aplicaciones en la industria, medicina, aeronáutica, electrónica [16]. Muchos investigadores han trabajado en métodos de diseño de Controladores Lógico Difusos (CLD), métodos sistemáticos para sistemas lineales [9,8,10], para no lineales [15], considerando criterios de estabilidad [1,4,12]. También, se ha reportado el desarrollo de herramientas para el diseño de sistemas difusos para control, su implementación varía según el objetivo de la aplicación [5,6,7].

El enfoque presentado en [2] se centra en la construcción de las reglas difusas utilizando la información de las variables del error y cambio del error, las cuales son mapeadas en el espacio 2D. De esta forma, el controlador difuso controla cinco (5) sistemas diferentes. Asimismo, en [14], se desarrolla una herramienta, la cual brinda la posibilidad de interactuar con los sistemas difusos tipo Mamdani, permitiendo su implementación como sistema embebido y con características limitadas sobre un microcontrolador PIC. Donde su objetivo principal es la generación automática del código en lenguaje C para el sistema lógico difuso deseado por el usuario. Sin embargo, en el presente trabajo se crea una familia de controladores lógico difusos, los cuales responden al control de un sistema dinámico. De esta forma se selecciona un CLD que cumpla con los parámetros en la respuesta del sistema deseados.

En la síntesis de CLDs, con reglas fijadas, los parámetros de las funciones de membresía influyen en la respuesta del sistema en lazo cerrado [3]. Al realizar una modificación uniforme de los parámetros de las funciones de membresía hasta cierto límite, el sistema en lazo cerrado mantiene algunas propiedades como la

estabilidad, pero varía en algunos otros tal como el rendimiento [3]. Realizar ajustes de los parámetros para conseguir el desempeño deseado puede resultar no trivial, por ello, considerar herramientas de software que asistan en el diseño de CLDs pueden ser una alternativa que agiliza y apoya en etapas del proceso de desarrollo, tales como detectar aquellas relaciones entre variables que signifiquen relaciones de compromiso que se contraponen a los requerimientos de diseño.

El problema considerado en este artículo esta relacionado con ajustar CLDs con procedimientos computacionales, para cumplir requisitos de desempeño. La problemática obedece a resolver problemas de optimización no lineal, lo cual obliga a acotar el problema como sigue: Dada una estructura de sistema difuso definida en términos de variables de entrada–salida, funciones de membresía y reglas, obtener de forma automatizada una familia de CLDs mediante la manipulación de los parámetros de las funciones de membresía. En este artículo se propone una metodología basada en las ideas de [3], que es implementada en software para producir un conjunto de CLDs a través de modificaciones a las funciones de membresía.

El resto del artículo se presenta de la siguiente manera: En la Sección 2 se establecen para beneficio del lector definiciones relevantes en el tema de conjuntos y sistemas difusos; una metodología para la manipulación de funciones de membresía y modificación de CLDs implementada en software se proporciona en la Sección 3; la verificación de la propuesta a través de un caso de estudio se muestra en la Sección 4; finalizando en la Sección 5 con conclusiones.

2. Marco teórico

Los sistemas basados en lógica difusa son buenos para la representación de conocimiento parcial o incompleto, imitando el razonamiento humano mediante reglas. Las aplicaciones de la lógica difusa a problemas de ingeniería, se han ido acumulando experiencias y resultados que se pueden agrupar en [13]; (i) representación del conocimiento aproximado, (ii) aproximación universal de funciones, (iii) auto-organización. A continuación, se presentan conceptos básicos necesarios en la solución al problema planteado.

2.1. Conjuntos difusos

Definición 1 *Sea X un conjunto (en el sentido clásico). Un subconjunto difuso A de X está definido por una función $\mu_A : X \to [0,1]$ llamada* función de membresía *de A.*

El valor $\mu_A(x) \in [0,1]$ indica el grado en el cual cada elemento $x \in X$ pertenece al conjunto difuso A. Un conjunto difuso A de X puede ser visto de manera estándar como un subconjunto (en sentido clásico) del producto cartesiano $X \times [0,1]$. Además, podemos identificar a un subconjunto difuso A de X como el conjunto de todas las parejas ordenadas $\{(x, \mu_A(x)) : \text{ con } x \in X\}$, el cual define a la gráfica de μ_A. Un conjunto difuso se dice ser *vacío* si y solo si $\mu_A(x) = 0$ para cada $x \in X$.

Los conjuntos difusos que usaremos son funciones de membresía triangular de la forma

$$\mu_A(x, a, b, c) = \text{máx}\left(\text{mín}\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right), \tag{1}$$

donde $a < b < c$ (siendo b el núcleo del triángulo). Por otra parte, la definición de Altura de un conjunto difuso:

Definición 2 *Sea μ_A la función de membresía de un subconjunto difuso A sobre un subconjunto $X \subset \mathbb{R}$. Se define como altura del conjunto difuso A como $h(A) = \sup\{\mu_A(x) \mid x \in X\}$, donde* sup *es el supremo del conjunto de valores de μ_A en X.*

2.2. Controladores lógico-difusos

Para el presente trabajo, nos enfocaremos en sistemas difusos tipo Mamdani, los cuales describiremos a continuación.

Definición 3 *Un CLD tipo Mamdani es una terna* $I = (E, R, S)$ *donde:*

- E *es el conjunto de variables lingüísticas de la entrada.*
- R *conjunto de reglas que relacionan la entrada con la salida.*
- S *es el conjunto de variables lingüísticas de la salida.*

En este trabajo se usa el método de defuzificación media de los máximos (MOM), el cual está dado por

$$v = \frac{\sum_{x_i \in M}(x_i)}{|M|}, \tag{2}$$

donde $M = h(A)$ es la altura del conjunto difuso A, descrito en la Definición 2, y $|M|$ se refiere a la cardinalidad de M.

Una variable puede tener una o más etiquetas lingüísticas, matemáticamente estas se representan a través de funciones de membresía y en consecuencia por conjuntos difusos. La relación de las variables lingüísticas de entrada y salida se hace a través de reglas **SI** − **ENTONCES**, en este trabajo se asume que el CLD tiene una entrada y una salida. Por ello, las reglas tienen la forma:

$$R_i : \textbf{SI } x_n \text{ es } M_{p_k} \textbf{ ENTONCES } y_m \text{ es } N_{l_j}, \tag{3}$$

donde $R_i \in R$ representa las reglas, las $x_n \in E$ es la variable lingüística de entrada, M_{p_k} son las etiquetas de la variable de entrada, $y_m \in S$ es la variable lingüística de salida y N_{l_j} sus respectivas etiquetas lingüísticas.

3. Propuesta de solución

La generación automática de CLDs se realiza a través de una herramienta de software que se compone de dos módulos o subsistemas: (i) el primer módulo, el cual está implementado en lenguaje de programación Java™, recibe el CLD diseñado por el usuario y aplica la modificación de parámetros de funciones de membresía, cuyo resultado lo utiliza para crear nuevos CLDs que son representados con una estructura apropiada y posteriormente almacenados en archivos de texto; (ii) El segundo módulo esta implementado en Matlab™, se encarga de ejecutar simulación del sistema en lazo cerrado que usa cada nuevo CLD, cada respuesta en lazo cerrado del sistema controlado se almacena para futuros procesamientos y se utiliza para generar reportes gráficos. En la Fig. 1 se representa el funcionamiento general de la propuesta.

En la Fig. 2 se muestra el diagrama de clases con sus atributos y sus métodos, que conforman al primer módulo. La clase *fisCreator* es la encargada de llevar el objeto instanciado de la clase *FISFILE* (que representa cada CLD) a objeto archivo con extensión .fis; Los métodos de la clase *Controller* generan la familia de archivos (nuevos CLDs), en ella esta implementado el de método de variación de parámetros de funciones de membresía que se describe en la subsección 3.1 y participan en la construcción de la base de reglas.

La interfaz gráfica de usuario (GUI) permite recibir los datos que describen al CLD, los cuales son utilizados para hacer instancias de un objeto de la clase *FISFILE*. La Fig. 3 corresponde a la GUI que permite introducir el nombre

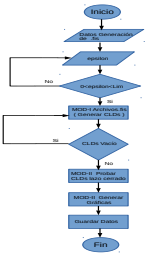


Figura 1. Diagrama de Flujo.

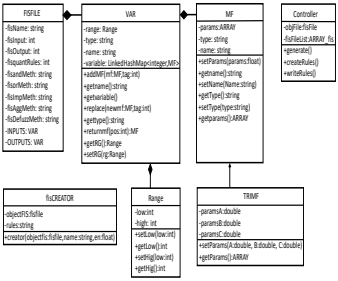


Figura 2. Diagrama de Clases

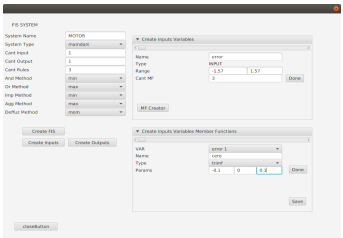


Figura 3. GUI del software para definir variables y MFs del CLD.

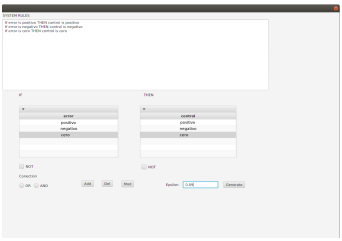


Figura 4. GUI del software para definir reglas del CLD.

del CLD, variables lingüisticas de entrada y salida, funciones de membresía, los métodos que se aplicarán para la inferencia y la defuzificación; La GUI mostrada en la Fig. 4 permite capturar cada una de las reglas y un número nombrado *Epsilon* (ϵ) que el usuario decide e indica el grado de variación que se aplica en los parámetros de funciones de membresía cuando se crean la familia de CLDs.

El segundo módulo consiste de la implementación en software del modelo matemático en ecuaciones diferenciales del sistema a controlar, el sistema en lazo cerrado; funciones que evaluan numéricamente el sistema en lazo cerrado con cada controlador; y funciones que almacenan los resultados numéricos y generan gráficas de resultados.

3.1. Variación de parámetros de funciones de membresía y generación de CLDs.

La implementación en el primer modulo de la modificación de los parámetros en funciones de membresía considera que en cada variable lingüística, sean

Tabla 1. Parámetros del motor de corriente directa.

Constante	Parámetro	Valor
J_m	Momento de inercia del motor	0,01 $kg \cdot m^2$
f_m	Fricción viscosa	0,1 $N \cdot m \cdot s$
K_b	Fuerza electromotriz	0,01 $v \cdot s/rad$
K_a	Torque	0,01 $N \cdot m/Amp$
R_a	Resistencia eléctrica	1 Ohm
L_a	Inductancia eléctrica	0,5 H

cada conjunto difuso A_i definido sobre $X = [D_{min}, D_{max}]$ con función de membresía $\mu_{A_{i,j}}(x; a_{i,j}, b_{i,j}, c_{i,j})$ que satisfacen la Definición 1 y una cantidad $\epsilon = min\{c_{i,j} - b_{i,j}, b_{i,j} - a_{i,j}\} > 0$ y fija. Para el conjunto difuso A_i seleccionado, se utiliza ϵ para obtener de cada $\mu_{A_{i,j}}$ dos nuevas funciones de membresía tal que se determinan como $\mu_{A_{i,j}}(x; a_{i,j} - \epsilon, b_{i,j}, c_{i,j} + \epsilon)$ y $\mu_{A_{i,j}}(x; a_{i,j} + \epsilon, b_{i,j}, c_{i,j} - \epsilon)$.

La construcción de la familia de CLDs es de forma incremental dentro de un loop. En la primera iteración, al obtener las dos funciones de membresía para $\mu_{A_{i,j}}$ como lo describe párrafo anterior, se crean dos nuevos controladores difusos donde cada uno surge al actualizar el controlador difuso proporcionado por el usuario con una de las modificaciones recien aplicadas a $\mu_{A_{i,j}}$. Para las iteraciones subsecuentes, se modifican los parámetros de otra función de membresía y se crean los nuevos CLDs utilizando los CLDs de la iteración anterior, generando tantos CLDs como la combinaciones de las nuevas funciones de membresía recien modificadas lo permitan. En cada iteración, después de obtener el CLD, se transfiere la estructura a un archivo con extensión `.fis`. Este proceso concluye hasta que se han modificado parámetros a todas las funciones de membresía del CLD que fueron seleccionadas.

4. Caso de estudio: Descripción del Sistema y CLD

Considerese el modelo de un motor de corriente directa (CD):

$$\begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} -\frac{f_m}{J_m} & \frac{K_a}{J_m} \\ -\frac{K_b}{L_a} & -\frac{R_a}{L_a} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{L_a} \end{pmatrix} u, \qquad y = (1\ 0) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{4}$$

donde $[x_1, x_2]^T$ es el vector de estados que representa la posición y la velocidad angular, los parámetros $f_m, K_a, J_m K_b, R_a$ y L_a se muestran en la Tabla 1.

El objetivo es determinar un comando de control u para el sistema descrito en (4) tal que el sistema sea llevado a $\theta_d = \pi/2$ rad. Para ello, se define la variable $\xi = y - \theta_d$, siendo θ_d la posición deseada, por lo que el problema de control a resolver es $\lim_{t\to\infty} \|\xi(t)\| = \lim_{t\to\infty} \|y(t) - \theta_d\| = 0$.
El controlador difuso diseñado por el experto, usa como variable lingüística de entrada a ξ y de salida a u. A cada variable lingüística se asocian tres etiquetas lingüísticas de nombre "*negativo*", "*cero*", "*positivo*", cada una se representa por una función de membresía tipo triangular cuyos parámetros de función de membresía se muestran en la Tabla 2, y el conjunto de reglas son:

R_1 : `Si` ξ `es` *negativo* `Entonces` u `es` *negativo*,
R_2 : `Si` ξ `es` *cero* `Entonces` u `es` *cero*,
R_3 : `Si` ξ `es` *positivo* `Entonces` u `es` *positivo*,

Tabla 2. Parámetros de funciones de membresía.

i	Variable	$[D_{min}, D_{max}]$	**negativo**			**cero**			**positivo**		
			$a_{i,1}$	$b_{i,1}$	$c_{i,1}$	$a_{i,2}$	$b_{i,2}$	$c_{i,2}$	$a_{i,3}$	$b_{i,3}$	$c_{i,3}$
1	ξ	[-1.57, 1.57]	-2.827	-1.571	0.0	-0.1	0.0	0.1	0.0	1.571	2.827
2	u	[-1.00, 1.00]	-1.800	-1.000	0.0	-0.1	0.0	0.1	0.0	1.000	1.800

4.1. Resultados

El CLD diseñado para el motor de CD se toma como base para construir de forma automática un conjunto de CDL's a través del software descrito en la sección 3 usando las reglas, funciones de membresía con los parámetros de la Tabla 2. El software aplica un mismo valor fijo de $\epsilon = 0{,}07$ para toda función de membresía que se modifica en cada ejecución y verifica numéricamente el sistema en lazo cerrado usando (4), u es la salida de cada CLD obtenida del método de defuzificación MOM dado en Definición 2. Los resultados que arrojan las pruebas del software se reportan a través de dos casos de estudio.

Caso 1. Las ejecuciones del software usan lo descrito en subsección 3.1 modificando solo los parámetros de las funciones de membresía "*cero*" de la variable de entrada ξ y de salida u. La evolución en el tiempo de las trayectorias de la posición se muestran la Fig. 6 (izquierda) para un conjunto de 20 CLDs que fueron generados por el software, se puede observar que todos convergen cercanamente a la posición deseada, se puede identificar que el desempeño en estado estable de unos controladores es peor que otros, esto debido a que la modificación de la función de membresía "*cero*" de la variable u en la regla R_2 no fue favorable, pues influye en aplicar un comando de control que le aleja un poco más del objetivo de control con respecto al resto de los resultados.

Caso 2. El procedimiento descrito en subsección 3.1 se aplica a todas las funciones de membresía de las variables lingüísticas de entrada ξ y salida u del CLD. En la Fig. 5 se muestran las modificaciones que se hicieron a cada función de membresía durante la ejecución del software. En la Fig. 6 (derecha), se muestran las trayectorias que resultan de los 20 CLDs generados. Las trayectorias permiten identificar que tienen un desempeño inicial muy similar que varia levemente desde el 0,5 seg., y es notorio que se cumple con el objetivo de control. La configuración de las modificaciones de todas las funciones de membresía lograron mejorar la respuesta en el sentido de que se tuvo un sobreimpulso menor previo a 0.5 seg. en comparación con los resultados del primer caso de estudio en el mismo intervalo de tiempo.

5. Conclusiones

Se propone un algoritmo computacional que implementa la idea de [3] para construir CLDs a través la variación de los parámetros de las funciones de membresía. La propuesta da como resultado la generación una familia de CLDs que resuelven el problema del diseño de un control para un motor de corriente

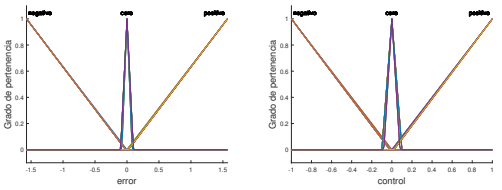


Figura 5. Variaciones en funciones de membresía que resultan del Caso de estudio 2.

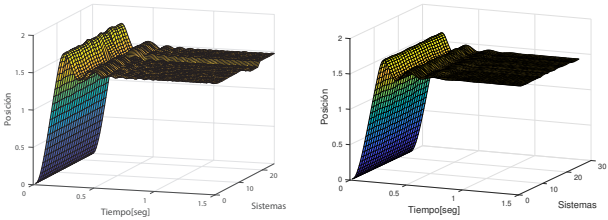


Figura 6. Trayectorias de posición usando los CLDs: Caso de estudio 1 (izquierda), Caso de estudio 2 (derecha).

directa, que estabiliza su sistema de errores asociado a una trayectoria predeterminada. La variación de los parámetros arrojó una familia de soluciones que preservan las trayectorias de control, esta afirmación es apoyada con simulaciones donde la familia de CLDs generada logra el control del sistema probado. Los resultados obtenidos muestran que el método de solución propuesto para construir una familia de CLDs con base en el conocimiento de expertos, representado a través de sistemas difusos, incorpora una gama de soluciones al problema de control de los cuales el usuario podría seleccionar según determinados criterios que se siguen para medir respuestas de control.

6. Agradecimientos

Esta investigación fue apoyada por CONACyT con el proyecto A1–S–32341 y por el Tecnológico Nacional de México mediante el proyecto 8058.20-P.

Referencias

1. Alcaraz-Rodriguez, M.A., Cazarez-Castro, N.R., Cardenas-Maciel, S.L., Coria, L.N., Contreras-Hernandez, S.: Estabilización basada en sistema difuso para estructuras afectadas por perturbaciones externas. In: AMCA (ed.) Memorias del

Congreso Nacional de Control Automático 2017. pp. 504–5009. Monterrey, Nuevo León (Ocubre 4–6 2017)
2. Altas, I.H., Sharaf, A.M.: A generalized direct approach for designing fuzzy logic controllers in matlab/simulink gui environment. International journal of information technology and intelligent computing **1**(4), 1–27 (2007)
3. Castillo, O., Aguilar, L., Cázarez, N., Cárdenas, S.: Systematic design of a stable type-2 fuzzy logic controller. Applied Soft Computing Journal **8**(3), 1274–1279 (2008)
4. Castillo, O., Aguilar, L., Cázarez, N., Rico, D.: Intelligent control of dynamical systems with type-2 fuzzy logic and stability study. Proceedings of the 2005 International Conference on Artificial Intelligence, ICAI'05 pp. 400–405 (2005)
5. Gonzalez, R., Torralba, A., Franquelo, L.: Afan, a tool for the automatic synthesis of neural and fuzzy controllers with architecture optimization. In: Proceedings of 1997 IEEE International Symposium on Circuits and Systems. Circuits and Systems in the Information Age ISCAS'97. vol. 1, pp. 637–640. IEEE (1997)
6. Hollstein, T., Halgamuge, S.K., Glesner, M.: Computer-aided design of fuzzy systems based on generic vhdl specifications. IEEE Transactions on Fuzzy Systems **4**(4), 403–417 (1996)
7. Kim, D.: An implementation of fuzzy logic controller on the reconfigurable fpga system. IEEE Transactions on industrial Electronics **47**(3), 703–715 (2000)
8. Lee, J.: On methods for improving performance of pi-type fuzzy logic controllers. IEEE transactions on fuzzy systems **1**(4), 298–301 (1993)
9. Li, H.X., Gatland, H.: A new methodology for designing a fuzzy logic controller. IEEE transactions on systems, man, and cybernetics **25**(3), 505–512 (1995)
10. Li, H.X., Gatland, H.: Conventional fuzzy control and its enhancement. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **26**(5), 791–797 (1996)
11. Mamdani, E.H.: Application of fuzzy algorithms for control of simple dynamic plant. In: Proceedings of the institution of electrical engineers. vol. 121, pp. 1585–1588. IET (1974)
12. Margaliot, M., Langholz, G.: New approaches to fuzzy modeling and control: Design and analysis, vol. 38. World Scientific (2000)
13. Olivares, M., Rojas, R.: Modelado y control difuso. Dpto. Electrónica, Universidad Técnica Federico Santa María pp. 6–9 (2002)
14. Reyes, S.F.P., Charry, C.A.P., García, E.E.G., et al.: Generación automática de sistemas lógicos difusos (sld) tipo mamdani sobre microcontrolador de 8 bits. Tecnura **17**, 93–108 (2013)
15. Ying, H.: Practical design of nonlinear fuzzy controllers with stability analysis for regulating processes with unknown mathematical models. Automatica **30**(7), 1185–1195 (1994)
16. Young, D., Utkin, V., Ozguner, U.: A control engineer's guide to sliding mode control. In: Proceedings., 1996 IEEE International Workshop on Variable Structure Systems, VSS'96. pp. 1–14. IEEE (1996)
17. Zadeh, L.A., et al.: Fuzzy sets. Information and control **8**(3), 338–353 (1965)

14

Optimization of the GTAW manual welding process, considering random covariates

YANETH GUADALUPE ROBLES DE LEÓN, DAVID SALVADOR GONZÁLEZ GONZÁLEZ AND ARGELIA FAVIOLA MIRANDA PÉREZ

Optimization of the GTAW manual welding process, considering random covariates

Yaneth Guadalupe Robles de León [1], David Salvador González-González[1-2] and Argelia Fabiola Miranda Pérez[1-2]

[1] Corporación Mexicana de Investigación en Materiales (COMIMSA) Calle Ciencia y Tecnología #790, Col. Saltillo 400 C.P. 25290, Saltillo, Coahuila, México. Phone: (+52) 01 844 411 32 00.
[2] Facutad de Sistemas, Universidad Autónoma de Coahuila, Ciudad Universitaria, Carretera a México Km. 13, Arteaga, Coahuila, México. Phone: (+52) 01 844 689 10 30.

Abstract. In the GTAW manual welding process (gas tungsten arc welding) three uncontrollable factors are presented called voltage, feed rate (welding), and feed speed (filler material) which, due to their nature, they are identified as random factors. Necessarily, in the modeling of manual welding processes only fixed or controllable factors are considered (as in the case of amperage); The other factors involved in the process are left to the considerations of the welder and his experience. In this sense, the use of a generalized linear model of mixed - effects is proposed, with the purpose of predicting the penetration and crown height considered as quality characteristics of the welded joint, fixed and random factors. In addition, it is proposed to perform optimization profiles using the NSGAII multiobjective algorithm, in order to find the region of process parameters where the desired penetration and Crown height quality is obtained. Results indicating random factors - voltage, displacement speed and wire feed speed have a significant effect on Crown height and penetration, as random factors increase, crown and weld joint geometry increases. An optimization profile is exposed.

Keywords: NSGAII multi-objective algorithm, Generalized linear Mixed Effects Model, Optimization Profile, Manual GTAW.

1 Introduction

Food industry manufacturing has manual welding processes that are used to leave a cleaner, finer, and more uniform finish suitable for the containers and connections involved in the manufacturing processes. These manual processes have the ability to do what sometimes a robot or an automatic welding machine cannot do. The manufacture of the food industry is mostly made of thin metal parts, which is why the GTAW process is ideal for this type of material, since one of its characteristics is to easily weld low-thickness metal parts. Despite the experience of trained welders with
in the GTAW welding process, welding defects continue to occur; not having a detailed procedure to correctly set the parameters, an unstable quality continues to be generated.

One of the initial techniques to detect surface defects and imperfections is visual inspection; discovering and repairing any defect in time is an essential cost reduction (AWS, 2010).
A product with a desired quality is obtained primarily through the proper programming or calibration of the process parameters. In some cases, there are variables that are presented randomly and cannot be controlled, therefore, it is necessary to think of a model that considers controllable process factors and random factors. Such is the case of generalized linear models with mixed effects, useful for making adequate predictions of the process and avoiding trial-and-error adjustment.
In relation to modeling and optimizing welding processes, Tarng et al. (1999) describe a model for optimizing the TIG welding process. Using a back propagation neural network to construct the relationship between weld parameters and weld position geometry characteristics. Likewise, Benyounis, Olabi (2008) conclude that by combining techniques such as Genetic Algorithms and Response Surface, good results are obtained to establish optimal welding conditions. Sathiya et al. (2010) show the optimization of a GMAW process using the Taguchi and "Gray-based" methods, considering as output variables or quality characteristics: grain width, heel height and penetration depth. The welding parameters were: voltage, displacement speed and wire feed speed. Optimal process conditions were identified in order to increase productivity and minimize the total cost of operation. Chaki et al. (2015) proposes the combination of neural networks, genetic algorithms and the simulated annealed name algorithm, for the modeling, prediction and optimization of the resistance of the hybrid CO2-MIG laser welding process in joints welded with aluminum alloy. They show the development of three models.
In the present work, the modeling of a manual butt joint GTAW welding process is performed considering the voltage, displacement speed and wire feed speed as random covariates, in addition to the fixed or controllable process parameter (Amperage). Regarding the process, Gang Zhang et al. (2017) notes the current benefits of manual welding processes, mentions that the intelligence and skills of the human welder are irreplaceable by a robot used for less critical components, also Y.K. Liu, et al. (2014) mentions that the experiences and skills of the human welder are critical to produce quality welds.

2 Methods

2.1 Generalized linear mixed effects model

Linear regression models contain only one random element, the error. Many situations require models in which there is more than one random term. Mixed models encompass both fixed effects and various random effects. Within the mixed models, the hypothesis test based on generalized least squares (GLS), maximum likelihood (ML) and maximum restricted likelihood (REML) is analyzed. A mixed effects model consists of both fixed effects and random effects. Fixed effect terms are usually the conventional linear regression part of the model. The terms of random effects are associated with individual experimental units drawn at random from a population, and explain variations between

groups that could affect the response. Random effects have prior distributions, while fixed effects do not.
Generalized linear mixed effects models (GLME) describe the relationship between a response variable and independent variables using coefficients that can vary with respect to one or more grouping variables, for data with a response variable distribution that is different from normal. Examples of such grouped data include longitudinal data, repeated measurement data, multi-level data, and block designs. The mixed general linear model is given by:

$$Y = X\beta + Zv + \epsilon \quad (1)$$

Where X is a matrix of known Nxp dimension, β is a vector px1 of fixed parameters, Z is an Nxq matrix of known constants, v is a vector qx1 of random effects, and ε is an Nx1 vector of random errors. Assuming that:

$$\begin{matrix} v \\ \varepsilon \end{matrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}\right) \quad (2)$$

Where G and R are matrices in a known way, but they depend on an unknown parameter θ, considering also that Var (Y) = ZGZ´ + R.
The estimation of the mixed generalized models not only implies the β parameters, but also the variance of the θ parameters. Furthermore, you may be interested in the hypothesis test not only about β, but also about θ, considering the expression given by

$$(Y - X\beta)'V_Y^{-1}(Y - X\beta) \quad (3)$$

2.2 Estimation by Maximum Likelihood

REML is used as a method for fitting linear mixed effects models and random field theory. In contrast to the earlier maximum likelihood estimation, REML can produce unbiased estimates of variance and covariance parameters. Penetration and Crown height models were estimated by REML.

2.3 Non-dominated Sorting Genetic Algorithm (NSGA-II)

The NSGA-II algorithm (Non-dominated Sorting Genetic Algorithm, version II) was presented by K. Deb and colleagues at the Kanpur Institute of Genetic Algorithms Laboratory in India in 2000 (Deb et al., 2000). It emerged as an improved version of the NSGA algorithm (Srinivas and Deb, 1994.
In this algorithm, the descendant population Qt (size N) is first created using the parent population Pt (size N). After this, the two populations are combined to form Rt of size 2N. After the aforementioned, utilizing a non-dominated ordering, the Rt population is classified on different Pareto fronts, it is justified by the fact of allowing a global verification of dominance between the population of parents and descendants. Once the non-dominated ordering process has finished, the new population is generated from the

configurations of the non-dominated fronts. This new population begins to be built with the best non-dominated front (F1), continues with the solutions of the second front (F2)... and so on. Each solution is assigned a fitness function according to its level of non-dominance (1 is the best level) and it is understood that this function must decrease during the process. The fitness assignment is oriented to provide a better possibility of perpetuating those individuals with better characteristics, considering the values of the objective functions and the results of the metric used to assess diversity. Those fronts that cannot be accommodated disappear. It is important that the non-rejected configurations are of good quality and chosen using a methodology that guarantees diversity. The idea is that configurations that ensure diversity within the same Pareto front are always promoted. When the entire population converges to the optimal Pareto front, the algorithm ensures that the solutions are distanced from each other.

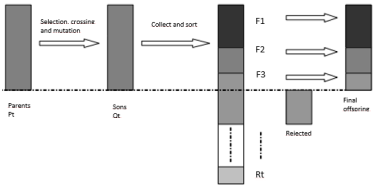


Fig. 1. Scheme of the NSGA-II algorithm. Source: Own elaboration.

3 Methodology.

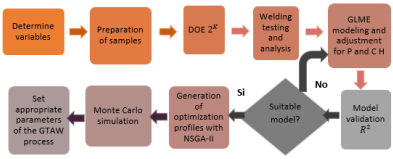


Fig. 2. GTAW process optimization methodology. Source: Own elaboration.

4 Application.

In the GTAW welding process, a 304 stainless steel base metal is used, and filler material type 308-L, two 304 stainless steel plates 10 cm long by 6.5 cm wide are selected. The union of these plates is a butt, with a 45 ° Bisel; The welding procedure was developed using a Miller Syncrowave 350 CC AC / DC welder. The measured quality characteristic is the penetration and crown height of the weld bead. The controllable process parameter is only Amperage (Amp), and the random factors are Voltage (Volt),

displacement speed (m / min), and wire feed speed (rpm). Variables voltage, displacement speed, and wire feed speed are referred to as uncontrollable and therefore must be modeled as random covariates.

Data were taken directly from the welding process, firstly taking into account data suggested by the expert welder and then verified with preliminary tests, selecting the best result, which was taken as the basis to carry out a kind of design of 2^k experiments, and in this way obtain the data to carry out the definitive tests. In each test data is taken from the same voltage that is automatically determined by the machine by giving it an amperage value, the displacement speed and wire feed speed are measured using video and stopwatch, using the speed formula v= d/t . Both are the most significant random covariates within the manual GTAW process.

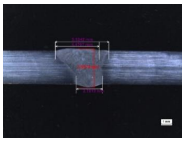


Fig. 3. Test specimen 1, analyzed by stereoscope. Source: Own elaboration.

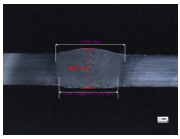


Fig. 4. Test specimen 2, analyzed by stereoscope. Source: Own elaboration.

After the welding process, each sample of the welded plates is cut, assembled, polished and etched with a reagent to reveal the weld zone and thus measure the penetration and the crown of the weld bead, these are analyzed in a in a stethoscope. Figures 1 and 2 show the analysis of two of the process samples. Table 1 shows the parameters used to make the unions, the random covariates and the result of the penetration and the crown in each case.

Table 1. Samples of the GTAW welding process.

Sample number	Amperage	Voltage	Application Speed	Forward speed	Penetration	Crown height
1	120	15.1	1.9830	1.5932	3.7074	0.25
2	125	15.6	2.6862	1.8627	3.5882	0.48
3	115	15.5	2.7623	1.4814	3.7722	0.82
4	127	15.4	3.4186	2.1395	3.7422	0.85
5	113	15.2	3.4084	1.4264	3.7804	1.19
6	113	14.1	1.9526	0.7761	3.6975	0.53
7	115	14.7	2.9738	1.0239	3.6316	0.65
8	120	15.1	3.2494	1.1151	3.7009	1.31
9	125	15.5	2.5370	1.7222	3.8045	0.45
10	127	15.7	4.6825	1.4603	3.6436	0.80

Source: Own elaboration.

Once the information on the welded joints has been obtained, it is necessary to establish the distribution of the random covariates, since when modeling the data with mixed linear regression, the random covariates must follow a normal distribution; An

Anderson-Darling test was performed, such assumption was not fulfilled, so it was necessary to define an alternative model called Generalized Model with Mixed-effects, the data fit this model, Maximum Restricted Partial Likelihood (REMPL) is used as the adjustment method and as a covariance method the 'conditional' method. A linear correlation test is performed (Pearson's), to see if there is a linear correlation between the response variables (Penetration and Corona), the test indicates that there is no correlation between them, therefore it is necessary to establish two mixed generalized linear models for each answer. These models were as follows:

Penetration: Amperage + (1| Voltage) + (1| Feed speed) + (1| Displacement speed)
Corona: Amperaje +(1| Displacement speed)+ (Displacement speed | Feed speed) + (1 | Feed speed)

With a likelihood value of $Lk = 8.0527$ for penetration and $Lk = -6.6956$ for Crown height, the coefficients of determination $R^2 = 0.997$ for penetration were obtained, indicating that 99.7% of the variability is being explained with welding parameters and random variables for penetration. Considering the P-value of the significance test, it can be concluded that the displacement speed, wire feed speed, voltage, and amperage are significant and have a considerable effect on penetration. On the other hand, the crown determination coefficient is $R^2 = 0.99301$, which indicates that 99.301% of the variability of the parameters involved is being explained. Considering the P-value of the significance test, it is concluded that the amperage and the voltage are not significant, but the displacement speed and the wire feed speed are very significant and have a considerable effect on the crown height and it is proved that the interaction between them is significant.

Table 2. Penetration Adjustment Results.

Fixed effects coefficients	**Estimate**	**P Value**	
Intercept	3.3227	1.63e-10	
Amperage	0.0031197	0.0013	
Random effects coefficients	**Estimate**	**Lower**	**Upper**
Voltage	0.079368	0.0468	0.13453
Feed speed	0.0018226	5.44e-78	6.10e+71
Displace Speed	0.0018225	1.00e-67	3.31e+61

Source: Own elaboration

Table 3. Crown height Fit Results.

Fixed effects coefficients	**Estimate**	**P Value**	
Intercept	2.1604	0.3841	
Amperage	-0.0119	0.5552	
Random effects coefficients	**Estimate**	**Lower**	**Upper**
Displace Speed	0.0002	4.67e-22	1.18e+14
Displace Speed* Feed Speed	0.5404	0.1449	2.0142
Feed Speed	0.0059		1.44e+34

Source: Own elaboration

Once the usefulness of the adjusted model has been identified, the process is optimized. It is worth mentioning due to the nature of the model, the use of an NSGA-II algorithm is proposed: Random values were generated in the random covariates of the process (Voltage, displacement speed, and wire feed speed), optimization profiles were set and At the same time that the algorithm was giving ideal amperage values for each generation and in turn identify which amperage parameter values would be the best to obtain a quality penetration and crown height based on the restriction data that was offered to said algorithm (penetration 3.4 mm and Crown height 0 mm). A region of generations with acceptable data was obtained. The objective of the optimization was

to achieve a penetration of 3.4 mm since the specification according to the ASME code mentions that there must be a total penetration of the thickness of the container to avoid leaks of the contained liquid. For the crown, as an objective it should be preferably null, to avoid defects in the weld bead of the container. A table was established with maximums and minimums of the data of the best generations obtained by the algorithm, the results are shown in Table 4.

Table 4. Optimal region NSGA-II.

		Amperage	Voltage	Application Speed	Forward Speed	Penetration	Crown
Optimal Parameters	High	127	15.7	1.9830	1.0239	3.5953	0.0072
	Least	125	15.4	1.9830	0.7761	3.5938	0.0013
		High	Automatic	Slow	Slow		

Source: Own elaboration.

Finally, with the optimal parameters established, these data were validated with the help of a Monte Carlo simulation: 5,000 validation runs were performed, randomly generating Voltage values between 15.4 and 15.7, Application Speed of 1.98 and Advance Speed between 0.77 and 1.02. With the obtained fusion values, a confidence interval was performed, which establishes that with 95% confidence, penetration values between 3.5938 mm and 3.5953 mm and a crown between 0.0013 and 0.0072 will be obtained if the parameters of the GTAW process are maintained. within established limits and welds at a slow speed and applies filler material slowly.
In the analysis of the results, it was observed that the application speed has a greater effect on the crown and is a very sensitive parameter in the process since its value variation is little. As a conclusion of the Monte Carlo simulation runs, it was shown that when the application speed and the forward speed are decreased, the crown height sold percentage decreases and the penetration percentage is closer to the objective value of 3.4 mm. A summary of the simulation is shown in Table 5.

Table 5. Summary of the simulation.

Penetration			Crown height			Amperage	Voltage	Application Speed	Forward Speed
3.5799	3.6907	3.7961	0	0.0121	0.2986	125.1	15.415	1.9831	0.7880
3.5869	3.6928	3.7987	0	0.0380	0.3336	126.02	15.548	1.9831	0.8993
3.5939	3.695	3.8016	0	0.0637	0.3684	126.9	15.685	1.9831	1.0107

Source: Own elaboration.

5 Conclusions.

It is common for the food industry to work empirically, without defining parameters or variables adequately. Obtaining a reliable model with solid mathematical bases can be the solution to problems in welded joints, reducing costs, times and obtaining quality final products. In this case study, a model was obtained to represent the behavior of the GTAW welding process considering random covariates; therefore, within the industry it is difficult to control due to process situations. In addition to modeling the GTAW welding process with random covariates, an optimization profile is presented, in this

case several profiles can be presented, as required. In the proposed profile, a confidence interval is shown that establishes with 95% confidence that penetration values will be obtained between 3.5938 mm and 3.5953 mm, very close to the thickness measurement of the welded plates that make up the container themselves. Are within the specifications of the ASME code, if the parameters of the GTAW process are kept within the established limits and a Penetration greater than or equal to 3.4 mm is presented. And a Crown less than or equal to 0. It was clearly observed and based on the model, the high correlation that the penetration and crown presents with the random variables advance speed and application speed. If the advance speed and the application speed become faster, especially the latter, the penetration percentage increases, showing over penetration and the crown percentage also increases and a low welding quality is obtained. As future work, more data could be obtained from the welding process and modeling with a larger number of samples, so that the model is improved in terms of prediction, and it could also be modeled with a greater number of random variables, such as the angle of the torch at the time of welding to the existing climate at the time of welding.

References

1. Chaki, S., Shanmugarajan, B., Ghosal, S., Padmanabham, G.: Application of integrated soft computing techniques for optimization of hybrid CO2laser–MIG welding process. Applied Soft Computing. Volume 30, Pages 365–374, May (2015).
2. Correa C., & Bolaños R. (2008). Algorímo multiobjetivo NSGA-II aplicado al problema de la mochila. Scientia et Technica Año XIV, No39. Universidad Tecnológica de Pereira.
3. Gang Zhang, Yu Shi, YuFen Gu, Ding Fan.: Welding torch attitude-based study of human welder interactive behavior with weld pool in GTAW. Robotics and Computer Integrated Manufacturing 48, 145 – 156, March (2017).
4. Montgomery, D. C., Peck, E. A., Vining, G. G.: Introducción al análisis de regresión lineal. México: Grupo editorial patria. (2007).
5. Peñuela C, & Granada E. (2007). Optimización multiobjetivo usando un algorítmo genético y un operador elitista basado en un ordenamiento no-dominado (NSGA-II). Scientia et Technica Año XIII, No35. Universidad Tecnológica de Pereira.
6. Sathiya, P., Aravindan, S., Jeyapaul, P. M., Ajith, A., Haq, N.: Optimizing the weld bead characteristics of super austenitic stainless steel (904L) through grey-based Taguchi method. Multidiscipline Modeling in Materials and Structures. Volume: 6 Issue: 2. (2010).
7. Tarng, T. S., Tsai, H. L., Yeh, S. S.: Modeling, optimization and classification of weld quality in tungsten inert gas welding. International Journal of Machine Tools and Manufacture. Vol 39, Issue 9, Sep 1999, Pages 1427–1438. (1999).

15

Turisteando Ando: una App con Kernel Evolutivo

MISAEL LOPEZ SANCHEZ, ANABEL MARTINEZ VARGAS, MA. COSÍO LEÓN Y JUAN PABLO GARCÍA VÁZQUEZ

Turisteando Ando: una App con Kernel Evolutivo

Misael Lopez-Sanchez[1][0000−0002−0097−5022], Anabel Martínez-Vargas[1][0000−0001−6885−3833], MA. Cosío-León[1][0000−0003−4407−1495], and Juan Pablo García-Vázquez[2][0000−0003−1787−7223]

[1] Universidad Politécnica de Pachuca
misael.lopez@micorreoupp.edu.mx,
{anabel.martinez, ma.cosio.leon}@upp.edu.mx
http://www.upp.edu.mx
[2] Facultad de Ingeniería, MyDCI, Universidad Autónoma de Baja California (UABC), Mexicali
pablo.garcia@uabc.edu.mx
http://www.uabc.mx

Resumen El turismo se ha convertido en un motor de desarrollo económico y cultural, representando en México una de las principales actividades económicas. Esto ha generado la necesidad de desarrollar una industria turística más sofisticada y sostenible, que se adapte a la actual sociedad digital. En esta dirección los avances en tecnologías de la información y comunicaciones, particularmente las aplicaciones móviles, revolucionan la forma en como las personas consumen servicios turísticos. En este trabajo presentamos Turisteando Ando, una aplicación móvil para la planificación de viajes turísticos personalizados con base en el contexto y preferencias personales. Para ello utiliza un planificador optimizado para dispositivos móviles, basado en un algoritmo genético para maximizar la satisfacción del turista en su viaje, adhiriéndose a la limitación de tiempo impuesta por el mismo. Los resultados indican que los lugares seleccionados por Turisteando Ando en el itinerario personalizado son aquellos mejor valorados por los turistas que previamente los han visitado.

Keywords: Turismo · Algoritmo Genético · Itinerario · Crowdsourcing.

1. Introducción.

El sector turístico ha experimentado una continua expansión y diversificación. Las estadísticas de la Organización Mundial del Turismo muestran que en el 2018 tuvo un crecimiento del 6 % a nivel mundial con respecto al año anterior [3]. Lo que repercutió positivamente en el crecimiento y desarrollo económico de los países; por ejemplo, en México el Instituto Nacional de Estadística y Geografía (INEGI), indicó que en el 2018 la actividad turística aportó un 8.7 % del producto interno bruto (PIB) y que generó 2.3 millones de puestos de trabajo [8].

El continuo crecimiento de este sector refuerza aún más la necesidad de desarrollar una industria turística más sostenible, que pueda adaptarse a las necesidades de una sociedad digital y que pueda hacer frente a flujos turísticos sin

poner en peligro los recursos naturales y culturales o la calidad de vida de las comunidades anfitrionas. Las nuevas tecnologías de la información y comunicación (TIC) han sido consideradas como el elemento clave para lograrlo; específicamente los dispositivos móviles, debido a su gran aceptación e incorporación a la vida cotidiana de las personas [1] y al éxito de su aplicación en diversas industrias [9,2].

En la industria turística las aplicaciones móviles han estado revolucionando la manera en como se ofrecen los servicios a los viajeros [10]. Esto ha cambiado la forma en que los viajeros planifican su viaje (itinerario) y el disfrute del mismo. El itinerario de un viaje es un plan para la programación de actividades. El estilo y contenido de un itinerario de viaje depende de su naturaleza laboral o personal pero ambos tipos de viajes tiene algo en común, los recursos son limitados en tiempo y dinero. Así optimizar el itinerario del viajero de tal forma que maximice la experiencia en tanto que se acota a las restricciones antes citadas es un nuevo reto del sector turístico [16,10].

Este trabajo se enfoca en la resolución del problema del *Diseño de Itinerarios Turísticos Personalizados* (por su sigla en inglés, TTDP) [5]. Para ello, se presenta una aplicación móvil llamada *Turisteando Ando* (TA) que basa su diseño en el problema de la mochila resuelto con un algoritmo genético. Su resultado es un conjunto de lugares donde el turista tiene la flexibilidad de seleccionar cuales visitar.

El resto de este artículo se organiza y presenta de la siguiente manera. La Sección 2 muestra el trabajo relacionado. La Sección 3, describe la arquitectura del sistema de TA. La Sección 4 muestra los resultados. Finalmente, la Sección 5, presenta las conclusiones del trabajo realizado y los posibles nichos de investigación.

2. Trabajo Relacionado

Los trabajos en la literatura para la generación de itinerarios en su mayoría corresponde a generadores de rutas que maximizan la cantidad de Puntos de Interés (PdI) visitados por jornada. Un PdI, es un lugar de interés con una ubicación definida. Por ejemplo, en el trabajo [7] se analiza el impacto del aspecto de los caminos que conectan los PdI en una ruta turística, y como esto influye en la decisión de los turistas al elegir un camino para llegar a un PdI. Se concluye que es relevante crear itinerarios con esta característica más allá de solo buscar el camino más corto. Mientras que en [6] se explica el uso del problema de la mochila para resolver el TTDP, cuando los costos por viajar entre los PdI se desconocen. El resultado es un conjunto de sitios sin un orden, pero si es necesario tener una ruta, entonces se sugiere utilizar un solver del problema del agente viajero para que la genere considerando los PdI seleccionados. En el trabajo [12] se resuelve el TTDP mediante el problema de Orientación Capacitada de Equipo (por su sigla en inglés, CTOP) y un conjunto de casos de prueba sintéticos. Esto mediante un algoritmo propuesto por los autores llamado Algoritmo Evolutivo Diferencial basado en Distancia (por su sigla en inglés, DRDE). El DRDE es una

combinación del Algoritmo Evolutivo Diferencial (por su sigla en inglés, DE) y un proceso de codificación/decodificación llamado la Distancia Relacionada (por su sigla en inglés, DR) que es la principal aportación de este trabajo. Otra propuesta identificada para abordar el TTDP en la literatura es propuesto por [12] en donde se propone utilizar el CTOP para formar un conjunto de itinerarios, que representen una visita secuencial de nodos. Cada itinerario está limitado por la capacidad del vehículo y el tiempo total de viaje.

3. Arquitectura del Sistema.

La arquitectura de TA contiene tres módulos internos: *vistas, gestor de planes de viaje y el planificador basado en un algoritmo genético.* TA reúne conocimiento del perfil de turista que incluye: tiempo de viaje, ubicación y preferencias personales. A partir de ese conocimiento, TA solicita información a fuentes de datos externas sobre los PdI que cumplan con las características indicadas por el turista. Posteriormente, utiliza la información recabada para generar el plan de viaje personalizado, a través de un algoritmo evolutivo. Finalmente, el plan de viaje personalizado es mostrado al turista. En las subsecciones siguientes se describen con más detalle cada módulo.

3.1. Vistas.

Las vistas es la interfaz gráfica de usuario (por su sigla en inglés, GUI). A través de la GUI, el turista interacciona con TA. Esta contiene los controles y elementos suficientes para que el turista pueda ingresar la información solicitada por TA. También mediante la GUI se le muestra al usuario el plan de viaje personalizado.

3.2. Gestor de Planes de Viaje.

Este módulo prepara la información que será usada para generar el plan de viaje personalizado. Recibe y acondiciona los datos que componen el contexto del turista; y recopila en servicios en línea la información de los PdI. Se necesitan tres datos fundamentales del turista:

- *Duración del viaje:* Cantidad de tiempo que durará el viaje. El usuario la puede ingresar en días u horas. Si es ingresado en días, se transforma en una jornada de 12 horas, y posteriormente se realiza la conversión a segundos; por otro lado, si es ingresado en horas, solo se realiza la conversión a segundos. La conversión a segundos se realiza dado que esa unidad de medida es usada por TA.
- *Ubicación del turista:* Es la ubicación del turista que tomará TA para recolectar la información de los PdI.
- *Preferencias personales:* Es el tipo de lugares que le interesan visitar al turista, los cuales están clasificados por parques, museos, zoológicos y sitios históricos.

Otra parte importante que controla este módulo, son los procesos para recolectar y acondicionar la información referente a los PdI. La información de los PdI, se recolecta de fuentes de datos en línea como *Places API* de Google y *Directions API* de MapBox. TA, en el primer paso, consulta Places API para recuperar el conjunto de los PdI y sus propiedades, que se encuentren a un máximo 30 km radiales desde la ubicación registrada por el turista. *Places API* mantiene una base de datos creada y alimentada por *crowdsourcing*, es decir, los usuarios añaden y actualizan la información de los PdI con fotos, opiniones, votos para calcular la valoración (o *rating*) de cada PdI. Después, se consulta *Directions API*, para calcular el tiempo aproximado de desplazamiento desde la ubicación registrada por el turista, hasta cada PdI.

3.3. Planificador basado en un Algoritmo Genético.

El TTDP considera un conjunto de los PdI con sus respectivas características como ubicación, tipo, valoración y duración de la visita. Entonces, la solución consta de una lista de los PdI que el turista puede recorrer durante su viaje, maximizando su satisfacción sin violar las restricciones establecidas.

Para resolver el TTDP, este trabajo considera el modelo de optimización del *Problema de la Mochila* (por su sigla en inglés, KP) [13]. El KP trata con un conjunto de artículos, cada uno con un peso y una ganancia. Entonces la tarea es incluir artículos en la mochila para maximizar su ganancia (o beneficio) sin exceder su peso máximo (o capacidad). El modelo de optimización del KP es como sigue:

$$Maximizar \;\; P(x) = \sum_{i=1}^{n} x[i] \cdot P[i] \tag{1}$$

s.a.

$$\sum_{i=1}^{n} x[i] \cdot W[i] \leq C \tag{2}$$

$$x \in \{0, 1\}, i = 1, 2, ..., n \tag{3}$$

La ecuación (1) es la función objetivo que maximiza el beneficio, donde $x[i] = 1$, si el objeto i es incluido en la mochila, de otro modo $x[i] = 0$. $P[i]$ es el vector de ganancias. Por otro lado, la ecuación (2) es la restricción de peso máximo de la mochila donde $W[i]$ es el vector de pesos y C es la capacidad máxima de la mochila.

Para adaptar el KP al contexto del TTDP, se considera a un turista quien tiene que preparar su itinerario (mochila) para su viaje y tiene que decidir por los PdI (objetos) a incluir en el. El turista, cuenta con un conjunto de los PdI disponibles para visitar. Cada PdI cuenta con una valoración (beneficio) y un tiempo (peso). Para este último, se considera el tiempo de desplazamiento de la ubicación del usuario a la ubicación del PdI, y viceversa; así como el tiempo que se necesita para disfrutarlo. El problema consiste en elegir un subconjunto de los PdI para maximizar su experiencia de viaje, sin exceder su tiempo de estancia (capacidad de la mochila).

Para encontrar el vector binario $x = \{x[1],\ldots,x[n]\}$ que represente el mejor itinerario personalizado, se aplica un algoritmo genético estacionario (AGE) [11]. Un AGE es una versión simplificada del algoritmo genético clásico y por lo tanto menos costosa. Esta última característica, hace adecuado su implementación en teléfonos inteligentes. Un AGE genera uno o dos hijos en cada generación.

El AGE implementado en TA se muestra en Algoritmo 1. Se inicializa la población aleatoriamente en el PASO 1. Enseguida en el PASO 3 se seleccionan dos individuos aleatoriamente de la población para competir con base en su aptitud, el mejor de ellos se convierte en padre. Se repite el mismo proceso para seleccionar al segundo padre. Las soluciones no factibles son reparadas con el método aleatorio descrito en [15]. En el PASO 4, si se cumple la probabilidad de cruce entonces los dos padres se cruzan por el método de dos puntos [4] para generar dos hijos. Los hijos resultantes se mutan por el método de cambio de bit [4] como se muestra en el PASO 10. Entonces, en el PASO 11 se decide quienes serán los dos sobrevivientes de la siguiente iteración. Para ello, de los 4 individuos (2 padres y 2 hijos), se selecciona al mejor individuo con base en su aptitud. Este es el primer sobreviviente. Enseguida para seleccionar al segundo sobreviviente, de los 3 individuos restantes, se selecciona el que tenga la mayor distancia del mejor individuo [14]. El mejor individuo encontrado a lo largo de las iteraciones es la solución al problema del TTDP.

Algorithm 1: Algoritmo genético estacionario para resolver el TTDP

Data: Tamaño de población, probabilidad de cruce, probabilidad de mutación, y número de iteraciones. Tiempo de estadía, vector de ganancias (valoraciones), y vector de pesos (tiempo de los PdI).
Result: El mejor itinerario personalizado

```
1  Generar población aleatoriamente;
2  repeat
3  |  Seleccionar dos padres mediante método de torneo;
4  |  if numero_aletorio <= probabilidad_cruce then
5  |  |  Cruzar padres por método de dos puntos;
6  |  else
7  |  |  hijo_1 = padre_1;
8  |  |  hijo_2 = padre_2;
9  |  end
10 |  Mutar los dos hijos con el método cambio de bit;
11 |  Usar el método de reemplazo de selección familiar correlativa;
12 |  Seleccionar al mejor individuo de la población y guardarlo;
13 until numero_de_iteraciones > numero_total_de_iteraciones;
```

4. Resultados.

TA se desarrolló bajo el sistema operativo Android, y se codificó en Java, usando Android Studio (versión 3.5.3) como entorno de desarrollo integrado. Así, las ejecuciones se llevaron a cabo usando un teléfono inteligente con procesador Qualcomm Snapdragon 625 Octa-core de 2.0 GHz, con 3 GB de RAM y con un sistema operativo MIUI 9.5 basado en Android 7.1.2 Nougat. Respecto a la configuración del AGE, se usó una de población de 100 individuos. Para el métodos de selección por torneo se usaron 2 contrincantes. La probabilidad de cruce fue de 0.65 y la probabilidad de mutación fue de .01. El algoritmo se detiene al alcanzar las 500 iteraciones. Se realizaron 30 ejecuciones.

Para realizar las ejecuciones, se creó una instancia basada en datos obtenidos de las herramientas Places API y Directions API. La instancia se compone de 236 elementos. Los datos de los PdI se obtuvieron usando como ubicación del usuario las coordenadas: (19.435726, -99.143893), y se incluyeron los cuatro tipos de lugares (parques, museos, zoológicos y sitios históricos) para la búsqueda. La capacidad de la mochila (duración del viaje), se definió en un día (jornada de 12 horas), la cual se transformo a 43200 segundos.

Los resultados de las ejecuciones se reportan en la Tabla 1 donde id es el número de ejecución. f es la aptitud de la mejor solución encontrada. x es la suma de los pesos (en segundos) de los PdI incluidos en la mochila (plan de viaje personalizado). n es el número de lugares incluidos en el plan de viaje personalizado. Finalmente, t es es el tiempo (en milisegundos) que tardo en ejecutarse el AGE.

Los resultados de los experimentos muestran que en promedio, la calidad de las soluciones es de 42.9 puntos, incluyendo de 9 a 10 lugares. Esto indica que las soluciones en su mayoría incluyen lugares con puntuación alta (de 4 a 5 puntos), lo que se traduce en una mayor satisfacción del usuario. Con respecto a la suma de los pesos, el promedio es de 41435 segundos. Si se considera que la duración del viaje que se definió en las ejecuciones es de 43200 segundos; esto indica que en su viaje, el turista siempre tendrá actividades que realizar. De esta forma, aprovecha la mayoría de su tiempo. Respecto al tiempo de ejecución del AGE, en promedio, el plan de viaje personalizado se genera en 23734 milisegundos (menos de 30 segundos). Esto es un tiempo aceptable y se traduce en eficiencia de la aplicación para proveer de forma eficaz el servicio de creación de planes de viajes personalizados para los turistas.

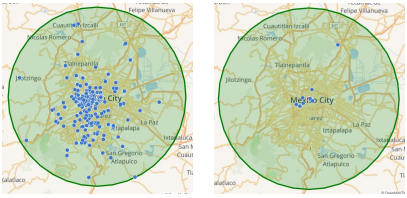
La Figura 1 representa la ejecución que tuvo la aptitud más alta de las ejecuciones reportadas en la Tabla 1. Esta fue la ejecución 30 con 46.3 puntos. De un total de 236 lugares alrededor del punto de origen (ver Figura 1a), el AGE obtiene el plan de viaje personalizado que cuenta con un total de 10 lugares (ver Figura 1b).

5. Conclusiones y Trabajo Futuro.

El AGE facilita el proceso de filtrar la información turística, y seleccionar las atracciones para ofrecer una alternativa de recorrido turístico que se adecue a las

Tabla 1: Ejecuciones realizadas para Turisteando Ando. Fuente: Elaboración propia.

id	*f*	*x*	*n*	*t*	*id*	*f*	*x*	*n*	*t*	*id*	*f*	*x*	*n*	*t*
1	41	42385.4	9	24403	11	44.5	42916	10	23482	21	40.8	41364.4	9	22650
2	41.4	41976.2	9	23463	12	40.2	40048	9	24625	22	45.5	41838	10	23201
3	44.3	40436.2	10	23967	13	41.6	42014.4	9	23738	23	42	42866	9	23549
4	45.6	38708.4	10	23880	14	40.9	40124.4	9	23901	24	45.6	41767.4	10	24015
5	44.6	42134.6	10	23737	15	41.6	42072.6	9	23424	25	45.1	40804.6	10	23885
6	40.8	41619.4	9	23169	16	42.1	42971	9	23888	26	44.6	42317	10	23440
7	40.8	43088.2	9	23736	17	40.7	42352.8	9	24395	27	41.8	39359.8	9	24192
8	44.1	40513.4	10	23706	18	45.5	41760.6	10	23518	28	41.1	41418.6	9	22961
9	41.2	41256.6	9	23470	19	41.5	41842	9	23711	29	45	41405	10	24100
10	40.4	41231.8	9	23373	20	45.7	38137.2	10	24158	30	46.3	42318.4	10	24293



(a) Escenario inicial, antes de ejecutar el AGE. Fuente: Elaboración propia.

(b) Escenario final, después de ejecutar el AGE. Fuente: Elaboración propia.

Figura 1: Escenarios de la mejor solución encontrada. Fuente: Elaboración propia.

necesidades y preferencias del turista. TA, modifica la forma como las personas consumen servicios turísticos, ya que desde cualquier ubicación mientras tengan conectividad a internet puede ser generado un plan de viaje personalizado.

En lo que respecta a trabajo futuro, se planea realizar un estudio piloto de la usabilidad y aceptación a la aplicación. Considerando invitar a sujetos (usuarios) que viajan frecuentemente. Por otro lado, se contemplan modificaciones en el módulo gestor de planes de viaje, para que TA seleccione los PdI con conciencia del contexto y de políticas de salud.

Referencias

1. Ansari, M.S., Channar, Z.A., Syed, A.: Mobile phone adoption and appropriation among the young generation. Procedia-Social and Behavioral Sciences **41**, 265–272 (2012)

2. Bidin, S., Ziden, A.A.: Adoption and application of mobile learning in the education industry. Procedia-social and behavioral sciences **90**, 720–729 (2013)
3. Council, W.T.O.: Global and regional tourism performance. url https://www.unwto.org/global-and-regional-tourism-performance (2020)
4. Eiben, A.E., Smith, J.E.: Introduction to Evolutionary Computing. Springer Publishing Company, Incorporated, 2nd edn. (2015)
5. Gavalas, D., Kasapakis, V., Konstantopoulos, C., Pantziou, G., Vathis, N.: Scenic route planning for tourists. Personal and Ubiquitous Computing **21**(1), 137–155 (2017)
6. Herzog, D., Dietz, L.W., Wörndl, W.: Tourist trip recommendations—foundations, state of the art and challenges. Personalized Human-Computer Interaction pp. 159–182 (2019)
7. Herzog, D., Sikander, S., Wörndl, W.: Integrating route attractiveness attributes into tourist trip recommendations. In: Companion Proceedings of The 2019 World Wide Web Conference. pp. 96–101 (2019)
8. INEGI: Estadísticas a propósito del día mundial del turismo datos nacionales. url https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2019/turismo2019$_N al.pdf$(2020)
9. Kalem, G., Turhan, Ç.: Mobile technology applications in the healthcare industry for disease management and wellness. Procedia-Social and Behavioral Sciences **195**, 2014–2018 (2015)
10. Kim, D., Kim, S.: The role of mobile technology in tourism: Patents, articles, news, and mobile tour app reviews. Sustainability **9**(11), 2082 (2017)
11. Lozano, M., Herrera, F., Cano, J.R.: Replacement strategies to preserve useful diversity in steady-state genetic algorithms. Information Sciences **178**(23), 4421 – 4433 (2008). https://doi.org/https://doi.org/10.1016/j.ins.2008.07.031, `http://www.sciencedirect.com/science/article/pii/S0020025508002867`, including Special Section: Genetic and Evolutionary Computing
12. Marinaki, M., Marinakis, Y., Matsatsinis, N.: A novel solution encoding in the differential evolution algorithm for optimizing tourist trip design problems. In: Learning and Intelligent Optimization: 13th International Conference, LION 13, Chania, Crete, Greece, May 27–31, 2019, Revised Selected Papers. vol. 11968, p. 253. Springer Nature (2020)
13. Martello, S., Toth, P.: Knapsack Problems: Algorithms and Computer Implementations. John Wiley Sons, Inc., USA (1990)
14. Matsui, K.: New selection method to improve the population diversity in genetic algorithms. In: IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028). vol. 1, pp. 625–630 vol.1 (1999)
15. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs (3rd Ed.). Springer-Verlag, Berlin, Heidelberg (1996)
16. Yang, Y., Duan, Y., Wang, X., Huang, Z., Xie, N., Shen, H.T.: Hierarchical multi-clue modelling for poi popularity prediction with heterogeneous tourist information. IEEE Transactions on Knowledge and Data Engineering **31**(4), 757–768 (2018)

16

PSO in Ridge Regression to eliminate multicollinearity in a PTA welding process

JORGE SAÚL RODRÍGUEZ VÁZQUEZ, ROLANDO JAVIER PRAGA ALEJO AND DAVID SALVADOR GONZÁLEZ GONZÁLEZ

PSO in Ridge Regression to eliminate multicollinearity in a PTA welding process.

Jorge Saúl Rodríguez Vázquez[1], Rolando Javier Praga Alejo[1-2] and David Salvador González González[1-2]

[1] Corporación Mexicana de Investigación en Materiales (COMIMSA) Calle Ciencia y Tecnología #790, Col. Saltillo 400 C.P. 25290, Saltillo, Coahuila, México. Phone: (+52) 01 844 411 32 00.
[2] Facutad de Sistemas, Universidad Autónoma de Coahuila, Ciudad Universitaria, Carretera a México Km. 13, Arteaga, Coahuila, México. Phone: (+52) 01 844 689 10 30.

Abstract.
Nowadays industrial processes present variability in their processes making their control difficult. Linear regression is a statistical tool that can solve this problem, unfortunately if the process variables are highly related, the model obtained by Ordinary Least Squares (OLS) is not suitable to control or predict. This condition is called multicollinearity. Fortunately, there are statistical metrics capable of detecting linear dependence, such as the variance-covariance matrix, the VIF (Variance Inflation Factors) and the condition number. Ridge Regression (RR) is a method that eliminates the problem of multicollinearity. The basic idea of RR is generate a parameter of bias k that counteracts the dependency between the variables. There are methods that provide the value of bias k but affect model fit. For this reason, other alternatives should be choose. Therefore, in this article, a global optimization was performed applying the PSO metaheuristic on the parameter of bias k, to obtain the value that eliminates multicollinearity without affecting the fit of the model, taking the coefficient of determination R^2 as an objective function. The optimization was applied to a case study, the results were contrasted against the Ridge Regression method, obtaining better results for the Particle Swarm Optimization algorithm proposed in this article.

Keywords: PSO, OLS, Ridge Regression, Multicollinearity, VIF, Optimization.

1 Introduction

Regression remains a widely used statistical technique to model and predict industrial processes [1, 2]. There are different methods that can be used to obtain the regression model estimators, depending on the characteristics of the process. If the variables do not have collinearity, the method of Ordinary Least Squares (OLS) is used, however in most of the processes the variables have linear dependence and alternative methods have to be chosen [3,4] to generate adequate estimators of the regression model.

Ridge Regression (RR) is a viable option to use to treat multicollinearity, this method proposed by Hoerl et al. [5] has been applied in other investigations to eliminate the linear dependence between the input variables [6,7] by means of a parameter of bias k that eliminates the collinearity problem. This value can be obtained using the technique called Ridge trace or an iterative method proposed by Hoerl et al. [8] Currently, there is recent research that has modified the iterative method to calculate the bias value, obtaining good results to eliminate multicollinearity [9, 10, 11]. However, although the previous articles work on this problem, none of them proposes an alternative to avoid reducing the model fit, which can be measured with R^2 statistic.

In addition, optimization is a tool that has been used in various investigations (see Strömberg [12], Munner et al. [13]) in order to find the values of the input variables that generate the optimal solution for the objective function. However, there are problems for which one cannot guarantee to find an optimal solution in a reasonable time and these are classified (according to the theory of computational complexity) as "difficult". When a difficult problem is addressed, nature-inspired algorithms such as PSO (Particle Swarm Optimization), produce good results (Zadeh et al. [14], BharathiRaja et al. [15], Liu et al. [16]) because they only evaluate a part of the solutions and not the whole set of solutions.

Therefore, ones of the objectives of this work are to optimize the choice of the bias K matrix parameters that eliminate multicollinearity without affecting too much the statistical metric R^2.

2 Methods

2.1 Multiple linear regression models

The multiple linear regression model is used when the response variable y depends on a set of return variables X, by matrix notation the model is expressed as:

$$y = X\beta + \varepsilon \quad (1)$$

Where y is a vector of nx1 observations, X is a matrix of nxp regressors, β is a vector of px1 regression coefficients, ε is a vector of nx1 random errors that are $NID(0, \sigma^2)$. To solve the regression model using Ordinary Least Squares (OLS) the coefficients of β are obtained with the equation:

$$\hat{\beta}_{MC} = (X'X)^{-1}X'y \quad (2)$$

If the columns of the matrix X are linearly independent then the inverse matrix $(X'X)^{-1}$ will exist, however when the regressors have a strong linear relationship, the inference of the estimators by OLS is wrong. This matrix problem is also called multicollinearity.

2.2 Multicollinearity

Multicollinearity can be defined in terms of the linear dependence between the columns of the matrix X considering that their column vectors are $[X_1, X_2, \dots, X_p]$ so the equation can be established as follows $\sum_{j=1}^{p} t_j X_j \approx 0$. This implies that the matrix $X'X$ will be

poorly conditioned because its determinant will be close to zero $|X'X| \approx 0$. Therefore, the method of Ordinary Least Squares will produce poor estimators of the individual parameters of the model. A serious problem involving multicollinearity is reflected in the hypothesis tests of the model, since the Type 1 error (reject H_0 when it should not be made) or the Type 2 error (accept H_0 when it should not be).

2.3 Multicollinearity detection

There are several techniques to detect multicollinearity, Variance Inflation Factors and the analysis of the eigensystem are the best diagnoses available today because they are easy to calculate, direct interpretation and useful to investigate the specific nature of multicollinearity [17].

1. Variance Inflation Factors: For each term of the model the VIFs measure the combined effect that the dependencies have between the regressors on the variance of each term, the Variance Inflation Factors are defined as:

$$VIF_j = \frac{1}{(1 - R_j^2)} \tag{3}$$

R_j^2 is the coefficient of determination obtained when the regression is made x_j compared to other $p - 1$ regressors. If the VIF obtained are higher than 10 there is multicollinearity.

2. Condition Number: If one or more eigenvalues are very small, it implies that there is almost linear dependence between the columns of the matrix. The condition number that is defined as:

$$\eta = \frac{\lambda_{max}}{\lambda_{min}} \tag{4}$$

Where λ_{max} is the maximum eigenvalue and λ_{min} is the minimum eigenvalue of the $X'X$ matrix. If $\eta > 1000$ it indicates strong multicollinearity.

2.4 Ridge Regression.

A review of the literature indicates that there is a wide range of research related to Regression Ridge (see [18,19]), this method proposed Hoerl and Kennard in 1970 suggests adding a parameter bias $k > 0$ in the diagonal of the matrix $X'X$ to eliminate multicollinearity coefficients β of Ridge Regression are obtained by solving:

$$\hat{\beta}_R = (X'X + kI)^{-1}X'y \tag{5}$$

The parameter of bias k can be obtained by means of the Ridge trace, this technique is subjective because a value must be chosen large enough to stabilize the coefficients of the model, but not unnecessarily large so as not to introduce too much bias affecting the Mean Square Error (MSE) obtained by the following equation:

$$MSE(\hat{\beta}_R) = \sigma^2 \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2\beta'(X'X + kI)^{-2}\beta \tag{6}$$

Where λ_j are the eigenvalues of the matrix $X'X$, the first term of the equation is the sum of the variances of the parameters in $\hat{\beta}_R$ and the second term is the square of the bias, from Eq. (7) it is observed that if $k > 0$ the bias in $\hat{\beta}_R$ increases, however the variance decreases. The covariance matrix of $\hat{\beta}_R$ is:

$$Var(\hat{\beta}_R) = \sigma^2(X'X + kI)^{-1}X'X(X'X + kI)^{-1} \quad (7)$$

Because Ridge Regression focuses on the choice of the bias parameter and since the Ridge trace procedure is considered subjective (because criteria are required for its choice), one analytical method that determines the k-value for the Ridge Regression is proposed in [8] as follows:

$$k = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \quad (8)$$

From Eq. (8) $\hat{\beta}$ and $\hat{\sigma}^2$ are determined from the Ordinary Least Squares solution. In a later publication [20] proposed an iterative procedure based on Eq. (8) the algorithm stops when $(k_{j+1} - k_j)/k_j \leq 20T^{-1.3}$ where $T = Tr(X'X)^{-1}/p$, at the end of the procedure the bias parameter k_j is used for the Ridge Regression. This procedure has been modified by other authors such as [21] where they compared 26 alterations made to the analytical method by various researchers in recent years.

2.5 Particle Swarm Optimization (PSO).

The Particle Swarm Optimization algorithm was proposed by James Kennedy and Russell Eberhart in 1995, the PSO algorithm is an algorithm in the area of artificial intelligence classified as swarm intelligence, the algorithm is stochastic and inspired by behavior social of some animals [22]. The algorithm allows the solution space to be explored in multiple directions, simultaneously, avoiding local minima.

The algorithm starts with a group of random particles (solutions) and then searches for optimal solutions by updating the generations. In each iteration, each particle is updated with two "best" values. The first value called pbest is the best solution you have achieved so far. The other "best" value called gbest is the best overall in the entire swarm. After finding the best two values, the particles update their velocity and positions applying equations (9) and (10).

$$v_i^{k+1} = w * v_i^k + c1 * rand1 * (pbest_i^k - x_i^k) + c2 * rand2 * (gbest_i^k - x_i^k) \quad (9)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (10)$$

Where v_i^{k+1}, v_i^k are the velocity of the particle in iteration $k+1$ and k respectively. Also x_i^{k+1}, x_i^k are the position of the particle in iteration $k+1$ and k respectively. Furthermore $rand1$, $rand2$ are random numbers evenly distributed between (0,1). The coefficients $c1$ and $c2$ indicate the degree of confidence in the best position found by the individual particle (cognitive parameter) and that of the entire swarm (social parameter), respectively, w is the weight of inertia used to achieve a balance in The exploration and exploitation of the search space and plays a very important role in the convergence behavior of PSO. The inertia weight is dynamically reduced from 1 to about 0 in each generation according to (11):

$$w_i = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}} * iter \quad (11)$$

Where $iter_{max}$ is the maximum number of iterations and $iter$ is the current number of iterations, w_{max} and w_{min} are the maximum and minimum inertia weight values. The position of each particle is updated using its velocity vector as shown in Eq. (10).

3 Application.

A PTA (Plasma Transfer Arc) welding process is modeled where 4 input variables are involved and the relationship against an output variable is studied. Table 1 shows the 33 observations of the process:

Table 1. Observations of the welding process PTA.

x_1	x_2	x_3	x_4	y
Powder feed rate [%]	Powder feed rate [g / min]	Process speed [cm / min]	Welding Current [Amps]	Area HAZ[mm^2]
70	26	80	60	4.908
60	21	100	60	5.184
50	18	80	80	4.859
60	21	100	80	3.693
60	21	100	70	3.623
70	26	100	70	3.504
70	26	120	80	3.711
60	21	100	70	3.737
70	26	120	60	3.371
50	18	100	70	3.345
60	21	100	70	3.541
60	21	100	70	4.092
50	18	80	60	4.674
60	21	100	70	4.023
60	21	100	70	3.47
60	21	100	70	3.478
50	18	120	80	3.202
50	18	120	60	3.081
60	21	100	70	3.447
60	21	100	70	4.016
60	21	120	70	2.719
60	21	100	70	3.526
70	26	80	80	4.584
60	21	80	70	4.662
60	21	100	50	3.498
60	21	80	50	4.449
60	21	120	50	3.197
50	18	120	50	2.985
60	21	120	50	3.028
70	26	120	50	2.099
70	26	160	50	1.813
70	26	140	50	2.406
70	26	180	50	1.945

Whit process data of Table 1 the bias is calculated to eliminate multicollinearity. Table 2 shows and compare the results of the bias k_i parameters obtained by the three techniques OLS, RR and PSO.

Table 2. Values k_j bias obtained with the three techniques.

OLS	Iterative Method [20]	PSO
$k = 0$	$k = 0.998$	$k = \begin{bmatrix} 0.5459 \\ 0.0708 \\ 0.0100 \\ 0.1522 \end{bmatrix}$

Table 3 shows the results of the Variance Inflation Factors to determine the elimination of multicollinearity. If any VIF value exceeds 10 then the linear dependency is still present among the regressors.

Table 3. VIF obtained based on the values k_j

Element	OLS	RR	PSO
VIF_1	33.77	0.11979	0.0842
VIF_2	34.86	0.11571	0.8571
VIF_3	1.48	0.22863	1.2786
VIF_4	1.27	0.23546	0.8888

Table 4 shows the coefficient of determination R^2 for OLS, RR and PSO. This statistical metric indicates the proportion of variance explained by the model. A value for R^2 close to 1 implies that most of the variability of prediction y is explained by the regression model.

Table 4. Adjusting the model based values k_j

	OLS	RR	PSO
Metric R^2	0.7451	0.4070	0.7367

4 Conclusions.

As shown in Table 2, there is more than one way to obtain the bias parameters that eliminate multicollinearity, so it is necessary to include other metrics that determine which of the techniques is better. Furthermore, considering that by means of the PSO metaheuristic a bias vector is obtained and not just a parameter as in Ridge Regression, it is necessary to indicate what effect it has on the regression.

A first metric considered was the Variance Inflation Factors since they indicate the degree of multicollinearity presented by input variables, as observed in Table 3 when applying Ordinary Least Squares the VIF_1 and VIF_2 have value of 33.77 and 34.86 respectively, this indicates that the data present collinearity problems because the

values should not be exceed than 10. When applying Ridge Regression the four VIF values were less than 10, thus indicating that the bias value k = 0.99812 obtained by the method [20] eliminates multicollinearity between variables. The same occurs with the bias vector generated by PSO, the four VIF are less than 10.

Table 4 shows the results where the coefficient of determination R^2 is considered as a metric. For OLS an R^2 of 74.51% is obtained, the result could be considered adequate, however, in the presence of multicollinearity, the model generated unstable regression estimators. In RR an R^2 of 40.70% is obtained, the percentage is far from 1 for the model to be considered representative of the process, even when multicollinearity has been eliminated. With the PSO method an R^2 of 73.67% is obtained, it is considered a better result than the previous two, because collinearity is eliminated and the model fit is not affected too much.

In conclusion, applying PSO to obtain the parameters of bias k is better than OLS and RR because PSO in addition to eliminating multicollinearity and having a model fit close to 1, more stable regression estimators will be obtained.

References

1. Shengzheng, W. et al. “Predicting ship fuel consumption based on LASSO regression”. Transportation Research Part D, 1-8. (2017).
2. Wang, X. et al. “Empirical models for tool forces prediction of drag-typed picks based on principal component regression and ridge regression methods”. Tunnelling and Underground Space Technology, 75-95. (2017).
3. Demirhan, H. “The problem of multicollinearity in horizontal solar radiation estimation models and a new model for Turkey”. Energy Conversion and Management, 334-345. (2014).
4. Blommaert, A. et al. “Data mining for longitudinal data under multicollinearity and time dependence using penalized generalized estimating equations”. Computational Statistics and Data Analysis, 667-680. (2014).
5. Hoerl, A., & Kennard, R. “Ridge regression:Applications to nonorthogonal problems”. Technometrics, 69-82. (1970b).
6. Liu, H. et al. “Robust modeling method for thermal error of CNC machine tools based on ridge regression algorithm”. International Journal of Machine Tools & Manufacture, 35-48. (2017).
7. Wang, X. et al. “Empirical models for tool forces prediction of drag-typed picks based on principal component regression and ridge regression methods”. Tunnelling and Underground Space Technology, 75-95. (2017).
8. Hoerl, A., Kennard, R., & Baldwin, K. “Ridge Regression: some simulations”. Commun State, 105-123. (1975).
9. Alkhamisi, M., & Shukur, G. “A Monte Carlo study of recent ridge parameters”. Communication in Statistic. Simulation and Computation, 535-547. (2007).
10. Dorugade, A., & D.N., K. “Alternative method for chossing ridge parameter for regression”. Applied Mathematical Sciences, 447-456. (2010).

11. Kibria, G. B. “Performance of some new ridge regression estimators”. Communication in Statistics. Simulation and Computation, 419-435. (2003).
12. Strömberg, N. “Reliability-based design optiization using SORM and SQP”. Structural and Multidisciplinary Optimization, 631-645. (2017).
13. Munner, K. M. et al. “Multiobjective optimization of Nd:YAG direct laser writing of microchannels for microfluidic applications”. International Journal Advanced Manufacture Technology, 1363-1377. (2015).
14. Zadeh, S. et al. “Optimal sliding mode control of a robot manipulator under uncertainty using PSO”. Nonlinear Dyn, 2227-2239. (2016).
15. BharathiRaja, S., & Baskar, N. “Particle swarm optimization technique for determining optimal machining parameters of different work piece materials in turning operation”. International Journal of Advanced Manufacturing Technology, 445-463. (2011).
16. Liu, Y. et al. “Parameter estimation of cutting tool temperature nonlinera model using PSO algorithm”. Journal of Zhejiang University SCIENCE, 1026-1029. (2005).
17. Montgomery, D. C., Peck, E. A., & Vining, G. G. “Introducción al análisis de Regresión Lineal”. México: Continental. (2006).
18. Muniz, G., & Golam, B. “On Some Ridge Regression Estimators: An Empirical Comparisons”. Communications in Statistics-Simulation and Computation, 621-630. (2009).
19. Clark, A., & C.G., T. “Ridge Regression-A Simulation Study”. Communications in Statistics-Simulation and Computation, 605-619. (2006).
20. Hoerl, A., & Kennard, R. “Ridge regression: Iterative estimation of the biasing parameter”. Commun Stat, 77-88. (1976).
21. Wong, K. Y., & Chiu, S. N. “An iterative approach to minimize the mean squared error in ridge regression”. Computation and Statistic, 625-639. (2015).
22. Kennedy, J., & Eberhart, R. “Partice Swarm Optimization”. IEEE International Conference on Neural Networks, 1942-1948. (1995).

17

Redesign of a back-propagation neural network in a manufacturing process

ROCIO NATALI CARRANZA SANTOY, ROLANDO JAVIER PRAGA ALEJO AND DAVID SALVADOR GONZÁLEZ GONZÁLEZ

Redesign of a back-propagation neural network in a manufacturing process

Rocío Natali Carranza-Santoy [1], Rolando Javier Praga-Alejo [1-2] and David Salvador González-González [1-2]

[1]Corporación Mexicana de Investigación en Materiales (COMIMSA) Calle Ciencia y Tecnología # 790, Col. Saltillo 400 CP 25290, Saltillo, Coahuila, México. Phone: (+52) 01 844 411 32 00.
[2]Facultad de Sistemas, Universidad Autónoma de Coahuila, Ciudad Universitaria, Carretera a México Km. 13, Arteaga, Coahuila, México. Phone: (+52) 01 844 689 10 30.

Abstract
One of the primary objectives in the industry is to control and improve some manufacturing processes, in addition to predicting the results, to achieve this, it is necessary to know the relationship between variables, this is achieved through modeling. In the diversity of modeling tools, there are the back-propagation neural networks that are part of artificial intelligence, which are capable of modeling relationships mainly of nonlinear and complex data. This work shows the modifications made to the back-propagation neural network implementing the Gauss-Newton method to model a PTA welding process. A comparison of the back-propagation neural network models is performed with the gradient method and with Gauss-Newton, applying statistical metrics to evaluate the efficiency of the models. The Gauss-Newton back-propagation neural network model is shown to be better than the gradient method, more accurately predicting and representing the PTA welding process.

Keywords: Back-propagation neural network, Gauss-Newton, Statistical metrics

1 Introduction

In the diversity of manufacturing processes is found, the PTA (Plasma Transfer Arc) welding process which is applied for coatings and for the recovery of components; this process presents a difficulty, that the parameters of the PTA welding process are not well known, and this is essential to obtain the desired quality characteristics [1]. In order to improve the process, it is necessary to know how the input variables are related to the quality characteristic (output variable) of the process. Through the use of modeling, this relationship can be known; artificial intelligence, such as artificial neural networks, is capable of modeling complex relationships between inputs and outputs or discovering patterns in the data. Models created by artificial neural networks are methods primarily for complex, nonlinear data systems, where predictions fall within

the range of training data. For this reason, artificial neural networks have been shown to be superior compared to conventional mathematics-based modeling techniques due to their ability to provide high response speed when classifying and approximating large data and noise-containing data [2]. In this document a redesign is made to the back-propagation neural network to improve its prediction capacity, for this modification the Gauss-Newton method is used, at the end comparison is shown by statistical metrics between the back-propagation neural network model with the gradient method and the back-propagation neural network model with Gauss-Newton.

2 Methods

2.1 Neural network back-propagation

Back-propagation neural network has at least three layers of neurons, the first is called the input layer, then there is the hidden layer (it can have more than one hidden layer) and finally the output layer. The essential idea of back-propagation artificial neural network training is to propagate the output layer error to the hidden layers. Since the error of the output layer is the only one that can be calculated exactly, the algorithm proposes to propagate this error backwards to estimate the error in the outputs of the neurons of the hidden layers, in order to modify the interconnection weights of these neurons [3].

Description of the artificial back-propagation neural network:

The vector $x_p = [x_{p1}, x_{p2}, \dots, x_{pi}, \dots, x_{pN}]^T$ is applied to the neurons of the input layer of the network, the input vector is transferred to the following neurons of the hidden layer in which the net input of the jth neuron of this layer is calculated, that is, the neurons of the hidden layer make the summation of the product of the w_{ji}^h so-called interconnection or synaptic weight and the input vector x_{pi}, then a bias θ_j^h is added, as shown in the equation (1):

$$Net_{pj}^h = \sum_i^N w_{ij}^h \; x_{pi} + \theta_j^h \tag{1}$$

Starting with the training of the artificial neural network, the weight of interconnection w_{ij}^h is a random matrix (at the start) of $m \times n$ dimensions. After having calculated the net input of the jth neuron of the hidden layer (eq. 2), the sigmoid activation function is used:

$$f_j^h\left(Net_{pj}^h\right) = \frac{1}{1+e^{-Net_{pj}^h}} \tag{2}$$

To get the output of the jth neuron:

$$i_{pj} = f_j^h(Net_{pj}^h) \tag{3}$$

Once the outputs of the neurons of the hidden layer are calculated, these become input signals for the neurons of the output layer, where the summation of the multiplication of w_{kj}^o interconnection weight and the input i_{pj} is performed, then the bias θ_k^o is added, that is (eq. 4):

$$Net_{pk}^{o} = \sum_{j=1}^{L} w_{kj}^{o}\ i_{pj} + \theta_{k}^{o} \tag{4}$$

In the same way that the weight of interconnection w_{ij}^{h} of the hidden layer; the weight w_{kj}^{o} the output layer is a random vector when network training starts.
Subsequently using the sigmoid activation function (eq. 5):

$$f_{k}^{o}\left(Net_{pk}^{o}\right) = \frac{1}{1 + e^{-Net_{pk}^{o}}} \tag{5}$$

The output estimated by the neural network is calculated (eq.6):

$$O_{pk} = f_{k}^{o}(Net_{pk}^{o}) \tag{6}$$

Then the error terms for the output layer are calculated (eq.7):

$$\delta_{pk}^{o} = (y_{pk} - O_{pk}) f_{k}^{o'}(Net_{pk}^{o}) \tag{7}$$

As previously mentioned, the error propagates backwards starting from the neurons of the output layer, towards all the neurons of the hidden layer, this is in order to estimate the error in the outputs of the neurons of the hidden layer (eq. 8):

$$\delta_{pj}^{h} = f_{j}^{h'}(Net_{pj}^{h}) \sum_{k=1}^{M} \delta_{pk}^{o}\ w_{kj}^{o} \tag{8}$$

The equation 9, is used to update the weights of the output layer:

$$w_{kj}^{o}(t+1) = w_{kj}^{o}(t) + \eta \delta_{pk}^{o} i_{pj} \tag{9}$$

The modification of the synaptic weights of the hidden layer is made with equation 10:

$$w_{ji}^{h}(t+1) = w_{ji}^{h}(t) + \eta \delta_{pj}^{h} x_{pi} \tag{10}$$

Finally, the mean square error is calculated (eq. 11):

$$E_{p} = \frac{1}{2} \sum_{k=1}^{M} (y_{pk} - O_{pk})^{2} \tag{11}$$

In each of the iterations that are carried out in the network, the synaptic weights are updated and the objective is that the error is minimal [3].

2.2 Gauss-Newton method

The Gauss-Newton method is a combination of Newton's method and least squares estimation, this method is used when the parameters in the model appear non-linear and

there is no transformation by which the non-linear can be linearized, the main idea one is to approximate the nonlinear least squares problem to a linear least squares problem. The objective is to minimize the squared distance between an observed value (eq. 12, 13) and an adjusted value of a model with adjustable parameters, that is:

$$\min_{\theta} S(\theta) = \sum_{\mu=1}^{n} (y_\mu - f(\theta, x_\mu))^2 \quad (12)$$

$$\min_{\theta} S(\theta) = \sum_{\mu=1}^{n} r_\mu(\theta)^2 \quad (13)$$

Where θ are the adjustable parameters of the model, y_μ is the observed value of the dependent variable (*assumed* to contain the error), x_μ are the observed values of the independent variables (free *assumed* error), and n is the total number of observed data [4]. A linear approximation is then made using the Taylor series [5] for $f(\theta, x_\mu)$ near to $\bar{\theta}$ to obtain (eq. 14):

$$f(\theta, x_\mu) \approx f(\bar{\theta}, x_\mu) + \frac{\partial f(\bar{\theta}, x_\mu)}{\partial \theta'} (\theta - \bar{\theta}) \quad (14)$$

Using the approximation (14) in (12), the minimization problem was converted to a linear least squares problem

$$Min \left\| \bar{r} - \bar{F}.(\theta - \bar{\theta}) \right\|^2 \quad (15)$$

where $\bar{r} = y - f(\bar{\theta}, x)$ and $\bar{F}.$ is $\frac{\partial f}{\partial \theta'}$ evaluated on $\bar{\theta}$. Now (15) has a solution and shown in (16):

$$\theta - \bar{\theta} = (\bar{F}.' \ \bar{F}.)^{-1} \bar{F}.' \bar{r} \quad (16)$$

leading to the Gauss-Newton algorithm (eq. 17, 18)

$$\theta^{(k+1)} = \theta^k + \delta^k, \quad (17)$$

where

$$\delta^k = \left(\bar{F}.' \ \bar{F}.\right)^{-1} \bar{F}.' \bar{r} \quad (18)$$

2.3 Redesigned back-propagation neural network

The redesigned retro propagation neural network structure is made up of an input layer and an output layer, the Gauss-Newton method is used to update the weights in each of the iterations (epochs), below are the steps for generate the retro propagation neural network using Gauss-Newton:

Step 1 Get weights with least squares (first iteration only).
Step 2 Apply the vector x_p to the neurons in the input layer of the network.
Step 3 Calculate the net input values for the output layer.
Step 4 Apply the activation function, to obtain the network output.

Step 6 Calculate the error terms.
Step 7 Update the weights using Gauss-Newton.
Step 8 Get the mean square error.
Step 9 Return to step 3 until you meet the condition of completing.

2.4 Statistical metrics

This section describes the statistical metrics used to evaluate and compare the models obtained from back-propagation neural network with the methods: gradient and Gauss-Newton

Coefficient of Determination R^2 . This value is often called (eq. 19) the proportion of the variation explained by the return x variables.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{19}$$

Where y_i represent the desired outputs, $\hat{y}_i$ are the answers obtained from the model, $\bar{y}$ is the average response and R^2 ranges from 0 to 1 ($0 < R^2 < 1$), the values of R^2 close to 1 imply that most of the variability of y is explained by the model [6], otherwise, R^2 values close to zero indicate the non-representativeness of the model, in other words, that the model does not explain the total variation of the variable y.

Adjusted R^2. It is the value obtained by (20)

$$R^2_{Adj} = 1 - \frac{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-p)}}{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{(n-1)}} \tag{20}$$

Where y_i represents the desired responses, $\hat{y}_i$ are the responses obtained from the model, $\bar{y}$ is the average response, $(n-p)\ and\ (n-1)$ are degrees of freedom and $n > p$, n is the total number of data, p is the number of input variables in the model. The adjusted R^2 penalizes the addition of terms that are not useful, and is also advantageous for evaluating and comparing models [6].

Prediction R^2. This statistic indicates the ability of a model to predict new data, this value is calculated with the following equation (21) [6]:

$$R^2_{prediction} = 1 - \frac{\sum_{i=1}^{n}\left(\frac{e_i}{1-h_{ii}}\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{21}$$

Where y_i represents the desired responses, $\bar{y}$ is the average response, h_{ii} is the diagonal of the hat matrix and e_i are the residuals.

3 Application

It is considered a PTA welding process (Plasma Transfer Arc), which has the characteristic of presenting high variability in the quality characteristic.This process uses a metallic material in powder form to bond to a workpiece and thus generate a hard, wear-resistant, corrosion-resistant and high-temperature coating [7]. Table 1 shows the observations of the PTA welding process, of which 75% of the observations are taken for training and to apply them in the neural network back-propagation models with the gradient method and the Gauss-Newton method. The rest of the observations (25%) are used to make new predictions.

Table 1. Observations of the welding process PTA.

x_1	x_2	x_3	x_4	y
Powder feed rate [%]	Powder feed rate [g / min]	Process speed [cm / min]	Welding Current [Amps]	HAZ area [mm^2]
70	26	80	60	4.908
70	26	180	50	1.945
70	26	120	50	2.099
60	21	100	60	5.184
50	18	80	80	4.859
70	26	140	50	2.406
60	21	100	80	3.693
60	21	100	70	3.623
70	26	160	50	1.813
70	26	100	70	3.504
70	26	120	80	3.711
60	21	100	70	3.737
70	26	120	60	3.371
50	18	100	70	3.345
60	21	120	50	3.028
60	21	100	70	3.541
60	21	100	70	4.092
50	18	80	60	4.674
60	21	100	70	4.023
50	18	120	50	2.985
60	21	100	70	3.47
60	21	100	70	3.478
50	18	120	80	3.202
50	18	120	60	3.081
60	21	100	70	3.447
60	21	80	50	4.449
60	21	100	70	4.016
60	21	120	70	2.719
60	21	120	50	3.197
60	21	100	70	3.526
70	26	80	80	4.584
60	21	80	70	4.662
60	21	100	50	3.498

Table 2 compares the results obtained from the different neural networks.

Table 2. Results of the comparison between neural networks.

Statistical metrics	Neural network back-propagation	
	Gradient method	Gauss-Newton
R^2	0.9992	0.8134
$R^2_{adjusted}$	0.9990	0.7867
$R^2_{prediction}$	0.9987	0.7056

Figure 1 shows the graphical comparison between the estimates of the neural networks.

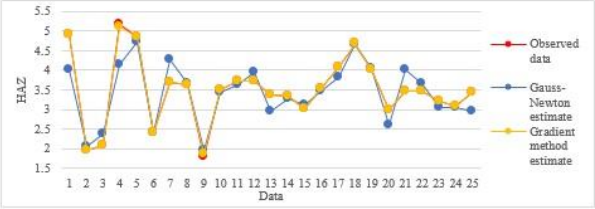


Fig. 1. The comparison between the observed data, the neural network with the gradient method and the neural network with Gauss-Newton is shown.

In Figure 2, the graphical comparison of the prediction of new data is made with both models of neural networks.

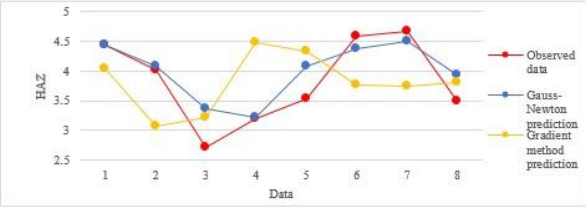


Fig. 2. Graphical comparison between the observed data of the HAZ and the prediction of the neural network with the gradient method and the prediction of the neural network with Gauss-Newton.

4 Discussion and conclusion

In Table 2 it shows, that the neural network back-propagation with the gradient method represents the variability of the PTA welding process by 99.92%, while the neural network with Gauss-Newton represents the variability in a 81.34% , also shown that the network Neural with the gradient method compared to the Gauss-Newton neural network has a greater ability to predict new values. In Figure 1, it is observed that the estimation of the area of the heat affected zone (HAZ) obtained by the back-propagation neural network with the gradient method, mostly coincide with the observed data, this phenomenon is called over fit. Over fitting indicates that the artificial neural network allows an almost perfect fit to the observed values, the problem of over fitting is that the neural network will poorly forecast new data [8], as shown in Figure 2, for the prediction of new Values 25% of the remaining data from Table 1 was used.

In Figure 2, the predictions of both networks are observed obtaining a mean square error of 0.6496, the neural network with the gradient method and the neural network with Gauss-Newton, obtain a mean square error of 0.1313.

Therefore the model that adequately represents the PTA welding process and is capable of making good predictions is the Gauss-Newton back-propagation neural network model.

References

1. Ming Der et al. “Analysis of Hard-Facing Appearance of Specific Powdered Superalloys for PTA-Coating Processes” Journal of Materials Engineering and Performance, 14, 307-314, (2005)
2. Asante Okyere et al. “Generalized regression and feed forward back in modelling flammability characteristics of polymethyl methacrylate” Thermochimica Acta, 79-92, (2018)
3. Caicedo, E. F., & Jesús, A. L. “Una aproximación práctica a las redes neuronales artificiales” Cali, Colombia: Universidad del Valle, 75-85, (2017)
4. Floudas , C. A., & Pardalos, P. M. “Encyclopedia of Optimization” Boston: Board Members, (2009)
5. Seber , G., & Wild, C. “NONLINEAR REGRESSION” Canada: WILEY-INTERSCIENCE, 619-623, (2003)
6. Montgomery, D. C. et al. “Introducción al análisis de regresión lineal” México: Compañía editorial continental, (2006)
7. Echavarría Figueroa, C. J. “Optimización de parámetros y análisis de resistencia al desgaste en recubrimientos base Ni sobre acero A36 mediante plasma transferido por arco (PTA)” Saltillo, Coahuila: Tésis de maestría, Corporación Mexicana de Investigación en Materiales, S.A. de C.V., (2016)
8. Piotrowski, A. P., & Napiorkowski, J. J. “A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling” Journal of Hydrology, (2013)

Parte IV

Sistemas Inteligentes en Internet de las Cosas

Introducción

La conectividad generalizada y los dispositivos inteligentes manifiestan que el internet de las cosas y la inteligencia artificial seguirán creciendo. Cada día se incorporan nuevos sensores en los objetos cotidianos, y a medida que las personas utilizan más tecnología, se crean negocios y aplicaciones cada vez más centrados en los datos que se generan, por lo que se requieren sistemas que perciban y respondan al mundo que los rodea.
Esta sección representa un esfuerzo por motivar los nuevos paradigmas de sistemas inteligentes en internet de las cosas, y se compone de cuatro artículos que se sometieron a un arbitraje doblemente ciego. El capítulo "Variante Algorítmica tendiente a dar Eficiencia a Búsquedas en el entorno de Internet" propone un algoritmo de búsqueda basado en Best First Search, para optimizar los motores de búsqueda y así proveer información en contexto. El capítulo "Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío" propone y evalúa un algoritmo metaheurístico basado en la metodología de algoritmos meméticos para optimizar las operaciones de compra en línea. El capítulo "Optimización del Problema Multi-Objetivo de la Selección de Proyectos con Parámetros Difusos" aborda el problema multiobjetivo en la selección de proyectos de inversión y propone un algoritmo con parámetros difusos. El capítulo "Optimizing the Selection of Mechanical Exoskeletons in the Industry: An Approach from TOPSIS and MOORA" presenta una evaluación de exoesqueletos con base en el método TOPSIS y MOORA como apoyo para la toma de decisiones.

Alicia Martínez Rebollar

18

Variante Algorítmica tendiente a dar Eficiencia a Búsquedas en el entorno de Internet

ALEJANDRO ARANDA DÍAZ Y MANUEL MEJÍA LAVALLE

Variante Algorítmica tendiente a dar Eficiencia a Búsquedas en el entorno de Internet

Alejandro Aranda Díaz[1], Manuel Mejia Lavalle[1]

[1] Tecnologico nacional de Mexico, campus CENIDET, Ciencias de la Computación, Cuernavaca, morelos.
{alejandro.aranda19ca, mlavalle}@cenidet.edu.mx

Resumen. El objetivo de este artículo es proponer un algoritmo de búsqueda basado en *Best First Search,* buscando optimizar las búsquedas que se realizan en contexto de Internet e Internet de las cosas. El algoritmo propuesto incluye una heurística basada en la heurística Manhattan y un umbral para limitar el tamaño de los nodos en la agenda. Aunque la intención del trabajo es aplicarlo en Internet e Internet de las cosas, por ahora se realizaron experimentos en un entorno controlado como lo es el juego del 8-Puzzle. Los resultados mostraron que se logró obtener un mejor tiempo (31.75% mejor) con el método propuesto y la heurística propuesta, en todas las configuraciones se consiguió reducir el tamaño de la agenda en un 44.99 %.

Keywords— WKS4, Métodos de Búsqueda, Best First Search, Internet of Things, Search Engine Optimization.

1 Introducción

La inteligencia artificial (IA) se ha integrado en varios aspectos de la vida humana, como han sido: educación, salud, inversiones, etc. [1]. La IA también ha sido aplicada en búsquedas de Internet, ayudando a obtener información confiable y en contexto a los deseos del usuario, a través de motores de búsqueda [2] [3].

En la actualidad se busca la optimización de estos motores en un área conocida como optimización de motores de búsqueda (SEO por sus siglas en inglés) donde el resultado considerará varios factores, como el área donde reside el usuario, historial de búsqueda, *websites* favoritos y lo que han seleccionado otros usuarios que han realizado el mismo *query*; los algoritmos deciden los factores mejor relacionados a tomar en cuenta en cada búsqueda [4].

Se utilizan Redes neuronales profundas de secuencia a secuencia en varios ámbitos de los cuales se podrían beneficiar los dispositivos móviles e Internet de las cosas, en términos de latencia, consumo de energía y seguridad. Desafortunadamente, el gran coste computacional de estos modelos propone la necesidad de utilizar técnicas de rendimiento y consumo de energía aceptable en procesadores de un solo núcleo, para esto se ha utilizado un algoritmo de búsqueda inspirado en *Beam Search,* reduciendo el coste computacional en tiempo de ejecución [5].

Hoy en día se utiliza IA para identificar el contenido de “calidad”, es decir aquel que presenta contenido más cercano al buscado por el usuario, y excluirlo del malo,

así mismo se ha aplicado con algoritmos de clasificación los cuales cumplen un papel muy importante a la hora de proporcionar de forma ordenada los sitos provistos por el motor de búsqueda; de forma más técnica, se espera que los algoritmos creen una lista con los mejores posibles resultados en orden óptimo a partir de un *query* [2].

Los algoritmos de búsqueda son procedimientos que permiten recorrer, en el peor de los casos, todas las posibles soluciones para encontrar una secuencia de operadores a un problema específico, hasta encontrar una solución [6].

Un algoritmo de búsqueda toma un problema como entrada y devuelve una solución en forma de una secuencia de acciones constituidos por los siguientes elementos: un estado inicial, estado del cual parte la búsqueda; estado final, estado objetivo; estados y operadores [3]. Muchos algoritmos de búsqueda se han desarrollado desde los inicios de la IA y aunque son muy utilizados en el ámbito de juegos [7] su utilidad se extiende más allá.

Los algoritmos de búsqueda se han aplicado en otros problemas del mundo real como direcciones de manejo en *websites* y sistemas integrados en automóviles [3]. El algoritmo *Best First Search* se ha aplicado a manejo de tiempo en áreas de trabajo, mientras que el algoritmo *Beam Search* se ha aplicado en el área de optimizado de empaquetado [8].

El principal problema con BFS es la gran cantidad de espacio que requiere para la agenda lo cual, en práctica, podría agotar la memoria de la máquina utilizada antes de alcanzar una solución [9].

En este artículo se propone una modificación para BFS que pretende reducir el problema que comprende la gran cantidad de nodos generados por el algoritmo, disminuyendo costo en memoria y tiempo, así mismo se propone una heurística basada en Manhattan.

2 Algoritmos de búsqueda

2.1 Descripción de Best First Search

En BFS se considera cualquiera de los nodos como el nodo siguiente. La idea es utilizar una heurística que decida cuál de los nodos en la lista, llamada agenda es el más prometedor a ser explorado. A continuación, una descripción del algoritmo:

1. Crear una Agenda vacía.
2. Insertar estado inicial
3. Hasta Agenda este vacía
   Si Agenda(0) es la Meta
   Salir
   SiNo
   Explorar Agenda(0)
   Agregar Hijos
   Marcar Agenda(0) “Visitado”
   Eliminar Agenda(0)

Reordenar Agenda
FinProceso

Dado el siguiente árbol en donde cada nodo representa el valor de una heurística **Fig. 1**, es posible observar la estructura que toma la agenda y los nodos expandidos:

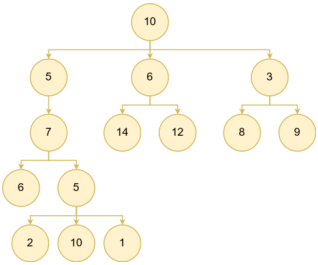


Fig. 1. Árbol de estados.

1. [10]
2. [3,5,6]
3. [5,6,8,9]
4. [6,7,8,9]
5. [7,8,9,12,14]
6. [5,6,8,9,12,14]
7. [1,2, 6,8,9,10,12,14]

2.2 Descripción del algoritmo propuesto

El algoritmo propuesto parte de BFS, utilizando un principio de solo admitir estados que posean una heurística no mayor a un umbral definido por la heurística de menor valor en la agenda. Esto con el propósito de reducir el tamaño de la agenda, disminuyendo el espacio en memoria utilizado y el costo en tiempo; a continuación, se describe el algoritmo llamado AAD.

Haciendo H una abreviación de Heurística y U de Umbral:

1. Crear una Agenda vacía.
2. Insertar estado inicial
3. U ← 10 (Ajustable)
4. MainH ← HInicial
5. Hasta Agenda este vacía

Si Agenda(0) es la Meta
Salir
SiNo
Explorar Agenda(0)
Comparar H con MainH+U
Agregar hijos no visitados en U
Marcar Agenda(0) “Visitado”
Eliminar Agenda(0)
Reordenar Agenda

```
            Si Agenda(0).H >MainH
                        MainH ← Agenda(0).H
FinProceso
```

3 Heurística

Son criterios, métodos o principios para decidir entre algunos cursos de acción para alcanzar el objetivo de la forma más efectiva [10]. Representan un valor cuyo significado es la cercanía que posee el estado actual de resolver el problema dado.

3.1 Manhattan

Consiste en obtener la distancia de las entidades con respecto a la posición meta; para 8-Puzzle se obtiene individualmente la distancia de los números a su posición meta, posteriormente se suma el total de los valores para de esta forma obtener el valor de Heurística **Fig. 2**.

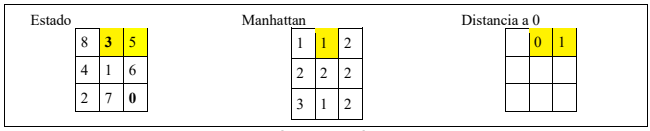


Fig. 2 Obtención de Heurística Manhattan

3.2 Heurística Propuesta

La heurística propuesta es una variación de la heurística Manhattan, a la cual se le multiplica la distancia de cada elemento hasta el elemento vacío o cero, de esta maera el resultado final tendrá un valor extra que representará una heurística de mayor valor.

Tomando en cuenta el estado dado: {8,3,5,4,1,6,2,7,0} **Fig. 3**, y se desea llegar al estado meta: {1,2,3,8,0,4,7,6,5} **Fig. 4**.

8	**3**	**5**
4	**1**	**6**
2	**7**	**0**

Fig. 3 Estado inicial

1	**2**	**3**
8	**0**	**4**
7	**6**	**5**

Fig. 4 Estado meta

Para cada elemento del estado se obtiene la distancia a 0 y finalmente se suman todos los elementos para obtener la heurística del estado. Esto se puede observar de manera gráfica en la **Fig. 5.**

Estado

8	**3**	5
4	1	6
2	7	**0**

Distancia a 0

	0	1
		2
		3

Manhattan * Distancia a 0

1***4**	1***3**	2***2**
2***3**	2***2**	2***1**
3***2**	1***1**	2***0**

Fig. 5 Funcionamiento de distancia a 0

4 Experimentacion

Para la experimentación se utilizó un procesador Intel Core i7 de séptima generación con 16GB RAM en el SO Windows 10, el código fue realizado en C# con Visual Studio 2019. Se probaron 100 casos de estados iniciales generados al azar realizando 55 movimientos.

El tiempo fue medido utilizando la clase *StopWatch* de C#, la cual se inicia al comienzo del código que se desea medir y se detiene al finalizar este, de esta forma *StopWatch* ahora porta el atributo *ElapsedMilliseconds* el cual pose el tiempo transcurrido en milisegundos.

Los 100 casos se ejecutaron de forma continua sin interrupciones, para esto, un preprocesamiento de los datos fue realizado con el fin de ajustarlos al manejo de los datos que utiliza el algoritmo.

A ambos algoritmos se les aplico una limitación en el nivel de profundidad siendo el máximo nivel de profundidad 55.

4.1 Experimentación del algoritmo propuesto con heurística propuesta y Manhattan

Los resultados obtenidos por el algoritmo propuesto utilizando la heurística propuesta y Manhattan se muestran en la **Tabla I**:

Tabla 1. Comparación de algoritmo propuesto (aad), heurística propuesta (hpro) contra Manhattan (man), profundidad (p), nodos visitados (v), tamaño de agenda (a), tiempo transcurrido en milisegundos (t).

	P	V	A	T
AAD+HPRO	**22.121**	**93.141**	**40.323**	**2.559**
AAD+ MAN	22.18	110.42	89.61	3.749

En la **Tabla 1** se muestra en **negritas** el mejor caso, en esta ocasión la heurística propuesta obtuvo un ligero mejor desempeño exceptuando en el tamaño de la agenda donde logro reducirla por la mitad.

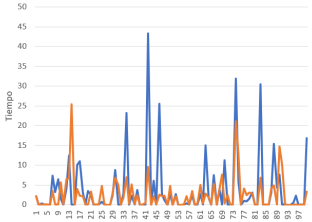


Fig. 6. Gráfica con la comparacion de heuristica prouesta (Naranja) y Manhattan (Azul) respecto al tiempo con AAD.

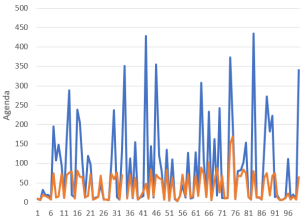


Fig. 7. Gráfica con la comparacion de heuristica prouesta (Naranja) y Manhattan (Azul) respecto al tamaño de la agenda con AAD.

En la **Fig. 6** podemos observar el comportamiento en los 100 casos de las heurísticas con respecto al tiempo, es posible observar que Manhattan muestra picos más elevados a lo largo de los casos.

En la **Fig. 7** es posible observar que la heurística propuesta logro una agenda más reducida y consistente, ya que manhattan muestra grandes elevaciones.

Algo más a agregar es que con la heurística propuesta, uno de los casos no fue alcanzado mientras que manhattan finalizo todos los casos.

4.2 Experimentación con algoritmo Best First Search con heurística propuesta y Manhattan

En la **Tabla 2** se muestran los resultados obtenidos con el algoritmo Best First Search (BFS) con la heurística propuesta y Manhattan. Los experimentos se realizaron con los mismos casos utilizados para el algoritmo propuesto.

Tabla 2. Comparación de algoritmo BFS, heurística propuesta (HPRO) contra Manhattan (MAN), Profundidad (P), Nodos Visitados (V), Tamaño de Agenda (A), Tiempo transcurrido en milisegundos (T).

	P	V	A	T
BFS+HPRO	23.68	**100.12**	**81.8**	**3.22**
BFS+ MAN	**22.18**	110.42	89.61	3.44

En la **Tabla 2** se muestra en **negritas** el mejor caso, en esta ocasión la heurística propuesta obtuvo un ligero mejor desempeño exceptuando la profundidad, en la cual fue derrotada por manhattan.

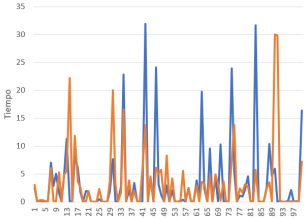


Fig. 8 Gráfica con la comparacion de heuristica prouesta (Naranja) y manhattan (Azul) respecto al tiempo con BFS.

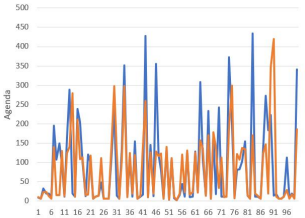


Fig. 9 Gráfica con la comparacion de heuristica prouesta (Naranja) y manhattan (Azul) respecto al tamaño de la agenda con BFS.

En la **Fig. 8** es posible observar que ambas heurísticas se comportaron de forma similar, por lo que se considera que HPRO logro tener mejor tiempo solo por el pico más elevado de MAN.

En la **Fig. 9** se muestra que HPRO presenta varias coincidencias con su contrincante, pero, HPRO presento menos picos en la agenda, de esta forma logro un mejor puntaje que MAN.

4.3 BFS contra AAD

Una comparación de las tablas mostradas anteriormente es mostrada en la **Tabla 3**, en negrita los mejores puntajes y cursiva los peores.

Tabla 3. Comparación de algoritmo BFS vs AAD, heurística propuesta (HPRO) contra Manhattan (MAN), Profundidad (P), Nodos Visitados (V), Tamaño de Agenda (A), Tiempo transcurrido en milisegundos (T).

	P	V	A	T
AAD+HPRO	**22.121**	**93.141**	**40.323**	**2.559**
AAD+ MAN	*22.18*	*110.42*	*89.61*	*3.749*
BFS+HPRO	23.68	100.12	81.8	3.22
BFS+ MAN	*22.18*	*110.42*	*89.61*	3.44

Estos datos muestran que el algoritmo propuesto es más eficiente con la heurística propuesta, así mismo se puede observar que AAD+MAN y BFS+MAN dan un resultado similar, se considera que esto se debe a el umbral utilizado, quizá si es reducido podría obtener un resultado diferente.

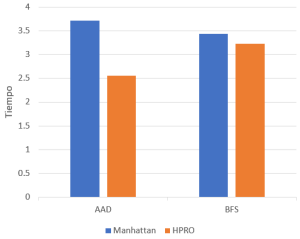


Fig. 10 Gráfica con la comparacion en tiempo de AAD v BFS.

Fig. 11 Gráfica con la comparacion en tamaño de agenda de AAD y BFS.

5 conclusión

Al finalizar el experimento con las configuraciones utilizadas podemos concluir que el algoritmo propuesto resultó ser más eficiente con la heurística propuesta, mejorando en un 31.75% en tiempo y en 49.99% el número de nodos en agenda; debido a los resultados de Manhattan, se considera ajustar el umbral, reduciendo su valor, para observar que diferencias en el comportamiento obtendrían las heurísticas con algoritmo propuesto.

Referencias

1 B. Thomas, “Two REITs for Our Artificial Intelligence Future,” Forbes, Nov. 25, 2019. [Online] Available: https://www.forbes.com/sites/bradthomas/2017/04/18/two-reits-for-our-artificial-intelligencefuture/#23b09815fc68.

2 M. Kaput, ‘How Search Engines Use Artificial Intelligence’ Marketing Artificial Intelligece Institute, Jan. 14, 2020. [Online] Available:
https://www.marketingaiinstitute.com/blog/how-search-engines-use-artificial-intelligence

3 P. Gupta, ‘Search algorithms in artificial intelligence’, hackernoon, Jan ,14 2020. [Online] available: https://hackernoon.com/search-algorithms-in-artificial-intelligence-8d32c12f6bea

4 S. R. Jaiswal, V.S. Belsare, ‘Theoretical Approach of Artificial Intelligence and Search Engine Optimization’, International Journal of Advance Innovative Technology in Engineering (IJAITE), Vol. 2, Special Issue 1, 2018.

5 D. J. Pagliari, F. Daghero y M. Poncino, ‘Sequence-To-Sequence Neural network Inference on Embedded Processors Using Dynamic Beam Search’, Electronics, MDPI, 2020

6 S. Lizma, A. Muñoz. ‘Algoritmos de consistencia para programación con restricciones.’ pontificia universidad católica de Valparaíso, 2013.

7 B. Nasarre, ‘Método de Monte-Carlo Tree Search (MCTS) para resolver problemas de alta complejidad: Jugador virtual para el juego del Go’ Escuela de ingeniería y arquitectura, universidad de Zaragoza, 2012.

8 J. Bennell, M. Cabo, A. Martinez-Sykora, ‘A beam search approach to solve the convex irregular bin packing problem with guillotine cuts’, European Journal of Operational Research 270 89–102, 2018

9 C. Mencia, M. Sierra, R. Varela, ‘An efficient hybrid search algorithm for job shop scheduling with operators’, International journal of production research, 2013.

10 [Pearl, 1984] J. Pearl, ‘Heuristics: intelligent search strategies for computer problem solving’ Los Angeles, CA: Addison-Wesley Longman Publishing Co. 1984

19

Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío (MAIShOP)

MIGUEL ANGEL GARCÍA MORALES, HÉCTOR JOAQUÍN FRAIRE HUACUJA, LAURA CRUZ REYES, CLAUDIA GUADALUPE GÓMEZ SANTILLÁN, NELSON RANGEL VALDEZ Y MARÍA LUCILA MORALES RODRÍGUEZ

Algoritmo Memético para el Problema de Ventas por Internet con Costos de Envío (MAIShOP)

Miguel Ángel García Morales, Hector Joaquín Fraire Huacuja, Laura Cruz Reyes, Claudia Guadalupe Gómez Santillán, Nelson Rangel Valdez, María Lucila Morales Rodríguez

Tecnológico Nacional de México / Instituto Tecnológico de Ciudad Madero, Cd. Madero, Tamaulipas, MÉXICO
soporteclusterlanti@gmail.com
automatas2002@yahoo.com.mx

Abstract. En este artículo se aborda el problema de Ventas por Internet con costos de envío (IShOP). En la literatura relacionada solo se encontró una solución metaheurística reportada. Esta solución es un algoritmo de procesamiento celular que simula el procesamiento paralelo de dos o más procesos de búsqueda través del espacio de soluciones y actualmente se le considera el mejor del estado del arte de la solución del IShOP. En este trabajo se propone un nuevo algoritmo metaheurístico basado en la metodología de algoritmos meméticos. Para validar este nuevo enfoque se realiza un estudio comparativo del desempeño del algoritmo propuesto contra el del algoritmo de procesamiento celular. En los experimentos computacionales se usó un conjunto amplio de instancias aleatorias estándar y los resultados muestran una clara superioridad del algoritmo propuesto. Para soportar estos resultados se aplicó la prueba no paramétrica de Wilcoxon, en la que se verificó la significancia de las diferencias observadas.

Keywords: WKS4, Algoritmos meméticos, Ventas por Internet, Optimización heurística.

1 Introducción

En la actualidad, el comercio electrónico se ha convertido en parte importante de la sociedad, la implementación de tecnología novedosa y en constante crecimiento ha hecho inevitable adaptarse a esta evolución [1].

Para los vendedores, la principal ventaja es que sus ofertas están disponibles para un público más amplio, sin la mayoría de los costos asociados, como alquiler, impuestos, mantenimiento y publicidad. Los compradores tienen la comodidad de poder realizar sus compras desde cualquier lugar, en cualquier momento, con mejores precios y una gama más amplia de productos, siempre que tengan acceso a Internet [2].

En el problema de optimización de compras por Internet (IShOP) [3] se supone que un cliente con una lista de compras requiere adquirir los productos en un conjunto de tiendas en línea al menor costo posible. En [4] se propuso el problema y se demostró que es NP-duro.

En [2] se propone un algoritmo de procesamiento celular que simula el procesamiento paralelo de dos o más procesos de búsqueda en el espacio de soluciones y actualmente se le considera el mejor algoritmo del estado del arte de IShOP.

En este artículo, se propone un algoritmo memético (MAIShOP) que a diferencia del algoritmo del estado del arte utiliza una representación vectorial para la representación de las soluciones candidatas, un mecanismo que acelera el cálculo de la función objetivo y para cada generación se actualiza la población dejando solo la mejor solución y recalculando el resto de la población. Estos elementos se diseñaron con el propósito de lograr una mejora significativa en el desempeño del algoritmo. Para validar este nuevo enfoque se realiza un estudio comparativo del desempeño del algoritmo propuesto contra el del algoritmo del estado del arte. En los experimentos computacionales se usó un conjunto amplio de instancias estándar y los resultados muestran una clara superioridad del algoritmo memético. Para soportar estadísticamente los resultados se aplica la prueba de hipótesis no paramétrica de Wilcoxon.

1.1 Definición del problema

Un cliente desea comprar por internet un conjunto de *n* productos *N*, los cuales se pueden adquirir en un conjunto de *m* tiendas disponibles *M*. El conjunto N_i contiene los productos disponibles en la tienda *i*, cada producto $j \in N_i$ tiene un costo de c_{ij} y un costo de envío d_i.

El costo de envío se cobra solo si se compran uno o más productos en la tienda *i*. El problema consiste en minimizar el costo total de la compra de todos los productos en *N* (costo de los productos más los gastos de envío).

Formalmente el problema consiste en determinar una partición de los productos a comprar en las diferentes tiendas $X = (X_1, \dots, X_m)$, de modo que $X_i \subseteq N_i$ y que se compren todos los productos $U_{i=1}^{m} X_i = N$ y que se minimice el costo total:

$$F(X) = \sum_{i=1}^{m} \left(\sigma(|X_i|) d_i + \sum_{j \in X_i} c_{ij}\right) \quad (1)$$

Donde: $|X_i|$ es la cardinalidad del contenedor X_i, y $\sigma(x) = 0 \; si \; x = 0 \; y \; \sigma(X) = 1 \; si \; x > 0$.

El valor objetivo de la solución *I* se calcula de la siguiente manera:

$$F(I) = \sum_{i=1}^{m} \sum_{j=1 \&\& I(j)=i}^{n} \left(d_i + c_{ij}\right) \quad (2)$$

Si la lista de productos que se va a comprar está formada por *N* productos y hay *M* tiendas disponibles, entonces una solución se representa en un vector de longitud *N*, el cual contiene para cada producto la tienda donde se va a comprar.

2 Estructura general del algoritmo memético

En el algoritmo se usa una representación vectorial de las soluciones que, a diferencia del algoritmo del estado del arte, reduce la complejidad del procesamiento de una solución de $O(n^2)$, a $O(n)$. La búsqueda local asigna cada producto a la tienda en que

resulta más conveniente su compra, este proceso se realiza de manera intensiva en el algoritmo. Con el propósito de hacer más eficiente la localización de la tienda más conveniente, se incorpora un mecanismo que acelera el cálculo de la función objetivo. Finalmente, la estructura tradicional del algoritmo memético se modificó de tal manera que al termino de cada generación la población se actualiza dejando solo la mejor solución y regenerando el resto de la población. La idea es que la mejor solución transmita sus características a las siguientes generaciones y conjuntamente con la regeneración del resto de los individuos, se logre un mejor balance de la intensificación y la exploración del algoritmo.

2.1 Selección por torneo

Se seleccionan aleatoriamente *p* individuos de la población *Pop* y el más apto es el que pasa a la siguiente generación.

En el algoritmo memético se realiza un torneo binario (con *p=2*) sobre la población Pop. Cada solución participa en dos torneos y la solución ganadora se selecciona y se inserta en *NewPop*. De esta manera, cualquier solución en *Pop* puede tener como máximo dos copias en *NewPop*.

2.2 Operador de cruza

Este operador se aplica a un par de soluciones $padre_1$ y $padre_2$, seleccionados de manera secuencial hasta que se hayan seleccionado un porcentaje de los individuos de la población [5]. Para generar la solución $hijo_1$, primero se toma la mitad inicial del $padre_1$ y se une con la segunda mitad del $padre_2$. Posteriormente para formar la solución $hijo_2$, se toma la mitad inicial del $padre_2$ y se une con la segunda mitad del $padre_1$ [6].

En el operador de cruza, el punto $\lfloor N/2 \rfloor$ ó $\lceil N/2 \rceil$ se selecciona como punto de cruce.

2.3 Operador de mutación

Se diseñó un método de mutación heurística, considerando el siguiente procedimiento:

Seleccionamos $ps - ps * er$ soluciones y por cada una de estas se genera un número al azar, si este valor generado al azar es menor que *ps*pm* se procede a generar todas las posibles combinaciones de asignación (producto, tienda), se evalúan todas estas combinaciones y se selecciona la mejor.

2.4 Búsqueda local

Este procedimiento comienza seleccionando una solución de *IntermediatePop*, siendo está un vector de valores asociados que contiene las tiendas *j* asignadas en la secuencia de selección de productos X_{old} con su costo, que incluye c_{ij} y d_i del multiconjunto N_i.

Posteriormente se verifica para cada producto, desde que tienda podríamos comprarlo para reducir el costo total de la selección de productos *X*.

Cuando finaliza la búsqueda de esa tienda, se mueve a un nuevo producto. Este proceso continúa hasta que todas las tiendas *j* en *X* hayan sido revisadas. Concluye cuando todas las *X* hayan sido asignadas en *ChildPop*.

2.5 Algoritmo memético (MAIShOP)

La Figura 1 describe la estructura general del algoritmo MAIShOP. El algoritmo, comienza definiendo los parámetros tales como tamaño de la población inicial *ps*, porcentaje de cruza *pc*, probabilidad de mutación *pm*, porcentaje de elitismo *er*, criterio de terminación *timelimit*, enseguida se genera una población inicial (Paso 2) utilizando una heurística de construcción. En la fase generacional (Pasos 5-15), se aplica el operador de torneo binario para la selección en el Paso 6. Las mejores soluciones *ps*er* (elitismo) de *pop* se copian en *IntermediatePop* en el Paso 7. En *ps*pc* de las soluciones restantes en *NewPop*, se aplica el operador de cruza para generar más soluciones para *IntermediatePop* y las soluciones restantes se copian tal como están en *IntermediatePop* en el Paso 8. Se supone que *pc* se selecciona de una manera que *ps*pc* es menor o igual que el número de soluciones restantes en *NewPop*. En el Paso 9, los individuos en *IntermediatePop* sufren mutación excepto aquellos individuos que fueron seleccionados en el Paso 6 usando elitismo. En el Paso 10, se aplica la búsqueda local a *IntermediatePop* para mejorar las soluciones. En el Paso 11, se obtiene la mejor solución local (*BestLocal*) de *ChildPop* y se compara con la solución global (*BestGlobal*) en el Paso 12 además se calcula el tiempo de ejecución (*timesolution*) una vez evaluado este paso, si *BestLocal* tiene un menor costo entonces reemplaza a *BestGlobal*, en el Paso 13 la mejor solución local (*BestLocal*) es copiada a la población inicial y todas las demás soluciones son recalculadas, se calcula el tiempo de ejecución total del algoritmo para cada generación en el paso 14 y finalmente se retorna la mejor solución global (*BestGlobal*) con el cálculo de tiempo de ejecución en la cual se encontró esta solución en el paso 16.

Input:

timelimit: Tiempo límite en la ejecución del algoritmo

Output:

BestGlobal: Mejor solución global

timesolution: Tiempo de la solución

1: Se definen los siguientes elementos: tamaño de la población inicial *ps*, porcentaje de cruza *pc*, probabilidad de mutación *pm*, porcentaje de elitismo *er*, criterio de terminación *timelimit*

2: Generar población inicial *pop* de tamaño *ps*

3: Determinar *BestLocal* en *pop*.

4: *BestGlobal = BestLocal*.

5: **while** no se cumpla el criterio de terminación *timelimit* **do**

6: Aplicar operador de torneo binario en *pop* para obtener *NewPop*

7: *ps.er* soluciones élite de *NewPop* se pasan a *IntermediatePop*

8: Entre los individuos restantes en *NewPop*, se aplica el operador de cruza a *ps*pc* soluciones seleccionadas al azar y las soluciones restantes se copian tal como están a *IntermediatePop*

9: Aplique el operador de mutación en *IntermediatePop* con probabilidad *pm* excepto a las soluciones élite.

10: Para cada individuo en *IntermediatePop,* se realiza una mejora local asociando a cada producto la tienda que reduce el costo del producto, el individuo mejorado se coloca en *ChildPop*.

11: Se obtiene la mejor solución (*BestLocal)* de *ChildPop*

12: Se determina si el valor objetivo de la mejor solución local (BestLocal) es menor que la solución global (BestGlobal), si es así se actualiza la solución global. Se calcula el tiempo de ejecución del algoritmo hasta este paso *(timesolution)*.

13: La mejor solución local (*BestLocal) se copia a pop* y todos los demás individuos son generados nuevamente.

14: Se calcula el tiempo de ejecución del algoritmo para cada generación *(totaltime)*.

15: **end while**

16: **return (***BestGlobal***,** *timesolution*)

Figura 1: Algoritmo memético (MAIShOP). Fuente: Elaboración propia

3 Experimentos computacionales

La plataforma de hardware y software en que se realizaron los experimentos incluye: procesador Intel Core 2 Duo a 2.53 Ghz. con 4 GB de RAM, Java versión 1.8.0_231.

El nombre de las instancias especifica su tamaño, *m* representa el número de tiendas, y *n* la cantidad de productos. Se crearon tres conjuntos de instancias de diferentes tamaños. Un conjunto de instancias pequeñas, con 3 subconjuntos: 3n20m, 4n20m y 5n20m. Un conjunto con instancias medianas, con tres subconjuntos: 5n240m, 5n400m y 50n240m. Un conjunto de instancias grandes con tres subconjuntos: 50n400m, 100n240m y 100n400m. Cada subconjunto contiene 30 instancias.

El objetivo de los experimentos computacionales que se realizaron fue comparar el desempeño del algoritmo memético (MAIShOP) y el algoritmo de procesamiento celular (Pcell). Cada instancia se resolvió realizando 30 corridas independientes con cada uno de los algoritmos y se determinó el valor objetivo de la mejor solución encontrada y el tiempo de dicha solución. Para determinar el desempeño de los algoritmos se calcularon los valores promedio del valor objetivo y del tiempo de solución; y las desviaciones estándar. Se aplicó la prueba de hipótesis no paramétrica de Wilcoxon para determinar la significancia estadística de las diferencias observadas en calidad y eficiencia.

La Tabla 1 muestra los resultados obtenidos con cada grupo de 30 instancias. Para cada algoritmo se muestra el promedio y la desviación estándar (subíndice) del valor objetivo y del tiempo de solución. En las celdas de Pcell se indica si hay diferencias significativas a favor de Pcell (↓) o a favor de MAIShOP (↑) y si no hay diferencias significativas entre los dos algoritmos (--). La significancia se determinó con una confiabilidad de 95%. Las celdas sombreadas corresponden al algoritmo que obtuvo el mejor resultado en calidad o eficiencia.

Tabla 1. Desempeño comparativo de los algoritmos. Fuente: Elaboración propia.

Instancia	Pcell (FO)	MAIShOP (FO)	Pcell (Time)	MAIShOP (Time)
3n20m	$63.03_{0.483}$ ↑	$62.76_{2.57E-14}$	$0.00049_{0.0006}$ ↑	$1.2222E-05_{6.21E-05}$
4n20m	$78.83_{0.40}$ ↑	$78.73_{2.45E-14}$	$0.00185_{0.0012}$ ↑	$6.6667E-06_{4.26E-05}$
5n20m	$102.56_{0.80}$ --	$102.09_{0.0043}$	$0.00273_{0.0014}$ ↑	$0.00016_{0.0002}$
5n240m	$75.43_{0.77}$ ↓	$76.32_{1.67}$	$0.27360_{0.0957}$ ↑	$0.0669_{0.0627}$
5n400m	$69.98_{1.72}$ --	$70.55_{1.68}$	$0.79002_{0.2882}$ ↑	$0.2425_{0.2139}$
50n240m	$1058.56_{103.06}$ ↑	$505.47_{17.60}$	$16.4441_{8.0266}$ ↓	$22.5010_{7.5014}$
50n400m	$951.08_{119.35}$ ↑	$438.74_{24.09}$	$48.6607_{25.3042}$ ↑	$33.1063_{3.8351}$
100n240m	$2349.94_{159.48}$ ↑	$993.70_{38.25}$	$73.8573_{39.7564}$ ↑	$50.4910_{5.6348}$
100n400m	$2213.90_{173.70}$ ↑	$801.51_{35.03}$	$181.9521_{95.3004}$ --	$176.1110_{41.5457}$

4 Conclusiones

En este artículo, se propone un algoritmo memético (MAIShOP) que incorpora una serie de mecanismos innovadores para resolver eficientemente el problema IShOP. Para validar este nuevo enfoque se realizó un estudio comparativo del desempeño del algoritmo propuesto contra el del algoritmo del estado del arte. Los resultados de los experimentos realizados muestran que el algoritmo propuesto supera claramente en calidad y eficiencia al algoritmo del estado del arte. En calidad mejora significativamente en 6/9 conjuntos y en eficiencia mejora significativamente en 7/9 conjuntos.

Referencias

1. Musial, J., Pecero, J., Dorronsoro, B., & Blazewicz, J. Internet Shopping Optimization Project (IShOP). European IST Projects, 16, (2014).

2. Lopez-Loces, M. C., Musial, J., Pecero, J.E., Fraire-Huacuja, H. J., Blacewicz, J., & Bouvry, P. Exact and heuristic approaches to solve the Internet shopping optimization problem with delivery costs. International Journal of Applied Mathematics and Computer Science, 26(2), 391-406, (2016).
3. Musial, J. Applications of Combinatorial Optimization for Online Shopping. NAKOM, Poznán, (2012).
4. Mitsuo Gen and Runwei Cheng. Genetic Algorithms & Engineering Optimization. Wiley Series in Engineering Design and Automation. John Wiley & Sons, New York, (2000).
5. Holland, J. H. Adaptation in Natural and Artificial Systems. University of Michigan Press, And Arbor, Michigan, (1975).
6. Umbakar, A. J., & Sheth, P. D. Crossover operators in genetic algorithms: a review. ICTACT journal on soft computing, 6(1), (2015)

20

Optimización del Problema Multiobjetivo de la Selección de Proyectos con Parámetros Difusos

Daniela López García, Héctor Fraire Huacuja, Laura Cruz Reyes, Nelson Rangel Valdez, Claudia Guadalupe Gómez Santillán, María Lucila Morales Rodríguez y Fausto Antonio Balderas Jaramillo

Optimización del Problema Multiobjetivo de la Selección de Proyectos con Parámetros Difusos

Daniela López-García, Héctor Fraire-Huacuja, Laura Cruz-Reyes, Nelson Rangel-Valdez, Claudia Guadalupe Gómez-Santillán, María Lucila Morales-Rodríguez, Fausto Antonio Balderas-Jaramillo

Tecnológico Nacional de México/Instituto Tecnológico de Ciudad Madero, Cd. Madero, Tamaulipas, MÉXICO.
automatas2002@yahoo.com.mx

Abstract. En este artículo se aborda el problema multiobjetivo de la selección de carteras de proyectos de inversión con parámetros difusos de tipo intervalo. Este problema es NP-Duro ya que se puede restringir al problema NP-Duro de la mochila. En la literatura revisada se reporta la solución de la versión difusa del problema con el algoritmo I-NSGA-II. En este artículo se propone utilizar como estimador de densidad la desviación espacial propagada en lugar de la distancia de agrupamiento para mejorar la distribución de las soluciones en los frentes. Adicionalmente se propone defuzzificar las soluciones en los frentes generados para medir el desempeño del algoritmo usando las métricas comúnmente usadas. En los experimentos computacionales se usaron un conjunto de instancias aleatorias del problema difuso y las métricas de hipervolumen y dispersión generalizada. Los resultados obtenidos son alentadores ya que confirman la factibilidad del enfoque propuesto.

Keywords: WKS4, Optimización multiobjetivo, problema de selección de cartera de proyectos, Algoritmo NSGA-II.

1 Introducción

El problema de optimización de la cartera de proyectos es un problema clásico de asignación de recursos que considera un conjunto de proyectos en los que impactan una empresa puede invertir sus recursos, los cuales normalmente son limitados. El problema consiste en seleccionar un subconjunto de proyectos que se puedan realizar con los recursos disponibles y maximicen beneficios generados por dichos proyectos. Este problema ya ha sido abordado en diferentes trabajos de investigación considerando que los recursos disponibles los beneficios derivados de los proyectos son conocidos de manera precisa. Sin embargo, en muchos casos este supuesto no siempre se cumple, dichos parámetros solo se conocen de manera imprecisa o aproximada. La incertidumbre en los valores de los parámetros se puede derivar de diferentes aspectos tales como la arbitrariedad, la imprecisión y mala determinación de dichos valores [1]. Una forma de modelar la incertidumbre de los valores de los parámetros del problema de selección de carteras de proyectos es usar parámetros difusos.

1.1 Números difusos de tipo intervalo

Una alternativa para representar números difusos [2] es utilizar intervalos, los cuales determinan un rango de valores en la recta real. Dados $a_1, a_2 \epsilon R$, el intervalo $[a_1, a_2] = \{x \varepsilon R | a_1 \leq x \leq a_2\}$, en lo sucesivo a los números de tipo intervalo se les representará con letras negritas cursivas, en este caso $\boldsymbol{A} = [a_1, a_2]$. F denota el conjunto de todos los números difusos de tipo intervalo. Por otra parte, dados $\boldsymbol{A} = [a_1, a_2]$ y $\boldsymbol{B} = [b_1, b_2]$ dos números difusos entonces se definen las operaciones siguientes [3]:

$$\boldsymbol{A} + \boldsymbol{B} = [a_1 + b_1, a_2 + b_2]$$
$$\boldsymbol{A} - \boldsymbol{B} = [a_1 - b_2, a_1 - b_1]$$
$$\boldsymbol{A} * \boldsymbol{B} = [a_1, a_2] * \left[\frac{1}{b_1}, \frac{1}{b2}\right]$$
$$\frac{\boldsymbol{A}}{\boldsymbol{B}} = [\min\{a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2\}, \max\{a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2\}]$$

Para introducir una relación de orden entre los números de tipo intervalo, en [4], se utiliza la siguiente medida:

$$P(\boldsymbol{A} \geq \boldsymbol{B}) = \begin{cases} 1 & si \quad p > 1 \\ p & si \quad 0 \leq p \leq 1 \\ 0 & si \quad p < 0 \end{cases} \qquad (1)$$

donde

$$p = \frac{a_2 - b1}{(a_1 | - b_2) + (b_1 - b_2)}.$$

La relación de orden entre $\boldsymbol{A}$ y $\boldsymbol{B}$ se define siguiente manera:

Si $P(\boldsymbol{A} \geq \boldsymbol{B}) = 0.5$, entonces $\boldsymbol{A} = \boldsymbol{B}$. Si $P(\boldsymbol{A} \geq \boldsymbol{B}) = 1$, entonces $\boldsymbol{A} > \boldsymbol{B}$.
Si $P(\boldsymbol{A} \geq \boldsymbol{B}) = 0$, entonces $\boldsymbol{A} < \boldsymbol{B}$. Si $P(\boldsymbol{A} \geq \boldsymbol{B}) > 0.5$, entonces $\boldsymbol{B} > \boldsymbol{A}$. Si $P(\boldsymbol{A} \geq \boldsymbol{B}) < 0.5$, entonces $\boldsymbol{B} < \boldsymbol{A}$.

Ahora se define la dominancia de Pareto entre vectores de valores de tipo difuso.
Dados los siguientes vectores de valores difusos de tipo intervalo:
$\hat{x} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots\ldots, \boldsymbol{x_n})$ e $\hat{y} = (\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots\ldots, \boldsymbol{y_n})$, se dice que $\hat{x}$ domina a $\hat{y}$, si y solo si $\boldsymbol{x_i} \geq \boldsymbol{y}_i$ para toda $i = 1,2, \ldots., n$ y $\boldsymbol{x_i} > \boldsymbol{y_i}$ para alguna $i = 1,2, \ldots., n$.

2 Modelo del problema multiobjetivo de la selección de carteras de proyectos con parámetros difusos de tipo intervalo y restricciones de presupuesto, regiones y áreas

Sean n el número de proyectos a considerar, $\boldsymbol{C}$ el presupuesto total disponible, O el número de objetivos, $\boldsymbol{c}_i$ el costo del proyecto i, $\boldsymbol{b}_{ij}$ el conjunto de contribuciones del proyecto i en el objetivo j, K el número de áreas a considerar, M el número de regiones a considerar, $\boldsymbol{A}_k^{min}$ y $\boldsymbol{A}_k^{max}$ los límites inferior y superior del presupuesto disponible para el área k, respectivamente, $\boldsymbol{R}_m^{min}$ y $\boldsymbol{R}_m^{max}$ los límites inferior y superior del presupuesto disponible para la región m, respectivamente. Los arreglos $a_i y b_i$ contienen el área y la región asignadas al proyecto i. $\hat{x} = (x_1, x_2, \ldots\ldots, x_n)$es un vector binario en el que se especifica la cartera de proyectos seleccionada, en donde $x_i = 1$ si el proyecto i es seleccionado, y $x_i = 0$ si el proyecto no se selecciona. El problema multiobjetivo de la selección de carteras de proyectos se define como:

$$\text{Maximizar} \qquad \hat{z} = (\boldsymbol{z}_1, \mathbf{z}_2, \ldots\ldots, \mathbf{z}_O) \tag{2}$$

donde

$$\boldsymbol{z}_j = \sum_{i=1}^{n} \boldsymbol{b}_{ij}\, x_i \quad j = 1,2, \ldots O \tag{3}$$

Sujeto a las siguientes restricciones:

$$\sum_{i=1}^{n} \boldsymbol{c}_i x_i \leq \boldsymbol{C} \tag{4}$$

$$\boldsymbol{A}_k^{min} \leq \sum_{i=1, a_i=k}^{n} \boldsymbol{c}_i x_i \leq \boldsymbol{A}_k^{max} \qquad k = 1,2, \ldots, K \tag{5}$$

$$\boldsymbol{R}_k^{min} \leq \sum_{i=1, b_i=k}^{n} \boldsymbol{c}_i x_i \leq \boldsymbol{R}_k^{max} \quad k = 1,2, \ldots\ldots M \tag{6}$$

$$x_i \epsilon \{0,1\} \quad \text{para toda } i = 1,2, \ldots\ldots, n \tag{7}$$

La función objetivo trata de maximizar las contribuciones de cada uno de los objetivos (2), Cada uno de los objetivos considerados se calcula sumando las contribuciones de todos los proyectos seleccionados en el vector binario (3). La restricción (4) asegura la suma de los costos de todos los proyectos seleccionados no excedan el presupuesto disponible. El conjunto de restricciones (5) asegura que la suma de los costos de los proyectos de cada área este en el rango del presupuesto disponible para dicha área. El conjunto de restricciones (6) aseguran que la suma de los costos de los proyectos de cada región este en el rango del presupuesto disponible para dicha región.

3 Estrategia propuesta para medir la calidad del frente de Pareto y mejorar la distribución de las soluciones.

Para instancias pequeñas es posible encontrar el Frente de Pareto óptimo, utilizando un algoritmo exhaustivo que explore todo el espacio de solución. Sin embargo, generalmente este método no se puede aplicar para resolver instancias de interés en aplicaciones reales. Por lo tanto, se utilizan algoritmos genéticos como NSGA II, los cuales producen Frentes de Pareto aproximados [5]. Para el problema que se aborda, se realizó una implementación de dicho algoritmo para producir aproximaciones al Frente de Pareto. Estas aproximaciones están formadas por conjuntos de vectores de valores difusos de tipo intervalo, que se ubican en el espacio de los objetivos del problema. Suponiendo que se tiene el Frente de Pareto óptimo o una aproximación de este, ¿cómo se puede evaluar la calidad del frente generado con respecto al óptimo? Para responder esta cuestión se aplica una transformación del espacio de los vectores difusos de tipo intervalo a los reales. Con una transformación de este tipo, los Frentes de Pareto se pueden transformar en conjuntos de vectores reales. Este tipo de transformación defuzzifica un número difuso asociándole un número real. La transformación que se propone en este artículo es $\delta: F \rightarrow R$ tal que $\delta([a_1, a_2]) = \frac{a_1+a_2}{2}$. Básicamente la transformación asocia a cada intervalo el punto medio del mismo. Una propiedad importante de esta transformación es que si $A \subset F^n, entonces \delta(A) \subset R^n$. Esta propiedad hace posible que el proceso de defuzzificación se aplique en el caso de múltiples objetivos. Una vez transformado el Frente de Pareto, se pueden aplicar las métricas comúnmente usadas para determinar la calidad y la distribución de las soluciones de los frentes.

Para mejorar la distribución de las soluciones en el frente generado, se propone incorporar al NSGA-II un estimador de densidad de las soluciones difusas en lugar de usar la distancia de agrupamiento [5]. En particular se propone adaptar el estimador de densidad reportado en [6].

4 Resultados experimentales

La plataforma de hardware y software en que se realizaron los experimentos incluye: procesador Intel Core i7-7500U 2.70GHz RAM 8GB, lenguaje Java IDE Eclipse.

El propósito de los experimentos es evaluar el impacto del uso del estimador de densidad SSD en el desempeño del algoritmo I-NSGA II. Los algoritmos que se comparan son el I-NSGA-II-CD el cual es la adaptación del algoritmo clásico de NSGA-II para resolver el problema en el dominio de los intervalos. Este algoritmo utiliza como estimador de densidad la distancia de agrupamiento. El algoritmo propuesto es denominado I-NSGA-II-SSD en el cual se sustituye la distancia de agrupamiento por la desviación espacial propagada. Para medir el desempeño se utilizan dos métricas: hipervolumen [7] y dispersión generalizada [8]. El hipervolumen es el volumen del espacio n-dimensional dominado por las soluciones en el conjunto de referencia A. Si dicho espacio es grande, indica que el conjunto se aproxima al Frente de Pareto. Es deseable que se tengan valores grandes del indicador. La dispersión generalizada calcula el promedio

de las distancias los puntos del conjunto de referencia a su vecino más cercano. Si este indicador tiene valores pequeños, significa que las soluciones en el conjunto de referencia están bien distribuidas.

En el experimento se usaron 13 instancias del problema generadas aleatoriamente y para cada instancia se hicieron 30 corridas independientes con cada algoritmo y a cada uno de los frentes encontrados se les aplicó la transformación propuesta para convertirlos en conjuntos de pares de objetivo reales. Para cada uno de los 30 frentes de cada instancia se calcularon las métricas de desempeño, dichos valores se ordenaron en forma ascendente y se determinaron la mediana y el rango intercuartil. Se utilizan estas medidas de tendencia central y de dispersión de los datos ya que son menos sensibles a valores extremos que la media y la varianza. La Tabla 1, muestra los resultados de los experimentos. En cada celda se muestra la mediana y como subíndice el rango intercuartíl que se obtiene para cada instancia y cada algoritmo. En cada caso se señala con sombreado intenso el mejor valor y menos intenso el segundo mejor. En el caso del hipervolumen los mejores valores son los más grandes y para la dispersión generalizada los menores.

Tabla 1. Desempeño de los algoritmos en comparación

	Hipervolumen		**Dispersión Generalizada**	
Instancia	**I-NSGA-II-CD**	**I-NSGA-II-SSD**	**I-NSGA-II-CD**	**I-NSGA-II-SSD**
o2p250	1.00E000.0E00	1.00E000.0E00	2.79E-017.7E-02	2.72E-016.1E-02
o2p251	4.85E-011.4E-01	1.00E000.0E00	4.85E-013.9E-02	3.54E-012.5E-02
o2p252	1.00E000.0E00	1.00E000.0E00	2.65E-015.9E-02	2.56E-012.5E-02
o2p253	1.00E000.0E00	1.00E000.0E00	6.75E-015.8E-02	6.55E-015.7E-02
o2p254	1.00E000.0E00	1.00E000.0E00	4.58E-013.5E-02	3.15E-012.4E-02
o2p255	1.00E000.0E00	1.00E000.0E00	5.33E-012.3E-02	4.30E-011.8E-02
o2p256	1.00E000.0E00	1.00E000.0E00	3.08E-013.5E-02	2.75E-012.9E-02
o2p257	7.84E-012.4E-01	1.00E000.0E00	2.78E-017.8E-02	2.78E-017.8E-02
o2p258	1.00E000.0E00	1.00E000.0E00	2.78E-017.8E-02	2.71E-016.3E-02
o2p259	1.00E000.0E00	1.00E000.0E00	2.65E-015.9E-02	2.54E-012.5E-02
o2p2510	1.00E000.0E00	1.00E000.0E00	3.08E-013.4E-02	2.95E-013.1E-02
o2p2511	1.00E000.0E00	1.00E000.0E00	2.68E-018.2E-02	2.84E-018.6E-02
o2p2512	1.00E000.0E00	1.00E000.0E00	4.53E-012.4E-02	4.22E-011.8E-02

Fuente: Elaboración propia

En la Tabla 1, aparecen sombreadas las celdas con los mejores valores de hipervolumen y de dispersión generalizada. En el primer caso los mejores valores son los que están más cercanos a uno y en el caso de la dispersión son los más pequeños. Como se observa en la Tabla 1, ambos algoritmos producen prácticamente el mismo nivel de calidad, sin

embargo, en dispersión el algoritmo I-NSGA-II-SSD supera claramente al algoritmo I-NSGA-II-CD.

5 Conclusiones

En este artículo se abordó el problema multiobjetivo de la selección de cartera de proyectos con parámetros difusos de tipo intervalo. Se propone una estrategia para medir el desempeño de los algoritmos usando las métricas clásicas que se aplican en el caso de problemas multiobjetivo con valores reales. Experimentalmente se muestra que el algoritmo NSGA-II para intervalos, mejora su capacidad de distribución de las soluciones en el frente si se usa como estimador de densidad la desviación espacial propagada en lugar de la distancia de agrupamiento

Actualmente se está trabajando en un nuevo algoritmo para la versión difusa del problema con parámetros de tipo intervalo, en el que se usa como estimador de densidad la desviación espacial propagada.

Referencias

1. Salo, A., Keisler, J., Morton, A. (2011). Portfolio Decision Analysis. Improved methods for resource allocation, International Series in Operations Research & Management Science. Springer New York, 162 3-27.
2. Moore, R.E. (1979). Methods and applications of interval analysis. Society for Industrial and Applied Mathematics.
3. Fausto A. Balderas-Jaramillo, Modelando la imprecisión del problema de cartera de proyectos con filosofía gris. Instituto Tecnológico de Tijuana (2018).
4. Yao, S., Jiang, Z., Li, N., Zhang, H., & Geng, N. (2011). A multi-objective dynamic scheduling approach using multiple attribute decision making in semiconductor manufacturing. International Journal of Production Economics, 130(1), 125-133.
5. Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In Proceedings of the 6th International Conference on Parallel Problem Solving from Nature, volume 1917 of Lecture Notes in Computer Science, Berlin, Heidelberg. Springer.
6. A. Santiago, Bernabé Dorronsoro, A. J. Nebrod, J. J. Durillo, O. Castillo, H. J. Fraire. A novel multi-objective evolutionary algorithm with fuzzy logic based adaptive selection of operators: FAME. Elsevier Inc.,Volume 471, January 2019, Pages 233-251.
7. L. While , L. Bradstreet , L. Barone , A fast way of calculating exact hypervolumes, IEEE Trans. Evol. Comput. 16 (1) (2012) 86–95.
8. A. Zhou , Y. Jin , Q. Zhang , B. Sendhoff, E. Tsang , Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion, in: Congress on Evolutionary Computation, IEEE, 2006, pp. 3234–3241 .

21

Optimizing the selection of mechanical exoskeletons in the industry: An approach from TOPSIS and MOORA

GRISEL ALVAREZ, SAMUEL SANDOVAL AND ALBERTO OCHOA

Optimizing the Selection of Mechanical Exoskeletons in the Industry: An Approach from TOPSIS and MOORA

Grisel Alvarez[1], Samuel Sandoval[1], Alberto Ochoa[1]
[1]Universidad Autónoma de Ciudad Juárez, Cd. Juárez, Chihuahua, México.
{al176533, al167695}@alumnos.uacj.mx, alberto.ochoa@uacj.mx

Abstract. The constant advancement of technology has caused robotization and automation to be part of our daily life. Nowadays it is sought to mechanize most of the systems, in order to achieve the highest efficiency and of course, meeting the established requirements. Clear example of this, are exoskeletons. Which are designed, according to specialists, by the combination of various technologies; so that they can be a solution to protect workers in industries. Currently, multi-criteria selection methods are a basic tool for decision making in the industry. In the following work we will focus on the application of two models: Topsis and Moora; since they analyze the impact of the results in a somewhat more qualitative way. With the above, it is intended to obtain knowledge of the current state of the market for this type of products and the most recent specifications that the exoskeleton has to satisfy.

Keywords: Multicriteria analysis - Topsis - AMT - Robotization - Biometric sensor - Synergy.

1 Introduction

The increasing level of competitiveness that is managed today in the globalized world, is forcing manufacturing companies to consider investments in Advanced Manufacturing Technologies (AMT), because they allow achieving quality and competitive prices.
Organizations that invest in AMT face a decision problem when choosing the alternative that meets the characteristics desired by the organization [1]. There are various evaluation techniques that allow choosing an alternative that satisfies the characteristics desired by the company and in turn justifies the investment of large economic sums. For this particular job, he will focus on the selection of exoskeletons.
An exoskeleton or exosuit is a device that resembles armor, which consists of an enhancer system that allows the wearer to carry out activities that they could not do without it [2]. The exoskeletons work with a series of biometric sensors, these are in charge of detecting the nerve signals that the brain sends to the muscles of our extremities [3]. So when the exoskeleton processing unit receives them, it responds to those signals by processing them and making the exoskeleton act in a fraction of a second.

2 Proposed methodology

The multicriteria analysis consists of a method that allows to facilitate and guide decision-making based on multiple criteria. It is mainly used to solve decision problems through judgments that arise after making comparisons between various alternatives, so it can also be used as an evaluation method.

The Topsis tool (Technique for Order Preference by Similarity to an Ideal Solution) was developed by Hwang & Yoon (1981) which is based on the concept of selecting the alternative that is a shortest possible distance from an ideal solution and as far as possible from another anti-ideal solution [4].

In this process of selecting alternatives is carried out using the multi-criteria point of view, based on the ordering according to the preferences assigned to each alternative (weight of each factor considered). In this process each alternative A_k is considered a vector in Euclidean space.

$$A^k = (x_1^k \ldots\ldots x_n^k) \; for \; k = 1,2,\ldots\ldots k \quad \textbf{(1)}$$

This technique is based on the assumption that there is an alternative that must be better or worse than all the others. Thus, the alternative with the best nominal characteristics in the attributes is called the ideal solution and is given by the following equation:

$$A^+ = (x_1^+, x_2^+, \ldots\ldots x_n^+) \quad \textbf{(2)}$$

And in the same way the anti-ideal solution is generated, only replacing the "+" by "-". Which is made up of all the unwanted values in the attributes. In such a way that the direct method of carrying out the multi-criteria evaluation is to choose the alternative that has the shortest distance to the ideal solution; or in any case, the one furthest from the anti-ideal solution.

Table 1. Comparison of attributes.

Attribute	Units
Accessibility to get spare parts	Likert scale
Adjustability	Likert scale
Height	Centimeters
Supported areas	0-7
Lifting assistance	Lb
Comfort	Likert scale
Complexity of technology use	Likert scale
Cost	Pesos
Maintenance cost	Pesos
Design	Likert scale
Parts Availability	Likert scale
Charging duration	Hours
Warranty	Months
Weight	kg
Reduced Strength Percentage	Percentage
Resistance	kg
Use of mechanical sensors	0-1
Useful life	Years

The TOPSIS technique uses the Euclidean distance to measure the distances between the attributes (ideal and anti-ideal).

$$d_E(x,y) = \sqrt{(x_{i1} - x_1)^2 + (x_{i2} - x_2)^2} \quad \textbf{(3)}$$

In such a way that, in order to select the exoskeleton that best suits our needs, it is necessary to consider a series of attributes. Here we consider the attributes of table 1.

Now we just have to see which attributes in the previous table should be maximized and which should be minimized. For the elements that use the Likert scale, it should be noted that values between 1 and 7 were considered; depending on the characteristic, 1 will be better than 7 as long as the attribute is minimized. Whereas when looking to maximize, 7 is our ideal solution and 1 is anti-ideal.

In our selection process, we took as reference 5 models of mechanical exoskeletons with application in the industry (all oriented to the upper body). The mechanisms were from Eksobionics: Zero G and Ekso Vest, Laevo V2.5 and ATOUN Model Y from Panasonic and finally Guardian XO from the Sarcos brand.

Table 2. Attributes.

Atribute	ZeroG	EksoVest	Laevo V2.5	ATOUN Model Y	Guardian XO
Accessibility to get spare parts	4	3	5	4	4
Adjustability	2	5	6	5	5
Height	165	172.5	177.5	170	175
Supported areas	4	5	3	3	7
Lifting assistance	9	17.5	5	8.5	75
Comfort	2	4	5	3	3
Complexity of technology use	4	2	2	3	6
Cost	4995	5635	2390	4372	100000
Maintenance cost	900	1127	478	874	4000
Design	2	5	3	4	6
Parts Availability	6	2	5	4	4
Charging duration	8	8	8	4	7
Warranty	6	6	12	12	12
Weight	15.7	4.3	2.8	4.5	68
Reduced Strength Percentage	20	25	15	22	90
Resistance	16.3	6.8	5.6	10	52.5
Use of mechanical sensors	0	1	0	0	1
Useful life	25	22	15	10	30

Considering all the proposed attributes, we proceed to create the ideal and anti-ideal solution with the Topsis method. It should be noted that for some attributes, such as cost or maintenance cost, an approximation to the mean was used to determine the ideal. As well as for the anti-ideal, a value that narrowly approached the highest range of all was considered. Here there was a big problem with the Guardian X0 model because its price far exceeds the price of others.

Another attribute that should be highlighted is the Height factor, since for the ideal solution, the average height of a Mexican [5] was searched in IMCO (Public Policy Research Center) data. While in the anti-ideal solution, a lower and a higher height range were placed. See the results in table 3.

Table 3. Ideal and anti-ideal solution.

Attribute	Units	Ideal Solution	Anti-ideal solution
Accessibility to get spare parts	Likert scale	7	1
Adjustability	Likert scale	7	1
Height	Centimeters	158 a 164	140 a 158 y 164 a 190
Supported areas	0-7	7	0
Lifting assistance	Lb	100	0
Comfort	Likert scale	7	1
Complexity of technology use	Likert scale	1	7
Cost	Pesos	4000	1000000
Maintenance cost	Pesos	1000	4000
Design	Likert scale	7	1
Parts Availability	Likert scale	7	1
Charging duration	Hours	8	0
Warranty	Months	12	0
Weight	kg	20	100
Reduced Strength Percentage	Percentage	100	0
Resistance	kg	50	0
Use of mechanical sensors	0-1	1	0
Useful life	Years	35	5

After that, we will obtain the Euclidean distance both from the ideal solution to each of the models, and from the anti-ideal solution to them.
In tables 4 and 5 we can see the distances obtained from each model to the ideal solution. It must be considered that for the values in which there was a range defined by an upper and a lower limit, a mean was created to determine the Euclidean distance.

Table 4. Distance to ideal solution.

Distance to Ideal Solution				
ZeroG	BksoVest	Laevo V2.5	ATOLIN Model Y	Guardian XO
-3	-4	-2	-3	-3
-5	-2	-1	-2	-2
4	11.5	16.5	9	14
-3	-2	-4	-4	0
-91	-82.5	-85	-91.5	-25
-5	-3	-2	-4	-4
3	1	1	2	5
995	1635	-1610	372	96000
-100	127	-522	-125.6	3000
-5	-2	-4	-3	-1
-1	-5	-2	-3	-3
0	0	0	-4	-1
-6	-6	0	0	0
-4.3	-15.7	-17.2	-15.5	48
-80	-75	-85	-78	-10
-33.7	-43.2	-44.4	-40	2.5
-1	0	-1	-1	0
-10	-13	-20	-25	-5

Table 5. Distance to anti-ideal solution.

Distance to Anti-ideal Solution				
ZeroG	BksoVest	Laevo V2.5	ATOLIN Model Y	Guardian XO
3	2	4	3	3
1	4	5	4	4
-12	-4.5	0.5	-7	-2
4	5	3	3	7
9	17.5	5	8.5	75
1	3	4	2	2
-3	-5	-5	-4	-1
-995005	-994365	-997610	-995628	-900000
-3100	-2873	-3522	-3125.6	0
1	4	2	3	5
5	1	4	3	3
8	8	8	4	7
6	6	12	12	12
-84.3	-85.7	-97.2	-95.5	-32
20	25	15	22	90
16.3	6.8	5.6	10	52.5
0	1	0	0	1
20	17	10	5	25

As we can see in the ideal solution, almost all of them come out negative since the most common thing is that the values are below the ideal and when subtracting them the same (which is greater) gives the negative sign. While in the values of the anti-ideal solution the opposite happens.
Then, with these data we proceed to apply the Euclidean distance formula. Having these values already, now the only thing left to choose an alternative is to give a certain weight to each attribute, so that together they give 100%. Resulting in the table 6.
Here you can see that the value with the most weight of all is the percentage of reduced strength, because according to experts, it is what is considered most necessary when choosing an exoskeleton. Since it is the main function of this; help the worker reduce the force he needs to use to carry out his daily routines within the organization, thus preventing incorrect movements or injuries due to excessive effort or poor posture.

So it only remains to multiply these weights for each attribute respectively of each of the proposed models. The table with the multiplications would be as follows (table 7).

Table 6. Attribute weights.

Attribute	Weights
Accessibility to get spare parts	2%
Adjustability	8%
Height	7%
Supported areas	9%
Lifting assistance	7%
Comfort	5%
Complexity of technology use	3%
Cost	6%
Maintenance cost	5%
Design	4%
Parts Availability	2%
Charging duration	7%
Warranty	3%
Weight	5%
Reduced Strength Percentage	10%
Resistance	8%
Use of mechanical sensors	4%
Useful life	5%
TOTAL	100%

Table 7. Applied weights.

Applied Weights				
ZeroG	EksoVest	Laevo V2.5	ATOLIN Model Y	Guardian XO
0.0849	0.0894	0.0894	0.0849	0.0849
0.4079	0.3578	0.4079	0.3578	0.3578
0.8854	0.8844	1.1555	0.7788	0.9899
0.4500	0.4847	0.4500	0.4500	0.6300
6.4011	5.9035	6.6592	6.4326	5.5340
0.2550	0.2121	0.2236	0.2236	0.2236
0.0713	0.6530	0.6530	0.6342	0.6530
58700.3288	59868.9807	59858.8778	58727.8942	54306.3311
155.0806	143.7993	178.0237	156.4089	150.0000
0.2040	0.1789	0.1789	0.1697	0.2040
0.4020	0.4020	0.0894	0.0849	0.0849
0.5600	0.5600	0.5600	0.3960	0.4950
0.2548	0.2548	0.3930	0.3930	0.3930
4.2205	4.8490	4.9355	4.8375	2.8844
8.2462	7.9057	8.6203	8.1043	9.0554
2.9948	3.4986	3.5801	3.2985	4.2048
0.0400	0.0400	0.0400	0.0400	0.0400
1.1810	1.0700	1.1810	1.2748	1.2748

As a next step we add all the values of each model (per column) and the results are shown in the table 8.

Table 8. Results.

ZeroG	EksoVest	Laevo V2.5	ATOLIN Model Y	Guardian XO
59881.76	59832.29	60063.33	59921.14	54482.91

So if we order them in increasing order (smallest to largest), we can obtain our table of best models (table 9). It should be emphasized that it is ordered from least to greatest because the minors mean that the Euclidean distance was smaller and therefore they do not stray as far from the proposed ideal solution.

Table 9. Selection by Topsis.

1º	54482.91	Guardian XO
2º	59832.29	EksoVest
3º	59881.76	ZeroG
4º	59921.14	ATOLIN Model Y
5º	60063.33	Laevo V2.5

Fig. 1. Selection by Topsis

As we see the model chosen by this technique was Guardian X0 of the Sarcos brand. It was chosen despite its high price, as other even more relevant factors in the industries weighed more when making the decision.
Here is a graph of it to see how much difference there was between each of the proposals. As can be seen, there is a lot of difference between 1st place and the rest, so we can safely say that we have made a very good multi-criteria decision.

3 Compared to a larger group of experts

There are many methods for multicriteria selection, to see another point of view MOORA will be used since it is one of the most powerful and used in various selection processes.
The MOORA (Multi Objective Optimization on basis of Ratio Analysis) method was created by Brauers and Zavadskas in 2006 and allows taking into account various aspects that could not be compared by themselves, due to the difference in their respective measurement units (suppresses these units).
The method begins with the response matrix of the different alternatives and different criteria xij, represented as follows:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1l} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{jl} & \cdots & x_{jn} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{ml} & \cdots & x_{mn} \end{bmatrix} \quad \textbf{(4)}$$

Where i = 1, 2… n are the criteria and j = 1, 2 … m are the alternatives, xij is the response of alternative j.
Next, the response matrix is normalized, where each response xij is compared to a denominator, which is representative of all the alternatives that involve the criterion. The denominator is obtained from the square root of the sum of squares of each alternative by criterion, which is represented as:

$$\overline{x_{ij}} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \quad \textbf{(5)}$$

Then the weight vector of the criteria is defined as follows:

$$W = |W_1 \;\; W_2 \;\; W_3 \;\; \ldots \;\; W_n| \quad \textbf{(6)}$$

The normalized matrix by weights is calculated. It is weighted by multiplying the normalized decision matrix by the weights of each criterion. Then, the aggregation function is determined to evaluate each alternative S (xi), using the equation below.

$$S(x_i) = \sum_{i=1}^{h} \overline{x_{ij}} - \sum_{i=h+1}^{h} \overline{x_{ij}} \quad \textbf{(7)}$$

Where i = 1, 2, 3, ..., h corresponds to the maximum cataloged criteria; i = h + 1, h + 2, ... n corresponds to the criteria listed as a minimum. Finally, the preference ranking is determined. The best alternative is the one with the highest S (xi) value [6]. The variation of the parameters can be seen in the figure 2.
Due to the great variation that exists between the same parameters, we have to perform an aggregation function; such that a single value expresses the entire behavior of the device. In order to evaluate the best alternative.
To finish we only obtain the aggregation function to evaluate each alternative S (xi). For this we consider the minimized alternatives; that is, previously multiplied by 0.04, with a negative sign and the maximized ones (multiplied by 0.06), with a positive sign to add them and obtain their final value.
Now by ordering the answers from highest to lowest we can obtain their preference ranking. The result can be seen in the table 10.
In figure 3 we see how much difference there was between each of the proposals.

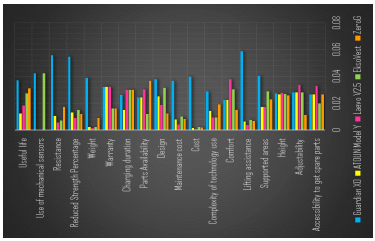


Fig. 2. Variation of the parameters with Moora

Table 10. Results by Moora.

	S(xi)	Ranking
ZeroG	0.225835	5
EksoVest	0.298891	2
Laevo V2.5	0.281937	3
ATOUN Model Y	0.233910	4
Guardian XO	0.369689	1

Fig. 3. Results by Moora

As we can see, the same result was reached as with the other method, the big difference here is the accommodation of the other positions. To better appreciate these differences, we have the table 11.

Table 11. Comparison between the two methods. Source: Own elaboration.

	Topsis	Moora
1°	Guardian XO	Guardian XO
2°	EksoVest	EksoVest
3°	ZeroG	Laevo V2.5
4°	ATOUN Model Y	ATOUN Model Y
5°	Laevo V2.5	ZeroG

The use of the Moora method has many advantages, for this reason it is one of the most used currently for multicriteria analysis. Furthermore, at the time of execution it is also very practical [7].

On the contrary, topsis uses a very predictive and rational logic, which makes it a friendly method; although according to studies, a great disadvantage of topsis is that if the values

are very close, objectivity is lost because the ranking of results can vary greatly compared to other multicriteria techniques.

4 Conclusions

It should be emphasized that there are an infinite number of methods for multicriteria analysis of variables, here we only analyze two (Topsis and MOORA).
the calculations can be performed on any computer that has Excel or that has any spreadsheet. In addition to that the algorithms developed for each of the techniques are already manufactured in the sheet used for this decision making.
It should be considered that any procedure is a methodological aid to decision-making, which must be considered with due caution, since in any case, in the multi-criteria evaluation process, as well as in other procedures used for the same purpose, they cannot be eliminated totally subjective factors.
In order to minimize the risk, a specific committee is generally formed, made up of the main actors involved in the subject, such as in our case, they must be experts in the field of automation.
Without a doubt we can say that the best exoskeleton model is Sarcos Guardian X0, as it best combines those requirements of the current manufacturing industry, despite the fact that the price is relatively high compared to the others.

References.

1. M, J. Mojarro, Zúñiga, (2007, Jan. 01), "Mejoramiento de la técnica Topsis mediante la integración de la distancia de Mahalanobis". [Internet]. Available at https://www.researchgate. net/publication/4812386_MEJORAMIENTO_DE_LA_ TECNICA_TOPSIS_MEDIANTE_LA _INTEGRACION_DE_LA_DISTANCIA_DE_MAHALANOBIS
2. A. Muñoz, (2017, Mar. 04), "¿Qué es exoesqueleto?". [Internet]. Available at https:// computerhoy.com/ noticias/hardware/que-es-exo esqueleto-59152
3. M. Ramírez, (2019, Mar.), "Exoesqueletos, el camino hacia la robotización humana". [Internet]. Available at http://www.tecnicaindustrial.es/TIFrontal/a-10162-exoesqueletos--camino-robotiza cion-humana.aspx
4. J C. Correa, L M. Ruiz, (2014, Feb 26), "Aplicación de una herramienta multicriterio para selección de proveedores en un hospital del norte del valle". [Internet]. Available at http:// bibliotecadigital.univalle. edu.co/bitstream/10893/ 10526/1/CB-0509949.pdf
5. IMCO Staff, (2016, Feb. 05), "¿CÓMO ES EL MEXICANO PROMEDIO? VÍA EL PAÍS". [Internet]. Available at https://imco.org.mx/como -es-el-mexicano-promedio-via-el-pais/
6. M. Chérrez, J. Martínez, D. Peralta, E A. Llanes, (2018, May. 14), "Métodos Multicriterio Aplicados En La Selección De Un Material Para Discos De Freno". [Internet]. Available at https:// www.redalyc.org/ jatsRepo/5055/505555586008 /html/index.html#gf48
7. J L. García, J. Romero, I. Canales, (2020, Apr. 14), "Selección De Proveedores Usando El Método Moora". [Internet]. Available at http:// erevistas.uacj.mx/ojs/index.php/culcyt/article/view/282/ 266

Parte V

Aplicaciones en el Sector Eléctrico

Introducción

En el sector eléctrico la demanda, producción y transformación de energía son factores preponderantes para el diseño, operación y mantenimiento de diversos sistemas y mecanismos que confluyen en dicha industria. El desarrollo y operación de sistemas computacionales que utilizan técnicas de inteligencia artificial para pronóstico y clasificación de variables, parámetros y mediciones ha permitido mejorar en gran medida el accionar de la industria, propiciando decremento de costos de producción o almacenamiento, así como se optimización de recursos que se requieren para su producción.

En años recientes, la aplicación de sistemas computacionales ha demostrado grandes avances y mejoras en los principales retos del sector eléctrico. En esta sección del libro, se presentan cuatro trabajos con aplicación en el sector eléctrico, los cuales representan algunos de los esfuerzos realizados en esta dirección en diversas instituciones nacionales. Cabe señalar que cada artículo fue evaluado por al menos dos revisores expertos del área en un proceso doble ciego, garantizando así la calidad e innovación de los trabajos que aquí se exponen.

Héctor Rodríguez Rangel

22

Predicción del Recurso Solar mediante Diferencias

JUAN JOSE FLORES, BALDWIN CORTES, JOSE CEDEÑO, ANTONY MORALES, HECTOR RODRIGUEZ Y FELIX CALDERON

Predicción del Recurso Solar mediante Diferencias

Juan J. Flores[1], Baldwin Cortes[1], José R. Cedeño González[1], Antony Morales[1], Hector Rodriguez[2], Roberto Tapia[1], and Felix Calderon[1]

[1]Universidad Michoacana de San Nicolás de Hidalgo
Facultad de Ingeniería Eléctrica
División de Estudios de Posgrado
[2]Division de Estudios de Posgrado e Investigación
Tecnológico Nacional de México Campus Culiacán
Juan de Dios Batiz 310 pte. Col Guadalupe C.P.80220, Culiacán, Mexico

Resumen. La radiación solar disponible en el lugar donde se encuentra o se planea instalar una planta fotovoltaica depende principalmente de la incidencia de nubes en el sitio, la cual depende del sistema climático de la región. En trabajos anteriores de predicción de la radiación solar, usando Redes Neuronales Artificiales, se ha observado que los modelos fallan al predecir los picos de radiación en condiciones de nubosidad intermitente. Para solucionar este problema, se propone no predecir la irradiancia solar directamente, sino la cantidad de radiación ocluida por las nubes; es decir, la diferencia entre la irradiancia medida por la estación meteorológica y la irradiancia teórica máxima. Esta última es obtenida utilizando el modelo *clear sky*, el cual proporciona la irradiancia ideal para una localidad y hora específica. Al modelo de predicción propuesto se le denomina SolarDiff, el cual pronostica esa diferencia utilizando Redes Neuronales Artificiales. En este artículo se demuestra empíricamente que el modelo SolarDiff exhibe un mejor comportamiento que los modelos basados en los datos directos, utilizando la métrica de error SMAPE.

Palabras claves: Predicción, Series de Tiempo, Radiación Solar, Redes Neuronales Artificiales

1 Introducción

Predicción de ST es un área de investigación muy importante hoy en día, debido a la importancia que tiene en varios campos de la actividad humana. Una de sus aplicaciones es el tema de este artículo: la radiación solar. La radiación solar y la temperatura son las dos variables más relevantes en la determinación de la cantidad de energía eléctrica que puede producir una planta fotovoltaica [1]. La importancia de la predicción de la radiación solar reside en la alta variabilidad de la cantidad de energía que produce una planta fotovoltaica. Estas variaciones causan problemas a la red eléctrica a la que se conecta la planta [2]. Adicionalmente, para los operadores de un sistema eléctrico, es necesario tener una estimación de la cantidad de energía que será producida por las plantas basadas en

energías alternativas, para así poder planear la asignación de la generación de plantas convencionales y poder en todo momento, satisfacer la demanda de la red.

Este artículo representa un reporte del trabajo de investigación desarrollado a la fecha por este equipo interdisciplinario de investigadores.

1.1 Series de Tiempo

Una ST es una secuencia ordenada en tiempo de observaciones de una variable – X, realizadas en intervalos de tiempo equiespaciados, representadas como un conjunto de valores discretos $\{x_1, x_2, x_3, \cdots, x_N\}$ [3]. Las ST se encuentran muchas actividades del ser humano: ingeniería, economía, finanzas, oceanografía, astronomía, etc. [4]. El análisis de ST ayuda a detectar regularidades en las observaciones de una variable, determinar un modelo adecuado, explotar la información incluida en los datos que la representan, de manera de determinar y explicar su comportamiento futuro [5].

Existen diferentes técnicas de predicción, las cuales pueden ser clasificadas en tres tipos: modelos estadísticos, inteligencia artificial e híbridos [6]. Los estadistas y economistas tienden a utilizar modelos del tipo ARIMA (Autoregresive Integrated Moving Average) y modelos derivados [7]. Por otro lado, los modelos basados en técnicas de Inteligencia Artificial y Aprendizaje de Máquina se basan principalmente en Redes Neuronales Artificiales (RNA - Perceptrones Multicapa o Redes Neuronales Recurrentes) [8].

1.2 Trabajo relacionado

Las condiciones de cielo despejado representan la condición más favorable (i.e., de rendimiento máximo) en la operación de los sistemas solares [9]. En la actualidad, existen dos enfoques principales para el pronóstico de la producción de energía solar. La primera opción consiste en usar ecuaciones analíticas para modelar los sistemas fotovoltaicos y la segunda opción consiste en predecir directamente la potencia de salida utilizando métodos estadísticos y de aprendizaje automático [10]. En contraste al enfoque paramétrico, para análisis generalmente se requiere un gran conjunto de datos meteorológicos y de potencia. La selección de un conjunto de datos de entrenamiento es crucial en la precisión del modelo desarrollado [11].

Los métodos de pronóstico que utilizan técnicas de inteligencia artificial (AI) y aprendizaje de máquina (ML) se basan en técnicas como redes neuronales artificiales (ANN), vecinos cercanos (kNN), aprendizaje de máquina extremo (ELM) y máquina de soporte vectorial (SVM), entre otros [12]. Algunos modelos utilizan ANNs para estimar directamente la energía producida utilizando diferentes entradas como humedad, temperatura, precipitaciones, radiación solar y datos meteorológicos, sin embargo, los modelos mejorados consideran la correlación de datos y las redes neuronales artificiales para seleccionar las entradas más importantes, para poder entregar mejores resultados [13]. Es importante mencionar

que en la literatura hay muchos estudios de evaluación comparativa para determinar las técnicas y las condiciones óptimas en las que funcionan mejor las estimaciones en diversas situaciones [14].

1.3 Medición del error

Un modelo de pronóstico para ST se puede evaluar con medidas estadísticas, que generalmente se basan en el error. La diferencia entre el valor pronosticado y el valor observado en el mismo instante se conoce como error de pronóstico. Para evaluar la cantidad de este error de pronóstico en un conjunto de datos se utiliza el error porcentual absoluto simétrico medio (SMAPE), definido por la ecuación (1). Esta versión de SMAPE varía de 0 a 100 %.

$$SMAPE = \frac{100}{n} \sum_{t=1}^{n} \frac{|y_t - \hat{y_t}|}{y_t + \hat{y_t}} \tag{1}$$

donde n es la longitud de la serie de tiempo, y_t representa el valor real, y $\hat{y_t}$ el valor pronosticado.

2 Predicción de la Irradiancia Solar

Esta sección presenta la propuesta de modelo de pronóstico de irradiancia solar. La radiación solar incidente en un punto depende principalmente de la variabilidad de la nebulosidad. Predecir la nebulosidad hace que la predicción de la radiación solar sea considerablemente compleja. En trabajos anteriores de predicción de la radiación solar, usando RNA, se ha observado que los modelos fallan al predecir los picos de radiación en condiciones de nubosidad intermitente. Existen modelos para determinar la radiación solar máxima en cualquier punto, para una fecha determinada, llamados de cielo despejado (clear sky models en inglés). La información disponible (irradiancia solar ideal y medida) ha sido tomada en cuenta para la formulación del modelo propuesto.

2.1 Modelo de Cielo Despejado

Los modelos a cielo despejado son aquellos modelos parametrizados que estiman la irradiancia solar sin presencia de nubes para un punto determinado [15]. Existe una amplia variedad de modelos que resuelven este problema; en este trabajo se utilizó el modelo SOLIS simplificado (sSOLIS). En el modelo sSOLIS se hacen los cálculos basados en modelo de transferencia radiativa (*radiative transfer model* – RTM). La base del modelo SOLIS, y por consiguiente del modelo sSOLIS, es la relación Lambert-Beer [16]. Esta relación debe ser modificada para extender la expresión de bandas de longitud de onda y para considerar carga alta de aerosoles. Al ser extendida la expresión a bandas de longitud de onda, es posible utilizar la irradiancia de haz, global o difusa de manera indistinta. En este trabajo se utiliza la irradiancia global, la cual se describe en su forma extendida en la Ecuación (2).

$$I_{\mathrm{gh}} = I'_o \cdot \exp\left(-\frac{\tau_g}{\sin^g h}\right) \cdot \sin h \tag{2}$$

donde I'_o es la irradiancia extraterrestre mejorada para alta carga de aerosoles, τ_g es la profundidad óptica total global, h es el ángulo de elevación solar y g es el parámetro de ajuste.

En el modelo SOLIS original, se requiren como parámetros la columna de vapor de agua y el contenido de aerosoles. Los cálculos RTM son costosas computacionalmente y son difíciles de implementar para procesos en tiempo real. En el modelo simplificado los cálculos RTM se realizan para diferentes rangos de altitudes y condiciones atmosféricas y el mejor ajuste se toma con respecto al aerosol atmosférico y al contenido de vapor de agua así como la altitud [17]. En el artículo original se incluyen los coeficientes para las ecuaciones con alta carga de aerosoles.

2.2 Redes Neuronales Artificiales

Una RNA consiste de un conjunto de neuronas interconectadas. Éstas pueden ser descritas usando un grafo dirigido, donde cada neurona i (un nodo del grafo) contiene una función de transferencia f_i de la forma

$$y_i = f_i\left(\sum_{j=1}^{n} w_{i,j} x_j - \Theta_i\right) \tag{3}$$

donde y_i es la salida del nodo i, x_j is the $j-$ésima entrada a la neurona y $w_{i,j}$ es el peso de la conexión entre los nodos i y j. Θ_i es el sesgo de la neurona [18].

Una RNA conectada hacia adelante puede considerarse como un mapeo entre sus conjuntos de entrada y salida. La arquitectura de conexión hacia adelante propaga de la capa de entrada a las capas ocultas y a la de salida. Durante el entrenamiento, BP se utiliza para llevar los errores hacia atrás para ajustar los pesos.

Una consecuencia del Teorema de Kolmogorov (1957) indica que cualquier mapeo continuo $y(x)$ de d variables de entrada x a una variable de salida y puede ser representado exactamente por una RNA de alimentación hacia adelante de tres capas [19]. Este teorema, unido con la falta de hardware suficientemente rápido para realizar el entrenamiento de RNA profundas, hizo que por mucho tiempo solo se utilizaran las RNA no profundas. Con el desarrollo de hardware y algoritmos de entrenamiento más eficientes, el uso de las RNA Profundas ha crecido mucho en los últimos años [20,21].

El modelo SolarDiff utiliza RNA para modelar y predecir la cantidad de radiación solar que incide sobre un punto sobre la superficie de la tierra (una planta fotovoltaica).

2.3 Modelo SolarDiff

El modelo propuesto es muy simple, pero no se encuentra reportado en ningún trabajo (del conocimiento de los autores) de investigación relacionado.

Sea X la serie de tiempo de irradiancia solar, medida en el sitio de interés y CS la irradiancia solar ideal del sitio. CS es calculada por medio del modelo de Cielo Despejado y se considera como la máxima irradiancia que se tendrá en el sitio en un momento determinado. La diferencia de estas dos series de tiempo está expresada por la Ec. (4).

$$Y = CS - X \tag{4}$$

La propuesta es predecir Y en lugar de X directamente. Una vez predicha la diferencia Y, volvemos al dominio original usando la misma ecuación, resolviendo para la variable X, como se muestra en la Ec. 5

$$X = CS - Y \tag{5}$$

La predicción de la serie de tiempo Y se realiza mediante un modelo ANN profundo. Para resolver el problema de predicción por medio de ANN, la serie de tiempo se convierte en una base de datos, tal y como lo proponen Flores et al. [22,23].

3 Resultados

La serie de tiempo de irradiancia solar utilizada en este artículo fue obtenida en una estación meteorológica ubicada en la ciudad de Amarillo, Texas[1]. La serie de tiempo está compuesta de 350,400 muestras, las cuales tienen una resolución de 30 minutos y abarcan el periodo de tiempo de 1998 a 2018.

Debido a que la irradiancia solar contiene un componente de aleatoreidad, el cual depende de la nubosidad presente en la zona en cuestión, es necesario utilizar un *Test Set* representativo de la serie de tiempo. Para la selección de éste se obtuvo la cantidad total de irradiancia incidente diaria (proporcional a la suma de las mediciones en un día), y posteriormente se realizó la clasificación en grupos por niveles de irradiancia. Finalmente se eligió de manera aleatoria un día por cada categoría; el conjunto de días seleccionados para el *Test Set* se puede apreciar en la Fig. 1, donde el número de categoría incrementa con la cantidad de irradiancia solar en ese día.

La tarea de predicción es de un paso en adelante a partir de 96 muestras anteriores (i.e, 2 días completos). Como referencia se utiliza un modelo de pronóstico basado en las mediciones directas, y se compara con el modelo SolarDiff aquí propuesto. Ambos modelos son formulados por RNA profundas, y constan de una capa de entrada de 96 neuronas, 3 capas ocultas de 100 neuronas y una capa de salida de una sola neurona, la cual produce el pronóstico a un solo paso. Para el entrenamiento, se utilizó un *learning rate* de 1×10^{-3} y el método de regularización *early stopping*.

El desempeño de los modelos, promedio de 10 experimentos independientes, se muestra en la Tabla 1. El modelo SolarDiff tiene un mejor desempeño en 20

[1] disponible en *National Solar Radiation Dabatabase* https://maps.nrel.gov/nsrdb-viewer/

métricas de error en las diferentes categorías, mientras que el modelo normal lo hace en 12 métricas.

Los resultados de la tarea de predicción para los días 1 y 3 del test set se muestran en la Figura 2. Estos días exhiben picos en su comportamiento, que son los que deseamos mejorar con el método propuesto. Note que en la categoría 1 el modelo por diferencias predice irradiancias negativas, lo cual evidentemente es incorrecto. El tratamiento de este problema en trabajos futuros puede mejorar aún más el desempeño del modelo SolarDiff.

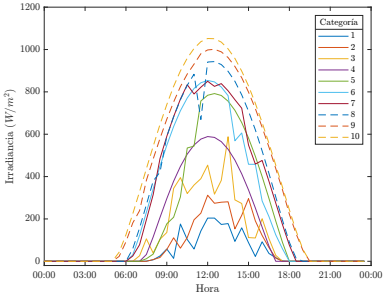
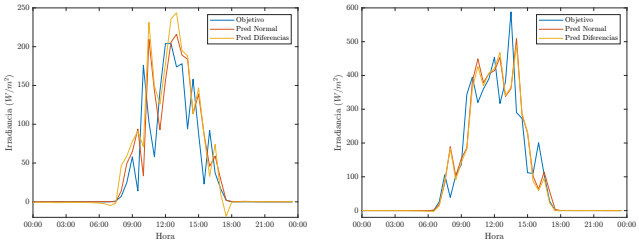


Figura 1: *Conjunto de días seleccionados para el Test Set.*

Tabla 1*: Comparación del desempeño de los modelos de predicción en diferentes métricas*

	Normal			Diferencias		
Categoría	RMSE	MAE	SMAPE	RMSE	MAE	SMAPE
1	**39.0761**	**19.8898**	**68.77**	42.5678	23.2226	72.70
2	32.6770	15.8060	**66.48**	**32.6166**	**14.2917**	66.65
3	67.1654	31.3412	63.73	**66.9748**	**31.1939**	**63.47**
4	8.4171	3.6965	61.64	**3.9937**	**2.1706**	**61.43**
5	**35.5953**	**12.2047**	**58.97**	38.8990	14.3533	59.98
6	**29.0159**	11.3749	54.18	29.4377	**10.7805**	**53.68**
7	**26.5729**	**12.5967**	**51.88**	28.3339	13.6117	52.24
8	**49.7112**	19.7051	50.70	50.3168	**15.3928**	**49.25**
9	13.6439	8.7428	45.01	**13.3790**	**7.8326**	**42.52**
10	15.5061	10.2231	43.97	**4.0973**	**2.8317**	**42.01**
Promedio	**38.4557**	16.2658	57.49	39.3304	**15.6527**	**57.05**



a) Categoría 1 b) Categoría 3

Figura 2: *Desempeño de los modelos en la tarea de predicción en diferentes categorías*

4 Conclusiones y Trabajo Futuro

Las experiencias anteriores de nuestro equipo de trabajo demuestran que es difícil para una RNA predecir los picos de irradiancia solar para días nublados, pero el resto de la predicción no es tan mala. Sin embargo, esas diferencias en los picos hacen que el error total de predicción se incremente.

El modelo SolarDiff propone no predecir la irradiancia solar directamente, sino la diferencia entre la irradiancia teórica máxima y la irradiancia medida por la estación meteorológica. El modelo de cielo despejado proporciona la irradiancia solar máxima que se puede obtener en un día soleado, para una localidad y hora específica.

Los resultados de los experimentos mejoran la predicción para los picos. El modelo SolarDiff resuelve el problema de los picos, pero no es mejor para los otros casos. La dirección que tomará nuestro trabajo de investigación en el futuro próximo es producir un modelo que nos indique en qué casos utilizar el modelo normal y en cuáles utilizar el de diferencias.

Bibliografía

1. B. Cortés, R. Tapia Sánchez, and J. J. Flores, "Characterization of a polycrystalline photovoltaic cell using artificial neural networks," *Solar Energy*, vol. 196, pp. 157 – 167, 2020.
2. S.-E. Razavi, E. Rahimi, M. S. Javadi, A. E. Nezhad, M. Lotfi, M. Shafie-khah, and J. P. Catalão, "Impact of distributed generation on protection and voltage regulation of distribution systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 157 – 167, 2019.
3. A. Palit and D. Popovic, *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*. Advances in Industrial Control, Springer London, 2006.

4. C. Granger, P. Newbold, and K. Shell, *Forecasting Economic Time Series*. Elsevier Science, 2014.
5. G. Kirchgässner and J. Wolters, *Introduction to modern time series analysis*. Springer Science & Business Media, 2007.
6. S. Xu, H. K. Chan, and T. Zhang, "Forecasting the demand of the aviation industry using hybrid time series sarima-svr approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 122, 02 2019.
7. G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
8. C. Lemke and B. Gabrys, "Meta-learning for time series forecasting and forecast combination," *Neurocomputing*, vol. 73, no. 10-12, pp. 2006–2016, 2010.
9. J. A. Ruiz-Arias and C. A. Gueymard, "Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface," *Solar Energy*, vol. 168, pp. 10–29, 2018.
10. J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016.
11. M. P. Almeida, O. Perpinan, and L. Narvarte, "Pv power forecast using a non-parametric pv model," *Solar Energy*, vol. 115, pp. 354–368, 2015.
12. S. Russel, P. Norvig, *et al.*, *Artificial intelligence: a modern approach*. Pearson Education Limited, 2013.
13. Y. Tao and Y. Chen, "Distributed pv power forecasting using genetic algorithm based neural network approach," in *Proceedings of the 2014 International Conference on Advanced Mechatronic Systems*, pp. 557–560, IEEE, 2014.
14. M. Zamo, O. Mestre, P. Arbogast, and O. Pannekoucke, "A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. part ii: Probabilistic forecast of daily production," *Solar Energy*, vol. 105, pp. 804–816, 2014.
15. F. Antonanzas-Torres, R. Urraca, J. Polo, O. Perpiñán-Lamigueiro, and R. Escobar, "Clear sky solar irradiance models: A review of seventy models," *Renewable and Sustainable Energy Reviews*, vol. 107, pp. 374–387, 2019.
16. D. F. Swinehart, "The beer-lambert law," *Journal of chemical education*, vol. 39, no. 7, p. 333, 1962.
17. P. Ineichen, "A broadband simplified version of the solis clear sky model," *Solar Energy*, vol. 82, no. 8, pp. 758–762, 2008.
18. X. Yao, "Evolving artificial neural networks," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1423–1447, 1999.
19. C. M. Bishop *et al.*, *Neural networks for pattern recognition*. Oxford university press, 1995.
20. D. A. Winkler and T. C. Le, "Performance of deep and shallow neural networks, the universal approximation theorem, activity cliffs, and qsar," *Molecular informatics*, vol. 36, no. 1-2, p. 1600118, 2017.
21. I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
22. J. Flores, J. R. Cedeno Gonzalez, H. Rodríguez, M. Graff, R. Lopez-Farias, F. Calderon, *et al.*, "Soft computing methods with phase space reconstruction for wind speed forecasting—a performance comparison," *Energies*, vol. 12, no. 18, p. 3545, 2019.
23. J. J. Flores, J. R. C. González, R. L. Farias, and F. Calderon, "Evolving nearest neighbor time series forecasters," *Soft Computing*, vol. 23, no. 3, pp. 1039–1048, 2019.

23

Metodología para la estimación del diámetro en líneas de transmisión mediante reconstrucción 3D

VICTOR ALEJANDRO GONZALEZ HUITRON, ABRAHAM EFRAIM RODRIGUEZ MATA Y MODESTO MEDINA MELENDREZ

Metodología para la estimación del diámetro en líneas de transmisión mediante reconstrucción 3D

Victor Gonzalez-Huitron[1], A.E. Rodriguez-Mata[1], and Modesto Medina Melendrez[2]

[1] CONACyT-Tecnológico Nacional de México/Instituto Tecnológico de Culiacán, Sinaloa, México
victor.gonzalez@itculiacan.edu.mx
[2] Tecnológico Nacional de México/Instituto Tecnológico de Culiacán, Sinaloa, México

Abstract. En las redes de transmisión eléctrica, las tareas de mantenimiento y reparación son especialmente riesgosas para el personal que las lleva a cabo, además de impactar en las pérdidas económicas generadas por la pausa al suministro. En este trabajo se presenta una metodología que puede reducir los riesgos que conllevan dichas tareas, además de automatizar la obtención de datos referentes al cableado. La propuesta consiste en un conjunto de algoritmos que sean capaces de estimar el calibre de las líneas de transmisión a partir de un par de imágenes estéreo, la metodología propuesta consiste en la calibración de un par de cámaras estéreo de alta resolución, segmentación de la región de interés, estimación del mapa de disparidad y reconstrucción de la escena para determinar el diámetro del cable.

Keywords: WKS6 Aplicaciones en el sector eléctrico · Procesamiento de imágenes · Lineas de transmisión · Reconstrucción 3D.

1 Introducción

En los sistemas de distribución eléctrica, la obtención de parámetros en sitio es clave para una rápida atención y solución a operaciones previstas o no previstas, donde debido a la premura, altura y altos voltajes, es considerado un trabajo de alto riesgo [1]. A pesar de los avances en las operaciones realizadas en campo, aun existen operaciones que son llevadas a cabo de forma manual por personal en sitio, ya sea en zonas urbanas o rurales. En ocasiones dichas operaciones involucran en primeras instancias la obtención de los parámetros de las líneas de transmisión, como lo son el diámetro o calibre (American Wire Gauge, AWG) a utilizar para el reemplazo, reconexión o mantenimiento.

En la actualidad, los vehículos aéreos no tripulados (UAV) o drones, han comenzado a ser parte de tareas especializadas enfocadas al sector eléctrico, en tareas donde se busca reducir el riesgo del personal y/o acceder a ubicaciones o posiciones complicadas, lo cual ha generado varias investigaciones basadas

en el procesamiento de imágenes e inteligencia artificial, utilizando cámaras de video montadas en los drones con el fin de obtener imágenes desde distancias seguras para el operador, con aplicaciones como la detección de irregularidades en estructuras [2], estimación de la longitud y continuidad del tendido eléctrico [3], monitoreo de sistemas de distribución [4] o la estimación del calibre en una línea de transmisión [5, 6].

En este trabajo se presenta una metodología general para la estimación del diámetro, el cual está relacionado directamente con el calibre AWG de un cable, a continuación se explica brevemente los siguientes pasos para dicho proceso: 1) Calibración de cámaras en arreglo estéreo 2) Segmentación de la región de interés, 3) Estimación de disparidad y 4) Reconstrucción 3D de la región de interés.

2 Metodología general

La propuesta presentada consiste en los pasos necesarios para la manipulación de las imágenes con un lenguaje de programación apropiado para operaciones con matrices, en la Fig. 1 se muestra un diagrama de la propuesta y los pasos intermedios que serán descritos a continuación

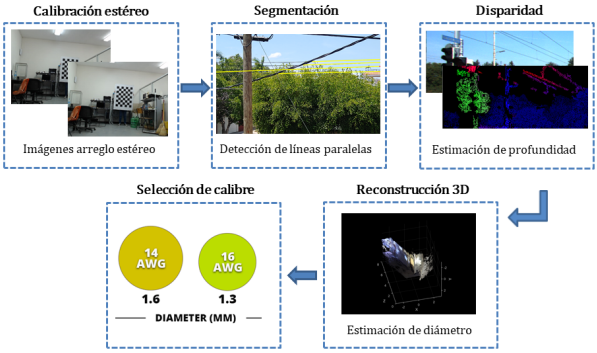


Fig. 1. Diagrama general de la metodología propuesta.

2.1 Calibración de arreglo estéreo

La etapa de calibración consiste en determinar los parámetros geométricos e intrínsecos de un arreglo estéreo de cámaras para lograr una rectificación y pos-

teriormente una reconstrucción 3D a partir del mapa de disparidad. La calibración es un paso fundamental para la extracción de información relacionada a las dimensiones reales de la escena a partir de imágenes 2D [7]. La calibración se realiza a partir de un objeto en un plano cuyas dimensiones son conocidas, a partir del cual pueden establecerse planos ortogonales relacionados con los ejes principales de las cámaras.

Para la realización de dicho proceso se deben tomar varias fotografías con el arreglo estéreo, para obtener una base de datos con escenas desde ambas perspectivas, done se utiliza un patrón conocido con conocimiento de sus medidas reales, el patrón más común de utilizar es un tablero con patrón de cuadros blanco y negro similar a un tablero de ajedrez. Esto logra una obtención rápida de los puntos principales y origen de dicho tablero en las escenas izquierda y derecha, obteniendo mediante relaciones geométricas los parámetros para efectuar una corrección a la distorsión radial de cada cámara, y para obtener los parámetros que se necesitan para la rectificación de las cámaras, garantizando que exista una condición de epipolaridad, necesaria para la estimación de disparidad. En la Fig. 2 se observan los planos estimados a partir de los parámetros de calibración de una escena, cada plano indica la posición del patrón de calibración respecto a las cámaras para cada fotografía.

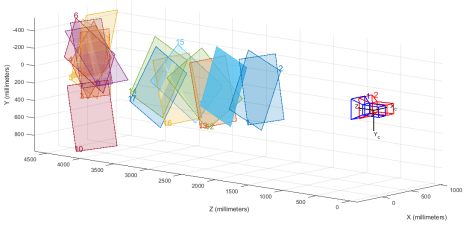


Fig. 2. Reconstrucción de escena a partir de calibración estéreo

2.2 Segmentación de región de interés

Para reducir el costo computacional derivado del alto número de operaciones que requiere procesar imágenes de alta resolución, es indispensable delimitar solamente las regiones que sean de interés, en este caso el cableado del tendido eléctrico, cabe destacar que con una cercanía de tan solo unos cuantos metros o incluso menor a un metro, el tendido eléctrico es prácticamente una línea recta, lo cual implica que con un detector de líneas, como lo es la transformada

de Hough, el cual tiene como paso previo la detección de bordes mediante el método de Canny [8], puede establecerse un criterio para la selección de una región de interés [9] estableciendo así una imagen binaria que contenga solamente los cables como región a trabajar para la siguiente etapa. En la Fig. 3 se muestra un ejemplo del desempeño de estos algoritmos para la detección de un cableado.

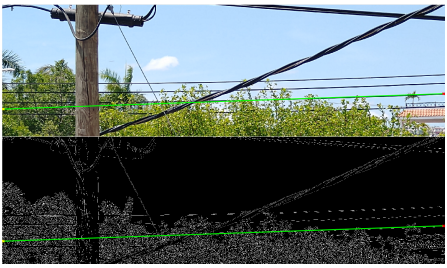


Fig. 3. Detección de tendido eléctrico a partir de transformada de Hough

2.3 Estimación de disparidad

La semejanza estéreo que se lleva a cabo para este proceso, consiste esencialmente en determinar la diferencia en la posición relativa para un punto cualquiera, que se encuentre en ambas imágenes, izquierda derecha, donde esta diferencia en posición relativa se conoce como disparidad, y se encuentra directamente relacionada con la profundidad de la escena [10]. Para esta etapa se considera que existen a la entrada 2 imágenes, izquierda y derecha delimitada en las filas que abarcan la región de interés, esto debido a que los algoritmos de disparidad dependen de que se cumpla una condición de epipolaridad, que establece una rectificación previa, para que todo valor en una determinada fila, tenga su correspondiente en la otra imagen [11], dentro de la misma fila, limitando la búsqueda a solamente el eje horizontal. Como podemos ver en la Fig. 4 el proceso para la estimación de disparidad se lleva a cabo para cada pixel de una imagen de referencia, siendo lo más común que esta sea la escena izquierda.

2.4 Reconstrucción 3D

Una vez generado el mapa de disparidad es posible utilizar el algoritmo Structure From Motion [12] el cual va a relacionar la disparidad y la posición de cada pixel, con su valor correspondiente en el espacio X,Y,Z, donde la distancia al objeto estará dada por la magnitud de estas variables, y a partir de estas, si el cable

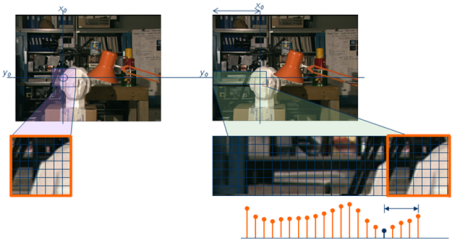


Fig. 4. Proceso para la búsqueda de correspondencia estéreo

está trazado sobre una linea horizontal, su diámetro estará dado por la suma de los valores Y para los pixeles que sean ortogonales a la dirección del cable, pudiendo realizar la estimación directa al calibre del cable procesado.

3 Conclusiones

En este estudio se ha presentado una metodología que debe ser capaz de establecer una aproximación adecuada para el calibre de tendidos eléctricos a partir de imágenes estéreo, las cuales pueden ser capturadas con un vehículo aéreo no tripulado, siendo la mayor dificultad del proceso el establecimiento de los múltiples parámetros relativos a las dimensiones de la escena, por lo que se recomienda el uso de un arreglo estéreo totalmente fijo, para que al realizar la calibración, esta sea generalizada para dicho modelo y resolución específica. En cuanto a una futura implementación, se debe considerar la relación que existe entre una reconstrucción mas fiel a la escena y el costo computacional que esto conlleva, debido a que a mayor resolución y niveles de disparidad en el área de búsqueda, mayor el número de operaciones que deben llevarse a cabo, siendo esta característica una limitante para su aplicación en hardware con poca capacidad para trabajar con imágenes de alta resolución o bien datos en paralelo.

Como trabajo a futuro se plantea llevar a cabo en una primer etapa el desarrollo de algoritmos y su ajuste para la reconstrucción 3D, las medidas de calidad corresponderán para cada tarea de visión artificial, así para la segmentación del tendido eléctrico se utilizará Intersection Over Union, que consiste la comparación entre la región segmentada y la región etiquetada manualmente. El algoritmo de disparidad será evaluado mediante el criterio de cantidad de pixeles mal relacionados, el cual se entrega como un porcentaje en relación al número total de disparidades estimadas incorrectamente. Por último, la reconstrucción 3D para el cableado se realizará primero en condiciones controladas, con fondo e iluminación controlado y previo conocimiento del diámetro de distintos cables,

para obtener una comparación con el diámetro estimado por el proceso. La segunda etapa consiste en la implementación en hardware específico y la estimación de diámetros en condiciones no controladas, con conocimiento del diámetro del cable a estimar sus dimensiones.

References

1. Apagón CFE en Yucatán, https://www.elfinanciero.com.mx/peninsula/peninsula-de-yucatan-sufre-tercer-apagon-del-2019. Last accessed 23 May 2020
2. Shajahan, N.M. et al.: Automated Inspection of Monopole Tower Using Drones and Computer Vision. In: 2019 2nd International Conference on Intelligent Autonomous Systems (ICoIAS). pp. 187–192 (2019).
3. Menendez, O.A. et al.: Vision based inspection of transmission lines using unmanned aerial vehicles. In: 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI). pp. 412–417 (2016).
4. Zhou, G. et al.: Robust real-time UAV based power line detection and tracking. In: 2016 IEEE International Conference on Image Processing (ICIP). pp. 744–748 (2016).
5. Lozano Guambaña, L.G. et al.: Aerial Power Lines Measurement Using Computer Vision Through an Unmanned Aerial Vehicle. In: Botto-Tobar, M. et al. (eds.) Applied Technologies. pp. 517–530 Springer International Publishing, Cham (2020).
6. D. Jones, "Power line inspection - a UAV concept," 2005 The IEE Forum on Autonomous Systems (Ref. No. 2005/11271), London, UK, 2005, pp. 8 pp.-, doi: 10.1049/ic:20050472.
7. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. 22, 11, 1330–1334 (2000). https://doi.org/10.1109/34.888718.
8. Yan, G. et al.: Automatic Extraction of Power Lines From Aerial Images. IEEE Geosci. Remote Sens. Lett. 4, 3, 387–391 (2007). https://doi.org/10.1109/LGRS.2007.895714.
9. Zhang, J. et al.: High Speed Automatic Power Line Detection and Tracking for a UAV-Based Inspection. In: 2012 International Conference on Industrial Control and Electronics Engineering. pp. 266–269 (2012).
10. Scharstein, D. et al.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001). pp. 131–140 (2001).
11. Gonzalez-Huitron, V. et al.: Parallel framework for dense disparity map estimation using Hamming distance. Signal, Image Video Process. 12, 2, 231–238 (2018). https://doi.org/10.1007/s11760-017-1150-3.
12. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, USA (2000).

24

Arquitectura de un Lago de Datos para Sistemas de Medición Inteligente

JUAN CARLOS OLIVARES ROJAS, LUIS ENRIQUE RUIZ MARTINEZ, ENRIQUE REYES ARCHUNDIA, JOSÉ ANTONIO GUTIÉRREZ GNECCHI Y ISMAEL MOLINA MORENO

Arquitectura de un Lago de Datos para Sistemas de Medición Inteligente

Juan C. Olivares-Rojas[0000-0001-5302-1786], Luis E. Ruíz-Martínez[0000-0003-3093-0156], Enrique Reyes-Archundia[0000-0003-3374-0059], José A. Gutiérrez-Gnecchi[0000-0001-7898-604X] and Ismael Molina-Moreno[0000-0002-8310-6040]

Tecnológico Nacional de México / Instituto Tecnoógico de Morelia, Morelia, Michocán, 58120, México
{juan.or, m11121552, enrique.ra, jose.gg3, ismael.mm}@morelia.tecnm.mx

Abstract. Con la inminente llegada de la cuarta revolución industrial, la red eléctrica se está moviendo hacia una red eléctrica inteligente, lo que le permite un sin fin de aplicaciones nuevas como lo son los sistemas de medición inteligente. Estas nuevas aplicaciones generan una cantidad inmensa de datos de diversas fuentes y para brindar la "inteligencia" necesaria se necesita del análisis y procesamiento de los datos. Para ello, es necesario de contar con repositorios de datos adecuados que permitan el almacenamiento, procesamiento y aplicación de los datos. Este trabajo presenta una arquitectura escalable de lagos de datos para redes eléctricas inteligentes utilizando un enfoque de cómputo borde-niebla-nube. Los resultados parciales muestran que es factible su implementación y que puede ser trasladada a otros contextos de sistemas ciber físicos o de IoT.

Keywords: Data Lake, Edge-Fog-Cloud, IoT, Smart Meter, Smart Grid.

1 Introducción

Uno de los principales motores en nuestra sociedad moderna es la red eléctrica dado que hace posible un sinfín de actividades humanas. Recientemente las tecnologías de información y comunicaciones han permeado en diversas áreas creando sistemas ciber físicos. Una de las principales aplicaciones de los sistemas ciber físicos es la Red Eléctrica Inteligente (REI) [1]. La REI está contribuyendo a combatir los tres principales retos del sector eléctrico: descarbonización, digitalización y descentralización (conocidos como las tres Ds) [2]. La tecnología que ha ayudado a revolucionar la REI es el Medidor Eléctrico Inteligente (MEI), el cual es un sistema embebido capaz de realizar mediciones de consumo de energía y es capaz de comunicar dichos datos a través de tecnologías de telecomunicación, motivo por el cual también se le considera un dispositivo de Internet de las cosas IoT (por sus siglas en inglés) [3].

Para que los MEI puedan comunicar y procesar datos se requiere de toda una infraestructura tecnológica denominada Sistemas de Medición Inteligente, siendo la implementación mejor conocida como AMI (infraestructura de medición avanzada, por sus siglas en inglés) [4]. Los SMI permiten a los usuarios finales conocer el consumo y precio de la energía eléctrica en tiempo real a través de sitios Web o aplicaciones móviles, mientras que a las empresas eléctricas les permite a automatizar procesos como los cortes y reconexiones del suministro eléctrico sin la intervención de un operador. Recientemente se ha extendido el uso de los SMI a nuevas funcionalidades que implican el análisis de los datos en aplicaciones muy diversas como la realización de pronósticos de demanda, monitoreo de la calidad de la señal eléctrica, detección de fraude y robo de energía entre otras aplicaciones [5].

Todas estás nuevas funcionales para SMI y REI, al igual que prácticamente cualquier sistema ciber físico de IoT, implican el manejo de gran cantidad de datos, los cuales necesitan ser almacenados y procesados [6]. Recientemente han surgido distintos enfoques que permiten una mejor forma de estructurar los repositorios de datos denominado lago de datos [7], los cuales permiten estructurar datos de distintas fuentes y en distintos formatos. Por otra parte, después de la creación del cómputo en la nube, nuevos paradigmas de cómputo distribuido como el cómputo en el borde [8], cómputo en la niebla [9], o combinación entre ellas (borde-niebla) [10] han aparecido en vistas de mejorar las arquitecturas de cómputo existente.

El presente trabajo muestra el diseño de una arquitectura de un lago de datos basada en un enfoque de borde-niebla escalable a la nube para almacenar y procesar datos en la REI particularmente en SMI.

El artículo se encuentra estructurado como sigue. En la sección 2 se muestra el estado del arte y se da una revisión crítica a la literatura existente. En la sección 3 se presenta el diseño de la arquitectura propuesta. Finalmente, en la sección 4 se discuten los resultados con los que se cuentan y se dan conclusiones.

2 Estado del Arte

Los SMI además de MEI están compuestos por otros dispositivos electrónicos inteligentes IED (por sus siglas en ingles). Dichos dispositivos están divididos en cuatro grandes áreas dependiendo de la localización geográfica de las redes de datos en las que se encuentran, así como los voltajes utilizados.

La primera área corresponde a una pequeña red de área local conocida de tres formas distintas dependiendo de donde se encuentra localizada. Si se encuentra en un hogar se denomina como HAN (red de área hogareña por sus siglas en inglés), si encuentra en un edificio se le conoce como BAN (red de área de edificio), o si se encuentra en ambientes industrial se le denominada IAN (red de área industrial). Los dispositivos que así se encuentran son dispositivos eléctricos tradicionales, dispositivos eléctricos inteligentes y recientemente los recursos de energía distribuida DER (por sus siglas en inglés). Tanto los dispositivos eléctricos tradicionales como inteligentes realizan consumos de energía, mientras que los DER se encargan de producir energía a través de fuentes como la solar (paneles fotovoltaicos) y eólica

(aerogeneradores). El hecho de que los usuarios finales puedan producir energía eléctrica está revolucionando los mercados eléctricos cambiando el rol de los usuarios a prosumidores [11], usuarios que producen y consumen electricidad. En esta área los voltajes son bajos y coincide con el área de consumo de la REI. La comunicación entre los dispositivos se hace generalmente por medio de redes inalámbricas de sensores o bien por el mismo cable eléctrico a través de PLC (comunicaciones por líneas de potencia) [12], siendo el MEI el dispositivo de recolectar la información de consumo/producción de los dispositivos.

La segunda parte está conformada por una red de alcance metropolitana denominada NAN (red de área vecinal) caracterizada por la comunicación de datos de los MEI hacia dispositivos intermedios denominados Concentradores de Datos (CD). Los CD concentran datos de varios MEI y sirven para abaratar costos de transmisión. Los CD tienen una capacidad promedio de unos 2000 nodos de MEI, en algunas ocasiones son menos dependiendo de las distancias en las que se encuentran los MEI. La comunicación generalmente se hace por enlaces de Radio Frecuencia (RF). En algunas ocasiones los MEI se comunican con otros MEI en redes de malla tratando de garantizar una mejor comunicación. Está área corresponde con el sector eléctrico de distribución y maneja voltajes medios.

La tercera área se denominada red de área de campo FAN (por sus siglas en inglés) y es básicamente una red de área amplia en la cual se interconectan entre sí distintos CD. La comunicación se realiza generalmente por enlaces de fibra óptica, dado que se requiere mayor ancho de banda. Debido a que estos equipos generalmente se encuentran en el sector eléctrico de transmisión, los equipos deben ser de uso rudo para soportar altos voltajes.

Finalmente, la cuarta área está formada por una red local de la empresa eléctrica conformada por los servidores del centro de datos que se encargan de almacenar todos los datos de medición denominada sistema de gestión de datos de medición MDMS (por sus siglas en inglés), el cual se interconecta con otros sistemas de la empresa eléctrica, como ERP, CRM, entre otros. Esta parte corresponde al extremo final de los SMI y los voltajes vuelven a ser bajos.

En los últimos años diversos autores se han enfocado en resolver el almacenamiento y procesamiento de grandes cantidades de datos para REI en particular para SMI [13]. Particularmente lo trabajos se han centrado más en la analítica de datos [14] así como principalmente el manejo de marcos de trabajos y aplicaciones para servidores en la nube [15 - 18]; sin embargo, dichas herramientas no son de todo adecuadas cuando las capacidades de procesamiento y almacenamiento son más limitadas como en los MEI y CD. Por tal motivo, se han diseñado soluciones más acordes a dichas limitaciones. Por ejemplo, en [19] se muestra un marco de trabajo diseñado para leer datos de dispositivos asilados en REI para poder ser almacenados y procesados. Por otra parte, el uso de lago de datos en REI y SMI ha sido poco utilizado en la literatura [20], al igual que el uso de enfoques de borde-niebla [21].

3 Diseño de la arquitectura propuesta

Para una adecuada construcción de un sistema de lago se necesitan de tres elementos fundamentales: sistema para adquisición de datos en tiempo real (ingesta), sistema de almacenamiento y consulta de datos; así como un sistema de visualización de datos (tablero de control) [22].

Tradicionalmente los componentes de un lago de datos se han realizado de forma manual, en donde dicha integración suele ser en algunas ocasiones complejas y consumidora de recursos como el tiempo y el dinero. Recientemente han surgido combinaciones de software denominadas pilas que nos permiten tener sistemas de adquisición, almacenamiento y visualización de datos de forma ágil y sencilla, las dos plataformas más importantes son ELK+ [23] y TIG [24].

ELK+ está formado por el software Elastiksearch, Logstack, Kibana y recientemente se ha integrado Beats. En el caso de la pila TIG está conformada por Telegrax, Influxdb y Grafana. Ambas pilas son gratuitas y pueden utilizarse para conformar un lago de datos. Se selecciono ELK+ debido a que cuenta con una mayor madurez derivada de tener más tiempo en el mercado, una comunidad de desarrolladores de software libre más amplia, así como una mejor implementación para el hardware utilizado.

El diseño de la arquitectura propuesta consiste en un lago de datos para HAN (nivel borde) y otro para NAN (nivel de niebla), los cuales se pueden visualizar en la Figura 1 y Figura 2 respectivamente.

El cómputo en el borde se realiza en los dispositivos en los extremos, en nuestro caso los medidores Inteligentes y en otros IDE. El cómputo en la niebla se realiza en dispositivos intermedios de mayor capacidad que los dispositivos del borde pero menor y más cerca que los dispositivos en la nube (servidores y clusters de computadoras).

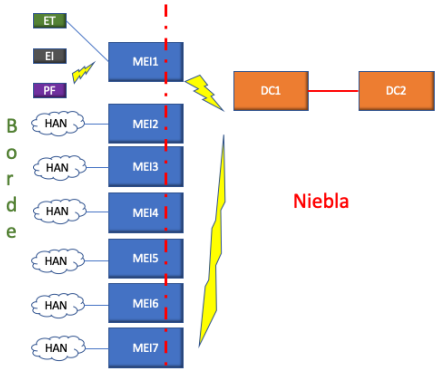


Figura 1: Arquitectura general propuesta.

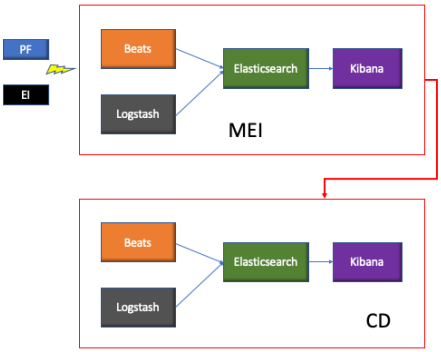


Figura 2: Arquitectura ELK en el nivel de borde y niebla.

En la figura 1, se puede visualizar la arquitectura general del lago de datos, en donde el hardware utilizado corresponde a siete MEI basados en una placa Raspberry Pi Model 3B+ con un módulo hardware de monitoreo de energía denominado SmarPi. Los CD están formados por una placa Lattepanda Alpha 864s con 8MB en RAM y 64 GB de almacenamiento. En MEI se conectaron Electrodomésticos Tradicionales (ET) como una plancha, licuadora y horno de microondas, así como Electrodomésticos Inteligentes (EI) como un aire acondicionado y un Panel Fotovoltaico (PF) con un inversor de la marca SolarEdge. En los demás MEI de momento solo se colocaron ET para formar una HAN. Los enlaces inalámbricos se dan entre los EI y los MEI, así como de éstos hacia los CD. El enlace de CD a CD es a través de un enlace Gigabit Ethernet. Los ET pueden tener comunicación a través de un plug de PLC.

En la figura 2 se aprecia la arquitectura de la pila ELK+ para la conformación del lago de datos tanto a nivel de borde (MEI) como de niebla (CD). En el nivel de borde se muestra que el PF se conecta con Beats a través del protocolo MQTT (Message Queuing Telemetry Transport, Transporte de telemetría de mensajes en cola); mientras que el aire acondicionado se conecta a través de escribir sus resultados en un archivo en formato CSV o JSON que es leído por Logstash. El MEI puede conectarse al CD usando el protocolo MQTT vía NodeRed o bien puede utilizarse un servicio Web para generar un archivo de resultados en formato CSV o JSON que puede ser leído por Logstash. Finalmente, la conexión entre CD se hace de la misma forma, gracias que tanto el MEI como CD tienen una versión del sistema operativo Linux embebida y a que la pila ELK permite dicha modularidad. A su vez, la arquitectura propuesta puede ser implementada en el nivel de nube debido a la portabilidad de la pila ELK para servidores de centros de datos.

4 Resultados preliminares, conclusiones y trabajo futuro

Cómo puede observarse en la sección anterior, se cuenta ya con la implementación de un lago de datos para sistemas de medición inteligente que trabaja de forma adecuada. Sin embargo, es necesario del desarrollo de una aplicación que permita el análisis y procesamiento de los datos para su posterior visualización.

Como trabajo futuro se pretende el desarrollo de un sistema de monitoreo que permita mostrar en tiempo real variables de la señal eléctrica, precios de consumo, producción, así como visualizaciones gráficas de predicciones de la demanda, detección de robo y fraude de energía, entre otras.

Con el presente trabajo se demuestra que es factible contar con un lago de datos que se puede implementar de manera sencilla para dispositivos IoT con capacidades de procesamiento y almacenamiento más limitadas. Particularmente este lago de datos está enfocado en aplicaciones de sistemas de medición inteligente.

En el futuro inmediato con la creciente proliferación de dispositivos IoT, el contar con lagos de datos se hará esencial para aplicaciones de analítica de datos y aprendizaje automático. Por tal motivo, es sumamente importante contar con herramientas y metodologías que sean fácilmente implementables en este tipo de entornos.

Agradecimientos

Este trabajo está parcialmente apoyado por el Tecnológico Nacional de México bajo los proyectos 7948.20-P y 8000.20-P.

Referencias

1. Yu, X., and Xue, Y. (2016). Smart Grids: A Cyber–Physical Systems Perspective," in Proceedings of the IEEE, vol. 104, no. 5, pp. 1058-1070, May 2016, doi: 10.1109/JPROC.2015.2503119.
2. Di Silvestre, M., et al. (2018). How Decarbonization, Digitalization and Decentralization are changing key power infrastructures, Renewable and Sustainable Energy Reviews 93, pp. 483-498, doi: https://doi.org/10.1016/j.rser.2018.05.068.
3. Kaur, M., et al. (2018). Implementation of Smart Metering based on Internet of Things. IOP Conference Series: Materials Science and Engineering 331, 3rd International Conference on Communication Systems (ICCS-2017), Rajasthan, India, doi: https://iopscience.iop.org/article/10.1088/1757-899X/331/1/012015.
4. Le, T., et al. (2016). Advanced Metering Infrastructure Based on Smart Meters in Smart Grid, Smart Metering Technology and Services - Inspirations for Energy Utilities, Moustafa Eissa, IntechOpen, doi: http://doi.org/10.5772/63631.
5. Dileep, G. (2020). A survey on smart grid technologies and applications, Renewable Energy 146, pp. 2589-2625, doi: https://doi.org/10.1016/j.renene.2019.08.092.
6. Wang, Y., et al. (2019). Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges, IEEE Transactions on Smart Grid 10: 3, pp. 3125-3148, doi: http://doi.org/10.1109/TSG.2018.2818167.

7. Llave, M. (2018). Data lakes in business intelligence: reporting from the trenches, Procedia Computer Science 138, pp. 516-524, doi: https://doi.org/10.1016/j.procs.2018.10.071.
8. Samie, F., et al. (2019). Edge Computing for Smart Grid: An Overview on Architectures and Solutions. In: Siozios K., Anagnostos D., Soudris D., Kosmatopoulos E. (eds) IoT for Smart Grids. Power Systems. Springer, doi: https://doi.org/10.1007/978-3-030-03640-9_2.
9. Okay, F., and Ozdemir, S., et al. (2016). A fog computing based smart grid model, 2016 International Symposium on Networks, Computers and Communications (ISNCC), Yasmine Hammamet, pp. 1-6, doi: http://doi.org/10.1109/ISNCC.2016.7746062.
10. Gilbert, G., et al. (2019). A Critical Review of Edge and Fog Computing for Smart Grid Applications. 15th International Conference on Social Implications of Computers in Developing Countries (ICT4D), Dar es Salaam, Tanzania. pp.763-775, doi: http://doi.org/10.1007/978-3-030-18400-1_62.
11. Espe, E., et al. (2018). Prosumer Communities and Relationships in Smart Grids: A Literature Review, Evolution and Future Directions. Energies 2018:11, doi: https://doi.org/10.3390/en11102528.
12. Artale, G., et al. (2018). A new PLC-based smart metering architecture for medium/low voltage grids: Feasibility and experimental characterization, Measurement 129, pp. 479-488, doi: https://doi.org/10.1016/j.measurement.2018.07.070.
13. Zhang, Y., et al (2018). Big data analytics in smart grids: a review. Energy Informatics 1:8, doi: https://doi.org/10.1186/s42162-018-0007-5.
14. Rossi, B. and Chren, S. (2020). Smart Grids Data Analysis: A Systematic Mapping Study, IEEE Transactions on Industrial Informatics 16: 6, pp. 3619-3639, June, doi: http://doi.org/10.1109/TII.2019.2954098.
15. Munshi, A., and Mohamed, Y. (2018). Data Lake Lambda Architecture for Smart Grids Big Data Analytics, IEEE Access 6, pp. 40463-40471, doi: http://doi.org/10.1109/ACCESS.2018.2858256.
16. Ansari, M., et al. (2019). Evaluation of big data frameworks for analysis of smart grids. Journal of Big Data 6:109, doi: https://doi.org/10.1186/s40537-019-0270-8.
17. Wilcox, T., et al. (2019). A Big Data platform for smart meter data analytics, Computers in Industry 105, pp. 250-259, doi: https://doi.org/10.1016/j.compind.2018.12.010.
18. Sandrini, S. (2019). Smart meters, data lakes, and serverless for Utilities. Amazon Web Services. https://d1.awsstatic.com/events/reinvent/2019/Smart_meters,_data_lakes,_and_serverless_for_Utilities_PNU301.pdf
19. Sha, K., et al. (2017). A Secure and Efficient Framework to Read Isolated Smart Grid Devices, IEEE Transactions on Smart Grid 8:6, pp. 2519-2531, doi: http://doi.org/10.1109/TSG.2016.2526045.
20. Chen, T., et al. (2019). On Construction of a Power Data Lake Platform Using Spark. In: Hung J., Yen N., Hui L. (eds) Frontier Computing. FC 2018. Lecture Notes in Electrical Engineering vol 542, doi: https://doi.org/10.1007/978-981-13-3648-5_11.
21. Sirojan, T., et al. (2019). Embedded Edge Computing for Real-time Smart Meter Data Analytics, 2019 International Conference on Smart Energy Systems and Technologies (SEST), Porto, Portugal, pp. 1-5, doi: http://doi.org/10.1109/SEST.2019.8849012.
22. Rooney, S., et al (2019). Experiences with Managing Data Ingestion into a Corporate Datalake, 2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC), Los Angeles, CA, USA, 2019, pp. 101-109, doi: http://doi.org/10.1109/CIC48465.2019.00021.

23. Tsung, C., et al. (2020). Visualizing Potential Transportation Demand from ETC Log Analysis Using ELK Stack, IEEE Internet of Things Journal, doi: http://doi.org/10.1109/JIOT.2020.2974671.
24. Kychkin, A., et al. (2019). Architecture of Compressor Equipment Monitoring and Control Cyber-Physical System Based on Influxdata Platform, 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russia, pp. 1-5, doi: http://doi.org/10.1109/ICIEAM.2019.8742963.

25

Diseño e implementación de hardware propio de un interruptor de energía eléctrica con comunicación WiFi para aplicaciones en Industria 4.0

ROBERTO MORALES, Y. GUADALUPE LAZARO, ADRIAN REYES, J. FEDERICO CASCO Y H. PATRICIA MARTÍNEZ

Diseño e implementación de hardware propio de un interruptor de energía eléctrica con comunicación WiFi para aplicaciones en Industria 4.0

Roberto Morales[1], Y. Guadalupe Lázaro[1], A. Salvador Reyes[1], J. Federico Casco[1] and H. Patricia Martinez[1]

[1] Tecnológico Nacional de México/Instituto Tecnológico de Apizaco
División de Estudios de Posgrado e Investigación
Av. Instituto Tecnológico s/n. Col. Centro. C.P. 90300. Apizaco, Tlaxcala, México
roberto.mc@apizaco.tecnm.mx

Abstract. Se presenta la metodología de diseño y desarrollo de un interruptor eléctrico basado en un relevador con comunicación WiFi para aplicaciones de la Industria 4.0. Se explica la metodología de desarrollo y el proceso de diseño de cada módulo del dispositivo. El relevador y el módulo WiFi ESP12E que se encuentran embebido en la misma tarjeta de circuito impreso se manipulan por medio de una red local desde una página web que envía comandos MQTT. El interruptor de energía eléctrica diseñado tiene la capacidad de encender o apagar un rango de cargas de hasta 20 A. Por lo tanto, se puede utilizar para fines industriales, así como para el hogar o la oficina. Además, se muestra el diseño de la fuente para suministrar energía a todo el dispositivo. Resultados experimentales del funcionamiento del dispositivo muestran excelente desempeño, confiabilidad y fiabilidad del interruptor desarrollado para aplicaciones en la Industrial 4.0.

Keywords: Hardware design, industry 4.0, MQTT, WiFi technology.

1 Introducción

Dentro de las tecnologías que sustentan la Industria 4.0 (I.4.0), se encuentra: la fabricación aditiva, los sistemas de integración horizontal y vertical, la ciberseguridad, la realidad aumentada, el cómputo en la nube, los robots colaborativos, el big data y la analítica avanzada, los dispositivos móviles, las plataformas y aplicaciones tecnológicas, así como, la inteligencia artificial (IA) y las tecnologías de la información y comunicación (TICs) [1].

A diferencia de los países desarrollados, la I.4.0 en los países de América Latina y el Caribe (ALC) aún se encuentra en una fase temprana [2]. Ya que la región aún no ha alcanzado las capacidades necesarias en las 5 tecnologías básicas de la I.4.0, las cuales son: conectividad, capacidad de almacenamiento, computación en la nube, big data e internet de las cosas. La principal causa por la cual no se desarrollan estas capacidades es económica, además de que el costo por servicios de conectividad es más alto en los países de ALC respecto de los países desarrollados [3].

Si bien en la región, se ha comenzado a desarrollar y a utilizar tecnología digital para el sector industrial [4], [5], para las ciudades inteligentes [6] y el hogar [7], [8], existe aún el reto de continuar desarrollando tecnología digital propia sobre todo en lo que se refiere al diseño y desarrollo de hardware para evitar componentes de código "abierto" [9]. Es en este punto dónde incide el presente trabajo, ya que en este caso la estrategia para asimilar el *"know-how"* tecnológico se hace por medio del diseño, desarrollo e implementación de un interruptor de relevador inteligente (IRI) de capacidad industrial.

2 Metodología de Diseño del Hardware

2.1 Modelo V

La metodología que se sigue para implementar el IRI es el "Modelo V" para el desarrollo de sistemas embebidos. La Fig. 1 muestra el diagrama a bloques general del Modelo V y cada uno de sus etapas se explica a continuación:

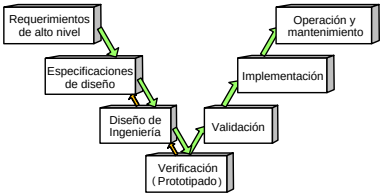


Fig. 1. Modelo V para el desarrollo de sistemas embebidos. Fuente: Elaboración Propia.

Requerimientos de alto nivel y especificaciones: Se plantean los requerimientos, arquitectura, viabilidad y funcionalidades principales del sistema, para posteriormente establecer las especificaciones de diseño de hardware y software.
Diseño de ingeniería: El proceso de diseño comienza con el fin de obtener las funcionalidades requeridas, se crea un concepto y arquitectura de alto nivel, así como la selección de los componentes. El resultado de la etapa de diseño de ingeniería se presenta a la etapa de *verificación*. En caso de reconocer algún elemento abierto, que impide la conclusión de la etapa presente, se regresa a la etapa anterior.
Verificación: El diseño de ingeniería, es sometido a diversas pruebas de operación y funcionalidad con ayuda de un prototipo para identificar errores ocultos, además en esta etapa se determinan posibles nuevas funciones y/o operaciones no previstas en las dos etapas anteriores de desarrollo.
Validación: En esta etapa se detalla el concepto de forma, operación y funcionalidad; si pasa con éxito la etapa se dice que la versión se libera por el área de diseño.
Implementación: Esta etapa se encarga de la planeación de la fabricación y de una posible producción en masa, tomando en cuenta la disponibilidad de materiales, tiempos de ensamble etc.
Operación y mantenimiento: Se genera un manual de operación y mantenimiento.

2.2 Requerimientos

Se ha propuesto que el dispositivo cumpla con los siguientes requerimientos:

- Comunicación WiFi: integrada al dispositivo, el cual debe ser capaz de recibir las órdenes desde una página web para manipular un aparato eléctrico.
- El usuario debe poder saber si el dispositivo esta encendido, apagado, o si no hay conexión a la red WiFi.
- Debe poder accionar una carga de hasta 20 A.
- Debe tener su propia fuente de energía para el control.

Durante el diseño y desarrollo del IRI se ha considerado tomar en cuenta las normas internacionales IPC-2221, IPC2251 e IPC 2252.

El dispositivo operará en dos modos, el modo de configuración y el modo de control. El modo configuración es utilizado cuando el IRI se energiza por primera vez o cuando el usuario desea cambiar la red local a la cual se conecta. Las responsabilidades del modo configuración se enlistan a continuación:

- Generar una red de acceso local para permitir la conexión a un dispositivo móvil con WiFi para su configuración.
- Generar una página web de configuración para el IRI.
- Recibir los datos de "nombre de red" y "contraseña de red" provenientes del dispositivo de configuración externa y guardarlos en una localidad de memoria.

En caso de no poder establecerse una comunicación con la red local, el dispositivo generará un protocolo de envío fallido. Por otro lado, el modo de control debe:

- Accionar o no la bobina del relé de potencia.
- En caso de que el dispositivo determine conexión de red inválida, el dispositivo entrará en modo de configuración y activará un indicador luminoso (led)

2.3 Especificaciones

La siguiente Tabla1 muestra las especificaciones generales del IRI.

Tabla 1. Especificaciones generales del dispositivo a desarrollar. Fuente: Elaboración Propia.

Parámetro	Símbolo	Unidad	Mínimo	Nominal	Máximo	Descripción
Temp. de Operación	Ta	°C	-10		50	Temperatura ambiente
Humedad Relativa	Hum	%	10		90	Humedad ambiente
Dimensiones	Dim	mm		95x70 x31		Largo, ancho y altura
V para el control	V_f	volt	3.3	3.3	5	V para control
Voltaje de carga	V_L	volt		120		V de red
Corriente de carga	i_L	A	0		20	Corriente alterna
Consumo	Con	A	0.5	0.7	1.0	Consumo de corriente
Tipo de comunicación	Com	GHz		2.4		WiFi. IEEE 802.11 b/g/n
Alcance de señal WiFi	Alc	m	10		18	Depende de la línea de vista
Lugar de uso						En interiores

3 Diseño del Hardware

3.1 Diagrama esquemático

Transceptor: Dentro de las comunicaciones inalámbricas, WiFi, y Bluetooth son los dispositivos más comunes que se pueden encontrar en los teléfonos inteligentes, pero existen otros dispositivos de comunicación que usan otro protocolos, tales como: LoRa, SigFox, Xbee, etc. Existen numerosos trabajos en donde se analiza cada uno de ellos [10]. En este caso, después de tomar en cuenta las consideraciones técnicas y las funcionalidades deseadas del IRI a desarrollar, se ha seleccionado el módulo WiFi ESP-12E. La Fig. 2 muestra el diagrama esquemático del módulo [11].

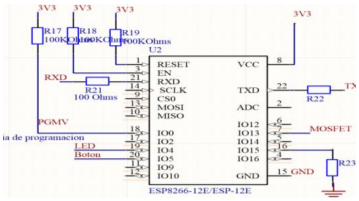


Fig. 2. Diagrama esquemático del modulo WiFi ESP-12E. Fuente: [11].

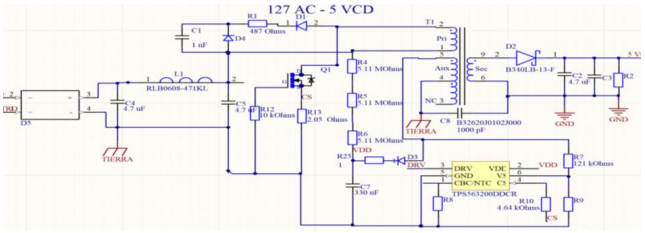


Fig. 3. Diagrama esquemático de la fuente de alimentación. Fuente: Elaboración Propia

Fuentes de alimentación: Se considera tomar energía de la misma red eléctrica que alimenta la carga para alimentar los dispositivos de control y comunicación. Primero se diseña la fuente de 120 Vca a 5 Vcd y a continuación la de 3.3 Vcd. La fuente de cd se ha diseñado para proveer 1A de salida, ya que todo el circuito de control consume 700mA. Se ha seleccionado un trasformador flyback TI 750312723 100/220 Vca a 5 Vca con un periodo estimado de frecuencia de resonancia en el nodo de conmutación de 2µs, lo que da un ciclo de trabajo máximo de 0.47. Con esta fuente se tiene un voltaje de salida regulado de 4.83 Vcd. Para obtener el voltaje de 3.3 Vcd para alimentar el ESP se seleccionó un regulador UCC28700DBV de baja caída, su diagrama de conexión se encuentra en su hoja de datos [12]. En la Fig. 3 se muestra el diagrama esquemático de la fuente diseñada.

3.2 Diseño de la PCB

Una vez realizados los diagramas esquemáticos, se crea un documento de PCB con el fin de generar las pistas de todas las capas físicas del proyecto. Se recomienda utilizar una herramienta CAD [*computer-aided design*]. También es importante seguir una metodología de diseño de la PCB [13]. Con ayuda de la herramienta CAD se pueden verificar las posiciones 2D (ver Fig. 4) y 3D (ver Fig. 5) de los dispositivos y el trazado de las rutas del interruptor desarrollado.

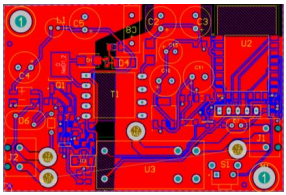


Fig. 4. Vista en 2D del diseño de la PCB.
Fuente: Elaboración Propia.

Fig. 5. Vista en 3D del diseño de la PCB.
Fuente: Elaboración Propia.

4 Implementación del Software

4.1 El firmware

La programación del firmware se hace con Arduino IDE a través de un puerto USB. Las ventajas de Arduino IDE son las bibliotecas proporcionadas para el ESP [14].

4.2 MQTT

La MQTT es un protocolo de mensajería instantánea abierto, simple y ligero. Está enfocado al envío de datos en plataformas de hardware limitado. La arquitectura de MQTT se divide en dos componentes principales: *cliente* y *broker* [15].
El cliente MQTT: Un cliente MQTT es cualquier dispositivo que ejecuta una biblioteca MQTT y se conecta a un agente MQTT a través de una red. Básicamente, cualquier dispositivo que hable MQTT a través de una agrupación TCP/IP (*Transmission Control Protocol/Internet Protocol*) puede llamarse cliente MQTT. Tanto un editor como un suscriptor son clientes MQTT.

4.3 Broker MQTT

En este proyecto se utiliza el *broker* MQTT de HiveMQ ya que facilita el traslado de datos. HiveMQ es una empresa que mueve datos desde y hacia dispositivos conectados a través de redes y plataformas en la nube. Conecta cualquier dispositivo y sistema backend de una manera fiable y segura [16]. La conexión MQTT es siempre entre un cliente y el *broker*. Los clientes nunca se conectan entre sí directamente. Para iniciar una conexión, el cliente envía un mensaje CONNECT al *broker*. El *broker* responde con un men-

saje CONNACK y un código de estado (ver Fig. 6). Una vez que se establece la conexión, el *broker* la mantiene abierta hasta que el cliente envía un comando de desconexión.

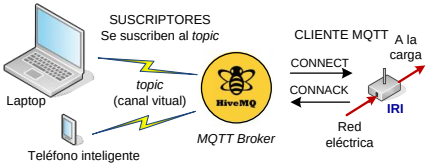


Fig. 6. Interacción cliente/suscriptor con el *broker* MQTT. Fuente: Elaboración Propia.

4.4 Seguridad

MQTT confía en tecnologías estándares para la seguridad, tales como:

- Autenticación Usuario/Password: Cada cliente MQTT tiene un identificador de cliente único. El cliente proporciona esta ID única al intermediario en el mensaje MQTT CONNECT. [17].
- Protocolos criptográficos SSL/TLS (*Secure Socket layer*/ *Transport Layer Security*): Si se desea cifrar la comunicación con el *broker*, debe emplear SSL/TLS en lugar de TCP/IP. Es responsabilidad del usuario abordar los problemas de seguridad de MQTT, y se utilizar certificados SSL/TLS con administración de claves de sesión [18].
- Tecnología *Push*: Es una forma de comunicación a través de Internet en la que la petición de envío tiene origen en el servidor, por oposición a la tecnología *pull*, en la que la petición tiene origen en el cliente. HiveMQ tiene esta opción de seguridad.

5 Resultados

La Fig. 7 muestra la PCB del interruptor antes de ser armada con los componentes electrónicos, la Fig. 8 muestra el dispositivo ya armado.
Se crean dos *topics* definidos en el *broker*. Se selecciona uno de los *topics* y se envía un mensaje, este mensaje se procesa en el ESP y dependiendo del mensaje se procesa un cambio de estado en el GPIO13 (salida a relé) o en el GPIO4 (salida a LED). La primera vez que interactúa el *broker* con el dispositivo se construye un mensaje con el ID del

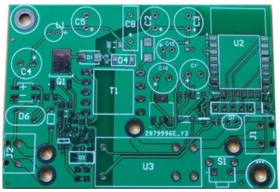


Fig. 7. Vista de la PCB fabricadar
Fuente: Elaboración Propia.

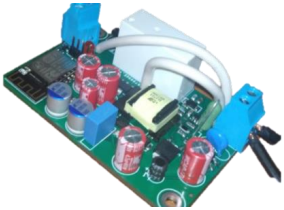


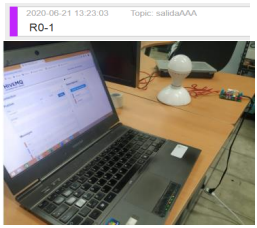
Fig. 8. Vista del interruptor implementado.
Fuente: Elaboración Propia.

dispositivo y se envía un carácter R para manipular el relé, o bien, un carácter L para manipular el LED indicador. En la Figura 9a se observa que el carácter R-1 se ha enviado del *broker* al IRI. Y la Figura 9b muestra como la lámpara "prenden". Después, al enviar el carácter R0-1 al interruptor desarrollado (ver Figura 10a), se "apaga" la lámpara que se encuentre conectada al dispositivo como se observa en la Figuras 10b.

a)

b)

Fig. 9. Operación. a) Mensaje enviado, b) "enciende" la carga.
Fuente: Elaboración Propia.



a)

b)

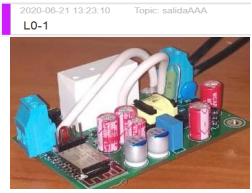
Fig. 10. Operación. a) Mensaje enviado, b) "apaga" la carga.
Fuente: Elaboración Propia.

Finalmente, las Figuras 11a y 11b muestran como al enviar una carácter L-1 el LED se servicio prende simulando una condición de falla de conexión o de software. Y si se envía el carácter L0-1, el interruptor desactiva el LED y este se "apaga", como se observa en las Figura 12a y 12b. Cabe mencionar que esta acción se hace de forma automática.

a)

b)

Fig. 11. Operación. a) Mensaje enviado, b) "enciende" el led.
Fuente: Elaboración Propia.



a)

b)

Fig. 12. Operación. a) Mensaje enviado, b) "apaga" el led.
Fuente: Elaboración Propia.

6 Conclusiones

Se mostró una metodología de desarrollo para un sistema embebido con aplicación a la industria 4.0. En este caso, por medio del desarrollo de un interruptor inalámbrico desde sus requerimientos y concepto hasta su puesta en operación. Se observa, que la etapa de diseño e implementación del hardware es clave y puede ser difícil de completar si no se sigue una adecuada metodología de diseño. También se debe mencionar que es necesario usar las normas nacionales e internacionales si se desea certificar el producto o de otro modo no podrá ser comercializado. Por otro lado, el uso de un broker simplifica

la manipulación inalámbrica del IRI desde un dispositivo móvil o laptop. El interruptor aquí desarrollado presenta dos diferenciadores respecto a otros interruptores inalámbricos existentes en el mercado: una capacidad mayor de carga (hasta 20A) y se puede implementar un protocolo de comunicación MQTT con alternativas de seguridad para lograr una manipulación segura.

Referencias

1. C.B Ynzunza-Costes, *et al*., "El entorno de la industria 4.0: Implicaciones y perspectivas futuras" *ConCiencia Tecnológica*. No. 54. pp. 33-45. Available: https://dialnet.unirioja.es/servlet/articulo?codigo =6405835 (2017).
2. G. Cimoli, M. Castillo, M. Porcile and G. Stumpo, "Políticas industriales y tecnológicas en América Latina", CEPAL (2017).
3. K. A. Sibrián-Sánchez and K. V. Amaya-Rodríguez, "Desafíos de la industria 4.0 y oportunidades de desarrollo sostenible para América Latina y el Caribe," in *XII Congreso de Economistas de ALyC*, pp. 1–30 (2019).
4. M. Casalet, "La digitalización industrial un camino hacia la gobernanza colaborativa" CEPAL. Available: https://repositorio.cepal.org/bitstream/handle/11362/44266/1/S1800941_es.pdf (2018)
5. R. Pita-Rolle, V. de Oliveira-Martucci and E. Paciencia Godoy, "Architecture for digital twin implementation focusing on industry 4.0," *IEEE Latin America Trans.*, vol. 18, no. 5, pp. 889-898 (2020).
6. R. ARenoo, et. al, "Ciudades inteligentes en Latinoamérica, el ciudadano vigilado" IV simposio Internacional LAVITS (2017).
7. J. E- Giral-Salas, R. Morales-Caporal, E. Bonilla-Huerta, J. J. Rodríguez-Rivas and J. de J. Rangel-Magdaleno, "A Smart switch to connect and disconnect electrical devices at home by using internet," *IEEE Latin America Trans.*, vol. 14, no. 5, pp. 1575-1581, (2016).
8. F. Garcia, W. Souza, and F. Marafao, "Embedded NILM as home energy management system: a heterogeneous computing approach" *IEEE Latin America Transactions.*, vol. 18, no. 2, pp. 360-367, Feb. 2020.
9. R. L. Katz, "El desafío de la cuarte revolución industrial en América Latina". Available: https://cic40.es/el-desafio-de-la-cuarta-revolucion-industrial-en-america-latina/
10. M. Pui-In, U. Seng-Pan, and M. Rui P. "Transceiver architecture selection: review, state of the art survey and case study" *IEEE Circuits and Sys. Mag.*, vol. 7, no. 2, pp. 6-25, (2007).
11. https://components101.com/sites/default/files/component_datasheet/ESP12E%20Datasheet.pdf
12. Texas Instruments, "Using the UCC28700EVM-068. User's Guide," Literature number SLUU968A. (2012). Revised (2015).
13. http://electronica.ugr.es/~amroldan/modulos/docencia/cursos/pcb_uhu_98/teoria%20de%20pcbs.html
14. C. Ubach-Granados, "Programando directamente un ESP8266". Escula Tecnica Superior d´Enginyeria Industrial de Barcelona (2017).
15. A. E. Andrada-Tivani, R. Martín-Murdocca, C. F. Sosa-Paez, and J. D. Dondo-Gazzano, "Didactic prototype for teaching the MQTT protocol based on free hardware boards and node-RED" *IEEE Latin America Transactions.*, vol. 18, no. 2, pp. 376-383, (2020).
16. https://www.hivemq.com/
17. https://aprendiendoarduino.wordpress.com/tag/hivemq/
18. M. Singh, M. A. Rajan, V. L. Shivraj, and P. Balamuralidhar, "Secure MQTT for internet of things (IoT)". In *Fifth Int. Conf. on Comm. Sys. and Network Techs.* pp. 746-751 (2015).

Parte VI

Trabajos en Desarrollo

Introducción

Esta sección es una presentación de trabajos que se encuentran en una etapa temprana de desarrollo principalmente por estudiantes de licenciatura, maestría y doctorado en ciencias de la computación, o campos relacionados.

Esta sección tiene objetivo principal permitir a los estudiantes compartir y obtener retroalimentación acerca de su trabajo por parte de colegas. Asimismo, fomentar la participación en proyectos de investigación y académicos en el área de la ciencias de la computación. Esta sección también tiene la intención de promover la colaboración y la comunidad de investigadores de la computación en México.

Para esta sección se recibieron un total de 17 trabajos resultado de las actividades académicas de 10 instituciones de educación superior. Mediante un proceso de evaluación doble ciego se seleccionaron los 7 trabajos mas avanzados para su presentación.

Héctor Efraín Ruiz y Ruiz
René Francisco Navarro Hernández

26

Generación de un grafo probabilístico en base a bitácoras de micro servicios RESTful

JAVIER IRACHETA Y SARA GARZA

Generación de un grafo probabilístico en base a bitácoras de micro servicios RESTful

Javier Iracheta y Sara E. Garza

Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, México
https://www.uanl.mx/ {javier.irachetavll,sara.garzavl}@uanl.edu.mx

Resumen La arquitectura basada en microservicios RESTFul expone múltiples ventajas respecto al modelo tradicional monolítico, entre ellas el aislamiento de los servicios, la compaginación de múltiples tecnologías, expansión de funcionalidad y mantenimiento con menor riesgo. Sin embargo, por su naturaleza, los microservicios son desacoplados y el diseño de pruebas punto a punto es complicado de definir y mantener. Teniendo esto en cuenta, la bitácora de uso registrada por entornos como Microsoft Application Insight permite obtener un gran conjunto de datos que representan el tiempo de registro y la operación, donde desde una perspectiva del usuario existe una lógica implícita que refleja los procesos de negocios implementados. El presente artículo describe un modelo de procesamiento de la firma de tiempo y operación registrada en la bitácora de microservicios con la intención de reconocer la conexión entre las operaciones REST y plasmarlas un grafo que represente la relación entre ellos. Este grafo puede ser utilizado para el diseño de pruebas más realistas y la identificación de riesgos ante la falla de un microservicio en relación con su entorno.

Palabras clave: reconocimiento de patrones · microservicios · grafo

1. Introducción

Los microservicios son un modelo arquitectónico derivado de las arquitecturas orientadas a servicios (SOA), los cuales se enfocan en servicios ligeros y autocontenidos. Esto a su vez implica una mejora a la ingeniería del software en agilidad, escalabilidad y autonomía. Los beneficios más importantes son una entrega más rápida, una escalabilidad mejorada y una mayor autonomía [10]. Sin embargo, el desacopla-miento natural de estos representa un reto para el área de calidad de software, ya que, si bien es relativamente fácil probar un microservicio tipo *Representational State Transfer* (REST), determinar la secuencia de operaciones y la relaciones entre microservicios, así como su relación a través del tiempo, requiere un esfuerzo significativo [8].

2. Trabajo relacionado

Existe una variedad de investigaciones que abordan problemáticas similares desde diferentes perspectivas o en condiciones particulares. S. Ma, C. Fan, Y. Chuang, W. Lee, S. Lee y N. Hsueh proponen la creación del grafo usando la propiedad de reflexión del código de los API REST [11], L. Berton, Valverde-Rebaza y Andrade proponen un modelo para predecir vértices usando aprendizaje supervisado y semisupervisado [7], D. Wang, L. Shi y J. Cao proponen un algoritmo rápido para la construcción aproximada del grafo k vecinos más cercanos, en el contexto de minería de datos [12], mientras que X. Xu, H. Jin, S. Wu, L. Tang y Y. Wang proponen un método basado en grafo generado por el comportamiento del usuario para web, usando bitácoras [13]. En el contexto de datos saturados, M. Khan, N. Krishnamoorthy, L. Jalali y R. Biswas proponen un grafo de identidad para Adobe basado en la identificación de millones de registros [9].

3. Objetivo de investigación

El objetivo de la investigación es desarrollar un método que, dada una bitácora de eventos, genere un grafo probabilístico de dependencias entre microservicios. Cada evento está representado por una operación de microservicios en un tiempo T_k, donde cada operación a su vez está compuesta de un método HTTP y una URL.

4. Metodología

La metodología se compone de tres fases principales: extracción de datos, representación y desarrollo.

4.1. Extracción de datos

Esta fase consta de dos capas que ejecutan múltiples procesos: el grafo de referencia y los datos de análisis. El grafo de referencia se necesita para comparar el grafo probabilístico que es extraído de los datos de análisis, los cuales se obtienen a partir de bitácoras de `Azure Application Insight`. Este grafo de referencia que está compuesto por el 23 % de los nodos potenciales del grafo probabilístico. Una simplificación del proceso para generar el grafo de referencia es la siguiente (también representada en la figura 1):

- Definir y ejecutar flujos funcionales en la aplicación (ejecución secuencial)
- Obtener las bitácoras a nivel navegador de las llamadas a los APIs por flujo
- Generar un diccionario de las llamadas a APIs registradas
- Utilizar la secuencia de llamadas para formar una matriz de adyacencia
- Generar el grafo en base a la matriz

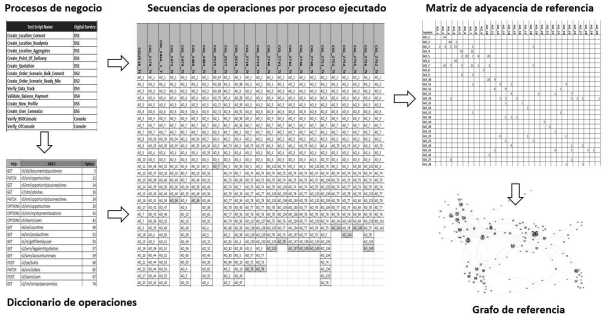


Figura 1. Generación del grafo de referencia

4.2. Representación de los datos

A partir de la bitácora de eventos se diseñó una estrategia para reducir el impacto de una de las variables que agregan complejidad: la concurrencia de usuarios. El efecto de esto es la potencial repetición y falta de una secuencia coherente de los datos. Una estrategia definida para reducir este impacto fue la creación de segmentos de operaciones en un tiempo dado fijo, es decir, escogemos un lapso de tiempo t (definido por la diferencia entre t_{final} - t_{inicial}) y creamos ventanas consecutivas de esta cantidad de tiempo, donde para cada ventana (segmento) se suma la cantidad de operaciones únicas, es decir, se genera un vector binario que representa si la operación O_i estuvo presente entre la ventana de tiempo t_{inicial} y t_{final}. Si la operación O_i se realizó durante la ventana (t_{inicial},t_{final}), entonces el valor de esta operación es 1 en el vector; de lo contrario, es 0. El siguiente segmento repite el proceso utilizando lapsos de tiempo iguales donde $t_{\text{final}(i)}$ y $t_{\text{inicial}(i+1)}$. El objetivo es agrupar las operaciones repetidas en un segmento. El proceso de extracción y representación de datos puede apreciarse en la figura 2.

4.3. Desarrollo y comparación del modelo de generación del grafo

Durante esta fase se conceptualiza el método para generar el grafo resultante esperado, mismo que es comparado contra el grafo de referencia. Durante la fase inicial de exploración y a modo de tener una línea de base se desarrolló una heurística basada en el coeficiente de correlación de Pearson, teniendo el grafo de referencia como una matriz $A_{n\times n}$ y la matriz formada por la agrupación de segmentos que denominaremos $B_{m\times n}$. Se tomó la matriz B y se obtuvo el coeficiente de correlación entre columnas donde $\text{Corr}(B_{(1:(n-1),m)}, B_{(2:n,m+1)})$; es importante remarcar que se hace un corrimiento intencional en n para mejorar

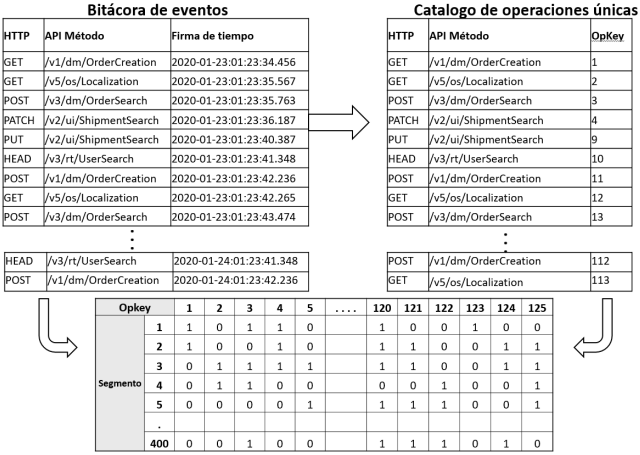


Figura 2. Generación de segmentos

el coeficiente de correlación buscado. Con los resultados de este coeficiente, se crea una matriz de correlación que denominaremos $C_{(n,n)}$. Dado que la matriz de adyacencia A es entera, se transforma en booleana A_{bol} al aplicar la operación considerando todo $A_i > 0$ como verdadero y $A_i <= 0$ como falso. Por otro lado, para la matriz de coeficientes de correlación C utilizamos un coeficiente de umbral f para determinar qué es verdadero y qué falso; de este modo, creamos la matriz C_{bol} usando un método análogo a A_{bol} donde considerando todo $C > f$ como verdadero y $C <= f$ como falso. Habiendo obtenido ambas matrices booleanas, procedimos a obtener las siguientes mediciones:

- Positivos Verdaderos (TP) = A_{bol} AND C_{bol}
- Negativos Verdaderos (TN)= No(A_{bol}) AND No(C_{bol})
- Falsos Negativos (FN) = A_{bol} AND No(C_{bol})
- Falsos Positivos (FP) = No(A_{bol}) AND C_{bol}

En base a estas medición se obtuvo la métrica de F_1:

$$\text{Presición} \qquad \frac{\text{TP}}{\text{TP+FP}}$$

$$\text{Exhaustividad} \qquad \frac{\text{TP}}{\text{TP+FN}}$$

$$F1_{score} \qquad 2 \times \frac{\text{Precisión} \times \text{Exhaustividad}}{\text{Precisión+Exhaustividad}}$$

Una ejemplificación de cómo fue implementada la operación de crear la matriz C se encuentra la lista 1.1 de código.

```
Matriz_Correlacion=Empty Matrix(n,n)
for x in range(n):
  for y in range(n):
    Matriz_Correlacion[x,y]=corrcoef(vector[:-1,x],vector[1:,
    y])
```

Listing 1.1. ciclos de correlación

5. Logros

Se ejecutaron varios ciclos de evaluación usando el método descrito en la sección anterior y se aplicaron algunas variantes tales como niveles de correlación (0.1 y .99, ya que valores menores a 0 indicarían una correlación negativa). El método de evaluación probó ser efectivo y repetible para las distintas variantes de experimentación utilizadas. La implementación se hizo inicialmente en `Octave` [1](variante de código abierto que soporta codificación en lenguaje de `MATLAB` [3]); sin embargo, posteriormente se migró e implementó el código en `Python` [6], debido a la flexibilidad de librerías como `numpy` [4] y `pandas` [5] para el cálculo numérico y la manipulación vectorial. Una ventaja adicional es la posibilidad de eventualmente acelerar la ejecución vía hardware usando la tecnología de aceleración NVIDIA Cuda [2].

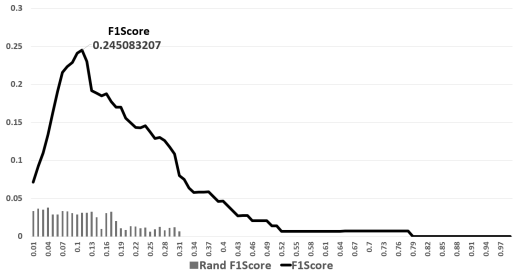


Figura 3. F_1 del método propuesto vs. F_1 con generación de grafos aleatorios

6. Resultados esperados

Dada la naturaleza de la métrica elegida como referencia (F_1), tenemos 0 como valor mínimo posible y 1 como máximo posible. Los experimentos ejecutados al momento de este reporte arrojan un F_1 máximo de 0.24. Como referencia se ejecutó una línea de base con más de 1000 grafos aleatorios con características similares al grafo de referencia donde el valor más alto obtenido fue de 0.03 (muy poca coincidencia). Si bien el valor obtenido con el método descrito en este trabajo es 800 % mayor que el mejor valor aleatorio, está muy lejos del resultado esperado que debería ser cercano al 0.7 (ver figura 3). Esto es particularmente notorio porque para fines esperados de la investigación, los cuales buscan un método para inferir la relación entre microservicios en una aplicación dada, un F_1 menor de 0.7 no sería relevante.

7. Conclusiones

Si bien aún existe espacio para mejorar los resultados en base a algoritmos dado los múltiples valores y combinaciones de las variables identificadas (tales como agrupación de segmentos, procedimientos de vectorización y otras variables más por identificar e implementar), existe un riesgo alto de que la solución que eventualmente se encuentre no se pueda generalizar en problemas similares y nos enfrentemos a un sobreajuste de la solución (*overfitting*). Aún en este panorama es posible evitar este riesgo usando la segmentación de muestras de datos bajo el modelo usado en aprendizaje supervisado (entrenamiento, pruebas y validación) e incluso evaluar modelos de aprendizaje no supervisado. Sin embargo, una evaluación bibliográfica reciente reveló la existencia de una disciplina llamada minería de procesos, cuya área de estudio comprende la inferencia y validación de procesos en base a datos generados por los mismos. Revisiones tempranas han mostrado que podría ser redituable a esta investigación ahondar en algunos algoritmos particulares de esta disciplina.

Referencias

1. Gnu octave project. https://www.gnu.org/software/octave/, last accessed 2020/07/26.
2. High-performance python with cuda acceleration. https://developer.nvidia.com/, last accessed 2020/07/26.
3. Matlab. https://www.mathworks.com/products/matlab.html, last accessed 2020/07/26
4. Numpy open source project. https://numpy.org/, last accessed 2020/07/26.
5. Pandas open source project. https://pandas.pydata.org/, last accessed 2020/07/26.
6. Python official site. https://www.python.org/, last accessed 2020/07/26.
7. Berton, L., Valverde-Rebaza, J., de Andrade Lopes, A.: Link prediction in graph construction for supervised and semi-supervised learning. In: 2015 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (2015)

8. Cerny, T.: Aspect-oriented challenges in system integration with microservices, soa and iot. Enterprise Information Systems **13**(4), 467–489 (2019)
9. Khan, M., Krishnamoorthy, N., Jalali, L., Biswas, R.: Adobe identity graph. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 5354–5356 (2018)
10. Lewis, J., Fowler, M.: Microservices (Mar 2014), martinfowler.com/articles/microservices.html
11. Ma, S., Fan, C., Chuang, Y., Lee, W., Lee, S., Hsueh, N.: Using service dependency graph to analyze and test microservices. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). vol. 02. pp. 81–86 (2018)
12. Wang, D., Shi, L., Cao, J.: Fast algorithm for approximate k-nearest neighbor graph construction. In: 2013 IEEE 13th International Conference on Data Mining Workshops. pp. 349–356 (2013)
13. Xu, X., Jin, H., Wu, S., Tang, L., Wang, Y.: Urmg: Enhanced cbmg-based method for automatically testing web applications in the cloud. Tsinghua Science and Technology **19**(1), 65–75 (2014)

27

Equipos de Laboratorio a Escala Piloto con Realidad Aumentada a partir de modelados 3D, como Apoyo en el Proceso de Aprendizaje en el Nivel Superior

ELENA ORTIZ, ESBEIDY GOMEZ, JULIO CÉSAR SANDRIA Y HUGO DANIEL GONZÁLEZ

Equipos de Laboratorio a Escala Piloto con Realidad Aumentada a partir de modelados 3D, como Apoyo en el Proceso de Aprendizaje en el Nivel Superior

Elena Ortiz Hernández[1]✉ [0000-0003-0959-8435], Esbeidy Gómez Manuel[1] [0000-0003-0765-3402], Julio César Sandria Reynoso [1] [0000-0002-2810-1052] y Hugo Daniel González Dávila[1] [0000-0001-5545-4147]

[1] Universidad Tecnológica del Sureste de Veracruz, Nanchital de Lázaro Cárdenas del Río, Ver. México

elenaortiz.hz@gmail.com, esbeidy.gomez@utsv.edu.mx, julio.sandria@utsv.edu.mx, hugodgdavila@gmail.com

Resumen. El diseño 3D (Tercera Dimensión), la Realidad Virtual (RV) y la Realidad Aumentada (RA) son alternativas prometedoras para dar soporte al aprendizaje. Centrándose en su definición, el uso de la RA permite la visualización directa o indirecta de elementos del mundo real combinados o aumentados con elementos generados desde el ordenador [1].

El presente trabajo se enfoca en el desarrollo de aplicaciones de realidad aumentada de equipos de laboratorio a escala piloto incorporando gráficos generados a través de modelos 3D, con el fin de facilitar el proceso de aprendizaje mediante la información visual a través de las tecnologías emergentes. Los resultados preliminares exponen el avance de las aplicaciones en los laboratorios con las nuevas tecnologías dentro del campo de la educación superior, ubicándolo como una tendencia que podría transformar significativamente las prácticas docentes.

Palabras clave: Realidad aumentada, Laboratorio a escala piloto, Diseño 3D, aprendizaje.

Abstract. 3D design, Virtual Reality (VR) and Augmented Reality (AR) are promising alternatives to support learning. Focusing on its definition, the use of allows AR direct or indirect visualization of real-world elements combined or augmented with elements generated from the computer [1].

This paper focuses on the development of augmented reality applications of laboratory pilot scale equipment incorporating graphics generated through 3D models (third dimension), in order to facilitate the learning process through visual information through technologies emerging. Preliminary results show the progress of applications in the laboratories with new technologies in the field of higher education, placing it as a trend that could significantly transform teaching practices

Keywords: Augmented reality, Laboratory pilot scale, 3D Design, learning.

1 Introducción

La Realidad Virtual se usa ampliamente en diferentes campos, especialmente en ciencias de la computación, ingeniería y medicina. En ingeniería, la realidad virtual está siendo usada efectivamente para la educación, ya que representa ventajas como la simulación en tiempo real de equipos de laboratorio costosos; evita potenciales daños a los equipos causados por mal uso por parte de los estudiantes durante las prácticas; permite atender prácticas virtuales con grupos grandes; previene riesgos durante el trabajo; y permite a los estudiantes interactuar con procesos de manufactura completos, lo cual sería prácticamente imposible de otra forma [2].

La Realidad Aumentada no reemplaza al mundo real por uno virtual, al contrario, lo que ve el usuario lo complementa con información digital.

El usuario nunca pierde contacto con el mundo real que puede visualizar y al mismo tiempo le permite interactuar con la información virtual superpuesta [3].

Mediante esta tecnología, se puede incorporar contenido en forma de texto, imagen, audio, vídeo, modelos 3D, entre otros elementos, en la percepción del mundo real del usuario. Estos "aumentos" de la realidad ayudan a mejorar el conocimiento del individuo y permitirle un mayor grado de comprensión sobre lo que sucede a su alrededor [4].

Por su parte Cabero y García [5], señalan sus propiedades más significativas: ser una realidad mixta, integrada en tiempo real, que posee una diversidad de capas de información digital, que es interactiva y que, mediante su utilización, enriquecemos o alteramos la información. La RA ofrece numerosas posibilidades educativas y un inmenso potencial para mejorar el aprendizaje y la enseñanza [6]. Además, proporciona a los usuarios el acceso a un contenido multimedia variado y significativo, facilitándoles un contexto relevante y con el que poder interactuar de manera inmediata.

Este trabajo está enfocado en la segunda etapa del modelado 3D y entorno virtual aumentado de las instalaciones y equipos de laboratorio a escala piloto del Nivel Superior. La importancia de este proyecto está en las aplicaciones que fortalezcan el proceso de aprendizaje, faciliten la interacción de los involucrados, promocionando al mismo tiempo el interés por la universidad.

2 Motivación

La motivación principal de este trabajo de investigación es a raíz del continuo interés por la evolución de las nuevas tecnologías que han surgido en los últimos años, sobre todo en el ámbito de la educación con la finalidad de optimizar el proceso de aprendizaje.

La Realidad Virtual es posiblemente una de las tecnologías actuales más conocidas por la sociedad, pero la Realidad Aumentada ha resurgido nuevamente y ha creado un gran interés no solo entre los usuarios comunes sino también entre investigadores, empresas e instituciones educativas que han visto en esta tecnología una oportunidad para mejorar la calidad en el proceso de enseñanza-aprendizaje en diferentes niveles educativos.

El hecho de observar como la Realidad Aumentada presenta características de interactividad y tridimensionalidad captando el interés por parte de los estudiantes y pensar en las aplicaciones que podría tener en diferentes asignaturas, hace que aumente la perspectiva sobre esta tecnología.

3 Trabajo relacionado

Esta tecnología de visualización ha sido especialmente empleada en Europa y en países desarrollados como Estados Unidos, donde se encuentran los simuladores de realidad aumentada más avanzados hasta el momento. Sin embargo, a continuación se describen algunos proyectos relevantes que han sido desarrolladas y aplicadas en distintas universidades.

El proyecto ARLAB (Laboratorio con Realidad Aumentada), se implementó en el 2013, en el Instituto Tecnológico y de Estudios Superiores de Monterrey. Que consiste en un laboratorio basado en Realidad Aumentada y dispositivos móviles, que combinados con los elementos físicos clásicos del laboratorio permite desarrollar experimentos para reforzar y complementar el conocimiento teórico, así como las habilidades prácticas [7].

En el año 2017, se describen 21 experiencias derivadas del Proyecto de Innovación ACRA, impulsado por la Consejería de Educación Asturiana e implementado en aulas de Infantil, Primaria, Secundaria y Bachillerato de 9 centros escolares, que promueven la integración de la Realidad Aumentada (RA) en el proceso de enseñanza-aprendizaje de ciencias [8].

Por otro lado, en la Universidad de Santiago de Chile, en ese mismo año, se llevó a cabo el proyecto del “Aula virtual en la asignatura de Física: un apoyo a las Prácticas de Laboratorio" [9].

En los últimos años se ha experimentado un mejor avance de esta tecnología, llegando a ofrecer nuevas posibilidades de uso. Por otra parte, gracias a los dispositivos actuales que disponen de mayor capacidad de procesamiento e incluyen sensores de última generación, cámaras digitales y sistemas de localización global, ha sido posible desarrollar sistemas de realidad aumentada más precisos y asequibles.

4 Objetivo de la investigación

Como objetivo principal de esta investigación se pretende estudiar y demostrar la utilidad en la tecnología de RA en el proceso de aprendizaje para los estudiantes del Nivel Superior, mediante aplicaciones móviles o software para escritorio, partiendo de modelos 3D.

Con el propósito de alcanzar este objetivo, se propone desarrollar un entorno para los equipos de laboratorio a escala piloto pertenecientes a la universidad, mediante las tecnologías de modelado 3D y RA, en el cual, los alumnos podrán interactuar en tiempo real con el entorno que se les presenta, creando una experiencia única en el proceso de aprendizaje.

5 Metodología

La metodología de trabajo quedó dividida en dos partes fundamentales: la primera referida al diseño (explicación, experimentación y diseño de la propuesta) y una segunda dedicada a la implementación (contacto, puesta en práctica y valoración de los resultados), se basó fundamentalmente en el concepto didáctico *learning by doing*, es decir, aprender haciendo. El proceso queda reflejado en las siguientes líneas:

Fue necesario investigar para tener los conocimientos básicos sobre el área de RA con base a las siguientes cuestiones: ¿Qué es? ¿Para qué sirve? ¿Dónde se puede aplicar? ¿Cómo se puede utilizar? Y por último, el conocimiento sobre prácticas y experiencias didácticas basadas en la implementación de la misma en los equipos de laboratorio a escala piloto

También se construyeron las bases teóricas sobre los *códigos QR (del inglés Quick Response code, "código de respuesta rápida")* en el campo de la educación. Siendo un sistema que nos permite el almacenamiento de información variada, de manera que posibilita al usuario acceder a la misma de manera rápida mediante un lector de códigos QR [10]. De igual modo, se indago sobre las distintas aplicaciones móviles y páginas web teniendo opciones a elegir, seleccionando la más adecuada para la generación del QR.

Se experimentó con distintas aplicaciones de RA, observando sus entornos con la finalidad de seleccionar lo más óptimo, fácil de aprender y utilizar. Por mencionar un ejemplo a través de RA con marcador (Son los dibujos que ofrecen para su lectura en las aplicaciones) o sin marcador, es decir, creando la RA a partir de un objeto físico inmóvil o bien otorgando el acceso a las aplicaciones mediante códigos QR.

De manera colaborativa y continua se diseñaron con modelado 3D los objetos de los laboratorios del Nivel Superior: “Gestión de la producción”, “Metrología y Resistencia de Materiales” y “Electricidad y Electrónica”, cuidando cada detalle para no perder de vista la realidad con la nueva percepción de la RA.

Finalmente se realizó la exportación de todos los diseños en 3D en las aplicaciones de RA para móvil y escritorio.

6 Logros (avances en el trabajo de tesis)

- Diseño de modelos 3D de los equipos de laboratorio a escala piloto de Mantenimiento Industrial y Mecatrónica. Que fueron construidos a través de un software de diseño gráfico arquitectónico con apoyo de los planos oficiales de la Universidad, fotografías y videos (Ver Fig. 1 y 2).
- Desarrollo de objetos en RA a través de distintas aplicaciones para escritorio y teléfono celular, disponibles para los sistemas operativos de diferentes plataformas. Dependiendo el nivel de RA y la utilización de marcadores o códigos QR.
- Mostrar los entornos aumentados de los equipos de laboratorio a escala piloto de las carreras de Mantenimiento área industrial y Mecatrónica.

- Cursos de introducción al modelado 3D, Realidad Virtual y Realidad Aumentada, para capacitación, logrando motivar a utilizar las tecnologías emergentes en futuros proyectos en el Nivel Superior.
- Promover una cultura de indagación por parte de los docentes y del alumnado de la universidad en la promoción de nuevas experiencias de aprendizaje haciendo uso de la RA y los códigos QR

Fig. 1. Modelado 3D del edificio de Laboratorio pesado

Fuente: elaboración propia

Fig. 2. Modelado 3D del Laboratorio de gestión de la producción

Fuente: elaboración propia

7 Resultados previstos

El proyecto “Equipos de Laboratorio a Escala Piloto con Realidad Aumentada a partir de modelados 3D, como Apoyo en el Proceso de Aprendizaje a Nivel Superior”, está desarrollado para ser implementado en la universidad, con la finalidad de fortalecer la percepción, interacción de los universitarios en la utilización de las instalaciones por medio del modelado 3D y la RA, así mismo dar una nueva experiencia con el uso de las tecnologías de vanguardia.

Se ha constatado la experiencia de aprendizaje como un modelo para fomentar prácticas curriculares inmersivas en el campo educativo, promoviendo un acercamiento mayor entre lo real y lo virtual [11].

7.1 Aplicación RA de Escritorio

Una vez instalado el software de RA que lleva asignado el reconocimiento del marcador, basta con iniciar la aplicación y enfocar con la cámara del dispositivo al marcador, visualizara de manera automática el objeto en RA. En este caso concreto, se ha optado por configurar la visualización del objeto 3D de manera común, es decir, representando el contenido de forma paralela al marcador (Ver Fig. 3 y 4), puesto que resulta más eficaz si el marcador se encuentra en posición vertical, siempre y cuando se encuentre con la configuración del software adecuado.

Fig. 3. Código QR usado como marcador.

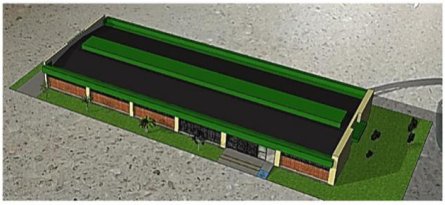


Fig. 4. Vista previa de RA del Edificio

Fuente: Elaboración propia

7.2 Aplicación móvil

Al momento de enfocar el marcador en forma de imagen a través de un dispositivo móvil, ésta es reconocida por la aplicación donde se encuentra el modelo 3D, dando

lugar a la visualización de los equipos de los laboratorios y su entorno en realidad aumentada (Ver Fig. 5).

Cabe mencionar se encuentra disponible en distintos sistemas operativos, además de una aplicación de pago, pero su utilización es posible gracias a la obtención de una licencia académica gratuita.

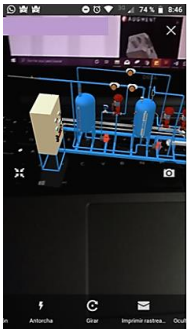


Fig. 5. Visualización RA en app móvil
Fuente: Elaboración propia

El análisis de la experiencia didáctica de aplicación de apps de RA con el alumnado universitario nos permite aseverar como primera conclusión que la utilización de objetos con base en la tecnología aumentada en la Educación Superior, despierta verdadero interés entre los estudiantes. Este extremo, también ha sido corroborado por diferentes estudios que muestran alto niveles de satisfacción cuando el alumnado utiliza esta tecnología y el aumento significativo de los niveles de motivación cuando los estudiantes están inmersos en acciones formativas con RA [12].

8 Conclusiones

Finalizada la investigación sobre el funcionamiento de la Realidad Aumentada, sus niveles, campos de aplicación y su puesta en práctica en el desarrollo e interacción con los estudiantes, se puede concluir que la utilización de objetos con base en la tecnología de modelado 3D y RA en la Educación Superior, despierta verdadero interés entre los estudiantes, debido a que les permite visualizar, interactuar y aprender con los equipos de laboratorio a escala piloto, sin embargo, aún se encuentra en fase de implementación, su evolución depende de las tecnologías emergentes, el conocimiento y la accesibilidad de los usuarios.

También se observó que la combinación de modelado 3D y RA ofrece una visualización tridimensional de los detalles singulares de las instalaciones de los laboratorios incluso a escala real, facilitando la comunicación y el entendimiento en el proceso de aprendizaje.

Cabe destacar que la utilización de la RA como tecnología emergente da un aspecto competitivo favoreciendo la promoción del programa educativo en los alumnos del nivel medio superior, ya que fortalece la oferta mediante una fuente visualmente atractiva. Y en el nivel superior es de suma importancia, sobre todo en lo que respecta al enriquecimiento del proceso educativo.

Referencias

1. Cobo, C. y Moravec, John. (2011). Aprendizaje Invisible. Hacia una nueva ecología de la educación. Col·lecció Transmedia XXI. Laboratori de Mitjans Interactius / Publicacions i. Edicions de la Universitat de Barcelona.
2. Vergara, D., Rubio, M.P. y Lorenzo, M. (2017). On the Design of Virtual Reality Learning Environments in Engineering. Multimodal Technologies and Interaction 1 (11), http://dx.doi.org/10.3390/mti1020011.
3. Basogain, X., Olabe, M., Espinosa, K., Rouèche y Olabe, J. (2007). Realidad Aumentada en la Educación: una tecnología emergente. Tesis Doctoral. Escuela Superior de Ingeniería de Bilbao, España.
4. Reinoso, R. (2014). Curso de Introducción a la Realidad Aumentada. CEPCantabria –Plan de formación permanente del profesorado.
5. Cabero, J., García, F. y Barroso, J. (2016). The production of learning objects in "Augmented Reality": the experience of the SAV of the Universidad de Sevilla. International Journal of Educational Research and Innovation (IJERI), volumen (6), pp. 110-123.
6. Bacca, J., S. Baldiris, R. Fabregat, S. Graf y J. Kinshuk, (2014). Augmented Reality Trends in Education: A Systematic Review of Research and Applications, Educational Technology & Society, 17 (4), pp. 133-149.
7. Rodriguez R. y Santillana R. (2013). ARLAB: Laboratorio con Realidad Aumentada. Tesis Doctoral. Instituto Tecnológico y de Estudios Superiores de Monterrey, Morelia.
8. Villalustre, L., Del Moral, M., Neira, M. y Herrero, M. (2017). Proyecto ACRA: experiencias didácticas en ciencias con realidad aumentada en los niveles pre-universitarios. Edutec. Revista Electrónica De Tecnología Educativa, (62), 1-18 (a369). https://doi.org/10.21556/edutec.2017.62.1009.
9. Calvo, I. (2018). El aula virtual en la asignatura de Física: un apoyo a las Prácticas de Laboratorio. Investigación, Innovación y Tecnologías, la triada para transformar los procesos formativos. Chile: USACH.
10. Román, P. y Méndez, J. (2014). Experiencia de innovación educativa con curso MOOC: los códigos QR aplicados a la enseñanza. Profesorado, 18(1), pp. 113-136.
11. Cabero, J. y Barroso, J. (2016). The educational possibilities of Augmented Reality. NAER, 5(1), 44-50.
12. Lee, M. (2016). 32 Augmented Reality Apps for the Classroom. Recuperado de <http://www.teachthought.com/the-future-of-learning/technology/32-augmented-reality-apps-for-the-classroom-from-edshelf/>.

28

Seguimiento de Objetos en la cadena de suministros de la industria de manufactura

ANA KAREN HERNANDEZ, ALEXIS JAIR LÓPEZ Y PERFECTO MALAQUÍAS QUINTERO FLORES

Seguimiento de Objetos en la cadena de suministros de la industria de manufactura*

Ana Karen Hernández Vázquez[1][0000−0002−8265−1303], Perfecto Malaquías Quintero Flores[1][0000−0001−7651−4364], and Alexis Jair López Gutiérrez[1][0000−0001−7755−5109]

Tecnológico Nacional de México , Av. Instituto Tecnológico s/n, 90300 Apizaco, México
{anakarenhv26,alexistegno,parfait.phd}@gmail.com

Resumen En la actualidad, el uso de redes de sensores es de gran interés para la investigación. Una de las aplicaciones más desafiantes de tal tecnología es el seguimiento de objetos, el cual consiste principalmente en detectar, monitorear y rastrear tales objetos en movimiento. Las principales áreas de oportunidad se encuentran en vida silvestre, agricultura, dirección de cadena de suministros de la industria de manufactura de diferentes bienes de capital y producción, sin dejar de lado que existen muchos otros campos de aplicación. El siguiente trabajo ataca el seguimiento de objetos mediante: 1.- Estructura Jerárquica de Agrupación de sensores para red de sensores inalámbricos, 2.- Seguimiento de ubicación basado en agrupación para redes de sensores de seguimiento de objetos. La simulación arrojan resultados que muestran la disminución de sobrecarga del uso de energía invertida en los nodos-sensores y se visualiza el logro de extensión de vida útil de la red.

Keywords: Sensor Networks · Object Tracking · Energy Efficiency · Lifetime Prolonging.

1. Introducción

En México hay una gran cantidad de sectores de manufactura que forman parte fundamental de la economía. Entre los más destacados se encuentran el automotriz, el de las maquinas y equipos y el de las industrias metálicas, así como el de productos de alimentos, bebidas y tabacos. La industria de manufactura en México sigue en la mira de inversionistas y empresas de diversas partes del mundo. De hecho, continúa empujando como motor de la economía mexicana, pero la producción ya no es la misma, de hecho los empresarios cada día buscan nuevas herramientas para la innovación [4].
El seguimiento de objetos tiene muchas aplicaciones en la vida real, como la vigilancia en redes de vehículos, monitoreo de vida silvestre, propagación de incendios, observación de pacientes, localización, etc. De hecho, existe una gran

* Supported by CONACYT

demanda de seguimiento de ubicación de objetos en movimiento desde diversos servicios basados en ubicación [15]. Una de las aplicaciones más desafiantes de tal tecnología es el seguimiento de objetos, el cual consiste principalmente en detectar, monitorear y rastrear tales objetos. La energía de las redes de sensores se puede conservar cambiando el nodo de un sensor a modo de suspensión cuando no hay ningún otro objeto en el rango de detección del nodo y activando un nodo del sensor cuando un objeto debe de ingresar al rango de detección [12]. Se han publicado un gran número de propuestas para el seguimiento de objetos, la mayoría de estas propuestas no abordan el equilibrio entre la precisión del seguimiento y la eficiencia energética. Con estas motivaciones proponemos la implementación de [8,9]. A nuestro criterio la implementación de estos trabajos no corresponde a las categorías de seguimiento existente. En México una industria principal, dinámica y en los últimos años muy competitiva es la industria de manufactura, sin embargo existen malas prácticas en las empresas causando deficiencias en la productividad, en la entrega de productos lo que produce una pérdida de capital humano, capital monetario y multas de millones de pesos en la industria. Por ello nace la necesidad de detectar, rastrear y monitorear objetos en las líneas de producción de la industria de manufactura. Teniendo como fundamento la perdida de capital monetario, de capital humano y las millonarias infracciones que año con año sufre la industria se implementan los trabajos [8,9] como posible solución a esta problemática.

2. Trabajo relacionado

Las aplicaciones de seguimiento de objetos en redes de sensores han atraído la atención de los investigadores debido a su amplia aplicación. Hasta la fecha, se ha publicado un número significativo de propuestas relacionadas con el seguimiento de objetos en redes de sensores inalámbricas, la mayoría de estos estudios no pueden abordar el equilibrio entre la eficiencia energética y la precisión del seguimiento. Los autores en [12] han propuesto un esquema de seguimiento de objetos OTSN (redes de sensores de seguimiento de objetos) utilizando el enfoque de minería de datos. En este esquema se ha mejorado el algoritmo de Apriori para las reglas de asociación minera y se ha aplicado a OTSN. De otra manera en [1,5] se presenta un modelo integral para el seguimiento de objetivos y el análisis y comparación de diferentes algoritmos para la ubicación de sensores basados en el uso de una red inalámbrica de sensores. Hawbani et al [8] proponen un modelo de seguimiento novedoso llamado *Seguimiento de ubicación basado en agrupación (GLT por sus siglas en inglés)*, que consta de dos niveles, el primer nivel, que se llama árbol de notificaciones (NT por sus siglas en inglés), y el segundo nivel se llama árbol de expansión jerárquica (HST por sus siglas en inglés). GLT está basado en [9] (GHS por sus siglas en inglés), en GHS los nodos son divididos en grupos (no clústeres) de acuerdo con su región máxima cubierta de modo que cada grupo contenga un número de nodos y un número de líderes. GTL garantiza un alto rendimiento mediante el soporte del NT y HST

para agregar e informar la fecha al nodo receptor utilizando el número mínimo de saltos considerando el balance de energía.

TS Chen et al. [2] proponen un método de seguimiento de objetos distribuidos utilizando la estructura de la red de poligonos convexos, llamados caras, los nodos en las caras cooperan para encontrar las trayectorias de un objeto y luego estas trayectorias se utilizan para predecir el movimiento de los objetos. El método propuesto basado en la construcción del árbol de trayectoria puede reducir tanto el espacio de almacenamiento de las trayectorias recolectadas como el tiempo dedicado al análisis de predicción de trayectorias. Las simulaciones muestran que el método propuesto puede reducir el consumo de energía de los nodos y hacer una predicción de la dirección de movimiento de los nodos con precisión.

3. Propuesta de Solución

Con base en los supuestos, este trabajo abordara el problema de seguimiento de objetos en redes de sensores, con siete algoritmos para el seguimiento de los objetos, ahorro de energía, un correcto enrutamiento de los datos, etc.:

- Agrupamiento / grupos asociados: El algoritmo de agrupamiento se utilizará para dividir los nodos en grupos de acuerdo con las regiones superpuestas en el campo. La idea principal es dividir los nodos de la red en grupos según sus regiones geográficas máximas cubiertas, de tal manera que cada grupo contenga un número de nodos y un número de líderes.
- Enrutamiento de datos: Cuando un nodo origen tiene nodos por enviar, primero comenzara por encontrar el grupo de distancia mínima a la estación base y luego desde el grupo seleccionado elige el siguiente salto para reenviar los datos.

  Cuando un nodo fuente tiene datos que enviar, primero comienza buscando el grupo de distancia mínima a la estación base y luego desde el grupo seleccionado elige el siguiente salto para reenviar los datos. El siguiente salto se puede seleccionar de acuerdo con la distancia mínima al siguiente nodo destino o la energía máxima del nodo destino. Después de mantener el siguiente salto (nodo destino), el líder del grupo seleccionado transmitirá los datos directamente al nodo destino. Una vez finalizada la operación de recepción de datos, el nodo destino encuentra su grupo de distancia mínima y luego transmite los datos de forma recursiva hasta que los datos lleguen a la estación base.

  El flujo de datos del proceso de enrutamiento de datos se puede observar en la figura 1.
- Modelo de seguimiento de objetos: Se propone que cualquier $s_j \in S$ es consciente de los grupos asociados a los que pertenece, de manera que $G^*(s_j) = G^x \mid G^x \in G^* \, y \, s_j \in G^x$.

  – Árbol de notificaciones/NT: El objetivo principal del árbol de notificaciones es distribuir la disipación de la energía entre los nodos de $G^*(s_j)$ de la manera más uniforme posible, de modo que cada nodo en $G^*(s_j)$ reciba una

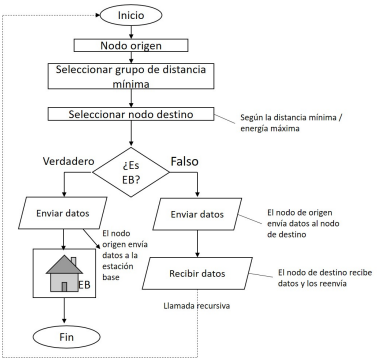


Figura 1. Diagrama de flujo de enrutamiento de datos.
Fuente: elaboración a partir de [9].

notificación y consuma la misma cantidad de energía. Cada nodo s_j tendra un árbol de notificación NT (s_j) y la raíz será s_j. Cuando haya un mensaje de notificación para ser enviado desde s_j a sus grupos asociados $G^*(s_j)s_j$ enviará $\mid G_j - s_j \mid$, es decir enviará un mensaje a cada componente de $G_j - s_j$. Cada $G_{j,i} \in G_j - s_j$ construirá un árbol de expansión $T_{j,i}$ que se unirá al nodo raíz de NT.

– Árbol de expansión jerárquica/HST: Cada subgrafo o componente $G_{j,i} \in G_j - s_j$ construira un árbol de expansión T_j, teniendo en cuenta que cualquier sensor de nodo en T_j, consumira la misma cantidad de energía para transferir el mismo mensaje de notificación que otros. Se supone que el conjunto de vértices del subgrafo $G_{j,i}$ es $V_{j,i}$

- Tabla de objetos detectados y tabla de mensajes de notificación: El objetivo de la creación de estas tablas se enfoca principalmente en el seguimiento del objeto de manera que cada objeto se asigne a un nodo proxy. Para $s_j \in S$ se asume que tiene una Tabla de objetos detectados TOD (s_j) para guardar la información del objeto detectado, (es decir, posición actual o imagen), y tiene una Tabla de mensajes de notificación TMN(s_j) para almacenar la notificación enviada/ recibida hacia/desde otros nodos.
- Reporte de datos mediante HST: El árbol de notificaciones agrega los datos y el nodo proxy comenzara el proceso de reenvío. Cualquier nodo en la ruta del proxy al sink (sumidero) reenvía los datos recopilados a su padre en HST o a su padre en NT. El padre en HST es el predeterminado. Si la energía del padre en HST es mayor que el valor umbral, entonces se selecciona el padre en HST, de lo contrario se selecciona el padre en NT.

- Mecanismo de prolongación de vida: Cada nodo puede tener dos padres: un padre para el árbol de notificaciones y un padre para el árbol de expansión jerárquica. La tarea principal del árbol de notificaciones es evitar la redundancia y la duplicación de datos. Cuando el HST padre del nodo informador tiene una potencia menor (es decir, inferior al valor umbral).

4. Resultados

Para la evaluación del desempeño de [8] y [9], desarrollamos un simulador en Java. Se decidió en desarrollo en Java ya que es un sistema multiplataforma y hay innumerables bibliotecas y/o utilidades creadas para ampliar su funcionalidad, además de la experiencia previa con este lenguaje.

4.1. Simulación:

Las unidades de medida son el tiempo de simulación (segundos), el retraso de transmisión (milisegundos), el consumo de energía (julios) y la velocidad del objeto (metros/segundo). La Tabla 1 muestra los parámetros utilizados para evaluar el desempeño.

Tabla 1. Parámetros de simulación. Fuente: elaboración a partir de [8,9]

Parametro	Valor
Tamaño de red	500 x 500 m^2
Nodos	coberturas grid 1 y zigzag
Radio	25 metros
Tamaño de datos	1024 bits
Energía inicial	0.5 julios
E_{elec}	50 nJ/bit
ε_{mp}	0.001 pJ/bit/m^2
Estación base	Dentro

Donde:
E_{elec}=Disipación de energía.
ε_{mp}= modelo multitrayecto del ampliador transmisor.
BS= Estación base.

Cobertura de Topologías: El primer paso en la implementación de estas redes inalámbricas de sensores es determinar, con respecto a los criterios de rendimiento específicos de la aplicación, (i) en el caso de que los sensores sean estáticos, y (ii) en el caso de que los sensores sean móviles. Estos dos casos se denominan colectivamente como el problema de cobertura [11]. Como el rango de comunicación del nodo lejano es limitado, los nodos cercanos a la estación

base actuarán como enrutadores para reenviar paquete que se transmiten desde nodos muy lejanos.

Versión de cobertura grid: Para los esquemas de cobertura basados en grid (cuadrícula),se han seleccionado dos algoritmos proporcionados en [7], Versión de Cobertura Grid (1) y Versión de Cobertura Grid(2). Ver Figura 2 como referencia de representación cobertura grid en su versión (1).

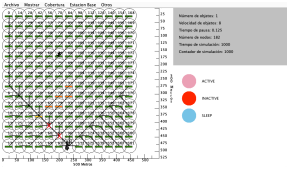


Figura 2. Árbol de expansión jerarquíca de 182 nodos desplegados utilizando el esquema grid 1. El rango de detección es de 25m y el rango de comunicación es de 50m. El proceso de la barra indica la capacidad de la bateria. La estación base se encuentra en el nodo 97.
Fuente: elaboración propia

Versión de cobertura zigzag: Para el esquema de cobertura zigzag [6], el área de interés dividida en múltiples patrones de zigzag con múltiples esquemas y segmentos de líneas, cada nodo se implementa en una esquina del patrón de zigzag. Ver la Figura 3.

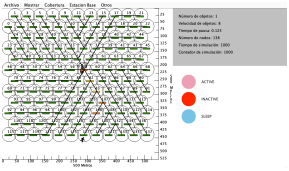


Figura 3. Árbol de expansión jerárquica de 138 nodos desplegados utilizando esquema zigzag, el rango de detección es de 25m y el rango de comunicación es de 50m. El proceso de la barra indica la capacidad de la batería. La estación base se encuentra en el nodo 58.
Fuente: elaboración propia

4.2. Evaluación

Para la evaluación, se realizó un análisis cuantitativo. Los parámetros utilizados para la evaluación son los que se visualizan en la tabla 1. La evaluación incluye la comparación con 4 protocolos de clústeres existentes (CAICS,DLC y MHC por sus siglas en inglés) [3, 13, 14]. La eficiencia en el envió de paquetes de la propuesta de solución en comparación con CAICS, MHC y DLC, se hace notorio, debido a que la estación base de ubica en el centro de la deteccion y la vida útil de la red es más duradera y el número de paquetes de operación de conmutación se reduce.

Consumo de energía : Se utiliza un modelo de energía llamado "Modelo de radio de primer orden" [10].

En la Figura 4, la energía residual (en joule) es mostrada después de 10 rounds de simulación.

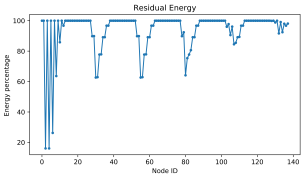


Figura 4. Energía residual de los nodos después de 10 rounds. La topología de red es zigzag.
Fuente: elaboración propia

5. Conclusiones

Las redes de sensores es una de las aplicaciones clave que está ampliamente adaptada debido a su gran cantidad de implementaciones y usos. El mayor obstáculo que enfrenta la mayoría de las técnicas de seguimiento en redes de sensores es el hecho de que la red de sensores carece de una fuente de alimentación, lo que a su vez restringe el tipo de tareas que debe y puede realizar la red. En base a los resultados experimentales, se concluye que es posible el seguimiento de objetos en movimiento en redes de sensores, gracias a que se tiene un seguimiento óptimo y al mismo tiempo se preserva la vida útil de las redes de sensores. Los algoritmos garantizan un alto rendimiento con el soporte del árbol de notificaciones y el árbol de expansión jerárquica para agregar e informar la fecha al nodo receptor utilizando el mínimo número de saltos. Para una futura investigación se propone el estudio del movimiento rápido de objetos.

Referencias

1. Calafate, C.T., Lino, C., Diaz-Ramirez, A., Cano, J.C., Manzoni, P.: An integral model for target tracking based on the use of a wsn. Sensors (Basel, Switzerland) **13**(6), 7250–7278 (Jun 2013)
2. Chen, T.S., Chen, J.J., Wu, C.H.: Distributed object tracking using moving trajectories in wireless sensor networks. Wireless Networks **22**(7), 2415–2437 (Oct 2016)
3. Chen, W.P., Hou, J., Sha, L.: Dynamic clustering for acoustic target tracking in wireless sensor networks. IEEE Transactions on Mobile Computing **3**(3), 258–271 (Jul 2004). https://doi.org/10.1109/TMC.2004.22, conference Name: IEEE Transactions on Mobile Computing
4. EFE: Industria manufacturera de México afronta reto de la digitalización, https://www.elsoldemexico.com.mx/finanzas/expo-manufactura-industria-mexico-afronta-reto-de-la-digitalizacion-4736869.html, library Catalog: www.elsoldemexico.com.mx
5. Faudo, I.E., Cota-Ruiz, J., Aguilera, F.J.E., Perea, P.R.: Análisis y comparación de algoritmos de localización en redes inálambricas de sensores. Pistas Educativas, Instituto Tecnológico de Celaya **108** (2014)
6. Hawbani, A., Wang, X.: Zigzag Coverage Scheme Algorithm & Analysis for Wireless Sensor Networks. Network Protocols and Algorithms p. 19 (Dec 2013). https://doi.org/10.5296/npa.v5i4.4688, http://www.macrothink.org/journal/index.php/npa/article/view/4688
7. Hawbani, A., Wang, X., Husaini, N., Karmoshi, S.: Grid Coverage Algorithm & Analyzing for wireless sensor networks. Network Protocols and Algorithms **6**(4), 1 (Dec 2014)
8. Hawbani, A., Wang, X., Karmoshi, S., Kuhlani, H., Ghannami, A., Abudukelimu, A., Ghoul, R.: GLT: Grouping Based Location Tracking for Object Tracking Sensor Networks. Wireless Communications and Mobile Computing (2017)
9. Hawbani, A., Wang, X., Karmoshi, S., Wang, L., Husaini, N.: Sensors Grouping Hierarchy Structure (GHS) for Wireless Sensor Network. International Journal of Distributed Sensor Networks **11**(8), 650519 (2015)
10. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. pp. 10 pp. vol.2– (Jan 2000). https://doi.org/10.1109/HICSS.2000.926982
11. Hou, J., Yau, D., Ma, C., Yang, Y., Zhang, H., Hou, I.H., Rao, N., Shankar, M.: Coverage in Wireless Sensor Networks, pp. 47–79 (01 1970)
12. Liao, W., Chang, K., Kedia, S.P.: An object tracking scheme for wireless sensor networks using data mining mechanism. In: 2012 IEEE Network Operations and Management Symposium. pp. 526–529 (April 2012)
13. Mamun, Q., Ramakrishnan, S., Srinivasan, B.: Selecting Member Nodes in a Chain Oriented WSN. In: 2010 IEEE Wireless Communication and Networking Conference. pp. 1–6 (Apr 2010). https://doi.org/10.1109/WCNC.2010.5506377, iSSN: 1558-2612
14. Xue, W., Jun-Jie, M., Sheng, W., Dao-Wei, B.: Cluster-based Dynamic Energy Management for Collaborative Target Tracking in Wireless Sensor Networks. Sensors **7** (Jul 2007). https://doi.org/10.3390/s7071193
15. Zheng, Y., Zhou, X.: Computing with Spatial Trajectories (01 2011)

29

Detección de la distancia de objetos mediante el uso de disparidad estéreo en un vehículo autónomo

JOSUE EZEQUIEL JUAREZ MANJARREZ Y VICTOR ALEJANDRO GONZALEZ HUITRON

Detección de la distancia de objetos mediante el uso de disparidad estéreo en un vehículo autónomo

Josué Ezequiel Juarez Manjarrez[1] y Víctor Alejandro González Huitron[1]

[1] TecNM-Instituto Tecnológico de Culiacán Av. Juan de Dos Bátiz 310 pte, Col Guadalupe, CP 80220, Culiacán, Sin.
j_ezequieljm@hotmail.com, victor.gonzalez@itculiacan.edu.mx

Resumen. En este trabajo se presenta una metodología para el cálculo de las distancias de objetos por medio de disparidad estéreo, las cuales son utilizadas en un vehículo autónomo para la toma de decisiones. La propuesta fue implementada en una microcomputadora Raspberry Pi 3 b+ demostrando su capacidad de ser aplicada en dispositivos portátiles y de bajo consumo. La visión estéreo binocular emplea dos imágenes para obtener dicha estimación de la profundidad, la cual está relacionada con la diferencia relativa en las posiciones donde los puntos de la escena son proyectados sobre cada una de las imágenes, conocida como disparidad. Finalmente utilizando un mapa de disparidad, se propone un sistema de detección de obstáculos para un vehículo en movimiento, en el cual se crea una zona de riesgo a partir del modelo cinemático del mismo y se buscan todos aquellos objetos que se encuentran dentro de la zona (obstáculos).

Palabras clave: CLQ, Detección de distancia, Mapa de disparidad, Visión estéreo, procesamiento de imágenes, detección de obstáculos.

1 Introducción

El ser humano y otros seres de la naturaleza perciben el mundo en tres dimensiones gracias a su sistema de visión binocular, también conocido como visión estéreo, que les dota de un sentido de percepción de la profundidad debido a un fenómeno visual conocido como estereopsis. [10] La estereopsis es el fenómeno dentro de la percepción visual por el cual, a partir de dos imágenes ligeramente diferentes del mundo físico proyectadas en la retina de cada ojo, el cerebro es capaz de recomponer una tridimensional. [9] Aunque este proceso, completamente natural e instantáneo en los seres vivos, es extremadamente complejo, se puede emular hasta cierto punto mediante computadoras. Para tal fin, se utilizan cámaras estéreo: básicamente dos cámaras unidas cuyo desplazamiento mutuo es conocido. [10]

El objetivo principal de este trabajo es la generación de un mapa de disparidad, interpretando las correspondencias como la deformación continua entre el par de imágenes estéreo. En este sentido, el problema se trata en un dominio continuo y el método de solución se basa en las intensidades en escala de grises para la detección de la distancia de los objetos y su aplicación en un vehículo autónomo.

2 Trabajos relacionados

Las múltiples aplicaciones que tiene el uso de la disparidad estéreo es algo que ha llevado a la realización de múltiples estudios en el campo de la visión artificial. A continuación, se describen varios los siguientes trabajos dando múltiples aplicaciones.

En [1,2] se trabajó en la generación de un mapa de profundidad utilizando múltiples imágenes estéreo apoyándose de un mapa de disparidad global. En [3] se utilizó un algoritmo de visión estéreo en tiempo real implementado en GPU. En [4] se desarrolló e implemento un algoritmo de detección de obstáculos por visión artificial en la navegación de un vehículo utilizando el mapeo. En [6] se diseñó un sistema de evasión de obstáculos para una aeronave no tripulada usando visión estereoscópica. En [7] se realizó una comparación en el cálculo de la disparidad en distintas imágenes estéreo.. En [8] se trabajó en un sistema avanzado de detección de obstáculos y navegación autónoma para vigilancia y protección basado en la flota de vehículos aéreos no tripulados. En [10] se investigó la aproximación del mapa de disparidad estéreo mediante técnicas de aprendizaje automático.

3 Metodología general

SAD - El algoritmo base implementado (SAD) requiere de una función costo que mide la similitud entre los píxeles de la imagen tomada con la cámara izquierda (IL) y los de la imagen tomada con la cámara derecha (IR). Como principio básico del algoritmo SAD, es necesario medir la diferencia absoluta entre la intensidad de píxeles entre una subimagen y su área de búsqueda, como expresa la ecuación. [5]

$$SAD(x,y,d) = \sum_{(x,y)\in W} |I_L(x,y) - I_R(x+d,y)| \quad (1)$$

En la ecuación, *d* representa la disparidad o corrimiento de píxeles entre una imagen y la otra. Esta diferencia es función de la separación entre las cámaras y de la posición en profundidad de los objetos en cuestión. De este modo, para cada píxel (*x,y*) se obtendrá un valor de disparidad *d*, tal que minimice el costo SAD [3]. Una vez obtenidos todos los valores de disparidad para los diferentes conjuntos de puntos en los que se habían calculado las correspondencias, se puede estimar un mapa de disparidad [1].

Mapa de disparidad - Es una reproducción de la imagen original en la que están representadas todas las disparidades para los diferentes puntos de la imagen. Dicho mapa de disparidad es una herramienta que se utiliza para el cálculo de la distancia a las cámaras de los distintos objetos en la escena [1].

Posteriormente, es necesaria una etapa de filtrado para reducir los errores que presenten en la imagen al detectar objetos pequeños. Se realizará un proceso de consistencia, el cual retendrá solo las disparidades que sean conjuntos mayores a un valor predeterminado p, estos conjuntos de disparidades serán evaluados por subgrupos correspondientes a doce valores adyacentes. (fig. 1)

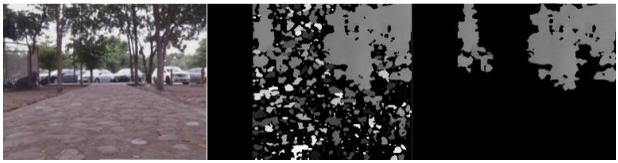


Fig. 1. – Proceso de generacion de un mapa de disparidad en escala de grises a partir de una toma de imagen, y aplicación de algoritmo de filtrado de objetos pequeños

4 Resultados

SGM – (Semi global Matching) fue el método utilizado para la generación del mapa de disparidad. La coincidencia estéreo precisa, es un requisito importante para muchas aplicaciones, como la reconstrucción 3D [5]. Los parámetros utilizados fueron un tamaño de bloque de 15x15, con un rango de disparidad de 64.

Finalmente, una vez obtenida una disparidad aceptable es posible realizar algunos ajustes visuales que permitan una mejor claridad, en este caso se realiza una representación con pseudocolor RGB, para posteriormente con el fin de simplificar la toma de decisiones en un vehículo autónomo, se procede a realizar un recorte de la imagen, lo que se obtiene es algo parecido a lo que se vería al mirar unas vías de tren que representan nuestro camino a seguir y mediante el cual el vehículo determinará hacia donde tiene que realizar un giro (fig. 2).

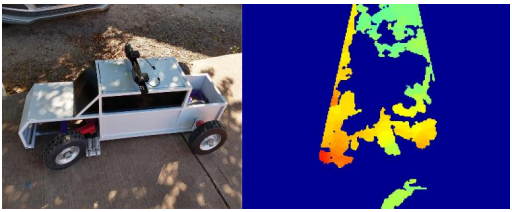


Fig. 2. – Mapa de disparidad para toma de decisiones en un vehículo autónomo

5 Conclusiones

Dado los resultados obtenidos se puede concluir que el descriptor propuesto más el filtrado en el tiempo obtuvo resultados satisfactorios en distancias entre 1 m y 15 m, con unos centros y desviaciones estándar cercanas a la media. Este resultado para un robot autónomo que se desplace en un ambiente interior es una distancia representativa

para identificar y evitar los obstáculos que se encuentran en el entorno.

Según los resultados obtenidos durante la etapa de pruebas, se determina que todo el proceso propuesto funciona de manera aceptable las veces que es ejecutado el programa. Se entiende por aceptable el hecho de encontrar un contraste claro de las formas de los objetos más cercanos respecto a los más lejanos. En imágenes como el último ejemplo presentado, es posible determinar a partir del mapa de disparidad solo la zona que representa riesgo de colisión.

References

1. Alcalde, J. (2014). Diseño de un protocolo de calibración de cámaras estéreo. Universidad Carlos III de Madrid. Madrid, España.
2. Calderón, D. (2012). Generación de mapas de profundidad a partir de imágenes estéreo utilizando registro no rígido. (Tesis de licenciatura). Universidad de Chile. Santiago de Chile, Chile.
3. Demski, A. J., Di Donato, A., Maudet, S., & Furfaro, A. (2015). Algoritmo de visión estereo en tiempo real implementado en GPGPU. http://sedici.unlp.edu.ar/handle/10915/52133
4. Escobedo, J. (2009). Desarrollo e implementación de algoritmos de detección de obstáculos por visión artificial en la navegación de un vehículo. (Tesis de maestría). Instituto Politécnico Nacional. Ciudad de México, México.
5. H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 328-341, Feb. 2008.
6. Layme, A. (2015). Diseño de un sistema de evasión de obstáculos para una aeronave no tripulada usando visión estereoscópica. (Tesis de licenciatura). Pontifica universidad católica del Perú. Lima, Perú.
7. Lecumberry, F. (2005). Calculo de disparidad en imágenes estéreo, una comparación. Congreso argentino de ciencias de la computación. Concordia, Argentina.
8. Merino, J. (2016). Sistema avanzado de detección de obstáculos y navegación autónoma para vigilancia y protección basado en flota de vehículos aéreos no tripulados. (Tesis de doctorado). Universidad de Málaga. Málaga, España.
9. Montañana, I. T., Fajarnés, G. P., & Carmona, P. L. (2006). EXPRESSIÓ GRÀFICA I INFOGRAFIA. Universitat Politècnica de València. https://books.google.com.mx/books?id=nVyXAS1BErAC
10. Peris, M. (2011). Aproximación del mapa de disparidad estéreo mediante técnicas de aprendizaje automático. (Tesis de maestría). Universidad politécnica de Valencia. Valencia, España.

30

Valoración de la salud mental de estudiantes universitarios a partir de datos de uso de teléfonos inteligentes

REGINA GONZÁLEZ GÓMEZ, VALERIA SOTO MENDOZA Y JESÚS ALEJANDRO NAVARRO ACOSTA

Valoración de la salud mental de estudiantes universitarios a partir de datos de uso de teléfonos inteligentes

Regina González Gómez[1], Valeria Soto-Mendoza[2], and Jesús Alejandro Navarro-Acosta[2]

[1] Facultad de Sistemas, Universidad Autónoma de Coahuila, Artega, Coahuila, México
regina.gonzalez@uadec.edu.mx
[2] Centro de Investigación en Matemáticas Aplicadas, Universidad Autónoma de Coahuila, Saltillo, Coahuila, México
{vsoto,alejandro.navarro}@uadec.edu.mx

Resumen La gran mayoría de estudiantes universitarios utilizan teléfonos inteligentes o *smartphones*. Durante sus estudios universitarios se ven sometidos a situaciones que pueden afectar su salud mental. La salud mental es un estado de bienestar donde la persona puede manejar adecuadamente el estrés normal que le produce su vida cotidiana para así poder trabajar de forma productiva y contribuir a su comunidad. La salud mental es algo más que la ausencia de trastornos o discapacidades mentales. Los trastornos mentales más frecuentes son la ansiedad y depresión. Se utilizó la técnica de regresión logística para analizar datos de uso de teléfonos inteligentes y diagnósticos de ansiedad, depresión y riesgo suicida en estudiantes universitarios. Los resultados de clasificación para la exactitud en los datos de validación arrojaron un 93% y 71% para los diagnósticos de ansiedad y depresión.

Keywords: regresión logística, smartphones, ansiedad, depresión, universitarios.

1 Introducción

Según los resultados de la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares [7], el número de usuarios de teléfonos inteligentes o smartphones es de 64.7 millones. El 72.2% de la población de seis años o más utiliza un teléfono celular. Con base en estas estadísticas se puede suponer que la gran mayoría de estudiantes universitarios utilizan teléfonos inteligentes o *smartphones*. La salud mental es un estado de bienestar donde la persona puede manejar adecuadamente el estrés normal que le produce su vida cotidiana para así poder trabajar de forma productiva y contribuir a su comunidad [12]. La salud mental es algo más que la ausencia de trastornos o discapacidades mentales. Dos de los trastornos mentales más frecuentes son la depresión y la ansiedad [24]. La salud mental es un tema de gran importancia que aún sigue siendo ignorado por muchas instituciones y personas. Entre los obstáculos que limitan a las personas a buscar ayuda cuando su salud mental se ve afectada, se encuentran la concepción de que ir a terapia es solo para enfermos mentales graves, la falta de recursos para acceder a un tratamiento y la evaluación clínica inexacta, entre otros. Por tanto, en este trabajo se propone una herramienta para analizar el

uso de *smartphones* mediante de técnicas de inteligencia artificial y detectar si los estudiantes universitarios presentan patrones que puedan ser relacionados con principios de ansiedad o depresión.

2 Trabajo relacionado

En [17] mencionan que la depresión en la vejez se asocia con la disminución de actividad física, por lo que recolectaron datos de acelerometría mediante una pulsera en la muñeca. Con estos datos, buscaron caracterizar los niveles de actividad en los participantes con depresión y sin ella. En [13] se habla sobre el trastorno de la fobia social. Los autores explican que se pueden utilizar sensores portátiles (*wearable sensors*) para detectar la ansiedad. Realizaron un experimento de ansiedad inducida, donde los participantes leyeron un discurso de 10 minutos frente a dos profesores y se midieron sus frecuencias de parpadeo espontáneo y cardíacas, para calcular su nivel de ansiedad. Los autores en [15] presentan un enfoque de aprendizaje automático para la detección de estrés, mediante la combinación de dos sensores que capturan respuestas fisiológicas y sociales. Se compararon los datos usando clasificadores como la máquina de soporte vectorial y K-vecinos más cercanos. Los autores en [14] proponen un modelo de cadena de Markov para la detección del estado de ansiedad de los cuidadores de personas con demencia utilizando datos fisiológicos. A partir de imágenes de la red social *Instagram*, en [20] detectaron marcadores de depresión mediante herramientas de aprendizaje automático. Las características estadísticas se extrajeron computacionalmente utilizando análisis de color, componentes de metadatos y detección de rostros.

3 Metodología

De manera general, la metodología empleada es la siguiente: primero se recolectaron datos generados por estudiantes universitarios, estos datos se pre-procesaron para posteriormente realizar su análisis. Los detalles se describen a continuación.

3.1 Datos.

Diecisiete estudiantes universitarios participaron voluntariamente en una campaña de recolección de datos firmando un consentimiento e instalando la versión gratuita de la aplicación *App Usage*. Los criterios de inclusión de los participantes fueron: i) ser mayor de edad, ii) contar con un teléfono inteligente con sistema operativo *Android* y iii) ser estudiante de licenciatura activo. La campaña tuvo una duración 10 semanas. Durante este periodo, se les solicitó a los participantes contestar encuestas para conocer su estado de ánimo aplicando la Escala de Autoevaluación para la Depresión de Zung [26] para diagnósticar el grado de ansiedad y depresión, y la Escala de Riesgo Suicida de Plutchik [19] para detectar las tendencias suicidas entre los participantes. La base de datos que se obtuvo consiste en una matriz de 17 columnas por 104,374 filas. Dentro de los diferentes tipos de datos que se obtuvieron se encuentran el historial de uso de las aplicaciones, historial de actividades, historial de consultas e historial de notificaciones

(ver Tabla 1). Para realizar la evaluación de la salud mental de los participantes, se contó con la participación de un psicólogo especialista en temas de salud mental en el entorno educativo. El psicólogo emitió un diagnóstico correspondiente a los niveles de ansiedad, depresión y riesgo suicida de cada participante. Dichos diagnósticos se obtuvieron a partir de las respuestas de los participantes a las encuestas que periódicamente se realizaron durante la campaña de recolección de datos.

Tabla 1: Tipos de datos recolectados durante la campaña de recolección de datos. Fuente: elaboración propia.

Tipo de dato	Ejemplo
Historial de uso	"App name","Usage time","Access count" "YouTube","0:47:31","2"
Historial de actividades	"App name","Date","Time","Duration" "Correo","2/3/20","11:00:52 AM","0:00:21"
Historial de consultas	"Date","Check phone count","Screen on time" "2/3/20","2","0:00:51"
Historial de notificaciones	"App name","Count" "Date","Count"

3.2 Limpieza de los datos.

La base de datos generada en la fase anterior de la metodología fue preparada para su análisis posterior. El pre-procesamiento de los mismos involucró las siguientes fases:

1. *Organización de datos.* Los archivos enviados por los participantes eran organizados en carpetas. Cada carpeta correspondía a un participante diferente y contenía todos sus archivos de datos.
2. *Lectura y agrupación de datos.* Se leyó el contenido de cada archivo de cada participante para agruparlo según su tipo de dato. De esta manera, se generó un solo archivo por tipo de dato que contenía información de todos los participantes.
3. *Procesamiento de contenido.* Se procesó el contenido de cada uno de los archivos que contenían los datos de todos los participantes por tipo de dato.
4. *Inclusión de etiquetas.* Las etiquetas, es decir, los datos que interesa clasificar, los cuales corresponden a los diagnósticos de ansiedad, depresión y riesgo de suicidio, se obtuvieron en un archivo separado. Por tanto, se incorporaron los diagnósticos de cada participante en los datos de cada tipo.
5. *Procesamiento de variables categóricas.* Las variables categóricas de las etiquetas fueron cambiadas a variables numéricas para facilitar el análisis. Parte de las variables categóricas fueron los nombres de las diferentes aplicaciones. Los nombres de 311 aplicaciones diferentes fueron buscadas en la *Play Store*® y según el tipo de aplicación fueron clasificadas en 25 categorías diferentes. En la Tabla 2 se pueden observar las categorías con algunos ejemplos.
6. *Eliminación de registros duplicados.* Se compararon todas las filas en la base de datos final para detectar que no hubiera duplicados. En caso de encontrar dos o más filas con exactamente la misma información en cada columna, se dejaba sólo la primera fila y el resto se eliminaba.

Tabla 2: Clasificación de las aplicaciones. Fuente: elaboración propia.

Categoría	Nombre de categoría	Nombre App	Categoría	Nombre de categoría	Nombre App
1	Juegos	Candy Crush Saga, Pokémon GO, WoodShop, Basta, Sopa de Letras.	14	Estilo de vida	Bodas.com.mx
2	Cómics	Manga PLUS.	15	Medios y Video	Vídeo, Video, Editor de video, Web Video Caster, Huawei Video
3	Compras	Mercado Libre, Wish, SHEIN, Amazon compras, Ofertia.	16	Música y audio	Spotify, Deezer, Grabadora, Música, Google Play Música, Dolby Audio
10	Finanzas	Banorte Móvil, BBVA México, Santander, Mercado Pago.	23	Clima	Hora y clima, Weather, Clima.
11	Fotografía	CameraMX, Fotos, AppGallery, Boomerang, Camera, Album.	24	Viajes y local	Maps, Notificación publicada (Maps).
12	Herramientas	Clean Master, Espejo, izzi, GoToolv2, Bluetooth, HiCare.	25	Sin categoría	Moto, Shell, Cyrus, Lite, Start, Funciones sugeridas.
13	Libros y referencias	Google Play Libros, Storytel.			

3.3 Análisis de datos.

La regresión logística (LR, *Logistic regression*, por sus siglas en inglés) es una técnica de clasificación y predicción de datos que se puede utilizar cuando las variables de salida son categóricas. Estas variables categóricas pueden ser variables binarias de decisión o variables con dos o más niveles [3]. A diferencia de otros enfoques multivariados, LR no requiere de supuestos como normalidad en los datos. Puede ser usada para evaluar la relación entre variables independientes y dependientes, siempre y cuando no existan restricciones para las variables explicativas o independientes [5]. Para tareas de clasificación, LR asigna las observaciones a un conjunto discreto de clases. Para ello, esta técnica de regresión utiliza la función de activación sigmoide con el fin de transformar datos de cualquier valor a un valor entre 0 y 1 [21]. Es decir, LR entrega la probabilidad de que una observación pertenezca a una clase o a otra. Un modelo de regresión estándar esta dado por la siguiente combinación lineal:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_M X_M$$

Donde $X_1, X_2, ..., X_M$ son variables independientes, $\beta_1, \beta_2, ..., \beta_M$ son los coeficientes de la regresión y β_0 es la intersección. Finalmente usando la función sigmoide, la probabilidad de pertenencia a determinada clase esta dada por:

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_M X_M)}}$$

Recientemente, la LR ha sido utilizada en diversas aplicaciones como clasificación [11] y asociación [6] [10] de genes, bio-informática [23], detección de enfermedades crónicas [16], procesamiento de señales cerebrales [18], evaluación de síntomas psicológicos [1], entre muchas otras.

Para el caso de variables de salida con múltiples niveles, el modelo se conoce como regresión logística multinomial (MLR) [8]. Suponga que Y es una variable aleatoria que puede tener un número finito de valores etiquetados de $1, 2, \ldots, L$. Sea

$$p_l = P(Y = l) \text{ y que } \sum_{l=1}^{L} p_l = 1.$$

El objetivo en MLR es encontrar un modelo que relacione las probabilidades p_l con las variables X_m. Similarmente,

$$\theta_l = \beta_{0l} + \beta_{1l}X_1 + \cdots + \beta_{Ml}X_M = \log \frac{p_l}{p_c}, l = 1, \ldots, L-1,$$

donde la categoría L se usa como referencia y

$$p_l = \frac{e^{\theta_l}}{1 + \sum_{l=1}^{L-1} e^{\theta_l}}, p_L = 1 - \sum_{l=1}^{L-1} p_l.$$

Cabe mencionar que no es relevante que categoría se elija como referencia; sin embargo, pueden existir categorías que resulten más convenientes para la interpretación de los resultados. El modelo MLR calcula para cada categoría un modelo de regresión lineal binario. El resultado son $L-1$ modelos de regresión lineal binarios. Cada modelo especifica el efecto de los predictores sobre la probabilidad de pertenecer a esa categoría, en comparación con la categoría de referencia. MLR el método de máxima verosimilitud para obtener la estimación de los parámetros.

Para valorar el desempeño del algoritmo, se utilizó la exactitud (Ec. 1) como métrica de evaluación [4]. La exactitud representa el porcentaje de elementos clasificados correctamente. La métrica se basa en el número de los verdaderos positivos (TP) y verdaderos negativos (TN), así como en el número de falsos positivos (FP) y falsos negativos (FN). Se calculan llevando a cabo el procedimiento llamado validación cruzada de K subconjuntos [22]. En este procedimiento el conjunto de datos es dividido en K grupos. De estos, K-1 se utilizan para entrenar al clasificador, y el restante se usa para evaluarlo. Este proceso es repetido K veces, entonces se promedian los resultados para obtener una métrica final. La ventaja de este método es que al contar con un conjunto de datos de entrenamiento limitado, es posible estimar el desempeño del clasificador sobre un conjunto de datos independiente [22].

$$\text{Ac} = \frac{\text{TP+TN}}{\text{TP+FN+FP+TN}} \tag{1}$$

4 Resultados

Los experimentos se ejecutaron en una computadora con sistema operativo macOS Catalina (versión 10.15.5), procesador 3.1 GHz 6-Core Intel Core i5 y memoria de 16 GB 2667 MHz. Se utilizó la librería *scikit-learn* de Python que incluye una gama amplia de algoritmos de aprendizaje automático. La base de datos limpia se dividió de manera aleatoria en dos: una parte para el entrenamiento y prueba (80%) del modelo y otra parte para su validación (20%). En la Tabla 3a se muestra la matriz de confusión de

la clasificación de diagnósticos para ansiedad y en la Tabla 3b para la depresión. Las etiquetas corresponden a valores de sin ansiedad (SA), con ansiedad (CA), sin depresión (SD), con depresión (CD) y diagnóstico dudoso (DD). Las filas indican las etiquetas de clasificación y las columnas los resultados de la predicción.

En el caso de *Sin ansiedad*, se clasificaron 7,917 elementos como SA, 170 como DD y cero como CA. Para la etiqueta *Diagnóstico dudoso* en asiedad, la clasificación arrojó 11,511 elementos como SA, cero como DD y tres como CA. Por último, para la etiqueta *Con ansiedad*, 1,274 elementos se clasificaron como SA, cero como DD y cero como CA. Estos resultados demuestran que el algoritmo no es capaz de clasificar aquellos diagnósticos de ansiedad etiquetados como *Diagnóstico dudoso* y *Con ansiedad*; sin embargo, el algoritmo clasifica correctamente la etiqueta *Sin ansiedad*. En cuanto a la depresión, los casos *Sin depresión* se clasificaron 572 elementos como SD y 11,079 como DD. Por último, los casos *Diagnóstico dudoso* se clasificaron 142 como SD y 9,082 como DD. Lo anterior indica que el algoritmo no es capaz de clasificar un gran número de elementos *Sin depresión*; pero con los elementos etiquetados como *Diagnóstico dudoso* en general los clasifica correctamente.

Tabla 3: Matriz de confusión.

(a) Ansiedad

	SA	DD	CA
Sin ansiedad	7917	170	0
Diagnóstico dudoso	11511	0	3
Con ansiedad	1274	0	0

(b) Depresión

	SD	DD	CD
Sin depresión	572	11079	-
Diagnóstico dudoso	142	9082	-
Con depresión	-	-	

Tabla 4: Resultados de precisión del algoritmo de regresión logística.

Variable	**Entrenamiento**	**Validación**
Ansiedad	0.936395	0.93418
Depresión	0.716512	0.715305

En la Tabla 4 se presentan los resultados de exactitud para la clasificación de diagnósticos de Ansiedad y Depresión. Para los datos de entrenamiento los porcentajes fueron alrededor del 94% y 72%, para los diagnósticos de ansiedad y depresió; mientras que para los datos de validación arrojaron un 93% y 71%, respectivamente. Cabe mencionar que para la clasificación del diagnóstico de riesgo suicida los datos presentaron una única clase *Sin riesgo*, por lo que la técnica no cubre este tipo de clasificación. Una solución es ampliar la base de datos con datos sintéticos que incluyan las demás clases o utilizar un clasificador uni-clase.

5 Conclusiones

La inteligencia artificial es un campo próspero con muchas aplicaciones prácticas y temas de investigación activos. El uso de técnicas de inteligencia artificial pueden ayudar a la detección temprana de los trastornos mentales como la depresión y la ansiedad. Si estos trastornos no se tratan adecuadamente, pueden convertirse en problemas serios de salud y causar gran sufrimiento. Además llegan a alterar actividades laborales, escolares y familiares. En el peor de los casos pueden llevar al suicidio. Alrededor de 800,000 personas en el mundo se suicidan cada año y es la segunda causa de muerte entre personas de 15 a 29 años [25]. Según datos del INEGI [2], en México ocurren aproximadamente 6 mil suicidios cada año, donde los jóvenes de 20 a 29 años son los que presentan las tasas más altas de suicidio. Coahuila es una de las entidades con más incidencia de suicidio en jóvenes en el país. Esto es relevante puesto que en la ciudad de Saltillo, Coahuila, el 80% de los estudiantes de nuevo ingreso a la Universidad, presentan algún tipo de problema psicológico [9]. Es importante concientizar a la sociedad sobre la importancia de estos temas, ya que los trastornos mentales pueden afectar de muchas maneras a la persona que lo padece y a la gente que la rodea. Cabe mencionar, que el análisis de datos de los teléfonos inteligentes, es una buena fuente para obtener información sobre los universitarios de forma oportuna, pues estos dispositivos los acompañan a lo largo de su día a día y su tiempo de uso puede ser un indicador de situaciones anormales por las que el estudiante esté atravesando, impactando su desempeño académico, físico y emocional, entre otros.

References

1. Descriptive and inferential statistics. In: Bellack, A.S., Hersen, M. (eds.) Comprehensive Clinical Psychology, pp. 243 – 285. Pergamon, Oxford (1998)
2. Las muertes de la tristeza: las cifras del suicidio en méxico (Oct 2019), https://www.excelsior.com.mx/nacional/las-muertes-de-la-tristeza-las-cifras-del-suicidio-en-mexico/1335558
3. Nonlinear regression models. In: Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C. (eds.) Encyclopedia of Bioinformatics and Computational Biology, pp. 731 – 737. Academic Press, Oxford (2019)
4. Abualigah, L.M., Khader, A.T., Hanandeh, E.S.: A new feature selection method to improve the document clustering using particle swarm optimization algorithm. Journal of Computational Science **25**, 456 – 466 (2018)
5. Aghaei, I., Sokhanvar, A.: Factors influencing sme owners' continuance intention in bangladesh: a logistic regression model. Eurasian Business Review (09 2019)
6. Cantor, R.M., Lange, K., Sinsheimer, J.S.: Prioritizing gwas results: a review of statistical methods and recommendations for their application. The American Journal of Human Genetics **86**(1), 6–22 (2010)
7. ENDUTIH: Encuesta nacional sobre disponibilidad y uso de tic en hogares (endutih) (2017), https://datos.gob.mx/busca/dataset/encuesta-nacional-sobre-disponibilidad-y-uso-de-tic-en-hogares-endutih/resource/bd89a4b7-d656-42ce-9d8f-662372b44de9
8. Faraway, J.J.: Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. CRC press (2016)

9. Gatica, P.: Con ideas suicidas 125 estudiantes de la uadec en saltillo, según encuesta (Feb 2020), https://vanguardia.com.mx/articulo/con-ideas-suicidas-125-estudiantes-de-la-uadec-en-saltillo-segun-encuesta?fbclid=IwAR1UPzf6WgWGC6NcaTv6foz6sUdPs9PO8V76zpLdWk01Ri4Y4scMD6bGejg
10. Jostins, L., McVean, G.: Trinculo: Bayesian and frequentist multinomial logistic regression for genome-wide association studies of multi-category phenotypes. Bioinformatics **32**(12), 1898–1900 (2016)
11. Liao, J., Chin, K.V.: Logistic regression for disease classification using microarray data: model selection in a large p and small n case. Bioinformatics **23**(15), 1945–1951 (2007)
12. Marten, W.d., Wilkerson, B.: Stress, work and mental health: a global perspective. Acta neuropsychiatrica **15**(1), 44–53 (2003)
13. Miranda, D., Calderón, M., Favela, J.: Anxiety detection using wearable monitoring. In: Proceedings of the 5th Mexican Conference on Human-Computer Interaction. pp. 34–41 (2014)
14. Miranda, D., Favela, J., Arnrich, B.: Detecting anxiety states when caring for people with dementia. Methods of information in medicine **56**(01), 55–62 (2017)
15. Mozos, O.M., Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N., Dobrescu, R., Ferrandez, J.M.: Stress detection using wearable physiological and sociometric sensors. International journal of neural systems **27**(02), 1650041 (2017)
16. Nusinovici, S., Tham, Y.C., Chak Yan, M.Y., Wei Ting, D.S., Li, J., Sabanayagam, C., Wong, T.Y., Cheng, C.Y.: Logistic regression was as good as machine learning for predicting major chronic diseases. Journal of Clinical Epidemiology **122**, 56 – 69 (2020)
17. O'Brien, J., Gallagher, P., Stow, D., Hammerla, N., Ploetz, T., Firbank, M., Ladha, C., Ladha, K., Jackson, D., McNaney, R., et al.: A study of wrist-worn activity measurement as a potential real-world biomarker for late-life depression. Psychological medicine **47**(1), 93–102 (2017)
18. Philiastides, M.G., Ratcliff, R., Sajda, P.: Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. Journal of Neuroscience **26**(35), 8965–8975 (2006)
19. Plutchik, R., Van Praag, H.M., Conte, H.R., Picard, S.: Correlates of suicide and violence risk 1: the suicide risk measure. Comprehensive psychiatry **30**(4), 296–302 (1989)
20. Reece, A.G., Danforth, C.M.: Instagram photos reveal predictive markers of depression. EPJ Data Science **6**(1), 1–12 (2017)
21. Rezapour, M., Mehrara Molan, A., Ksaibati, K.: Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models. International Journal of Transportation Science and Technology **9**(2), 89 – 99 (2020)
22. Rodríguez, J., Pérez, A., Lozano, J.: Sensitivity analysis of k-fold cross validation in prediction error estimation. Pattern Analysis and Machine Intelligence, IEEE Transactions on **32**, 569 – 575 (04 2010)
23. Tsuruoka, Y., McNaught, J., Tsujii, J.c., Ananiadou, S.: Learning string similarity measures for gene/protein name dictionary look-up using logistic regression. Bioinformatics **23**(20), 2768–2774 (2007)
24. Whiteford, H.A., Degenhardt, L., Rehm, J., Baxter, A.J., Ferrari, A.J., Erskine, H.E., Charlson, F.J., Norman, R.E., Flaxman, A.D., Johns, N., et al.: Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. The lancet **382**(9904), 1575–1586 (2013)
25. WHO: Suicide data (Sep 2019), https://www.who.int/mental_health/prevention/suicide/suicideprevent/en/
26. Zung, W.W.: Zung self-rating depression scale and depression status inventory. In: Assessment of depression, pp. 221–231. Springer (1986)

31

Una propuesta de sello digital para los CFDI utilizando criptografía de curvas elípticas

DIEGO ABRAHAM OLVERA MENDOZA, FIDEL GONZÁLEZ GUTIÉRREZ Y ARTURO GONZÁLEZ GUTIÉRREZ

Una propuesta de sello digital para los CFDI utilizando criptografía de curvas elípticas*

Diego Abraham Olvera-Mendoza[1], Fidel González Gutiérrez[1], and Arturo González Gutiérrez[1]

Universidad Autónoma de Querétaro; Querétaro, Qro.; México
www.uaq.mx

Resumen Este trabajo presenta una propuesta de sello digital para los comprobantes fiscales digitales por internet (CFDI) utilizando criptografía de curvas elípticas (ECC). El prototipo computacional fue desarrollado en el lenguaje de programación funcional Mathematica® y se utilizó la curva elíptica *secp256k1*. Se hace un comparativo con respecto a la seguridad, espacios de almacenamiento y tiempos de procesamiento entre los algoritmos RSA y la propuesta con ECC, teniendo como resultado una mayor eficiencia al sellar un CFDI con ECC.

Keywords: CLQ · criptografía de curvas elípticas · CFDI · sello digital.

1. Introducción

En México toda empresa o persona física está obligado a emitir Comprobantes Fiscales Digitales por Internet (CFDI). Los CFDI dan evidencia de la compraventa de un producto o servicio. También permite a los contribuyentes realizar sus deducciones para el pago de impuestos. Los CFDI son sellados digitalmente para evitar falsificaciones o manipulación de datos. El sello se realiza utilizando una *clave privada* asociada a un *Certificado de Sello Digital* (CSD). Actualmente el criptosistema asimétrico RSA es utilizado por la Secretaria de Administración Tributaria (SAT) en la generación de los CSD. Su seguridad se basa en el *problema de la factorización de números enteros muy grandes*.

El criptosistema RSA [6] utiliza un par de claves: la *clave pública* para cifrar un mensaje y la *clave privada* para descifrarlo. Se puede utilizar el proceso de descifrado para generar una *firma digital* y el cifrado para verificar su autenticidad. Es importante aclarar que desde un punto de vista criptográfico no es correcto utilizar el termino sello digital, lo que realmente se genera al *sellar* es una *firma digital*.

Para firmar documentos electrónicos oficialmente es necesario solicitar al SAT la *e.firma*. La e.firma es un medio de autentificación que permite realizar más

* Agradecimiento al CONACyT por el apoyo brindado para la realización de este trabajo.

de 400 trámites en línea [9]. Esta conformada por un par de claves RSA. Generalmente la clave pública esta contenida en un certificado digital.

Un *certificado digital* es un documento electrónico que expide una *autoridad certificadora* (CA) y avala la asociación entre una entidad y su clave pública. Una CA es un tercero de confianza con las tareas de: expedir certificados digitales, mantenerlos accesibles en un directorio público y revocarlos al terminar su periodo de validez o a petición del propietario. En México el SAT funge como CA y expide certificados digitales para la e.firma y el CSD.

Como parte de la información contenida en un *certificado digital* se encuentran: el algoritmo de firma utilizado, la función hash utilizada, el emisor del certificado, el periodo de validez del certificado, nombre de la entidad a quien se expide, la clave pública de la entidad y la firma de la CA. El estándar *X.509* es utilizado actualmente por el SAT para la generación de los certificados.

Para que un contribuyente pueda *sellar* un CFDI, el SAT genera un segundo par de claves distintas a la *e.firma*, el CSD. El CSD contiene una nueva clave pública. Su correspondiente clave privada es solo conocida por el contribuyente. La clave privada del CSD es la que se utiliza para sellar el *hash* de una cadena original. La *cadena original* de un CFDI se genera de acuerdo a las especificaciones que marca la Resolución Miscelánea Fiscal [7]. Se forma separando algunos campos del CFDI con el caracter *pipe* (|). Para ser sellada, toda cadena original es procesada por una función de *digestión* (mejor conocida como *Hash*) que mapea una entrada de longitud variable a un valor de longitud fija. Así, la mayoría de los contribuyentes tiene en su poder al menos dos pares de claves.

2. Antecedentes

Algunos autores [3], [1] han señalado graves fallos de seguridad en la infraestructura de clave pública (PKI) que soporta los CFDI en México. Entre los más graves están el uso de estándares obsoletos (funciones Hash MD5 o SHA1), la falta de una *lista de revocación de certificados* y el uso de tamaños de clave no recomendados.

Chávez y Henríquez en [1] sugieren el uso de *criptositstemas con curvas elípticas* (ECC) como reemplazo del cifrado RSA para proteger los CFDI. Los ECC [5] basan su seguridad en el *problema del logaritmo discreto sobre una curva elíptica*. Los ECC también son criptosistemas asimétricos, por lo tanto utilizan una clave privada (para descifrar y generar firmas digitales) y una clave pública (para cifrar y verificar firmas).

3. Objetivos

Este trabajo tiene como objetivo presentar una propuesta para generar sellos digitales de CFDI basados en ECC, a través del desarrollo de un prototipo computacional en Mathematica®.

4. Metodología

Se utilizó el lenguaje de programación funcional y simbólico *Mathematica® versión 12.0*, por contar con funciones criptográficas para generar claves y firmas digitales [2]. Mathematica® se ejecutó en una laptop Dell Inspiron 7537 con procesador Intel i7-4500U a 1.8 GHz, 8 GB de memoria RAM y sistema operativo Windows 10 Home de 64 bits.

El SAT a través de su página web proporciona facturas ficticias para pruebas. Una de ellas fue utilizada para obtener la cadena original mostrada a continuación:

||3.0|2010-03-06T20:38:12|ing|reso|PAGO EN UNA SOLA EXHIBICION|488.50|488.50|PPL961114GZ1|PHARMA PLUS SA DE CV|AV. RIO MIXCOAC|No. 140|ACACIAS—BENITO JUAREZ|MEXICO, D.F.|Mexico|03240|AV. UNIVERSIDAD|1858|OXTOPULCO|DISTRITO FEDERAL|Mexico|03910|PEPJ8001019Q8|JUAN PEREZ PEREZ|AV UNIVERSIDAD|16 EDF 3|DPTO 101|COPILCO UNIVERSIDAD|COYOACAN|DISTRITO FEDERAL|Mexico|04360|1.0|CAPSULAS|VIBRAMICINA 100MG 10|244.00|244.00|1.0|BOTELLA|CLORUTO 500M|137.93|137.93|1.0|TABLETAS|SEDEPRON 250MG 10|84.50|84.50|IVA|0.00|0.00|IVA|16.00|22.07|22.07||

A esta cadena original se le aplicó una función Hash256, la cual devuelve el valor numérico:

21159586336713373872137846405861447245017863645276105873016368938895716757977.

Después el procedimiento consistió en la creación de diez pares de claves RSA. Se utilizó cada una de las claves privadas para generar diez firmas digitales usando el hash de la cadena original mostrado anteriormente. La longitud del valor n utilizado en la operación modular fue de 2048 bits. Posteriormente cada firma se verificó utilizando su respectiva clave pública.

Finalmente se realizó el mismo procedimiento utilizando ECC. Se crearon diez pares de claves con la curva spec256k1 y diez firmas digitales. Cada una de las firmas generadas por el ECC fue verificada.

5. Avances en el trabajo de tesis

El presente trabajo muestra la comparación entre el criptosistema RSA y ECC al generar sellos digitales para un CFDI. De cada criptosistemas se utilizaron diez pares de claves para firmar y verificar el hash mostrado en la meto-

dología.

La Tabla 1 muestra que nuestra propuesta tiene un mayor nivel de seguridad. Los MIPS-año son una medida estándar en criptografía para cuantificar el esfuerzo computacional. Un MIPS-año es la cantidad de operaciones realizadas en un año por una computadora que ejecuta 1 millón de operaciones por segundo (1 MIPS). Un MIPS-año equivale aproximadamente a 2^{45} operaciones [4]. Un ataque por fuerza bruta requeriría 2^{128} operaciones para romper un nuestro sistema propuesto con ECC, por lo que un adversario tardaría aproximadamente $9{,}6 \times 10^{24}$ MIPS-años o $7{,}8 \times 10^{14}$ veces la edad del universo (14 billones de años). Mientras que para el sistema RSA el ataque aproximadamente tomaría $1{,}4 \times 10^{20}$ MIPS-año, por lo tanto el sistema ECC sería más resistente ya que se requerirían realizar 2^{16} veces 2^{112} operaciones.

También el espacio de almacenamiento ocupado por las claves y las firmas digitales es menor en comparación con RSA.

Una diferencia entre la firma que generan los ECC y la firma generada por RSA es que la primera está formada por la pareja (r, s) mientras que la segunda esta formada por un único valor (generalmente conocido como s).

Tabla 1: Nivel de seguridad y espacios de almacenamiento de los criptosistemas RSA y ECC usados.

Criptosistema	Nivel de seguridad	Clave Pública (bits)	Clave Privada (bits)	Firma Digital (bytes)
RSA-2048	112	17	2048	s=256
ECC(secp256k1)	128	256	512	r=32, s=32

El tiempo promedio de procesamiento de los algoritmos para generar el par de claves y el tiempo promedio para generar y verificar la firma digital usando ECC fueron notablemente menores respecto al método RSA (ver Tabla 2).

Tabla 2: Tiempo promedio de computo en Mathematica® para generar claves, firmas digitales y verificarlas.

Criptosistema	Generación de claves (s)	Generación de firma digital (s)	Verificación de firma digital (s)
RSA-2048	0.934375	0.81875	0.8
ECC(secp256k1)	0.0328125	0.21875	0.0015

Nuestra propuesta de sello digital puede verse en la Tabla 3. Se observa a simple vista la diferencia de tamaño entre el sello actual (firma digital) y el propuesto

al codificarlo en Base64.

Tabla 3: Firma Digital codificada en caracteres imprimibles Base64 para ser usada como sello digital en CFDI

RSA-2048	**ECC (spec256k1)**
mOrx99ocpD9dSCiFforIbM7M9CHUbp6 SNA8q+SvRjviK0q/PXscUcBklkSEk11 1sXkyKup1JkG7qvf0T41ngLNcnCln5US kwe2AKAojbvrYYR1q5JY+F/n9kbb5+ xu/Ajy7VngX1VaswQwfGolNR/9drm4lk ZklrUXBsjYLx191iA/CzY2r15MZw2Hn V/0WGOPucMHC1H3+ootqLAYVUg/l ajdIfBbe0QVf6HF5aY2Q8Fg3hWuwctPs EUlyWRpOe5nOjiQivDYeMfGumJsoBF Ta4WjPfJaGNFGhwYaYwaMxTdK1TJ wNSB01BvPOWbT23XgF9BmxDvHUC lVlJzC4Vlw==	BVah0V2AsAIRsHHT2qZqoPBO4MbEF AuxgxgxaDZHvLI=HRqakkgYCViqT2K Xba7xN6CannQidJrfEsGMk4WL/mU=

Actualmente se está definiendo una arquitectura de red utilizando tarjetas Raspberry Pi para medir el ancho de banda utilizado por un CFDI con este nuevo sello.

6. Resultados esperados

Como Vanstone ha señalado [8], los tiempos de cómputo y almacenamiento de claves y firmas digitales utilizando ECC son menores a los obtenidos con RSA. Nuestros resultados preliminares (ver Tabla 1 y 2), muestran que ECC es una mejor alternativa a RSA en cuanto a seguridad, espacio de almacenamiento y tiempo de cómputo. Así mismo, se espera que los CFDI basados en ECC utilicen un menor ancho de banda al ser transferidos por una red de computadoras.

7. Conclusiones

La principal ventaja que tienen los ECC sobre RSA es que utilizan claves de menor tamaño que ofrecen mayor seguridad. Los algoritmos implementados por Mathematica® obtuvieron un menor tiempo en la generacion de claves y firmas digitales con ECC. Presentamos como obtener un sello basado en ECC para un CFDI. Notamos que la PKI para los CFDI del SAT aún tiene graves fallos de seguridad y eficiencia. Revisando el proceso de sellado no encontramos la utilidad del *certificado de sello digital*. El utilizarlo solo genera al contribuyente un mayor número de claves que administrar.

Referencias

1. Chávez, León M., Henriquez, Francisco R. : Security vulnerabilities of the Mexican digital invoices by internet. In: 2015 International Conference on Computing Systems and Telematics (ICCSAT), pp.1–5, IEEE, (2015).

2. Cryptography-Wolfram Language Documentation, https://reference.wolfram.com/language/guide/Cryptography.html. Consultado el 13 Mar 2020
3. García, Vladimir G., Henríquez, Francisco R., Cortés, Nareli C.: On the Security of Mexican Digital Fiscal Documents. Computación y Sistemas **2**(5), 25–39 (2008)
4. Kaliski B.: MIPS-Year. In: van Tilborg H.C.A. (eds) Encyclopedia of Cryptography and Security. (2005)
5. Koblitz, N.: Elliptic Curve Cryptosystems. Mathematics in Computer Science **48**(177), 203–209 (1987)
6. Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM **21**(2), 120–126 (1978)
7. Servicio de Administración Tributaria: Anexo 20 de la Resolución Miscelánea Fiscal para 2017. Diario Oficial de la Federación, Ciudad de México (2017)
8. Vanstone, Scott A.: Elliptic curve cryptosystem—the answer to strong, fast public-key cryptography for securing constrained environments. Information Security Technical Report **2**(2), 78–87 (1997)
9. Blog e.firma, https://www.gob.mx/efirma/articulos/donde-puedo-usarla. Consultado el 29 Jul. 2020

32

Horizontal Fragmentation of Data Warehouses Using Decision Trees

NIDIA RODRÍGUEZ MAZAHUA, LISBETH RODRÍGUEZ MAZAHUA, ASDRÚBAL LÓPEZ CHAU AND GINER ALOR HERNÁNDEZ

Horizontal Fragmentation of Data Warehouses Using Decision Trees

Nidia Rodríguez Mazahua[1], Lisbeth Rodríguez Mazahua[1], Asdrúbal López Chau[2] and Giner Alor Hernández[1]

[1]Tecnológico Nacional de México/ IT Orizaba, División de Estudios de Posgrado e Investigación. Orizaba, Veracruz, México
[2]Universidad Autónoma del Estado de México, Centro Universitario UAEM Zumpango, Estado de México, México
nydirodriguez1@gmail.com, lrodriguezm@ito-depi.edu.mx, alchau@uaemex.mx, galor@ito-depi.edu.mx

Abstract. One of the main problems faced by Data Warehouse (DW) designers is fragmentation. Several studies have proposed data mining-based horizontal fragmentation methods, which focus on optimizing the query response time and execution cost to make the DW more efficient. However, to the best of our knowledge there not exist a horizontal fragmentation technique that uses a decision tree to carry out fragmentation. Given the importance of decision trees in classification, since they allow obtaining pure partitions (subsets of tuples) in a data set using measures such as Information Gain, Gain Ratio and the Gini Index, the aim of this work is to use decision trees in the DW fragmentation. For this, the requirements necessary to carry out horizontal fragmentation using decision trees will be determined, and the fragmentation method will be designed, which will consist of determining the most frequent OLAP (On-line Analytical Processing) queries, analyzing the predicates used by the queries, and based on this build the decision tree, from which the horizontal fragments will be generated. The method will be implemented and validated using a case study in tourism.

Keywords: CLQ, Fragmentation, Data warehouse, Decision Trees.

1 Introduction

A Data Warehouse is a theme-oriented, integrated, time variable, and non-volatile data collection in support of management's decision-making process. Data Warehousing provides architectures and tools for business executives to systematically organize, understand, and use the data to make strategic decisions [1].

On the other hand, fragmentation is a distributed database design technique that consists of dividing each database relation in smaller fragments and treating each fragment as an object in the database separately, there are three alternatives for that: horizontal, vertical, and hybrid fragmentation [2].

Decision tree classifiers are so popular because their construction does not require any domain knowledge or parameter setting, they can handle multidimensional data, the learning and classification steps of decision tree induction are simple and fast, and they have good accuracy [1]. Since horizontal fragmentation consists of dividing a database table into subsets of tuples and decision trees have been used to horizontally fragment relational database tables based on OLTP (On-line Transaction Processing) workloads [3], in this paper, we propose using a decision tree to horizontally fragment a Data Warehouse. This paper is made up of the following parts: Section 2 go through the related works on DW horizontal fragmentation, Section 3 set the research objective of work, in Section 4 the stages that conform the methodology to be followed are described, Section 5 reports the preliminary achievements in the work, and finally, the article is concluded in Section 6.

2 Related work

The approach proposed in [5] consisted of an incremental horizontal fragmentation technique of the DW through a web service. In [6], it was proposed a method called Chabok, which uses two-phase Map Reduce to solve DW problems with big data. As part of their ongoing work on workload–driven partitioning, Boissier and Kurzynski [7] implemented an approach called aggressive data skipping and extended it to handle both analytical and transactional access patterns. In [8] it was used linear programming to solve the NP-hard problem of determining a horizontal fragmentation scheme in relational DW. Also in [9], the authors proposed a graph-based database partitioning method called GPT that improves the performance of queries with less data redundancy. In [10] it was proposed a dynamic fragmentation strategy for OLAP cubes using association rule mining. Finally, in [11] it was presented a horizontal data partitioning approach tailored to a large DW, interrogated through a high number of queries, the idea was to fragment horizontally only the large fact table based on partitioning predicates, elected from the set of selection predicates used by analytic queries.

3 Research objective

The main objective of this work is to develop a horizontal fragmentation method that uses decision trees to obtain fragmentation schemes that achieve query optimization in a DW.

4 Methodology

The process to develop the fragmentation method which uses decision trees consists of four stages:

1. **Analysis.** In this stage, it is necessary to realize a comparative analysis of horizontal fragmentation methods and their cost models to know their advantages and disadvantages and select the most suitable for a comparison with the proposed method and cost model that will be performed in the evaluation stage.
2. **Design**. The decision tree-based horizontal fragmentation method for DW and the cost model to compare different horizontal fragmentation schemes will be designed in this stage, moreover, we will develop a tourism DW following the Kimball [12] methodology. This tourism DW will be useful for the evaluation stage. The tourism data will be integrated into the DW from different sources such as SECTUR (Secretaría de Turismo) and INEGI (Instituto Nacional de Estadística y Geografía).
3. **Implementation.** The proposed fragmentation method and cost model will be implemented in two data warehouses: the tourism DW and the SSB (Star Schema Benchmark).
4. **Evaluation.** The proposed method will be compared with the selected techniques of the comparative analysis to demonstrate the strengths of our method.

5 Achievements

The state of the art concerning to the horizontal fragmentation methods were analyzed to know their advantages and disadvantages, also the cost models used by methods was compared to find the most relevant costs considered by the authors when obtaining their fragmentation scheme in DW. 39 methods and 24 cost models formed the analysis. The horizontal fragmentation methods were analyzed by 6 criteria finding the following results: 1) Type of scheme used: the star scheme was the most multidimensional model used by methods; 2) Completeness: 34 methods presented the necessary information to implement them; 3) Validation: 30 techniques used benchmarks for their evaluation, the most used benchmarks were APB-1, APB-1 Release II, and TPC-H; 4) Implementation easiness: more than half of methods are classified as difficult to implement mainly because some of them focus on complex DW fragmentation, while others perform incremental (dynamic) fragmentation; 5) Cost model: 24 methods included a cost model; 6) Type and database management system used: the data model most used was the relational one. The most used DBMS was Oracle 11g. Whit respect to cost models, the criteria analyzed were costs considered and if they corresponded to local processing and/or communication. Almost all cost models only take into account the local processing cost, only [13] also considers the communication cost in its model.

6 Expected results

The expected results for this work are mainly to develop a horizontal fragmentation method that uses decision trees to obtain a partitioning scheme which optimizes the response time and the execution cost of the OLAP queries in a DW, to validate the method it will be applied in a case study of tourism in México.

7 Conclusions

In this work, a horizontal fragmentation method is being developed to take advantage of the potential of decision trees in the classification to adapt them in the process of DW horizontal fragmentation. Until now, interesting results have been found which are leading the investigation.

References

[1] J. Han, M. Kamber, y J. Pei, *Data Mining Concepts and Techniques*. USA: Morgan Kaufmann Publishers., 2012.

[2] M. T. Ozsu y P. Valduriez, *Principles of Distributed Database Systems*, Fourth edition. Cham, Switzerland: Springer Nature Switzerland AG, 2020.

[3] C. Curino, E. Jones, Y. Zhang, y S. Madden, "Schism: a workload-driven approach to database replication and partitioning", *Proc. VLDB Endow.*, vol. 3, núm. 1–2, pp. 48–57, sep. 2010, doi: 10.14778/1920841.1920853.

[4] R. C. Mateus, T. L. L. Siqueira, V. C. Times, R. R. Ciferri, y C. D. de Aguiar Ciferri, "Spatial data warehouses and spatial OLAP come towards the cloud: design and performance", *Distrib. Parallel Databases*, vol. 34, núm. 3, pp. 425–461, sep. 2016.

[5] E. Abdelaziz y M. Ouzzif, "Web Service for Incremental and Automatic Data Warehouses Fragmentation", *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, núm. 6, pp. 1–10, 2017.

[6] M. Barkhordari y M. Niamanesh, "Chabok: a Map-Reduce based method to solve data warehouse problems", *J. Big Data*, vol. 5, núm. 40, pp. 1–25, dic. 2018.

[7] M. Boissier y K. Daniel, "Workload-Driven Horizontal Partitioning and Pruning for Large HTAP Systems", en *2018 IEEE 34th International Conference on Data Engineering Workshops (ICDEW)*, Paris, abr. 2018, pp. 116–121.

[8] M. Barr, K. Boukhalfa, y K. Bouibede, "Bi-Objective Optimization Method for Horizontal Fragmentation Problem in Relational Data Warehouses as a Linear Programming Problem", *Appl. Artif. Intell.*, vol. 32, núm. 9–10, pp. 907–923, nov. 2018.

[9] Y.-M. Nam, M.-S. Kim, y D. Han, "A Graph-Based Database Partitioning Method for Parallel OLAP Query Processing", en *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, Paris, abr. 2018, pp. 1025–1036.

[10] K. Letrache, O. El Beggar, y M. Ramdani, "OLAP cube partitioning based on association rules method", *Appl. Intell.*, vol. 49, núm. 2, pp. 420–434, feb. 2019.

[11] M. Kechar y S. N. Bahloul, "Bringing Together Physical Design and Fast Querying of Large Data Warehouses: A New Data Partitioning Strategy", presentado en BDIoT'19: Proc. of the 4th Int. Conf. on Big Data and Internet of Things, Rabat Morocco, 2019.

[12] R. Kimball y M. Ross, *The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence*. Indianapolis, Indiana: John Wiley & Sons, Inc., 2016.

[13] Pedro Furtado, "Experimental Evidence on Partitioning in Parallel Data Warehouses", *Proc. 7th ACM Int. Workshop Data Warehous. OLAP*, pp. 23–30, 2004.

Índice de Autores

A

Abraham Efraim Rodriguez Mata, 204
Adan Hirales Carbajal, 104
Adrian Reyes, 220
Alberto Ochoa, 184
Alejandro Aranda Díaz, 160
Alejandro Flores Benitez, 95
Alexis Jair López, 248
Amaury Sanchez Negrin, 113
Ana Karen Hernandez, 248
Anabel Martinez-Vargas, 131
Antony Morales, 195
Argelia Faviola Miranda Pérez, 122
Armando Martinez Graciliano, 113
Arturo González Gutiérrez, 271
Asdrúbal López-Chau, 278

B

Baldwin Cortes, 195
Baruc de Los Santos Izquierdo, 48
Belgica Valencia Garcia, 104
Benjamin Oribe Mendieta, 69
Brenda L. Flores-Rios, 4, 13, 57

C

Claudia Guadalupe Gómez Santillán, 169, 177

D

Daniel Hernández-Balbuena, 57
Daniela López-García, 177
David Salvador González González, 30, 122, 140, 149
Diana Serrano-Delgado, 83
Diego Abraham Olvera-Mendoza, 271

E

Edgar O. Resendiz-Flores, 22
Eduardo Vazquez-Santacruz, 48
Elena Ortiz, 239
Elvia Elvira Morales Turrubiates, 48
Emmanuel López Hernández, 13
Enrique Reyes Archundia, 211
Esbeidy Gomez, 239

F

Fausto Antonio Balderas-Jaramillo, 177
Felix Calderon, 195
Felix F. González-Navarro, 4, 13, 57
Felix R. Saucedo-Zendejo, 22
Fernando Velázquez-López, 83
Fidel González Gutiérrez, 271
Flor Radilla López, 48

G

Giner Alor-Hernández, 278
Grisel Alvarez, 184

H

H. Patricia Martínez, 220
Héctor Fraire-Huacuja, 177
Héctor Joaquín Fraire Huacuja, 169
Hector Rodriguez, 195
Hugo Daniel González, 239
Hugo Gonzalez, 83

I

Ismael Molina Moreno, 211
Ivan Alejandro Encinas Monroy, 78

J

J. Federico Casco, 220
Javier Iracheta, 231
Jesús Alejandro Navarro-Acosta, 262
Jessica Beltran, 78
Jorge Antonio Lopez Renteria, 113
Jorge Eduardo Ibarra-Esquer, 57
Jorge Saúl Rodríguez Vázquez, 140
José Antonio Gutiérrez Gnecchi, 211
José Martín Molina Espinosa, 95
Jose Cedeño, 195
Josue Ezequiel Juarez Manjarrez, 257
Juan Carlos Olivares Rojas, 211
Juan Jose Flores, 195
Juan Pablo García-Vázquez, 131
Julio César Sandria, 239

L

Laura Cruz-Reyes, 169, 177
Lisbeth Rodríguez-Mazahua, 278
Luis A. Castro, 78
Luis Enrique Ruiz Martinez, 211
Luis Héctor Sánchez Gámez, 78
Luis Lizama, 69

M

Ma. Cosío-León, 131
Manuel Alvarez Madrigal, 95
Manuel Mejía Lavalle, 160
María Lucia Barrón Estrada, 39
María Lucila Morales Rodríguez, 169
María Lucila Morales-Rodríguez, 177
Marcial Sandoval-Gastelum, 4
Miguel Angel García Morales, 169
Misael Lopez-Sanchez, 131
Modesto Medina Melendrez, 204

N

Nelson Rangel-Valdez, 169, 177
Nidia Rodríguez-Mazahua, 278
Nohe Ramon Cazarez-Castro, 113

P

Perfecto Malaquías Quintero Flores, 248

R

Rafael Llamas-Contreras, 83
Ramón Zatarain Cabada, 39
Regina González Gómez, 262
Roberto Morales, 220
Rocio Natali Carranza Santoy, 149
Rolando J. Praga Alejo, 30
Rolando Javier Praga Alejo, 140, 149

S

Samuel Sandoval, 184
Sara Garza, 231
Selene Lilette Cardenas Maciel, 113

U

Ulises Uriel Cabrera Rubio, 39

V

Valeria Soto-Mendoza, 262
Victor Alejandro Gonzalez Huitron, 257
Victor-Alejandro Gonzalez-Huitron, 204

W

Wendy Elizabeth Corpus Rojas, 30

Y

Y. Guadalupe Lazaro, 220
Yaneth Guadalupe Robles de León, 122