

# **Advances in Intelligent Technologies and its Applications**

---

# Research in Computing Science

---

## Series Editorial Board

### Editors-in-Chief:

*Grigori Sidorov (Mexico)*  
*Gerhard Ritter (USA)*  
*Jean Serra (France)*  
*Ulises Cortés (Spain)*

### Associate Editors:

*Jesús Angulo (France)*  
*Jihad El-Sana (Israel)*  
*Alexander Gelbukh (Mexico)*  
*Ioannis Kakadiaris (USA)*  
*Petros Maragos (Greece)*  
*Julian Padget (UK)*  
*Mateo Valero (Spain)*

### Editorial Coordination:

*Alejandra Ramos Porras*

*Research in Computing Science* es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 128**, noviembre 2016. Tiraje: 500 ejemplares. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No. : 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Impreso en la Ciudad de México, en los Talleres Gráficos del IPN – Dirección de Publicaciones, Tres Guerras 27, Centro Histórico, México, D.F. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

**Editor responsable:** *Grigori Sidorov, RFC SIGR651028L69*

**Research in Computing Science** is published by the Center for Computing Research of IPN. **Volume 128**, november 2016. Printing 500. The authors are responsible for the contents of their articles. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research. Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

---

Volume 128

---

# **Advances in Intelligent Technologies and its Applications**

**Efren Mezura Montes (ed.)**



Instituto Politécnico Nacional, Centro de Investigación en Computación  
México 2016

**ISSN: 1870-4069**

---

Copyright © Instituto Politécnico Nacional 2016

Instituto Politécnico Nacional (IPN)  
Centro de Investigación en Computación (CIC)  
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal  
Unidad Profesional “Adolfo López Mateos”, Zacatenco  
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX and Periodica / Indexada en LATINDEX y Periódica

Printing: 500 / Tiraje: 500

Printed in Mexico / Impreso en México

## Editorial

This volume of the journal “Research in Computing Science” contains selected papers on intelligent and information technologies, which is an interdisciplinary area where several fields converge, such as machine learning, data mining, genetic algorithms, artificial vision, virtual learning environments, ontologies, mobile computing, natural language and text processing. The combination and cooperation between these fields has been the key of advance of Information Technologies.

The papers in this volume have been carefully chosen by the Editorial Board based on evaluation by at least two members of the reviewing committee. The main criteria for the selection were originality and technical quality of the papers. Submission, reviewing, and selection process was supported free of charge by the EasyChair system, [www.EasyChair.org](http://www.EasyChair.org).

This volume contains 15 papers related to various aspects of the development and applications of intelligent and information technologies, organized in the following sections:

- Ontologies,
- Genetic algorithms,
- Natural language and text processing,
- Machine learning and data mining,
- Applications development for mobile devices,
- Virtual learning environments,
- Artificial vision.

This volume is the result of hard work and collaboration of many people. First of all, we thank the authors of the papers included in this volume for their technical excellence, which made possible the high quality of this volume. We also thank the members of the Editorial Board and the reviewing committee for their hard work on selection of the best papers out of the forty submissions we received.

*Efrén Mezura Montes*  
Artificial Intelligence Research Center  
University of Veracruz, Mexico  
Guest Editor



## Table of Contents

	Page
<b>Implementación paralela de un algoritmo genético para el problema del agente viajero usando OpenMP .....</b>	<b>9</b>
<i>Edgar Martínez Vargas, Marcela Rivera Martínez, Luis René Marcial Castillo, Lourdes Sandoval Solis</i>	
<b>Método criptográfico simétrico utilizando teoría del caos, operaciones sobre ADN y raíces de funciones no lineales .....</b>	<b>21</b>
<i>Luis René Marcial Castillo, Erika Leonor Basurto Munguía, Marcela Rivera Martínez, María de Lourdes Sandoval Solis</i>	
<b>Introducción al problema inverso electrocardiográfico .....</b>	<b>35</b>
<i>Esteban Herrera Hernandez, Rafael Lemuz López, Carlos Guillén Galván, Ángel Ramos del Olmo</i>	
<b>Propuesta de arquitectura para un módulo de inteligencia de negocios basado en minería de datos .....</b>	<b>47</b>
<i>Yessica Thalia Apale Lara, Beatriz Alejandra Olivares Zepahua, Lisbeth Rodríguez Mazahua, Giner Alor Hernandez, Hilarión Muñoz Contreras</i>	
<b>Prototipo de un oxímetro de pulso con ESP8266 Wi-Fi .....</b>	<b>57</b>
<i>Gabriel Contreras Mota, Rafael Lemuz López, Carlos Guillén Galván, Blanca Bermúdez Juárez</i>	
<b>Aplicación multimedia para el entrenamiento en la certificación TOEFL mediante reconocimiento de voz .....</b>	<b>67</b>
<i>Miguel Hernández Ramos, Rafael Lemuz López</i>	
<b>Diseño de una ontología para el proceso de evaluación de las asignaturas técnico-científicas del Instituto Tecnológico de Orizaba .....</b>	<b>77</b>
<i>I. Colohua Cruz, L. A. Reyes Hernández, G. Hernández Chan, J. L. Sánchez-Cervantes</i>	
<b>Un algoritmo para detectar la polaridad de opiniones en los dominios de laptops y restaurantes.....</b>	<b>91</b>
<i>Karen L. Vazquez, Mireya Tovar, Darnes Vilariño, Beatriz Beltrán</i>	
<b>Avances en el desarrollo de un clasificador de imágenes termográficas de planta del pie diabético basado en una red neuronal de retropropagación .....</b>	<b>99</b>
<i>José Federico Ramírez Cruz, Edmundo Bonilla Huerta, Lauro Reyes Cocoltzi, José Crispín Hernández Hernández</i>	

<b>Evaluación de una ontología para el intercambio de información materno-infantil entre instituciones de salud en México</b> .....	109
<i>E. V. López Hernandez, Beatriz A. Olivares Zepahua, G. S. Hernández Chan, José L. Sánchez Cervantes</i>	
<b>Identificación de lengua de señas mediante técnicas de procesamiento de imágenes</b> .....	121
<i>Manuel Martínez Gutiérrez, José Rafael Rojano-Cáceres, Ismael Everardo Bárcenas Patiño, Fredy Juárez Pérez</i>	
<b>Breve revisión sobre el reconocimiento de la Lengua de Señas</b> .....	131
<i>Héctor Eduardo Sánchez Barrera, José Rafael Rojano-Cáceres, Guillermo Molero Castillo, Etelvina Archundia Sierra</i>	
<b>Sistema de planificación de trayectorias utilizando visión artificial</b> .....	141
<i>Jorge Cerezo-Sánchez, Griselda Saldaña-González, Mario M. Bustillo-Díaz, Apolonio Ata-Pérez</i>	
<b>RENTOL: Un algoritmo de agrupamiento basado en K-means</b> .....	149
<i>Erendira Rendon Lara, Itzel María Abundez Barrera</i>	
<b>Multi-sensory Tracking of Elders in Outdoor Environments on Ambient Assisted Living</b> .....	159
<i>Javier Jiménez Alemán</i>	

## **Implementación paralela de un algoritmo genético para el problema del agente viajero usando OpenMP**

Edgar Martínez Vargas, Marcela Rivera Martínez, Luis René Marcial Castillo,  
Lourdes Sandoval Solís

Benemérita Universidad Autónoma de Puebla,  
Facultad de Ciencias de la Computación, Puebla,  
México

edgar-93@outlook.com, {cmr, lmc}@solarium.cs.buap.mx  
maria.sandoval@correo.buap.mx

**Resumen.** En este trabajo se aborda el problema del agente viajero, resolviéndolo mediante la aplicación paralela de un algoritmo genético, implementado en lenguaje C, utilizando las directivas de OpenMP. Dichas soluciones al problema consisten en encontrar el camino más corto visitando todas las ciudades sin repetir alguna y con el menor costo posible. Las pruebas se llevaron a cabo en el Laboratorio Nacional de Supercómputo del Sureste de México y los resultados se comparan con problemas prueba reportados en la literatura.

**Palabras clave:** Algoritmos genéticos paralelos, agente viajero, OpenMP, lenguaje C.

## **Parallel Implementation of a Genetic Algorithm for the Traveling Salesman Problem Using OpenMP**

**Abstract.** In this paper explain the traveling salesman problem, solving it through the parallel application of a genetic algorithm, implemented in C language, using OpenMP directives. These solutions to the problem consists to find the shortest route visiting all the cities without repeating some and with the lowest possible cost. The tests were performed at “Laboratorio Nacional de Supercómputo del Sureste de México” and the results are compared with benchmark problems reported in the literature.

**Keywords:** Parallel genetic algorithms, traveling salesman problem, OpenMP, C language.

## **1. Introducción**

El problema del agente viajero es ampliamente estudiado ya que se considera como un problema difícil de resolver, denominándose en un lenguaje computacional como NP completo [1], es decir, es un problema para el que no se puede garantizar que se encontrará la mejor solución en un tiempo de cómputo razonable.

Dicho problema ha sido planteado a través de diversas técnicas, entre las que predomina la aplicación de algoritmos genéticos [2], este tipo de algoritmos son muy eficaces para problemas complejos de búsqueda y optimización, dado que se pueden adaptar de forma paralela para trabajos en los que se requiera manejar grandes volúmenes de datos. Algunos artículos que proponen soluciones al problema del agente viajero mediante algoritmos genéticos paralelos son los siguientes: GRISLAS [3], un algoritmo genético paralelo que combina los modelos de grillas e islas para encontrar buenas soluciones; y el modelo PPREGA (Parallel Pattern Reduction Enhanced Genetic Algorithm) [4], es un método escalable para reducir el tiempo de cálculo de los algoritmos genéticos.

También ha sido enfrentado mediante el algoritmo de recocido simulado [5], el objetivo general de este tipo de algoritmo es encontrar una buena aproximación al valor óptimo de una función en un espacio de búsqueda grande.

Otra técnica propuesta para resolver el problema es aplicando el algoritmo de colonia de hormigas [6], básicamente, este algoritmo menciona que las hormigas son capaces de seguir la ruta más corta en su camino de ida y vuelta a la colonia. Al desplazarse cada una va dejando un rastro de una sustancia química llamada feromona a lo largo del camino seguido, "transmitiéndose información" entre ellas.

En este trabajo se propone usar un algoritmo genético paralelo para encontrar soluciones a dicho problema, el cual se implementó en C utilizando las directivas de OpenMP [7] para paralelizarlo. Además, el propósito es reducir los tiempos de respuesta y encontrar soluciones buenas al aplicarlo en grandes problemas.

En la siguiente sección se presentan algunas notas preliminares sobre los orígenes y un diagrama general de los algoritmos genéticos, además de la definición del problema, en la sección 3 se explica brevemente el paradigma sobre los algoritmos genéticos paralelos, la sección 4 muestra y detalla el algoritmo propuesto, la sección 5 describe los experimentos y resultados de los mismos, así como del "speedup" y escalabilidad del algoritmo paralelo. La sección número 6 presenta las conclusiones y finalmente se listan las referencias utilizadas en el desarrollo de este trabajo.

## **2. Preliminares**

### **2.1. Algoritmos genéticos**

Los algoritmos genéticos fueron introducidos por John Holland a finales de los 60's inspirándose en el proceso observado en la evolución natural de los seres vivos. Son algoritmos de búsqueda basados en la mecánica de la selección natural y en la genética. Estos combinan la supervivencia de los individuos más aptos entre las cadenas de estructuras con un intercambio aleatorio para formar un algoritmo de búsqueda [8]. Los

algoritmos genéticos están determinados por tres características fundamentales: selección de parejas, cruce y mutación. La figura 1 presenta el esquema general de un algoritmo genético.

## 2.2. Definición del problema

El problema del agente viajero está dado de la siguiente forma: Sean  $n$  ciudades de un territorio, se especifica una ciudad  $k$ , que será el inicio y fin del recorrido, el objetivo es encontrar una ruta que pase una sola vez por cada una de las ciudades y minimice el costo realizado por el viajero, es decir, se debe minimizar:

$$\min \left\{ z = \sum_{i=1}^n \sum_{j=1}^n C_{ij} x_{ij} \right\}, \quad (1)$$

donde  $C_{ij}$  es el costo asociado de viajar de la ciudad  $i$  a la ciudad  $j$ , y  $x_{ij}$  es la variable de decisión de visitar la ciudad  $j$  después de visitar la ciudad  $i$ . La variable de decisión es 1 en caso de que se visite la ciudad  $j$  después de visitar la ciudad  $i$  y 0 en caso contrario.

## 2.3. Algoritmo genético para el problema del agente viajero

En la figura 1 se presenta el diagrama de flujo de un algoritmo genético secuencial, el cual es la base para el algoritmo paralelo expuesto posteriormente.

Se programó un algoritmo genético secuencial para resolver el problema del agente viajero, de acuerdo al diagrama de la figura 1. Enseguida se describe brevemente dicha implementación.

La función de aptitud a utilizar es la dada en la ecuación (1), y cada cromosoma es una permutación de valores enteros entre 1 y la cantidad de nodos  $n$ . Por ejemplo, una representación posible dado 5 ciudades es: (1, 2, 3, 4, 5).

**Generar la población inicial:** En este caso se hizo de manera aleatoria.

**Seleccionar parejas:** La selección de parejas para este problema en especial es ruleta por costo, la cual consiste en crear una ruleta en la que cada cromosoma tiene asignada una fracción proporcional, dicha fracción proporcional es con respecto al costo de la función de aptitud.

**Cruza:** Para el desarrollo del problema se considera la reproducción o cruce por medio de 2 puntos esto quiere decir que se toman 2 puntos distintos en cada padre, y se intercambian los alelos entre los padres para generar el mismo número de individuos, tomando en cuenta los valores antes del punto de cruce del primer padre, los valores medios del segundo padre y los valores finales del primer padre, para la generación de uno de los hijos, para el otro es inverso. Dado que para el problema del agente viajero no debe haber ciudades repetidas, se utiliza la cruce denominada PMX (Partially Mapped Crossover) [9], donde para los valores que faltan en los hijos, se copian los valores de los padres de la siguiente forma: Si el valor a copiar no está en la subcadena,

copiarlo, en caso contrario, se copia por el valor que tenga dicha subcadena en el otro padre.

**Mutación:** Busca cambiar aleatoriamente la información genética. La mutación que se utiliza es mutación por inversión, en ésta se generan dos números diferentes aleatorios entre 1 y el tamaño del cromosoma, esta subcadena se invierte posteriormente.

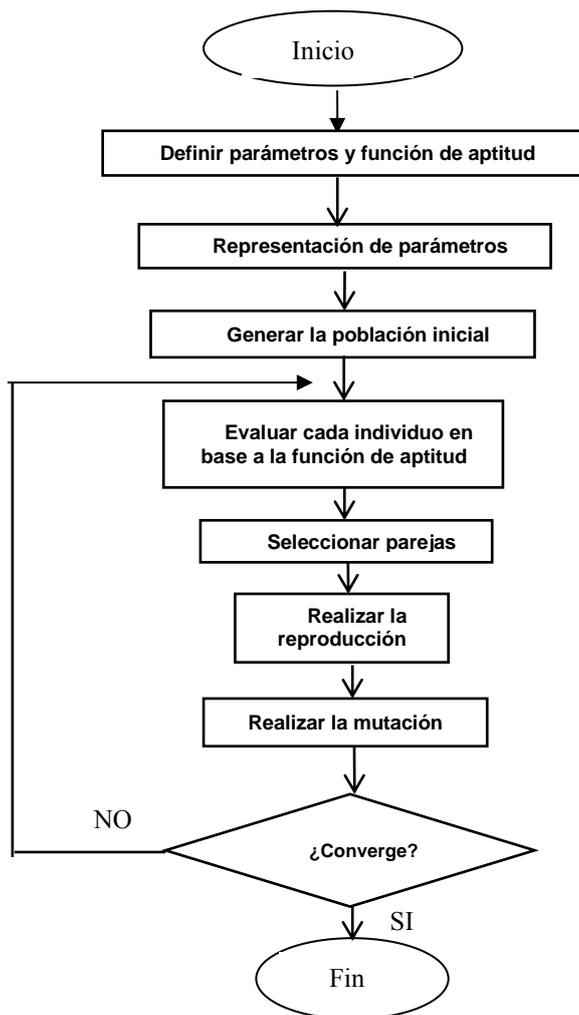


Fig. 1. Diagrama del funcionamiento de un algoritmo genético.

### 3. Algoritmos genéticos paralelos

Las técnicas de programación de alto rendimiento (performance), más comúnmente llamadas técnicas de procesamiento paralelo o técnicas de paralelismo, constituyen un

conjunto de métodos computacionales utilizados para la resolución de problemas sobre sistemas con una capacidad de procesamiento superior al tradicional modelo de computadora de Von Neumann. Estos sistemas, conocidos como multiprocesadores, permiten abordar la resolución computacional de problemas demasiado complicados de acuerdo a la complejidad computacional inherente al propio problema, de acuerdo a la dimensión de las instancias consideradas, o debido a que trabajan sobre un enorme volumen de datos que serían virtualmente imposibles de resolver mediante el paradigma tradicional. La característica principal de los sistemas multiprocesadores consiste en disponer de un conjunto de unidades de procesamiento interconectadas por algún medio que posibilita la comunicación de datos y de control entre ellos [10].

En la actualidad, el continuo aumento en el tamaño de los problemas ha encaminado a proponer distintas opciones a los enfoques tradicionales de programación para la resolución de problemas complejos. En este contexto, varias heurísticas se han aplicado a los problemas relacionados con búsqueda y optimización. Las técnicas de computación evolutiva, y los algoritmos genéticos en particular, se han manifestado como métodos propicios, ya que permiten un diseño paralelizable fácilmente, para resolver los problemas complejos, tal como es el problema del agente viajero.

### 3.1. Clasificación

A continuación se presentan las 3 maneras básicas para clasificar o paralelizar un algoritmo genético.

1. *Maestro-esclavo*: Estos algoritmos trabajan con una única población de individuos que será gestionada por el nodo maestro. La evaluación de la función de los individuos y/o la aplicación de los operadores genéticos puede ser realizada por los nodos esclavos.
2. *Grano fino*: Este tipo de algoritmos han sido diseñados para ser implementados usando computadores masivamente paralelos. Ahora, la población se encuentra dividida espacialmente entre los distintos procesadores e, idealmente, cada procesador debería albergar un único individuo. El cruce y la selección de individuos se hará entre individuos que pertenezcan a un mismo vecindario.
3. *Grano grueso*: La característica más importante de estos algoritmos es el uso de múltiples poblaciones y la inclusión de un nuevo operador denominado migración; la migración de individuos se debe aplicar entre las distintas poblaciones. Dado que cada una de las poblaciones evolucionan independientemente, el porcentaje de migración será muy importante de cara a obtener resultados satisfactorios.

### 3.2. Paradigmas

A su vez, este tipo de clasificación se puede llevar a cabo siguiendo tres estilos distintos de programación:

1. *Paralelismo en datos*: El compilador se encarga de la distribución de los datos guiado por un conjunto de directivas que introduce el programador. Para este

modelo de paralelismo destacan los lenguajes HPF (High Performance Fortran) y OpenMP.

2. *Programación por paso de mensajes*: Es el método más utilizado para programar sistemas de memoria distribuida. La forma básica consiste en que los procesos coordinan sus actividades mediante el envío y la recepción de mensajes. Las librerías más utilizadas son el estándar MPI (Message Passing Interface) y PVM (Parallel Virtual Machine).
3. *Programación por paso de datos*: Para este caso la transferencia de datos entre los procesadores se realiza con primitivas tipo put-get, lo que evita la necesidad de sincronización entre los procesadores emisor y receptor.

#### 4. Algoritmo propuesto

El algoritmo que a continuación se propone se basa en la configuración básica de maestro-esclavo, ya que existe un hilo maestro que asigna segmentos del algoritmo, a otros hilos para ejecutarse de manera paralela, cuando los hilos esclavos terminen de ejecutar el fragmento del algoritmo entonces se devuelven los datos al hilo maestro para que éste continúe ejecutando el programa. Se implementa a través del paradigma o modelo de programación de paralelismo en datos; la ventaja o aprovechamiento de realizar esta composición está dada en el tiempo de ejecución, además de que es fácil de implementar.

Enseguida se explican los segmentos (operadores) del algoritmo que se paralelizan, para esto se retoma el diagrama mostrado en la figura 1.

Se inicia generando y evaluando a la población en paralelo, seguidamente se seleccionan las parejas de manera secuencial, éste operador no se puede aplicar en paralelo ya que existe dependencia de datos. Posteriormente, se cruzan las parejas seleccionadas para formar nuevos cromosomas (hijos), este operador se trabaja sobre un 50 por ciento de la población en curso. También se aplica el operador de mutación sobre la población en curso, el porcentaje de la mutación se pasa como parámetro al inicio de la ejecución del programa. Estos dos últimos operadores se realizan de forma paralela. Y, finalmente, se actualiza la población para encontrar la mejor solución.

```
Inicio
  Hacer en paralelo
    Generar la población inicial
    Calcular la evaluación
  Fin paralelo
  Ordenar la población (menor a mayor)
  Mientras no se cumpla la condición de terminación
    Seleccionar parejas
    Hacer en paralelo
      Cruzar parejas (50% de la población en curso)
      Mutar a la población en curso
      Actualizar población
    Fin paralelo
  Fin mientras
Fin
```

**Algoritmo 1.** Algoritmo genético paralelo para resolver el problema del agente viajero.

Las directivas de OpenMP que se utilizaron para implementar el algoritmo son las siguientes:

- `#pragma omp sections` [cláusulas]  
Esta es una directiva que paraleliza trabajo no iterativo, es decir, paraleliza regiones o secciones de código que son independientes y NO participa ningún bucle. Para este caso se utiliza para la declaración de las matrices (Población inicial, población de los hijos) y para la declaración de otras estructuras de datos que se requieren para el programa.
- `#pragma omp for` [cláusulas]  
Esta directiva sirve para subdividir el ciclo for para que cada hilo trabaje por separado las iteraciones de éste, no funciona en el bucle “while”. La mayor parte de la paralelización del algoritmo utiliza esta directiva; se emplea para generar y evaluar a la población, en los operadores de cruce y mutación. Aunque parezca una directiva muy fácil de usar, se recomienda tener mucho cuidado al momento de manipular las variables que se pasen dentro de las cláusulas, ya que puede haber casos en los que la variable es global y no importa que la modifiquen todos los hilos o puede ser el caso que cuando una variable se modifique no le afecte al trabajo de los demás hilos. En pocas palabras, se debe tener cuidado con la dependencia de datos.

## 5. Experimentos y resultados

El algoritmo propuesto se validó utilizando problemas prueba reportados en la literatura [11], en la supercomputadora del Laboratorio Nacional de Supercómputo del Sureste de México [12], usando el lenguaje de programación C [13].

Para cada problema prueba se realizaron 15 ejecuciones del algoritmo con distintos parámetros sobre la cantidad de generaciones, tamaño de la población, porcentaje de mutación y la cantidad de hilos.

La tabla 1 reporta los costos menores obtenidos con las ejecuciones en el algoritmo genético secuencial y el algoritmo genético paralelo, así como el mejor reportado en la literatura en la actualidad.

La tabla 2 muestra una comparación de algunos tiempos de ejecución para cada problema variando la cantidad de hilos. El tiempo reportado está dado en segundos. Para los problemas de 10 y 20 ciudades se aplicaron los siguientes parámetros: 100 generaciones, 100 individuos y 0.3 porcentaje de mutación. Para los problemas restantes se aplicaron los siguientes parámetros: 500 generaciones, 500 individuos y 0.5 porcentaje de mutación, se trabaja con éstos parámetros con el propósito de comparar los tiempos de ejecución al operar con grandes volúmenes de datos. La última columna de la tabla 2 presenta la cantidad de hilos, realizando pruebas inicialmente con 2 y hasta 16 hilos.

Las figuras 2, 3 y 4 que se presentan son el resultado de hacer pruebas con el problema tsplib\_problem\_d1291, 1291 ciudades, con los siguientes parámetros: 100 generaciones, 100 individuos de población, 0.3 de porcentaje de mutación y 16 hilos de ejecución.

**Tabla 1.** Costos obtenidos del algoritmo genético secuencial, paralelo y los reportados en la literatura.

Problema	Cantidad Ciudades	AG secuencial	AG paralelo	Mejor reportado
TSP_small10	10	<b>1324.55</b>	<b>1324.55</b>	<b>1324.554</b>
TSP_small20	20	1847.55	<b>1836.92</b>	<b>1836.924</b>
data_100	100	<b>3746.40</b>	<b>3746.40</b>	<b>3746.40</b>
tsplib_problem_a280	280	4271.11	2828.80	<b>2579</b>
data_500	500	8862	<b>8456</b>	<b>8456</b>
tsplib_problem_d1291	1291	153672	98940	<b>50801</b>

**Tabla 2.** Comparación de los tiempos de ejecución del algoritmo genético secuencial y el paralelo.

Problema	Cantidad Ciudades	Tiempo AG secuencial	Tiempo AG paralelo	Cantidad Hilos
TSP_small10	10	0.0130	0.0134	2
TSP_small10	10	-	0.0255	5
TSP_small10	10	-	21.32	16
TSP_small20	20	0.0219	0.0287	2
TSP_small20	20	-	0.0317	5
TSP_small20	20	-	21.17	16
data_100	100	2.61	3.15	2
data_100	100	-	2.02	5
data_100	100	-	1.98	10
tsplib_problem_a280	280	13.42	6.11	5
tsplib_problem_a280	280	-	5.42	10
tsplib_problem_a280	280	-	37.31	16
data_500	500	36.78	25.81	2
data_500	500	-	12.11	10
data_500	500	-	28.08	16
tsplib_problem_d1291	1291	266.15	131.63	2
tsplib_problem_d1291	1291	-	77.18	5
tsplib_problem_d1291	1291	-	52.80	10
tsplib_problem_d1291	1291	-	93.40	16

La figura 2 representa los tiempos de ejecución del programa paralelo, el eje vertical simboliza el tiempo que esta dado en segundos y el eje horizontal indica el número de procesadores en los que se ejecutó el programa.

El factor de mejora del rendimiento o también conocido como “speedup” se define como:

$$S(n) = T(1) / T(n), \quad (2)$$

donde  $T(1)$  es el tiempo de ejecución del algoritmo secuencial (1 procesador) y  $T(n)$  es el tiempo de ejecución del algoritmo paralelo ejecutado sobre  $n$  procesadores. La figura 3 muestra el comportamiento de la función del speedup.

Por último, la figura 4 representa la escalabilidad del programa paralelo y su función está dada de la siguiente manera:

$$E(n) = S(n) / n = T(1) / (n * T(n)). \quad (3)$$

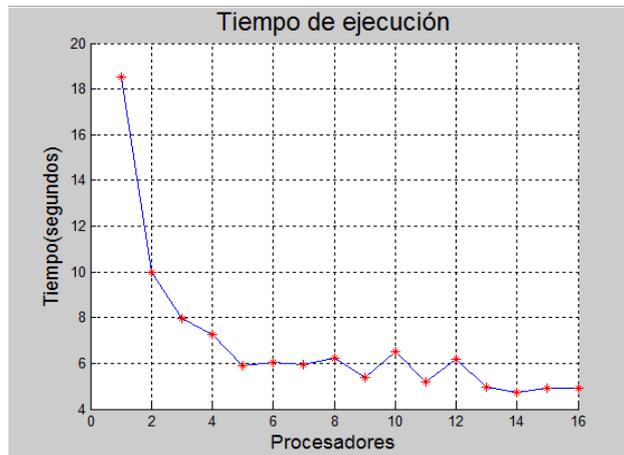


Fig. 2. Tiempos de ejecución del algoritmo genético paralelo para el problema tsplib\_problem\_d1291.

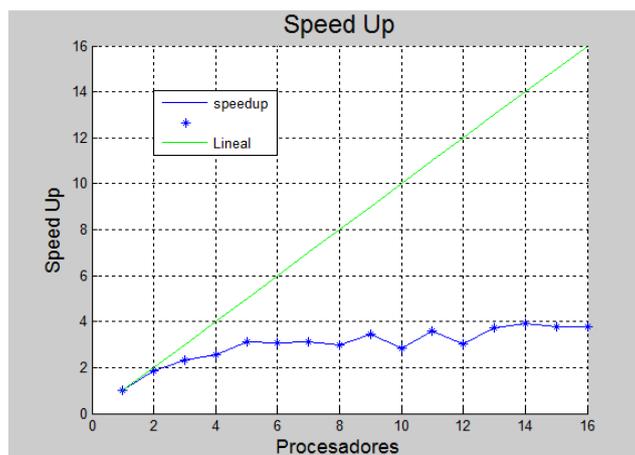


Fig. 3. Speedup del algoritmo genético paralelo para el problema tsplib\_problem\_d1291.

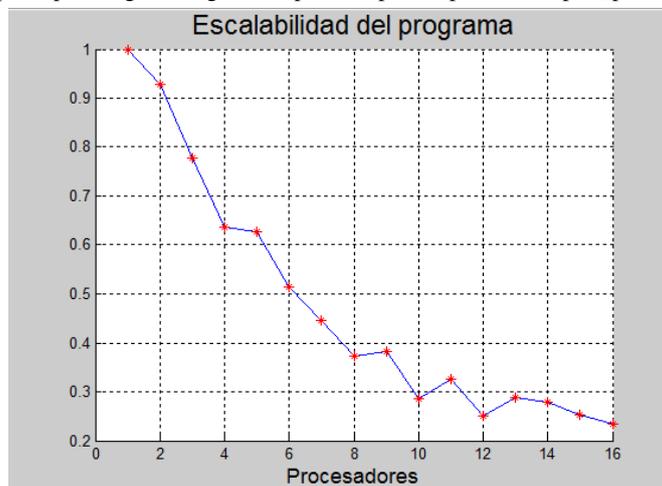


Fig. 4. Escalabilidad del algoritmo genético paralelo para el problema tsplib\_problem\_d1291.

## 6. Conclusiones

Los algoritmos genéticos tradicionales son de gran ayuda para problemas de búsqueda y optimización, además, estos algoritmos tienen la gran ventaja de ser propicios para poder paralelizarlos y así poder aprovechar o tener un mejor rendimiento sobre nuestro equipo de cómputo.

Se observa que con el algoritmo paralelo se alcanzan mejores resultados en comparación con el algoritmo secuencial, así mismo se obtiene un mejor rendimiento en cuestiones de tiempo de ejecución, logrando un porcentaje de mejora de tiempo de 24% para el caso de 100 ciudades y un 80% de mejora para el problema de 1291 ciudades. Con respecto a la cantidad de hilos, se nota que para los problemas de 10 y 20 ciudades es suficiente ejecutar el algoritmo con 2 hilos, mientras que para los problemas de igual o mayor a 100 ciudades se trabaja mejor con 10 hilos. Se concluye que es factible aplicar el algoritmo paralelo para problemas con cantidad de ciudades grandes.

El próximo objetivo a lograr en un futuro cercano es mejorar el algoritmo genético paralelo con la finalidad de encontrar mejores soluciones. Por ejemplo, aplicando distintas técnicas de mutación para explorar en diferentes espacios de búsqueda y generar nuevos individuos después de cada generación para sustituir a los peores individuos de la población.

**Agradecimientos.** Agradecemos el apoyo financiero de la Vicerrectoría de Investigación y Estudios de Posgrado de la Benemérita Universidad Autónoma de Puebla a través del proyecto SASM-ING16-G, de la misma manera agradecemos al Laboratorio Nacional de Supercómputo del Sureste de México por las facilidades prestadas para la realización de las pruebas.

## **Referencias**

1. Dasgupta, S., Papadimitriou, C., Vazirani, U.: NP-complete problems: Algorithms. McGraw-Hill Education, pp. 233–237 (2006)
2. Saloni, G., Poonam, P.: International Journal of Advanced Research in Computer Science and Software Engineering, Vol.3 (2013)
3. Poveda, R., Gómez, J., León, E.: Un algoritmo genético paralelo que combina los modelos de grillas e islas para encontrar soluciones óptimas cercanas al problema del agente viajero. Vol. 5, No. 3, pp. 13–19 (2008)
4. Tsai, C., Tseng, S., Chiang, M., Yang, C.: A Fast Parallel Genetic Algorithm for Traveling Salesman Problem. pp. 241–250 (2010)
5. Fang, Y.: Solving Traveling Salesman Problem Using Parallel Genetic Algorithm and Simulated Annealing. pp. 2 (2010)
6. Dorigo, M., Gambardella, L.: Ant colony system: a cooperative learning approach to the traveling salesman problem. Vol. 1, pp. 53–66 (2002)
7. OpenMP. <http://openmp.org/wp> (2016)
8. Haupt, R. L., Haupt, S. E.: Practical Genetic Algorithms. John Wiley & Sons, Inc. (1998)
9. Golberg, D., Lingles, R.: The travelling salesman problem. In: Proceedings of the First International Conference on Genetic Algorithms (1986)
10. Nesmachnow, S.: Algoritmos genéticos paralelos y su aplicación al diseño de redes de comunicaciones confiables. pp. 30 (2004)
11. Problemas benchmark. <https://sites.google.com/site/logisticslaboratory/lecture/tsp-program-benchmark-problems> (2016)
12. LNS. <http://www.lns.buap.mx> (2016)
13. Lenguaje C. <http://www.cprogramming.com> (2016)



## **Método criptográfico simétrico utilizando teoría del caos, operaciones sobre ADN y raíces de funciones no lineales**

Luis René Marcial Castillo, Erika Leonor Basurto Munguía,  
Marcela Rivera Martínez, María de Lourdes Sandoval Solís

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

{luis.marcial, marcela.rivera, maria.sandoval}@correo.buap.mx, iamdleonor@gmail.com

**Resumen.** En este trabajo se propone un algoritmo para criptografía simétrica el cual se basa en funciones caóticas, enmascaramiento basado en adición y sustracción sobre el ácido desoxirribonucleico y raíces de funciones no lineales. El algoritmo genera dos llaves; la primera es generada por la función caótica y la segunda por la función no lineal. La implementación computacional del algoritmo realizada en Octave permite cifrar y descifrar texto de cualquier dimensión con un alto nivel de seguridad apoyado esto en un análisis de sensibilidad que muestra la resistencia ante ataques exhaustivos.

**Palabras clave:** Criptografía simétrica, funciones caóticas, secuencias de ADN, funciones no lineales.

## **Symmetric Cryptographic Method Using Chaos Theory, Operations on DNA and Roots of Nonlinear Functions**

**Abstract.** In this work we propose an algorithm for symmetric cryptography based on chaotic functions, masking based on addition and subtraction on deoxyribonucleic acid and roots of nonlinear functions. The algorithm generates two keys; the first one is generated by the chaotic function and the second by the nonlinear function. The computational implementation of the algorithm realized in Octave allows to encrypt and decryption text of any dimension with a high level of security supported this in a sensitivity analysis that shows the resistance to exhaustive attacks.

**Keywords:** Symmetric cryptography, chaotic functions, DNA sequences, nonlinear functions.

### **1. Introducción**

La confidencialidad de las comunicaciones es de suma importancia en la sociedad moderna. La industria, gobierno y particulares confían en que la tecnología les garantice

que el intercambio de datos sea seguro de modo que no se permita a terceros acceder al contenido de tal comunicación, la cuestión de la confidencialidad ha sido dejada al campo de la criptografía [4]. Goldreich transcribe el planteamiento de la criptografía como "el problema de proveer comunicación secreta sobre medios inseguros" [8].

La criptografía es una herramienta muy útil cuando se desea tener seguridad informática, es decir, cuando se cuenta con un medio para garantizar las propiedades de confidencialidad, para lograrlo, se crean mecanismos que garanticen en cierta medida las propiedades de disponibilidad, integridad y confidencialidad. La disponibilidad se refiere a que la información siempre este presente, la integridad significa no perder información, la confidencialidad se puede lograr usando mecanismos que aunque sea robada la información, permita el no acceso a esta o garantice de alguna forma no poder llegar a ella, hasta que pierda su valor. Estos mecanismos permiten ver si la información ya creada ha sufrido o no alguna modificación no autorizada. El criptoanálisis, también llamado "criptología" es la disciplina contraria a la criptografía, se encarga de analizar la información cifrada para revelar la información original sin necesidad de las claves secretas y de esta forma romper los procedimientos previamente establecidos por la criptografía, el criptoanálisis lo usan los investigadores como una forma de probar las fortalezas o debilidades de los cripto sistemas [6].

Desde la década de los 90's muchos investigadores han notado que existe una importante relación entre el caos y la criptografía: muchas propiedades de los sistemas caóticos tienen sus correspondientes contrapartes en los cripto sistemas tradicionales. Los sistemas caóticos pueden conocer sus ecuaciones y sus condiciones iniciales fijas, sin embargo, la más mínima variación provoca una evolución radical en su comportamiento [4]. La teoría de funciones y mapas caóticos se presentan en varios trabajos. Fuan, Mengb, Zhanb, Zhuc, Laud, Tsed y Mae en el 2013 [5] proponen un esquema de protección de imágenes médicas basado en mapas caóticos; Gao y Chen en el 2008 [7] proponen un algoritmo nuevo de permutación de pixeles; Huang y Nien en el 2009 [10] proponen un sistema multi-caótico basado en el mismo principio de permutación de Gao y Chen; Patidar, Pareek y Sud en el 2009 [15] proponen un cifrado tipo sustitución-difusión basado en mapas logísticos y caóticos; Rhouma, Meherzi y Belghith en el 2009 [16] proponen el cifrado de imágenes a color basado en mapas caóticos; Sun, Liu y Li en el 2008 [17] proponen un esquema de encriptación basado en mapas caóticos espaciales; Tong y Cui en el 2009 [19] proponen un generador de cifrado de secuencias caóticas con componentes dinámicos; Wong, Kwok y Law en el 2008 [20] proponen un esquema de encriptación basado en el mapa caótico estándar; Xiao y Xia en el 2009 [21] proponen un esquema de encriptación usando mapas de permutación; Xu, Wang y Yang en el 2008 [22] proponen una mejora en el algoritmo de encriptación de imágenes que usan mapas caóticos y Ye en el 2009 [23] propone un cripto sistema basado en las matrices de Töplitz y Hankel.

En lo que se refiere a la teoría de secuencias de ácido desoxirribonucleico (ADN) se presenta en varios trabajos. Terec, Vaida, Alboaie y Chiorean, en el 2011 [18] proponen el uso de ADN para criptografía simétrica; Anwarl, Paul y Singh en el 2014 [2] hacen una revisión de la forma en que se ha realizado la transmisión de mensajes basada en el uso de ADN; Javheri y Kulkarni en el 2014 [13] proponen también un algoritmo criptográfico para la comunicación segura de datos basado en ADN; Anil y

Chirakkarottu en el 2014 [1] presentan un método de encriptación para el iris del ojo humano basado también en operaciones de ADN.

Como lo muestran los trabajos mencionados, las funciones caóticas y las operaciones de ADN han llamado la atención de varios investigadores y las han utilizado para diseñar sus algoritmos criptográficos.

La propuesta de este trabajo además de utilizar resultados sobre funciones caóticas para generar la primera llave como lo muestran los trabajos mencionados y operaciones de ADN para realizar el enmascaramiento, agrega el uso de raíces de una función no lineal, con lo cual se refuerza la seguridad, al proporcionarle al algoritmo una segunda llave, con un costo muy bajo.

En la siguiente sección se presentan los aspectos matemáticos usados en el algoritmo, en la sección 3 se muestra y detalla el algoritmo propuesto, la sección 4 proporciona las pruebas del cripto sistema. La sección número 5 presenta las conclusiones y finalmente se listan las referencias utilizadas en el desarrollo de este trabajo.

## **2. Aspectos matemáticos**

### **2.1. Teoría del caos**

Puede decirse que la dinámica caótica inició con el trabajo del matemático Francés Henri Poincaré a finales del siglo XIX. La motivación de Poincaré fue promovida por el problema de las orbitas de tres cuerpos celestes experimentando atracción gravitacional mutua (por ejemplo, una estrella y dos planetas). Poincaré fué capaz de mostrar que orbitas muy complicadas eran posibles (ahora llamadas caóticas). No obstante, la posibilidad de caos en sistemas físicos reales no fue ampliamente apreciada sino hasta la actualidad, mucho del crédito por este cambio es atribuido a la extensa solución numérica de sistemas dinámicos en computadoras digitales [14].

En los últimos años, los hilos del caos y la dinámica no lineal se han esparcido a través de disciplinas científicas como una intrincada red araña. Caos y dinámica no lineal han provisto de nuevas herramientas teóricas y conceptuales que permiten capturar, entender y enlazar los comportamientos complejos de sistemas simples (el tipo de comportamiento llamado caos en la ciencia contemporánea) [9]. Se puede decir, que una función caótica es una función matemática que describe un sistema dinámico no lineal complejo cuya evolución en el tiempo hace imposible la predicción a largo plazo, luciendo errático y casi aleatorio. Los ejemplos más notorios de las características del caos son el llamado efecto mariposa y la impredecibilidad de órbitas pseudo-aleatorias, generadas por ecuaciones deterministas. Estos fenómenos, así como otros relacionados con el caos, han sido tradicionalmente asociados a mecanismos de confusión y difusión los cuales son la base principal de un buen sistema criptográfico.

El elemento central de todos los sistemas caóticos es el concepto de iteración. El estado actual del sistema es una función determinística del estado o valor anterior. Formalmente, una correspondencia caótica se especifica por medio de la expresión:

$$x_{k+1} = f(x_k). \quad (1)$$

La fórmula (1) muestra la expresión de la correspondencia caótica, donde  $f$  es una función no lineal. En general, la teoría de caos surge de la necesidad de modelar mecanismos físicos tales como la predicción del tiempo atmosférico, la evolución de la población, la dinámica de fluidos, la teoría de gases, predicción del tiempo, etc.

Un sistema caótico es un sistema dinámico, no lineal, determinístico que muestra una dependencia muy sensible a las condiciones iniciales y presenta una evolución a través de un espacio de fase que parece ser aleatorio [8]. Estas propiedades proporcionan un potencial para aplicaciones en criptografía ya que las predicciones a largo plazo de los sistemas caóticos son muy difíciles [11].

Existen varias funciones caóticas usadas en el campo de la criptografía, pero de las analizadas por Tereee, Vaida, Alboae y Chiorea, 2011 [18] la *cross chaotic map* es la que les dio mejores resultados, razón por la cual es usada en este trabajo para generar la primera llave. Tal función es definida por las ecuaciones (2) y (3):

$$x_{i+1} = 1 - \eta y_i^2, \text{ donde } \eta \text{ es una constante,} \quad (2)$$

$$y_{i+1} = \cos(k \cos^{-1}(x_i)), \text{ donde } k \text{ es una constante.} \quad (3)$$

## 2.2. Secuencias de ADN

El ácido desoxirribonucleico es el material genético de todos los organismos celulares y de casi todos los virus. El ADN lleva la información necesaria para dirigir la síntesis de proteínas y la replicación. Cada molécula de ADN está constituida por dos cadenas o bandas formadas por un elevado número de compuestos químicos llamados nucleótidos. Estas cadenas forman una especie de escalera retorcida que se llama doble hélice. Cada nucleótido está formado por tres unidades: una molécula de azúcar llamada desoxirribosa, un grupo fosfato y uno de cuatro posibles compuestos nitrogenados llamados bases, las cuales son: Adenina (A), Guanina (G), Tiamina (T) y Citosina (C). Conforme al complemento presentado por Watson y Crick, la Adenina es complementada por Tiamina, y Guanina con Citosina [1].

**Tabla 1.** Regla de adición para el ADN.

+	A	C	G	T
A	A	C	G	T
C	C	G	T	A
G	G	T	A	C
T	T	A	C	G

En este método la Adenina se codifica con 00, la Tiamina con 11, la Guanina con 10 y la Citosina con 01 [18]. La regla de adición (suma binaria) para el ADN es realizada como lo muestra la tabla 1, por ejemplo, si se desea realizar la operación T + T el resultado sería G, ya que la suma binaria 11 + 11 es igual a 110 pero al desechar el bit más significativo (más a la izquierda) se obtiene el número binario 10 que corresponde a G. La regla de sustracción (resta binaria en complemento a 2) es mostrada en la tabla 2, por ejemplo, el resultado de la operación T – T es A, ya que el complemento a 2 de

T es: 01 y al sumar T + 01 resulta que 11 + 01 = 100, en donde se desecha el bit más significativo y el resultado es 00 que corresponde a la letra A.

**Tabla 2.** Regla de sustracción para el ADN.

-	A	C	G	T
A	A	T	G	C
C	C	A	T	G
G	G	C	A	T
T	T	G	C	A

Las operaciones de suma y sustracción de ADN son usadas en este trabajo.

### 2.3. Raíces en funciones no lineales

Cuando se desea encontrar algún  $x$  que resuelva la ecuación no lineal  $f(x) = 0$ , se puede usar algún método iterativo. Uno de los métodos más conocidos y eficientes es el método de Newton. El método de Newton se obtiene de la forma siguiente [3]:

Paso 1. Se realiza la expansión de Taylor de grado 1 a la función  $f(x)$  alrededor de un punto local  $x_k$  y se iguala a cero, es decir,

$$f(x) = f(x_k) + f'(x_k)(x - x_k) = 0. \quad (4)$$

Paso 2. Se despeja a la variable  $x$  de la ecuación (4) y se obtiene:

$$x = x_k - \frac{f(x_k)}{f'(x_k)}. \quad (5)$$

Paso 3. El siguiente punto  $x_{k+1}$  se obtiene simplemente al sustituir  $x$  por  $x_{k+1}$  en la ecuación 5, obteniendo:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \quad (6)$$

Una de las ventajas del método de Newton es que si  $r$  es una raíz simple de la función no lineal  $f(x) = 0$ , entonces, converge cuadráticamente, en caso contrario converge linealmente. Es importante que la aproximación inicial este cerca para que se cumplan las ventajas mencionadas. Una desventaja que tiene el método es que si no se conoce analíticamente la derivada de la función  $f(x)$  debe aproximarse de forma numérica, lo cual no pasa para la función  $f(x)$  que se usa en este trabajo.

**Algoritmo 1.** Algoritmo del método de Newton para encontrar alguna de las raíces de la función no lineal  $f(x)$ .

Inicio

Dado una aproximación inicial  $x_i$  a la raíz de la función no lineal  $f(x)$ .

$$x_n = x_i - \frac{f(x_i)}{f'(x_i)}.$$

Mientras no se cumpla la condición de paro

$$x_i = x_n$$

$$x_n = x_i - \frac{f(x_i)}{f'(x_i)}.$$

Fin del ciclo Mientras

Fin

Para el algoritmo criptográfico propuesto se encuentra la raíz de la función no lineal  $f(x) - \text{valorascii}$ , donde *valorascii* es el código ASCII asociado al carácter que se desea cifrar.

### 3. Algoritmo propuesto

El algoritmo que se propone se basa en las funciones caóticas, operaciones sobre ADN y raíces de funciones no lineales. El cripto sistema consta de los algoritmos de cifrado y descifrado sobre texto. A continuación se presentan tales algoritmos.

**Algoritmo 2.** Algoritmo de cifrado sobre texto.

Inicio

1. Se ingresa el texto a cifrar  $T$ , y se encuentra su longitud, dejando el resultado en  $n$ .
2. Se calcula el valor de  $m$  como:  $m = \lceil \sqrt{n} \rceil$ , donde  $\lceil x \rceil$  denota el menor entero que es mayor o igual al valor real  $x$ .
3. En caso de ser necesario, se agrega basura al final del arreglo  $T$  de modo que se obtenga un nuevo arreglo  $T_e$  cuya longitud es  $m \times m$ .
4. Se convierten a decimal los caracteres del arreglo  $T_e$  usando para ello los códigos ASCII de tales caracteres, obteniendo el nuevo arreglo  $T_d$ .
5. El arreglo  $T_d$  se apila en las columnas de una matriz  $I$  cuya dimensión será  $m \times m$ .
6. Codificar los elementos de  $I$  a binario usando 8 bits para representar cada dígito, dejando el resultado en la matriz  $I_b$ .
7. Se codifica la imagen binaria  $I_b$  a una secuencia de ADN y el resultado se deja en una matriz denotada con  $L$ .
8. Se construye la primera llave de la forma siguiente:  
Se calculan los valores  $k_1$  y  $k_2$  como:

$$k_1 = \frac{1}{126} \bmod \left( \sum_{i=1}^{m/2} \sum_{j=1}^m I_{i,j}, 256 \right),$$

$$k_2 = \frac{1}{126} \bmod \left( \sum_{i=m/2+1}^m \sum_{j=1}^m I_{i,j}, 256 \right).$$

Se generan de forma aleatoria los valores:  $x_0, y_0$  sobre el intervalo real  $[0,1]$ .

Se calcula el valor inicial  $x_0$  de la secuencia caótica como: (se repite el mismo proceso pero usando  $y_0$  en lugar de  $x_0$ ) [18]

$$x_0 = x_1 + k_1$$

Si  $x_0 > 1$  entonces  $x_0 = \text{mod}(x_0, 1)$ .

Se genera  $X = (x_1, \dots, x_m)$  como vector columna y  $Y = (Y_1, \dots, Y_{8m})$  como vector fila usando la fórmula caótica dada por las ecuaciones (2) y (3) con  $\eta = 2$  y  $k = 6$ .

Se multiplica el vector columna  $X$  por el vector fila  $Y$ , obteniendo una matriz  $M$  de dimensión  $m \times 8m$ .

Se convierte la matriz  $M$  a binario usando:

$$M_b(i, j) = \begin{cases} 0, & \text{Si } M(i, j) < 0 \\ 1, & \text{Si } M(i, j) \geq 0. \end{cases}$$

Se codifica  $M_b$  a cadenas de ADN obteniendo la matriz  $K$ .

9. Se aplica la adición de ADN como lo dicta la tabla 1.  
 $NADN = L + K$ .
10. La matriz obtenida  $NADN$  se pasa a binario obteniendo  $N_b$ .
11. Se convierte  $N_b$  a valores enteros agrupando cada 8 bits para formar un valor entero entre 0 y 255, obteniendo una matriz  $C_p$  de dimensión  $m \times m$ .
12. Generar de modo aleatorio en el rango  $[0, 1]$  los valores de las constantes que componen a la función no lineal  $f(x)$  seleccionada, estos valores constantes son la segunda llave del proceso de cifrado.
13. Para cada valor ASCII de  $C_p$ , se aplica el método de Newton para encontrar alguna  $c$  que sea raíz de la función no lineal  $f(x)$  — *valor ASCII*.
14. Pasar todos los valores reales  $c$  encontrados por el método de Newton a una cadena hexadecimal  $c_b$  que se compone de 64 bits si se usan reales dobles, o de 32 bits si se usan reales simples, dejando las cadenas hexadecimales en el arreglo  $C$ .

Fin

El cifrado final se encuentra en  $C$  y es una cadena en formato hexadecimal.

### Algoritmo 3. Algoritmo de descifrado sobre texto.

Inicio

1. Se pasan las cadenas hexadecimales almacenadas en  $C$  a números decimales, dejando el resultado en el arreglo  $D_d$ .
2. Los valores de  $D_d$  se evalúan en la función no lineal  $f(x)$  que contiene constantes que forman la segunda llave, las evaluaciones se dejan en  $D_{de}$ .
3. Se redondean al entero más cercano los valores de  $D_{de}$  dejando el resultado en  $D_{de}$ .
4. Se pasa la matriz  $D_{de}$  a valores binarios, obteniendo  $D_b$ .
5. Se codifica  $D_b$  a valores de ADN obteniendo  $DADN$ .
6. Se aplica la regla de sustracción de ADN dada por la tabla 2.  
 $L_2 = K - DADN$ , donde  $K$  es la primera llave generada en el proceso de cifrado.
7. Se convierte  $L_2$  a binario obteniendo  $L_{2b}$ .
8. Se convierte  $L_{2b}$  a decimal, obteniendo  $DI_d$ .
9. Se pasa  $DI_d$  a un vector de longitud  $m \times m$  desapilando por columnas, dejando el resultado en  $D_e$ .
10. Se elimina la basura tomando solo los primeros  $n$  elementos de  $D_e$ , dejando el resultado en  $DT_d$ .
11. Se convierte a carácter cada uno de los valores de  $DT_d$  dejando el resultado en  $T$ .

Fin

Al finalizar el algoritmo de descifrado se obtiene el texto original en  $T$ .

## 4. Pruebas

El algoritmo criptográfico simétrico propuesto se desarrolló bajo el lenguaje de programación OCTAVE [24] y las pruebas se realizaron en una computadora hp con 16 Gb de memoria RAM y un procesador intel core i7 a 2.6 GHz.

### 4.1. Cifrado y descifrado de texto

En la prueba que se presenta a continuación se usa como función no lineal a:

$$f(x) = ax^3 + b \cos(x), \quad (7)$$

donde las constantes reales  $a$  y  $b$  se generan de forma aleatoria sobre  $[0,1]$ . El algoritmo 2, obtiene los siguientes resultados cuando cifra el texto: “The 10th International Congress on Intelligent and Information Technologies 2016”.

El paso 1, obtiene  $n = 80$  y el paso 2, obtiene  $m = 9$ .

El paso 3, agrega al final del texto un carácter % como basura.  $T_e =$  The 10th International Congress on Intelligent and Information Technologies 2016%

El paso 4, obtiene (al pasar a decimal los caracteres de  $T_e$ )  $T_d =$  84 104 101 32 49 48 116 104 32 73 110 116 101 114 110 97 116 105 111 110 97 108 32 67 111 110 103 114 101 115 115 32 111 110 32 73 110 116 101 108 108 105 103 101 110 116 32 97 110 100 32 73 110 102 111 114 109 97 116 105 111 110 32 84 101 99 104 110 111 108 111 103 105 101 115 32 50 48 49 54 37.

El paso 5, apila los datos de  $T_d$  en una matriz  $I$  de  $9 \times 9$ .

```
I = 84 73 111 114 110 116 111 84 105
    104 110 110 101 116 32 114 101 101
    101 116 97 115 101 97 109 99 115
    32 101 108 115 108 110 97 104 32
    49 114 32 32 108 100 116 110 50
    48 110 67 111 105 32 105 111 48
    116 97 111 110 103 73 111 108 49
    104 116 110 32 101 110 110 111 54
    32 105 103 73 110 102 32 103 37.
```

El paso 6, pasa a binario los datos de  $I$  usando 8 bits para cada dato, y el paso 7 codifica las cadenas binarias a secuencias de ADN.

```
L = CCCACAGCCGTTCTAGCGTGCTCACGTTCCCACGGC
    CGGACGTGCGTGCGCCCTCAAGAACTAGCGCCCGCC
    CGCCCTCACGACCTATCGCCCGACCGTCCGATCTAT
    AGAACGCCCCGTAATCGTACGTGCGACCGGAAGAA
    ATACCTAGAGAAAGAACGTACGCACTCACGTGATAG
    ATAACGTGCAATCGTTCGGCAGAACGGCCGTTATAA
    CTCACGACCGTTCGTGCGCTCAGCCGTTTCGTAATAC
    CGGACTCACGTGAGAACGCCCCGTGCGTGCGTTATCG
    AGAACGGCCGCTCAGCCGTGCGCGAGAACGCTAGCC.
```

El paso 8, genera la primera llave usando la función caótica. La matriz final en cadenas de ADN es:

```
K = GAGCCAGGTCCAGATTATTGTTTATAGTTGGCAACC
    CTCGGTCCAGGTCTAATAACAAATATCAACCGTTGG
    CTCGGTCCAGGTCTAATAACAAATATCAACCGTTGG
    GAGCCAGGTCCAGATTATTGTTTATAGTTGGCAACC
    CTCGGTCCAGGTCTAATAACAAATATCAACCGTTGG
    CTCGGTCCAGGTCTAATAACAAATATCAACCGTTGG
    GAGCCAGGTCCAGATTATTGTTTATAGTTGGCAACC
    GAGCCAGGTCCAGATTATTGTTTATAGTTGGCAACC
    GAGCCAGGTCCAGATTATTGTTTATAGTTGGCAACC.
```

El paso 9, aplica la adición de ADN como lo indica la tabla 1.

```
NADN = TCTCGAATATATTTTCCC GAAGAAAGCGATTCCGTG
        GCTGTCATCACC GCCATCCAGATCGCGCTGTACTT
        GCGTTGGCCAGAGGATAGCGCGAACCACTCCAGGC
        GGGCGGTTATAATTTGCCGGACGGAGGAAAACAGCC
        CGCTTGCTAAGTCCAAGTCCGCTCGGACTAATGGA
        CGCGTCATCGGGGCTTAGGGAGATCCTCCTACTGGG
        TTTCGGGTATATTGGCCACATCCAGCGAACCATCG
        TGACGTTGATAGGGTTCATACGGAGCCAACAATGT
        GGGCGGATATGTTACACCGAACAGTGGTAATAAGGG.
```

El paso 10, pasa a binario el resultado *NADN* y el paso 11 pasa cada 8 bits a un número decimal.

```
Cp = 221 131 51 253 88 32 38 61 110
      158 211 69 149 53 35 102 123 31
      155 233 72 163 38 96 81 117 41
      169 175 48 254 90 26 40 1 37
      103 231 11 80 45 103 104 112 232
      102 211 106 159 42 35 93 113 234
      253 171 51 233 81 53 38 5 54
      225 190 50 175 83 26 37 4 59
      169 163 59 196 88 18 235 12 42.
```

El paso 12, genera los valores *a*, *b* de la segunda llave:

$a = 7.638979442864783e-001$ ,  $b = 7.593273831310963e-001$ .

El paso 13, encuentra las raíces de la función no lineal usando el método de Newton, y el paso 14 escribe los valores del paso 13 de doble precisión a formato hexadecimal.

```
C =
401a6d303b77b38f 40179ecda782ef8e 40177841d2962af6 40182836cc1ed798
40147dbfaaf0c981 40146cdd2d199c2b 401ba7149539747c 401a95feb0d786
40182836cc1ed698 401630f5884fd08a 401a0501c5ea7f4c 401ae6335d64bd7a
401870bb03277bd2 401ad2523455744b 401a0501c5ea7f4d 40184091bcaeacd3
```

```

40191f36f44d7ef5 4017de01cf6a6d59 40104672751f056e 4011f633ff5abedb
401236bb5cb5ef59 400fea5291610c10 4003cf1eeaae4b97 4014afc6f0595b09
40104672751f0acc 40102bb58d026495 401110e9c8ccec37 401ba7149539747e
401729b0e23f2caa 4017de01cf6c649d 401bb074bef5f47e 4012dab1fec3491e
4017ab8bf2451582 401ae6335d65643a 401870bb03277e89 4019627e794f1955
4013745c2b909c74 40107ae8bc0c3cc9 400d966499343aa9 4013995a5c45abf0
400f40ed7d8632bc 400e8feb2ed2e8c9 4012ee6e7f0fee9d 4013156fb5e5638a
4013745c2b909c73 400bfd084b717326 400ccf8fa5828904 4014053e2faf2c66
400a2c850e84dd64 40147dbfaaf0cbb2 400ccf8fa5873228 40107ae8bc0c35fa
400a2c850ea76eec 40073e9bc8110182 400d966499343877 40146cdd2d197dfc
4012ee6e7f0fdb6d 400e1538b7c2c99d 40148e872d073835 4013cfdc1a3f7818
400d96649934387e 400d55503bd38da2 401af9f7de35e7df 401140b086812c78
4015bb4c74527537 40155fc5b998242f 3fec0a9df4a76336 40151128dfa70a65
4015210f82ab020b 3ffe6999dfed1963 3ffc1ae17f876d80 4004614646a369c2
4014f114bd52a05d 400bb3ef591b1c04 400e53109ebf5cef 400d55503bbf82e5
401adc46643bdabc 401af0192c872d9b 401094a79cf91721 401110e9c8cecb79
400e8feb2ed2e9ed.
    
```

C contiene el mensaje cifrado en formato hexadecimal.

El algoritmo 3 de descifrado sobre texto, obtiene los siguientes resultados.

El paso 1, pasa los valores hexadecimales de C a decimales, el paso 2 evalúa los decimales usando la ecuación 7 y el paso 3 redondea los valores del paso 2 al entero más cercano obteniendo:

$D_{de} =$

```

221 131 51 253 88 32 38 61 110
158 211 69 149 53 35 102 123 31
155 233 72 163 38 96 81 117 41
169 175 48 254 90 26 40 1 37
103 231 11 80 45 103 104 112 232
102 211 106 159 42 35 93 113 234
253 171 51 233 81 53 38 5 54
225 190 50 175 83 26 37 4 59
169 163 59 196 88 18 235 12 42.
    
```

El paso 4, pasa  $D_{de}$  a binario y el paso 5 pasa a cadenas de ADN.

```

DADN = TCTCGAATATATTTTCCCGAAGAAAGCGATTCCGTG
      GCTGTCATCACCGCCCATCCAGATCGCGCTGTACTT
      GCGTTGGCCAGAGGATAGCGCGAACCACCTCCAGGC
      GGGCGGTTATAATTTGCCGACGGAGGAAAACAGCC
      CGCTTGCTAAGTCAAAGTCCGCTCGGACTAATGGA
      CGCGTCATCGGGGCTTAGGGAGATCCTCCTACTGGG
      TTTCGGGTATATTGGCCCACATCCAGCGAACCATCG
      TGACGTTGATAGGGTTCCATACGGAGCCAACAATGT
      GGGCGGATATGTTACACCGAACAGTGTAATAAGGG.
    
```

El paso 6, aplica la regla de sustracción de ADN  $L_2 = K - DADN$ .

$L_2 =$  CCCACAGCCGTTCTAGCGTGCTCACGTTCCCACGGC  
 CGGACGTGCGTGCGCCCTCAAGAACTAGCGCCCGCC  
 CGCCCTCACGACCTATCGCCCGACCGTCCGATCTAT  
 AGAACGCCCCGTAATCGTACGTGCGACCGGAAGAA  
 ATACCTAGAGAAAGAACGTACGCACTCACGTGATAG  
 ATAACGTGCAATCGTTCGGCAGAACGGCCGTTATAA  
 CTCACGACCGTTCGTGCGCTCAGCCGTTTCGTAATAC  
 CGGACTCACGTGAGAACGCCCCGTGCGTGCGTTATCG  
 AGAACGGCCGCTCAGCCGTGCGCGAGAACGCTAGCC.

El paso 7, pasa  $L_2$  a binario y el paso 8 pasa a valores decimales.

$DI_d =$  84 73 111 114 110 116 111 84 105  
 104 110 110 101 116 32 114 101 101  
 101 116 97 115 101 97 109 99 115  
 32 101 108 115 108 110 97 104 32  
 49 114 32 32 108 100 116 110 50  
 48 110 67 111 105 32 105 111 48  
 116 97 111 110 103 73 111 108 49  
 104 116 110 32 101 110 110 111 54  
 32 105 103 73 110 102 32 103 37.

El paso 9, pasa la matriz  $DI_d$  a un vector desapilando por columnas y el paso 10 elimina la basura obteniendo el vector:

$DT_d =$  84 104 101 32 49 48 116 104 32 73 110 116 101 114  
 110 97 116 105 111 110 97 108 32 67 111 110 103 114  
 101 115 115 32 111 110 32 73 110 116 101 108 108 105  
 103 101 110 116 32 97 110 100 32 73 110 102 111 114  
 109 97 116 105 111 110 32 84 101 99 104 110 111 108  
 111 103 105 101 115 32 50 48 49 54.

El paso 11, pasa  $DT_d$  a caracteres:

T = "The 10th International Congress on Intelligent and Information Technologies 2016".

## 4.2. Seguridad del cripto sistema

En esta sección se prueba la fortaleza del sistema criptográfico propuesto.

### 4.2.1. Ataque a fuerza bruta

El primer análisis que se presenta es el de ataque por fuerza bruta, que consiste en probar todas las posibles llaves hasta encontrar la llave con la cual se pueda recuperar el mensaje original. Para considerar que el espacio de llaves es adecuado para resistir un ataque de fuerza bruta, este espacio debe ser superior a  $2^{100} = 1.2677 \times 10^{30}$  [12]. El sistema criptográfico que aquí se presenta, depende de 2 llaves, la primera generada a partir de la secuencia caótica dada por las ecuaciones (2) y (3) dependiendo de dos números aleatorios iniciales  $x_0, y_0$  en el intervalo real  $[0,1]$ , al usar la función con menos

rango de posibles números aleatorios de Octave se tiene un total de  $2^{32}-1$  posibilidades para cada número, la segunda llave también depende de dos valores  $a$  y  $b$ , por lo que el total de posibles llaves es:  $(2^{32}-1) \times (2^{32}-1) \times (2^{32}-1) \times (2^{32}-1) = 3.4028 \times 10^{38}$ , dando seguridad al algoritmo criptográfico propuesto en este trabajo ante el ataque a fuerza bruta.

#### 4.2.2 Sensibilidad de los datos de entrada

Los sistemas caóticos tienen la característica de ser sensibles a las condiciones iniciales [14]. Es por ello, que el segundo análisis que se presenta es el de sensibilidad de los datos de entrada a un cambio muy pequeño en los valores iniciales de la generación de llaves.

Para la cross chaotic map [18] usada en esta propuesta, la sensibilidad se ilustra con un ejemplo, realizando la ejecución de la implementación computacional con los valores iniciales  $x_0 = 0.3$ ,  $y_0 = 0.6$  dejando sin cambio los valores iniciales de la segunda llave, se descifra de forma correcta el mensaje dado por la ecuación (8), pero al usar los valores iniciales modificados de  $x_0 = 0.30001$  y  $y_0 = 0.5999$  se obtiene el mensaje descifrado " *Wkd!2Iwk#mBXA^JEXM jila%Djib=«X6φ- -----  
ä Ñpô°ùPÖ[nYuH}{Z-³μ¶ièö-;½¹"* que claramente no corresponde al mensaje original de la ecuación (8).

También, dejando fijos los valores iniciales de la función caótica y con valores de  $a=0.3$  y  $b=0.6$  en la segunda llave se obtiene el mensaje original de la ecuación (8), sin embargo, modificando ligeramente los valores iniciales  $a$  y  $b$  como 0.30001 y 0.59999 se obtiene el descifrado " *Uif#03ui!eJPIVBMPE`cfk/N`chª «  
ø7X±- Ä°ñbÔ ÛpösFq]`Usr4-¼ª.¶³/æâü§μ·¾³"* que nuevamente no corresponde al mensaje original. Por lo tanto, se puede decir, que el sistema criptográfico es muy sensible a variaciones pequeñas en las llaves brindando una gran resistencia ante ataques exhaustivos.

### 5. Conclusiones

El uso de las funciones caóticas, junto con las operaciones de adición y sustracción sobre ADN y las raíces de funciones no lineales son útiles en el desarrollo de sistemas criptográficos simétricos. Nuestra propuesta usa dos llaves, la primera se genera a partir de las funciones caóticas como lo hacen los autores que se mencionan, la segunda se genera a partir de las constantes de una función no lineal, dándole por consiguiente más seguridad al método criptográfico y sólo se incrementa en dos datos más (las constantes reales  $a$  y  $b$  usadas en la función no lineal dada por la ecuación 7) los elementos usados como llaves.

El sistema criptográfico presentado es muy sensible como se muestra en la sección 4.2 a variaciones pequeñas en las llaves dando por consiguiente gran resistencia ante ataques exhaustivos.

**Agradecimientos.** Agradecemos el apoyo financiero de la Vicerrectoría de Investigación y Estudios de Posgrado de la Benemérita Universidad Autónoma de

Puebla a través del proyecto MACL-ING16-I, así como a los revisores de este trabajo por sus comentarios constructivos.

## Referencias

1. Anil, J., Chirakkarott, S.: Secure Encryption Method for Biometric Iris Pattern. *International Journal of Trends and Technology (IJCTT)*, Vol. 12, No. 5, pp. 217–219 (2014)
2. Anwar, T., Sanchita, P., Singh, S. K.: Message Transmission Based on DNA Cryptography: Review. *International Journal of Bio-Science and Bio-Technology*, Vol. 6, No. 5, pp. 215–222 (2014)
3. Burden, R., Douglas, F. J.: Numerical methods. 4th edition, International edition (2012)
4. Delfs, H., Maurer, U., Knebl, H.: Introduction to Cryptography Principles and Applications. Second Edition, Springer-Verlag Berlin Heidelberg (2007)
5. Fuan, Ch., Mengb, W., Zhanb, Y., Zhuc, Z., Laud, F., Tsed, Ch., Mae, H.: An efficient and secure medical image protection scheme based on chaotic maps. *Computers in Biology and Medicine*, Vol. 43, pp. 1000–1011 (2013)
6. Galende, J.: Criptografía, historia de la escritura cifrada. 1ra edición, editorial complutense (1995)
7. Gao, T. G., Chen, Z. Q.: Image encryption based on a new total shuffling algorithm. *Chaos Solutions & Fractals*, 38(1), pp. 213–220 (2008)
8. Goldreich, O.: Modern Cryptography, theory and practice, discrete mathematics and its applications. 3th edition, Chapman & Hall (1999)
9. Hilborn, R. C.: Chaos and nonlinear dynamics. 2da edition, Oxford University Press (2000)
10. Huang, C. K., Nien, H. H.: Multi chaotic systems based pixel shuffle for image encryption. *Opt. Commun.*, 282(11), pp. 2123–2127 (2009)
11. Inzunza, E., Cruz, C.: Double hyperchaotic encryption for security in biometric systems. *Nonlinear Dynamics and Systems Theory*, 13(1), pp. 5–68 (2013)
12. Jeevidha, S., Saleem, M. S., Dhavachelan, P.: Analysis on DNA based cryptography to secure data transmission. *International Journal of International Journal of Computer Applications*, Vol. 29, No. 8 (2011)
13. Javheri, S., Kulkarni, R.: Secure Data communication and Cryptography based on DNA based Message Encoding. *International Journal of Computer Applications*, Vol. 98, No. 16, pp. 35–40 (2014)
14. Ott, E.: Chaos in dynamical systems. Cambridge University Press (1993)
15. Patidar, V., Pareek, N. K., Sud, K. K.: A new substitution–diffusion based image cipher using chaotic standard and logistic maps. *Commun. Nonlinear Sci. Numer. Simulation*, 14(7), pp. 3056–3075 (2009)
16. Rhouma, R., Meherzi, S., Belghith, S.: OCML-based colour image encryption. *Chaos Solitons & Fractals*, 40(1), pp. 309–318 (2009)
17. Sun, F. Y., Liu, S. T., Li, Z. Q.: A novel image encryption scheme based on spatial chaos map. *Chaos Solitons & Fractals*, 38(3), pp. 631–640 (2008)
18. Teree, R., Vaida, M. F., Alboaie, L., Chiorea, L.: DNA security using symmetric cryptography. *International journal of new computer architectures and their applications*, IJNCAA, Vol. 1 (2011)
19. Tong, X. J., Cui, M. G.: Image encryption scheme based on 3D baker with dynamical compound chaotic sequence cipher generator. *Signal Processing*, 89(4) pp. 480–491 (2009)
20. Wong, K. W., Kwok, B. S., Law, W. S.: A fast image encryption scheme based on chaotic standard map. *Phys. Lett. A.*, 372(15), pp. 2645–2652 (2008)
21. Xiao, Y. L., Xia, L. M.: An Image Encryption Approach Using a Shuffling Map. *Commun. Theor. Phys*, 52(5), pp. 876–880 (2009)

*Luis René Marcial Castillo, Erika Leonor Basurto Munguia, Marcela Rivera Martínez, et al.*

22. Xu, S. J., Wang, J. Z., Yang, S. X.: An improved image encryption algorithm based on chaotic maps. *Chin. Phys. B*, 17(11), pp. 4027–4032 (2008)
23. Ye, G. D.: A chaotic image cryptosystem based on Toeplitz and Hankel matrices. *Imaging Sci. J.*, 57(5), pp. 266–273 (2009)
24. Octave: Disponible en <https://www.gnu.org/software/octave> (2016)

## Introducción al problema inverso electrocardiográfico

Esteban Herrera Hernandez<sup>1</sup>, Rafael Lemuz López<sup>1</sup>, Carlos Guillén Galván<sup>1</sup>,  
Ángel Ramos del Olmo<sup>2</sup>

<sup>1</sup> Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

<sup>2</sup>, Universidad Complutense de Madrid, Departamento de Matemática Aplicada  
Madrid, España

artanis\_ii@hotmail.com, rlemuz@cs.buap.mx, cguillen@fcfm.buap.mx, angel@mat.ucm.es

**Resumen.** En este trabajo se presenta una propuesta de adquisición de datos para el problema inverso en electrocardiografía utilizando un sistema de adquisición inalámbrico de bajo costo. Obtener una imagen de la actividad eléctrica en la superficie del corazón resolviendo el problema inverso es importante en el estudio y diagnóstico de enfermedades cardiacas. Además, se presenta una introducción al planteamiento matemático del problema describiendo dos métodos de regularización: el método de Tikhonov y la Descomposición de Valores Singulares Truncados (TSVD).

**Palabras clave:** Adquisición de datos, ecuaciones diferenciales parciales, problema inverso, electrocardiograma (ECG), EPOC headset.

## Introduction to the Inverse Electrocardiographic Problem

**Abstract.** In this work, a data acquisition approach for the inverse problem in electrocardiography using a low cost wireless acquisition system is presented. Obtaining an image of the electrical activity on the heart surface is important in the study and diagnosis of heart disease. In addition, an introduction to the mathematical problem is presented describing two regularization methods: the Tikhonov method and the Truncated Singular Values Decomposition (TSVD).

**Key words:** Data acquisition, partial differential equations, inverse problem, electrocardiogram (ECG), EPOC headset.

## **1. Introducción**

Las enfermedades cardiovasculares (ECV) son la principal causa de muerte en todo el mundo. Cada año mueren más personas por ECV que por cualquier otra causa. En México es el segundo lugar como causa de muerte, durante el año 2014 se registraron 82,334 decesos por enfermedades isquémicas del corazón lo cual es preocupante [4], ya que por nuestro ritmo de vida esto se incrementa año con año y muchos de los estudios que existen para diagnosticar este tipo de enfermedades son excesivamente caros al igual que sus tratamientos por lo que, resolviendo el problema inverso electrocardiográfico podría tener un impacto muy fuerte en la prevención de enfermedades cardíacas.

El problema inverso electrocardiográfico consiste en obtener potenciales miocárdicos a partir de los registros de superficie de manera no invasiva. Los registros superficiales se obtienen con la técnica de BSPM (Body Surface Potential Mapping) con alrededor de 32 a 300 electrodos alrededor del torso. Con ello se pretende encontrar regiones del corazón donde están ocurriendo patologías cardíacas, como son el bloqueo de rama, insuficiencia cardíaca, infarto del miocardio, potenciales tardíos ventriculares o el síndrome de Brugada, algunas de ellas causantes de la denominada muerte súbita.

En la literatura se dice que el problema inverso electrocardiográfico está mal planteado resultando en múltiples soluciones posibles e inestabilidad. Por tanto, es necesario el desarrollo de técnicas y algoritmos nuevos que simplifiquen y permitan aplicar estos resultados a una mayor cantidad de población.

En condiciones normales, existe una señal de naturaleza eléctrica que recorre el corazón de forma periódica y secuencialmente ordenada, que permite un bombeo eficaz de la sangre al resto de organismos del cuerpo a través del sistema circulatorio.

El corazón es una parte importante de este sistema, y suministra todos los demás órganos con sangre oxigenada. Este órgano consiste en cuatro cámaras, la aurícula izquierda y la derecha y el ventrículo izquierdo y derecho. La contracción, generada por un estímulo, provoca el bombeo de sangre a través de la circulación.

Este estímulo surge en una región llamada el nódulo sinusal situado en la aurícula derecha. El impulso llega a las células musculares por el sistema de conducción del corazón. La corriente eléctrica resultante controla la contracción del corazón y se puede medir en la superficie del cuerpo. Si esta corriente eléctrica se ve alterada puede provocar una arritmia.

## **2. Métodos de medición**

Existen dos métodos para medir el potencial eléctrico, unipolar y bipolar. Para medir un potencial eléctrico, es necesario utilizar al menos dos electrodos. En la medición unipolar, un electrodo se coloca cerca del corazón para medir la señal y el otro electrodo se coloca lejos del corazón como referencia. En la medición bipolar, dos electrodos tienen aproximadamente la misma distancia desde el corazón.

El estándar de 12 derivaciones de ECG, es un sistema que incluye tanto conductores bipolares como unipolares y es estándar en la práctica clínica actual. En este caso diez electrodos se colocan en el cuerpo y 2 son referencias. En ellas se registran doce señales eléctricas diferentes en aproximadamente al mismo tiempo [2]. Mientras más

electrodos superficiales se utilicen mejores resultados se obtendrán. Los electrodos se colocan directamente sobre la superficie del cuerpo y son capaces de tomar registros sobre la actividad eléctrica. Los materiales de los que se constituyen la mayoría de los electrodos son muy variables, entre ellos están la plata, el oro, el acero inoxidable, platino, entre otros [3].

### 3. Problema inverso

Para poder obtener datos y resolver el problema inverso en electrocardiografía, tenemos que interpretar los potenciales registrados en la superficie del cuerpo debido a la actividad eléctrica del corazón. El objetivo general de los problemas "inversos" llamados "hacia adelante" de electrocardiografía es una mejor comprensión cualitativa y cuantitativa de la actividad eléctrica del corazón [1].

Para poder medir la actividad eléctrica en la superficie del corazón utilizaremos un sistema de adquisición de bajo costo el cual es el EPOC Headset, ya que este sistema cuenta con sensores de humedad a base de solución salina (sin geles pegajosos) y 14 canales de EEG, más 2 canales de referencia y nos ofrece un posicionamiento óptimo para la adquisición de señales, es inalámbrico y recargable.

Este sistema nos ayudara a obtener la información necesaria sobre las señales P,T y el complejo QRS que mencionamos anteriormente, a través de electrodos superficiales puestos en 4 partes del torso, de los cuales 2, estarán cerca del brazo derecho y del brazo izquierdo, otro electrodo irá junto a la pierna izquierda y por último el electrodo que va a estar junto a la pierna derecha será nuestra referencia, el resto de los electrodos se colocarán para medir la actividad eléctrica de los ventrículos ver figura 1. Con esto se podremos obtener los datos necesarios para poder resolver el problema inverso mediante ecuaciones parciales y la descomposición de valores singulares regularizados y así obtener el mapeo de la actividad eléctrica en la superficie del corazón.



**Fig. 1.** 12 Derivaciones de ECG y EPOC Headset.

Se sabe que cuando existe alguna alteración en las señales P, T y el complejo QRS, el paciente puede sufrir de alguna enfermedad del corazón. En la figura 2 se muestra

los signos vitales de un corazón con ritmo sinusal, taquicardia ventricular, fibrilación auricular y fibrilación ventricular.



Fig. 2. Enfermedades del corazón.

#### 4. Métodos matemáticos

Para comprender el enfoque matemático del problema inverso de electrocardiografía se requiere un conocimiento básico. Aquí explicaremos algunas definiciones y teoremas, que se necesitan saber para manejar un problema mal planteado.

**Definición:** Un modelo matemático es un mapeo

$$A: X \rightarrow Y. \tag{1}$$

Del conjunto de causas (parámetros)  $X$  para el conjunto de la dirección (datos)  $Y$ . En el problema directo el efecto se calcula en función de la causa, es decir, para  $x \in X$  calculamos  $Ax \in Y$ . A la inversa, en el problema inverso tenemos que observar  $y \in Y$  y tratar de encontrar una  $x \in X$  cuando  $Ax = y$ .

La siguiente definición debido a Hadamard es una clasificación de los problemas bien planteados y mal planteados.

**Definición:** Sea  $A: X \rightarrow Y$  un mapeo entre espacios topológicos  $X$  y  $Y$ . El problema  $(A, X, Y)$  se llama bien planteado, si las siguientes condiciones son cumplidas:

- a) Para toda  $y \in Y$  la ecuación  $Ax = y$  es solución de  $x \in X$ .
- b) La solución es única.
- c) La aplicación inversa  $A^{-1}: Y \rightarrow X$  es continua, es decir la solución de  $x$  depende de forma continua de los datos en  $y$  (pequeños errores en  $y$  harán pequeños errores en  $x$ ).

#### 4.1. Inversa generalizada

Las dificultades en la resolución de un problema mal planteado son obvias. No podemos asumir que una solución es única, o que siquiera existe. Por lo tanto, tenemos que encontrar una manera de satisfacer las condiciones a) y c).

En lo que sigue, suponemos que  $X$  y  $Y$  sean espacios de Hilbert y  $A$  para ser un operador lineal,  $A \in L(X, Y)$ . Analizamos la solución de la ecuación [10]:

$$Ax = y. \quad (2)$$

Con un determinado arbitrario  $y \in Y$ . Si el rango de  $A: R(A)$  no es el espacio completo de la imagen, no es solucionable para  $y \in Y$ . Por otro lado, si  $A$  tiene un espacio nulo trivial  $N(A)$ , la solución no es única. Así que si hay varias soluciones, es razonable elegir el que tenga la norma mínima entre ellos. Este pensamiento lleva a la siguiente definición:

Un elemento  $x \in X$  es llamado:

- a) Mínimos cuadrados solución de la ecuación  $Ax = y$  si

$$\|Ax - y\|_Y = \inf\{\|Az - y\|_Y \mid z \in X\}. \quad (3)$$

- b) Solución de norma mínima de la ecuación  $Ax = y$  si

$$\|x\|_X = \inf\{\|z\|_X \mid z \text{ es una solución de mínimos cuadrados de } Ax = y\}. \quad (4)$$

Todavía no podemos asegurar que existe una solución de mínimos cuadrados para una  $y$  arbitraria, ya que  $R(A)$  no necesita ser cerrado. Pero por ahora, esto se puede suponer. En este caso, la solución de norma mínima es única, debido a la convexidad de la minimización relativa funcional. A continuación, se define la inversa generalizada Moore-Penrose, con la que podemos calcular la solución de la norma mínima [2].

**Definición:** Sea  $A$  un continuo funcional lineal entre espacios de Hilbert  $X$  y  $Y$ , es decir,  $A \in L(X, Y)$ . Sea  $A: N(A)^\perp \rightarrow R(A)$  sea su restricción. A continuación, la inversa generalizada de Moore-Penrose  $A^\dagger$  se define como la extensión lineal única de  $A^{-1}$  a su dominio

$$D(A^\dagger) := R(A) \oplus R(A)^\perp. \quad (5)$$

Con

$$N(A^\dagger) = R(A)^\perp. \text{ Así } (A^\dagger) \text{ son mapas } D(A^\dagger) \text{ para } N(A)^\perp. \quad (6)$$

Como se ha mencionado antes, la inversa generalizada se puede utilizar para calcular la solución mínima. Esto lo podemos ver en el siguiente teorema.

**Teorema:** La ecuación  $Ax = y$  tiene una solución única por la norma-mínima dada por

$$x^\dagger := A^\dagger y \quad \forall y \in D(A^\dagger). \quad (7)$$

El conjunto de todas las soluciones de mínimos cuadrados está dada por  $\{x^\dagger\} + N(A)$ . Con este Teorema, podemos encontrar una solución única para la ecuación  $Ax = y$  si  $y \in D(A)$ . Por lo tanto, las condiciones a) y b) de la definición anterior son

cumplidas. Si  $R(A)$  no está cerrado y por lo tanto no existe  $y \in Y \setminus D(A^\dagger)$ , porque se viola la tercera condición. Hay algunas técnicas de regularización, que ayudan a manejar este caso [2].

## 5. Técnicas de regularización

Como ya se ha mencionado, el problema inverso no posee una solución única y como hemos visto el problema inverso está mal planteado, es decir, el operador  $R^{-1}$  no tiene límites, por lo tanto, los pequeños errores en los datos medidos  $d$  pueden ser críticos e introducir grandes errores en  $g$ . Por esa razón, tenemos que encontrar una manera de manejar esta dificultad.

Hay algunas técnicas de regularización, que se aproximan al problema mal planteado mediante el uso de ecuaciones bien planteadas. A continuación, se presentan la Regularización de Tikhonov y la descomposición de Valores Singulares Truncados (TSVD) [10].

### 5.2. Regularización de Tikhonov

Este método de regularización se inspira en la formulación discreta de mínimos cuadrados [10]. Por esta razón, se introduce la función

$$J(g) = \|R(g) - d\|_{L^2(\sigma_1)}^2 \text{ para } g \in H^1(\sigma_3),$$

en la que usamos la misma denotación como antes.

Ahora volvemos a escribir la ecuación  $R(g) = d$  como el problema de minimización

$$\min_{g \in H^1(\sigma_3)} J(g). \quad (8)$$

Por supuesto, mediante la formulación del problema inverso de esta forma, el "mal planteado" aún está presente. Con el siguiente paso, nos aproximamos a (8) por un problema de buen comportamiento. Se define una nueva función  $j \in (g)$  mediante la adición de un término de regularización a  $j(g)$ :

$$J_\epsilon(g) = J(g) + \epsilon \|g\|_{L^2(\sigma_3)}^2 = \|R(g) - d\|_{L^2(\sigma_1)}^2 + \epsilon \|g\|_{L^2(\sigma_3)}^2 \quad (9)$$

para  $g \in H^1(\sigma_3)$ ,

con  $\epsilon > 0$  como el parámetro de regularización. Con (9) se formula una aproximación de (8):

$$\min_{g \in H^1(\sigma_3)} J_\epsilon(g). \quad (10)$$

Esto es un compromiso entre minimizar la norma de la residual y mantener  $g$  pequeña. Por último, se demuestra que la adición del término de regularización a nuestra función original hace cumplir la estabilidad.

Recordamos que  $g \in H^1(\sigma_3)$  se puede escribir en la forma

$$g(x) = \sum_{K=0}^{\infty} p_k \cos(k\pi x) \text{ para la adecuada } p_0, p_1, \dots \in \mathbb{R}.$$

Por otra parte con  $J_\epsilon = J_\epsilon(p_0, p_1, \dots)$ .

Y desde el primer orden la condición necesaria

$$\frac{\partial J_\epsilon}{\partial P_i} = 0 \text{ para } i = 1, 2, \dots,$$

Durante un mínimo, se deriva la formula

$$p_k = \frac{d_k \cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)},$$

para los coeficientes de Fourier de la solución  $g$  de (10), véase [9]. Por lo tanto, podemos concluir, que

$$g(x) = \sum_{K=0}^{\infty} \frac{d_K \cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)} \cos(k\pi x).$$

Resuelve (10):

$$d_k = 2 \int_{\epsilon} d(x) \cos(k\pi x) dx, \text{ para } k = 1, 2, \dots,$$

$$d_0 = \int_{\epsilon} d(x) dx.$$

siendo los coeficientes de Fourier los datos de medición  $d$ . Ahora obtenemos la siguiente aproximación  $R_\epsilon^{-1}$  del operador inverso  $R^{-1}$  del mapeo de  $R$ :

$$R_\epsilon^{-1}(d) = R_\epsilon^{-1} \left( \sum_{K=0}^{\infty} d_K \cosh(k\pi x) \right) = \sum_{K=0}^{\infty} d_K \frac{\cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)} \cos(k\pi x).$$

Y con

$$\lambda_K = \frac{1}{\cos(k\pi)} \text{ para } k = 0, 1, \dots$$

Siendo los valores propios de  $R$ , obtenemos:

$$R_\epsilon^{-1} = \sum_{K=0}^{\infty} d_K \frac{\lambda_k}{\lambda_k^2 + \epsilon} \cos(k\pi x). \tag{11}$$

Si comparamos  $R^{-1}$  y el  $R_\epsilon^{-1}$  aproximado, vemos que, por una parte, para los componentes de baja frecuencia de  $d$ , la acción de  $R^{-1}$  y  $R_\epsilon^{-1}$  son casi idénticos. Por otra parte, los componentes de alta frecuencia de  $d$  se amortiguan de manera eficiente por  $R_\epsilon^{-1}$ . Además de eso, podemos concluir, que los pares de valores propios de la función propia del  $R_\epsilon^{-1}$  son:

$$\lambda_{\epsilon,k}^{-1} = \frac{\cosh(k\pi)}{1 + \epsilon \cosh^2(k\pi)}, \epsilon_{\epsilon,k} = \cos(k\pi x) \text{ para } k = 0, 1, \dots \quad (12)$$

Mediante el uso de este, se muestra que  $R_\epsilon^{-1}$  es un operador lineal acotado para cada  $\epsilon > 0$  es fijo. Vemos que para  $k = 0, 1, \dots$

$$\lambda_{\epsilon,k}^{-1} = \frac{\cosh(k\pi x)}{1 + \epsilon \cosh^2(k\pi)} \leq \frac{\epsilon^{-1}(1 + \epsilon \cosh^2(k\pi))}{(1 + \epsilon \cosh^2(k\pi))} = \frac{1}{\epsilon},$$

$$\lambda_{\epsilon,k}^{-1} > 0.$$

Y de eso, podemos concluir

$$\lambda_{\epsilon,k} \geq \epsilon > 0.$$

En consecuencia, el cero no es un punto adherente de  $\lambda_\epsilon$  valores propios de  $R_\epsilon$ . El resultante  $R_\epsilon^{-1}$  es entonces un operador lineal acotado y por lo tanto continua, a condición de que  $\epsilon > 0$ . De esta manera, el problema (10) está bien planteado para cada  $\epsilon > 0$ . Todavía hay un reto en encontrar el equilibrio adecuado entre el tamaño de  $\epsilon$  y el orden de los errores presentes en  $d$ . Nosotros no nos extenderemos sobre este tema. Un ejemplo referente a esta materia se puede encontrar en [9].

### 5.3. Descomposición de valores singulares truncados

Otro método de regularización es la descomposición de valores singulares truncados (TSVD). La idea de esta técnica es la obtención de un nuevo problema con una matriz de transferencia deficiente con un rango bien acondicionado.

#### Matriz de transferencia

Sea  $R$  el mapeo se define en  $R: H^1(\sigma_3) \rightarrow L^2(\sigma_1)$  y  $R(g) = u(g)|_{\sigma_1}$ . También podemos decir que  $R$  es una función de transferencia entre la fuente y el  $g$  desconocido (ruido dañado) de superficie corporal potencial  $d$ . Es posible construir una matriz de transferencia explícita  $A$  de  $g$  a  $d$ , es decir,

$$Ag = d.$$

Ahora sólo discutiremos algunas propiedades de  $A$  y las utilizaremos para resolver el problema inverso.

#### Descomposición de valores singulares

Antes de pasar a TSVD, es necesario entender la teoría detrás de este método de regularización. La Descomposición en valores singulares (SVD) es una técnica para factorizar una matriz para determinar los componentes principales de la información que contiene.

Si utilizamos un SVD de la matriz de transferencia  $A$ , obtenemos la siguiente fórmula:

$$A = U\Sigma V^T, \quad (13)$$

con  $U$  y  $V$  de ser matrices ortogonales y  $\Sigma$ , una matriz diagonal, que contienen los valores singulares. Es obvio que:

$$A^T = U\Sigma V^T.$$

es un SVD para  $A^T$ . También podemos repetir la fórmula (13):

$$A \sum_{i=1}^N \sigma_i u_i v_i^t, \tag{14}$$

cuando  $\sigma_i$ , son los valores singulares para  $i = 1, \dots, N$  y  $u_i$  y  $v_i$  son vectores que forman las matrices ortogonales. Por lo tanto,  $A$  y  $A^T$  tienen los mismos valores singulares. Por lo general, los valores singulares son ordenados en orden decreciente de tamaño y si  $\sigma_1 \geq \sigma_2 \geq \dots \sigma_r > \sigma_{r+1} = \dots = \sigma_N = 0$ , podemos concluir que  $R(A) = r$ .

Ahora describiremos los vectores  $u_i$  y  $v_i$  con más detalle. Ellos pueden ser separados en dos grupos. Los vectores  $u_i$  y  $v_i$ , que corresponden a pequeños valores singulares, a menudo se descartan, porque se supone otorgar al espacio ruido aleatorio. Los vectores  $u_i$  y  $v_i$  que permanecen, son los principales componentes y también se conoce como los vectores propios. Determinamos, que valores singulares corresponden al espacio de ruido por el trazado de los valores singulares ordenados contra el rango de la matriz obtenida a partir del subconjunto apropiado de esos valores singulares.

**Descomposición de valores singulares truncados (TSVD)**

Ahora pasamos a la técnica de regularización TSVD. Sabemos que  $R$  describe un inverso y un problema mal planteado. Por lo tanto, la matriz de transferencia  $A$  está mal acondicionada, por lo tanto, nuestro objetivo es derivar un nuevo problema con una matriz de transferencia bien acondicionada. Para obtener una matriz de transferencia bien acondicionada, necesitamos omitir los valores singulares de  $A$  que contienen información relacionada con el ruido y los errores. Una matriz deficiente en rango, que no contenga esta información, se puede obtener una SVD de la matriz de transferencia  $A$ , es decir,

$$A_w(y, x) = \sum_{n=1}^w \sigma_n \sigma_n \sigma_n, \quad w \leq N. \tag{15}$$

Cuando  $u_n$  y  $v_n$  son los vectores  $n$  de un SVD de  $A$ , y  $n$  son los valores singulares correspondientes. El rango completo de la matriz es  $N$  y el tamaño de  $w > 0$  determina el nivel de regularización. La solución  $g$  de TSVD se puede calcular mediante la minimización

$$||A_w g - d||. \tag{16}$$

Sabemos que la solución de este problema de minimización está dada por la pseudo inversa de  $A_w$ :

$$A_w^\dagger = (A_w^T A_w)^{-1} A_w^T.$$

Escribimos  $A_w^\dagger$  en términos de una SVD por

$$A_w^\dagger = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}, \quad (17)$$

con  $\tilde{U}$  y  $\tilde{V}$  que contiene el primer  $w$  columnas de  $U$  y  $V$ , respectivamente.  $\tilde{\Sigma}$  es a  $w \times w$  una submatriz diagonal de  $\Sigma$

Para obtener la solución regularizada, calculamos

$$g = A_w^\dagger(d). \quad (18)$$

Si insertamos la ecuación (17) en (18), encontramos

$$g = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^T(d) = \sum_{n=1}^w \frac{\langle u_n, d \rangle}{\sigma_n} u_n,$$

donde  $\langle \cdot, \cdot \rangle$  denota el producto interno.

## 6. Conclusiones

En este trabajo se presentó una introducción al problema inverso en electrocardiografía y se describió como se puede resolver mediante el método de regularización de Tikhonov y la TSVD. Al mismo tiempo se mostró el diseño de un sistema de adquisición de señales electrocardiográficas a través de un sistema de adquisición de bajo costo para obtener las 12 derivaciones que en un trabajo futuro permitirán estudiar problemas asociados a algunas cardiopatías y síndromes que se manifiestan en el ventrículo izquierdo. El principal resultado del trabajo es la propuesta de diseño del sistema de captura de bajo costo de señales ECG. En un trabajo futuro se utilizará el sistema propuesto para resolver numéricamente el problema inverso electrocardiográfico que podrá servir para concientizar a la población de algunas de las principales enfermedades cardiacas y así poder ayudar en el entrenamiento médico. Además, se estudiará la estabilidad de ambos métodos de regularización.

## Referencias

1. Gulrajani, R. M.: The forward and inverse problems of electrocardiography. IEEE Engineering in Medicine and Biology Magazine, 17(5), pp. 84–101 (1998)
2. Kochinke, J.: Einfluss von Normierung auf die Analyse von EKG-BSPM Daten (2012)
3. Correa, J. L., Morales, E., Huerta, J. A., González, J. J., Cárdenas, C. R.: Sistema de Adquisición de Señales SEMG para la Detección de Fatiga Muscular. Revista Mexicana de Ingeniería Biomédica, 37(1), pp. 17–27 (2016)
4. Aguirre Botello, M.: Recopilación con datos de INEGI, OMS y SINAIS, México, Principales Causas de Mortalidad 1938–2014. Recuperado de: <http://www.mexicomaxico.org/Voto/MortalidadCausas.htm>
5. Brooks, D. H., Ahmad, G., MacLeod, R. S.: Multiply constrained inverse electrocardiography: combining temporal, multiple spatial, and interative regularization. In: Proc 16th Ann IntConf IEEE Eng Med Biol Soc, IEEE Press, NewYork, pp 137–138 (1994)

6. Goldman, L.: Approach to the patient with possible cardiovascular disease. In: Goldman L, Schafer A.I., eds., *Goldman's Cecil Medicine*, 25th ed. Philadelphia, PA: Elsevier Saunders, chap. 51 (2016)
7. Zemzemi, N.: An iterative method for solving the inverse problem in electrocardiography in normal and fibrillation conditions: A simulation Study. In: *International Congress on Electrocardiology, ICE*, pp. 53–56 (2014)
8. Berrier, K. L., Sorensen, D. C., Khoury, D. S.: Solving the inverse problem of electrocardiography using a Duncan and Horn formulation of the Kalman filter. *IEEE Transactions on biomedical engineering*, 51(3), pp. 507–515 (2004)
9. Lines, G. T., Cai, X., Nielsen, B. F., Mardal, K. A., Tveito, A.: *Computing the Electrical Activity in the Heart*. Monographs in Computational Science and Engineering, Springer (2006)
10. Kochinke, J.: *Einfluss von Normierung auf die Analyse von EKG-BSPM Daten* (2013)



## **Propuesta de arquitectura para un módulo de inteligencia de negocios basado en minería de datos**

Yessica Thalia Apale Lara, Beatriz Alejandra Olivares Zepahua,  
Lisbeth Rodríguez Mazahua, Giner Alor Hernandez, Hilarión Muñoz Contreras

Instituto Tecnológico de Orizaba, División de Estudios de Posgrado e Investigación,  
Orizaba, Veracruz, México

yapalelara@acm.org, bolivares@ito-depi.edu.mx,  
hmunozc189@msn.com, {lrodriguez,galor}@itorizaba.edu.mx

**Resumen.** Los constantes cambios en las preferencias de los viajeros, nacionales e internacionales influyen en el rendimiento de la industria hotelera creando así la necesidad de que los gerentes hoteleros continúen en la búsqueda de métodos que ayuden a predecir las preferencias y necesidades de los viajeros con el objetivo de obtener mejoras en su planificación y desarrollo de servicios estratégicos. Como solución a esta problemática el presente trabajo propone una arquitectura para el desarrollo e integración de un módulo de Inteligencia de Negocios en el sistema hotelero Softcalli, para que aplicando el proceso KDD se genere un modelo de Minería de Datos que permita identificar grupos de clientes con el fin de realizar la predicción de los servicios más idóneos de acuerdo al perfil de un nuevo cliente.

**Palabras clave:** Agrupamiento, hotel, inteligencia de negocios, KDD, minería de datos, WEKA.

## **Proposal of an Architecture for a Business Intelligence Module Based on Data Mining**

**Abstract.** The constant changes in the preferences of national and international travelers influences the performance of the hotel industry thus creating the need for hotel managers to continue in the search of methods that help to predict the preferences and necessities of the travelers with the objective to obtain improvements in their planning and the development of strategic services. As a solution to this problem, the present work proposes an architecture for the development and integration of a Business Intelligence module in Softcalli hotel system, so that by applying the KDD process, a Data Mining model will be generated to identify groups of clients in order to predict the most suitable services according to the profile of a new customer.

**Keywords:** Clustering, hotel, Business Intelligence, KDD, data mining, Weka.

## **1. Introducción**

La industria hotelera juega un papel muy importante dentro del entorno económico para la mayoría de los países en función de la demanda turística de cada país. La función principal de un Hotel es dar cabida a aquellos que se encuentran fuera de casa. Para muchos países los hoteles son un medio que ayuda a contribuir de manera significativa a la economía local, ya que son una fuente de ingresos en función de la demanda turística de cada país, por lo tanto, es conveniente que logre una ventaja competitiva ofreciendo los servicios de calidad que esperan los turistas [1].

El Banco de México registró que la llegada de turistas nacionales en México, en el mes de noviembre del 2015, fue de 7.6 millones, lo cual representa un incremento de 0.5 millones de viajeros nacionales respecto a noviembre de 2014, que corresponde a un crecimiento del 6.6%. Así mismo, la llegada de turistas internacionales a México, en noviembre del 2015 fue de 2.8 millones, lo que representa un incremento aproximado de 0.2 millones en comparación con noviembre del 2014, es decir 9.0% más.

Estos datos revelan un rápido crecimiento del Turismo en México y, como resultado, una fuerte demanda para el sector hotelero. En el Sistema Nacional de la Información Estadística del Sector Turismo de México (DATATUR) se establece que el porcentaje de ocupación hotelera en 70 centros turísticos, durante el periodo enero-noviembre de 2015, fue superior en 2.50 puntos porcentuales respecto al mismo periodo del año anterior, llegando a 59.6%. En enero-noviembre de 2015, la llegada de turistas nacionales a cuartos de hotel superó los 47.9 millones de turistas (76.2%); mientras que 15.0 millones fueron turistas extranjeros (23.8%) [2].

Uno de los problemas con los que se encuentran los gerentes hoteleros, en función de la alta demanda turística, es que cada vez es más difícil saber cuáles son las preferencias, necesidades y los servicios que esperan los turistas. Actualmente los hoteles en México cuentan con sistemas que permiten llevar la administración del hotel, almacenando información valiosa de las transacciones que se realizan dentro del hotel como son las reservaciones de habitaciones y los servicios extras que solicitan los huéspedes; esta información resultaría útil para solucionar los problemas de negocio si se contase con herramientas que identificaran comportamientos comunes en los huéspedes tal que el personal del hotel contase con una ayuda para ofrecer mejores servicios a sus clientes.

Microcalli del Golfo S.A de C.V, es una empresa que se dedica al desarrollo de sistemas hoteleros, Softcalli es un sistema hotelero integral desarrollado por esta empresa. El sistema Softcalli tiene como función principal administrar el hotel, almacenando altas cantidades de información relevante de los clientes, registra y administra en las bases de datos las reservaciones que realizan los clientes, también realiza el registros de información detallada de los clientes, consumos y servicios entre otros. Un problema detectado en Softcalli es que no cuenta con un módulo de Inteligencia de Negocios que sirva de apoyo para mejorar las estrategias del negocio hotelero.

Como solución a los problemas antes mencionados se propone una arquitectura para el desarrollo e integración al sistema Softcalli de un módulo de Inteligencia de Negocios, empleando técnicas de Minería de Datos como lo son la clasificación y el agrupamiento. Debido a que los hoteles no tienen identificados a sus tipos de clientes,

se aplicarán técnicas de agrupamiento (*clustering*) para formar conjuntos de clientes de acuerdo a una lista de características (variables relacionadas con consumo de servicios y reservaciones), seleccionadas con apoyo de un experto del negocio; una vez que el experto del negocio acepte los grupos identificados, se procederá a etiquetar a los clientes con dichos grupos. Después se aplicarán técnicas de clasificación para obtener un modelo capaz de predecir a qué grupo pertenece un nuevo cliente, de esta forma se conocerá el comportamiento en el consumo de los servicios y las características de la reservación que tendrá dicho cliente.

Este trabajo está estructurado de la siguiente forma: en la sección 2 se mencionan una serie de trabajos relacionados al tema propuesto, en la sección 3 se describen algunos conceptos relevantes a abordar en el desarrollo de este trabajo, en la sección 4 se describe la Arquitectura de la propuesta de solución, en la sección 5 se presentan las conclusiones y los trabajos a futuro.

## **2. Estado del arte**

Con el objetivo de tener una investigación más detallada acerca de necesidades emergentes de los viajeros, en [3] utilizaron una técnica de Minería de Patrones Emergentes, este trabajo se centró en identificar los patrones emergentes de los viajeros internacionales. La Minería de Patrones Emergentes ayudó a descubrir los cambios y tendencias en la atención de los viajeros, se identificaron una serie de características de interés específicas y su usuario objetivo.

En [4], se presentó un estudio comparativo de 4 herramientas de acceso público para Minería de Datos: KNIME (*Konstanz Information Miner*), Orange, RapidMiner y WEKA (*Waikato Environment for Knowledge Analysis*, Entorno para Análisis del Conocimiento de la Universidad de Waikato). Se evaluó el desempeño de dichas herramientas con el fin de proporcionar a los analistas la herramienta y las técnicas a utilizar que arrojaran resultados con mayor precisión y de forma rápida. Los resultados obtenidos indicaron que no existe una sola herramienta o técnica que sea mejor que otra y sea aplicable a cualquier tarea de clasificación. Individualmente, el mejor resultado se logró con Weka y la técnica de árboles de decisión.

En [5] propusieron una novedosa técnica de clasificación *Neuro-fuzzy* para Minería de Datos, con el fin de analizar el método propuesto y comparar su rendimiento con dos potentes algoritmos de clasificación supervisados: Función de Base Radial de Redes Neuronales (RBFNN) y Sistema de Inferencia Adaptativa Neuro-difusa (ANFIS). El método propuesto utilizó una matriz "*fuzzification*". Se evaluó el rendimiento de estos métodos de clasificación en función de las diferentes medidas de desempeño tales como la precisión, el error de raíz cuadrática media, estadística kappa, falsos positivos, falsos negativos y precisión. El NFS (*Neuro-fuzzy system*, Sistema *Neuro-fuzzy*) propuesto fue capaz de mejorar significativamente los métodos de clasificación convencionales.

En [6] propusieron un marco de trabajo que investiga los dos modelos de clasificación más comunes que se utilizan en Minería de Datos para predecir el comportamiento de los clientes. El estudio se realizó en una campaña de mercadotecnia de un banco. Los datos se extrajeron del banco de datos de la Universidad de California en Irvine, los algoritmos evaluados fueron MLPNN (*Multilayer Perception Neural Network*) y NB (*Naïve Bayes*), estos algoritmos se evaluaron utilizando la herramienta

WEKA. Como resultado se encontró que el modelo que obtuvo el mejor rendimiento predictivo fue MLPNN con una tasa de precisión de 88.63%.

En [7] presentaron diferentes técnicas de Minería de Datos para estimar el costo del Software y se evaluaron los resultados de cada técnica. Los modelos propuestos fueron LR (*Linear Regression*), ANN (*Artificial Neural Network*), SRV (*Support Regression Vector*) y KNN (*K-Nearest Neighbors*). Los estudios realizados demostraron que el costo estimado del software dentro de estos modelos tuvo mayor velocidad y la precisión de los modelos algorítmicos fue alta. Las técnicas de Minería de Datos mejoran la precisión de la estimación de los modelos en muchos casos.

En [8], con el fin de transformar adecuadamente los datos en algo útil para el desarrollo de negocio inteligente desde la perspectiva de maximizar los beneficios de la empresa, se diseñó y desarrolló un marco universal para apoyo a las empresas inteligentes. Este marco se forma por un conjunto de técnicas de fusión y un motor de minería de datos basado en el concepto del modelo KID (*Knowledge Information Data*) el cual fue responsable de realizar la recolección universal de datos e información valiosa. El motor desarrollado recogió información significativa extraída e interpretada por algoritmos fusionados, y toda la información se convirtió en conocimiento valioso acerca de la situación de los clientes y las situaciones de los negocios. Este marco se desarrolló con el objetivo de obtener conocimiento, a partir de los datos, para el desarrollo de futuros negocios inteligentes. El desarrollo del motor de minería, basado en técnicas de fusión, es el componente esencial en el marco de negocios inteligentes.

En [9] propuso una extensión de la metodología de Minería de Opiniones Basadas en las Expresiones de Aspectos, termino introducido por Bing Liu, aplicada al ámbito del turismo, con el fin de descubrir las preferencias de los consumidores sobre los productos turísticos, en particular hoteles y restaurantes, obteniendo opiniones y críticas desde la Web. La contribución de la extensión consistió en el desarrollo de nuevas reglas para cubrir la aparición de aquellos aspectos compuestos por más de un término y para abarcar también los aspectos que aparecen más de una vez en una frase. Además se incluyó la formalización de conceptos y la creación de conjuntos de datos especiales, para la evaluación de las propuestas.

Como resultado del análisis realizado en el estado del arte se concluyó que el uso de la Minería de Datos, dentro de la industria hotelera, aplicada en el desarrollo de negocios inteligentes, es de gran ayuda como apoyo para la toma de decisiones, así como para el desarrollo de mejores estrategias de negocio, para así llevar a la empresa a un nivel competitivo más alto.

### **3. Conceptos**

A continuación se describen los conceptos relevantes para el mejor entendimiento del presente documento.

#### **3.1. Minería de Datos**

La Minería de Datos (MD) es una herramienta que ofrece un conjunto de técnicas para extraer conocimiento a partir de grandes cantidades de datos. La MD se define como “el proceso de descubrir patrones interesantes y conocimiento desde grandes cantidades

de datos” [10]. Las fuentes de datos que nutren a la MD incluyen bases de datos transaccionales, almacenes de datos, la Web y otros repositorios de información como son archivos de texto, hojas de cálculo, archivos XML (*eXtensible Markup Language*, Lenguaje de Marcas Extensible) o datos que se transmiten al sistema dinámicamente.

### 3.2. Proceso KDD

El proceso KDD (*Knowledge Discovery from Data*, Descubrimiento de Conocimiento a partir de los Datos), es el proceso de extracción no trivial de información, previamente desconocida y potencialmente útil, a partir de los datos [10]. El proceso KDD implica una secuencia iterativa de una serie de pasos, en la figura 1 se presentan los pasos del proceso KDD.

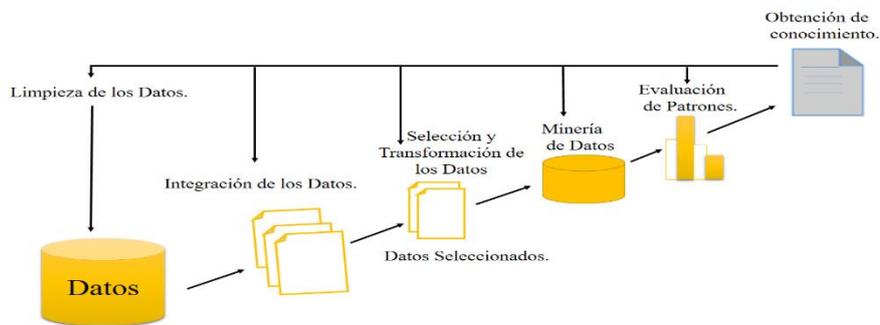


Fig. 1. Pasos del Proceso KDD.

### 3.3. Inteligencia de negocios

Se conoce como BI (*Business Intelligence*, Inteligencia de Negocios) al conjunto de estrategias, acciones y herramientas enfocada la administración y creación de conocimiento mediante el análisis de datos existentes dentro de una organización o empresa. Las operaciones que se definen como BI son aquellas que se involucran con los sistemas de información y con el desarrollo de métodos que generan un proceso de toma de decisiones inteligente [11].

### 3.4. Proceso ETL (Extracción, Transformación, y Carga)

ETL (*Extraction, Transformation & Load*, Extracción, Transformación y Carga), es el proceso que permite realizar la integración de la información en un Almacén de Datos (Data Warehouse, DW) o cualquier otro repositorio de datos. La Extracción se refiere a extraer la información de diferentes fuentes de información o Bases de Datos de distintos proveedores. En el proceso de Transformación, los datos extraídos pasan por varios procesos como son la discretización de los datos, limpieza de los datos, reemplazar valores faltantes, eliminación de ruido, entre otros procesos más, de modo que los datos se procesen correctamente y sean datos confiables. Por último, la Carga es el proceso de almacenar los datos ya transformados en el repositorio destino [12].

#### **4. Propuesta de arquitectura para la integración del módulo de inteligencia de negocios al sistema Softcalli**

Microcalli del Golfo es una empresa que desarrolla sistemas integrales para la administración de hoteles de cualquier tamaño. Muchos de sus clientes son pequeños hoteles que se encuentran en distintas ciudades del país y que no forman parte de grandes cadenas hoteleras, por lo que el presupuesto que destinan a la tecnología no es muy grande.

Softcalli es una suite de sistemas integrales hoteleros, desarrollada por la empresa Microcalli del Golfo, que cuenta con cuatro módulos con funciones específicas para cada área del hotel. Mk Front es el módulo que permite administrar las reservaciones del hotel, registrar clientes, realizar cambios de habitación, entre otras.

Esta suite se desarrolló con tecnología Microsoft .NET y utiliza una base de datos SQL Server. Con el fin de ofrecer los servicios más idóneos de acuerdo al perfil de cada tipo de turista se integrará a Softcalli un módulo de Inteligencia de Negocios que, aplicando técnicas de Minería de Datos, realice la predicción de los servicios que esperan los huéspedes.

Si bien SQL Server cuenta con herramientas propias para desarrollar Inteligencia de Negocio, específicamente Minería de Datos, dichas herramientas se venden por separado y tienen un costo muy alto, que en ocasiones lo hace prohibitivo para hoteles pequeños o mediados; por lo tanto, se ha considerado como alternativa el uso de la herramienta Weka, que es una aplicación y una biblioteca que contiene un gran número de algoritmos de Minería de Datos para diversas tareas como agrupamiento y clasificación entre otras; Weka trabaja bajo plataforma Java y que, en primer término, es incompatible de forma directa con la tecnología Microsoft.

La solución al manejo de plataformas heterogéneas (Microsoft y Java) será un módulo de Inteligencia de Negocios expuesto como WS (*Web Service*, Servicio Web) basado en SOAP (*Simple Object Access Protocol*, Protocolo Simple de Acceso a Objetos) que hará uso de la biblioteca Weka para aplicar algoritmos de agrupamiento y clasificación y así generar modelos de Minería de Datos que identifiquen los grupos de clientes específicos de un hotel y permitan predecir el comportamiento de nuevos clientes.

El módulo a desarrollar se alojará en un servidor de aplicaciones Glassfish y utilizará el lenguaje de programación Java. Por otra parte, se implementará en Softcalli un cliente del servicio Web como aplicación de escritorio, utilizando el lenguaje de programación C #, que consuma el WS para mostrar los resultados finales de la predicción.

En la figura 2 se muestra de forma general un esquema de solución para el desarrollo del módulo de Inteligencia de Negocios y su integración a Softcalli.

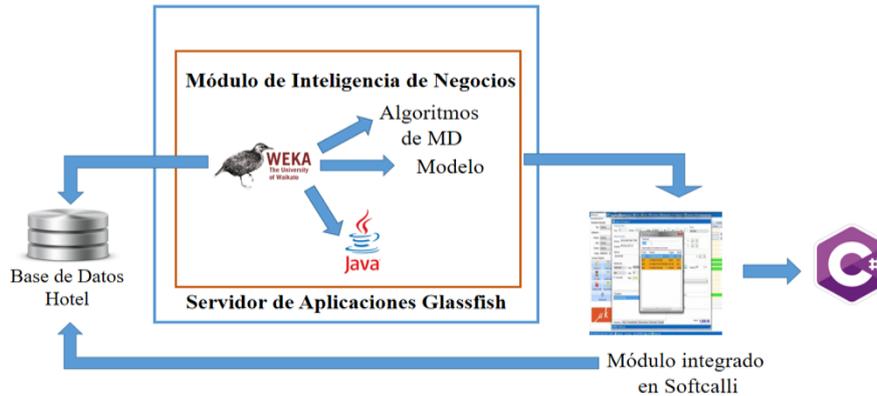


Fig. 2. Esquema de solución.

Es importante señalar que se cuenta con una base de datos (**Base de Datos Hotel** en la figura 2), que tiene almacenados registros históricos de las características y comportamientos de las reservaciones realizadas por los clientes; a partir de esa información, con la ayuda de los expertos de negocio, se identificaron, analizaron y seleccionaron aquellas variables que se consideraron importantes para realizar la identificación de los tipos de clientes y la predicción de los servicios.

En la figura 3 se muestra la estructura de la arquitectura planteada para el desarrollo de la aplicación y los componentes que la integran; esta arquitectura se basa en el Modelo Arquitectónico 4+1 Vistas y la figura corresponde a la vista de Despliegue. A continuación se describen los componentes de la arquitectura.

#### 4.1. Cliente Softcalli.exe

En el sistema Softcalli se implementará el Cliente encargado de consumir los servicios Web que entregará el Proveedor Web Java, será desarrollado bajo la tecnología C# y Microsoft Net Framework 4.5. El Cliente Softcalli mostrará al usuario los algoritmos disponibles para aplicar las técnicas de agrupamiento (*clustering*) y posteriormente las técnicas de clasificación (predicción). Por otro lado esta aplicación también permitirá introducir los valores necesarios para realizar y mostrar una lista con los resultados obtenidos de la predicción, para este caso mostrará una lista con las características de la reservación, consumos y servicios que ofrecerá el hotel a un cliente nuevo.

#### 4.2. Proveedor Web Java

Se trata del componente responsable de la Minería de Datos propiamente dicha, que hará llegar al cliente mediante un WS basado en SOAP. Entre sus elementos, el componente **Acceso a Datos** permite la comunicación con la **Base de Datos del hotel SQL** a través del uso de JDBC (*Java Database Connectivity*), el componente **Modelo** se auxiliará de él para acceder a los datos.

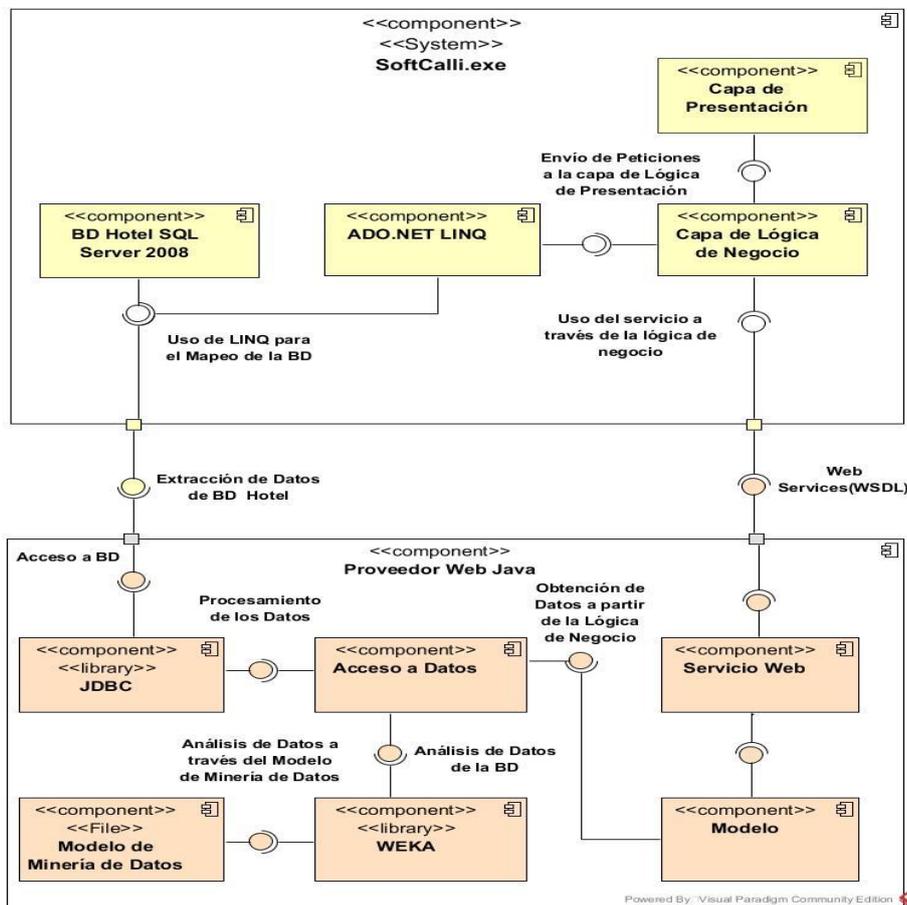


Fig.3. Diagrama de componentes.

En el **Modelo** se encuentran los Bean's que representan la lógica del módulo de Inteligencia de Negocio, donde se encontrarán a) el *ETL Bean*, que contendrá la lógica necesaria para aplicar el proceso ETL descrito en la sección anterior; b) el *Bean para el agrupamiento* y c) el *Bean para la clasificación* que implementarán la lógica necesaria para aplicar las técnicas de Minería de Datos con el objetivo de realizar la predicción de los servicios de acuerdo al perfil del cliente; por último d) el *Bean para aplicar algoritmos* y devolver las métricas de calidad asociadas para permitir al Gerente Hotelero realizar la elección de los algoritmos más óptimos para el proceso de Minería de Datos. Los algoritmos para las técnicas de agrupamiento y clasificación serán proporcionados por **WEKA**. La predicción se realizará en función de las diferentes variables de entrada que se hayan capturado en el módulo de Softcalli.

Para el caso de los algoritmos de agrupamiento se considerarán métricas de evaluación como la homogeneidad, el índice de confianza, la distribución del error y el índice David Bouldin. Para el caso de los algoritmos de clasificación, se considerarán métricas como el índice de certeza, la distribución del error y el error cuadrático. En

todos los casos, se presentarán tanto el valor de las métricas como su interpretación para que el gerente elija el algoritmo que dé los resultados que mejor reflejen la realidad de su negocio.

Para realizar la tarea de predicción, el Modelo de la Lógica de Negocio usará el componente **Modelo de Minería de Datos**, dicho componente contendrá un archivo con extensión .model que se obtiene como resultado de entrenar un conjunto de datos conformado por una lista de variables correspondientes a las características de los clientes cuya etiqueta de clase será el número de grupo, éstas variables son el país, estado y ciudad de procedencia, tarifa elegida, tipo de habitación, consumos, total de noches, total de días, origen de la reservación, si el huésped viene por asuntos de negocio, entre otras. El archivo .model se obtendrá bajo la aplicación del algoritmo que haya elegido el gerente hotelero. El Modelo de Minería de Datos contendrá el conocimiento de los patrones encontrados, y posteriormente se hará uso de él para realizar la predicción de ciertos servicios a ofrecer a un nuevo cliente; los servicios identificados son el tipo de habitación, tipo de tarifa, consumos como servicios de cuarto, restaurante, entre otros.

El Servicio Web proveerá todos los procesos necesarios que permitirán aplicar y regenerar, de ser necesario, los pasos del proceso KDD. Se propone el uso de Servicios Web para el desarrollo del módulo de Inteligencia de Negocios como mejor solución para la fácil integración del módulo con el sistema Softcalli. El servicio web se consumirá por un módulo integrado en Softcalli (Cliente) el cual como se mencionaba anteriormente está desarrollado en el lenguaje C#.

## **5. Conclusiones y trabajo a futuro**

El uso de Minería de Datos apoya a las empresas a plantear mejores soluciones de negocio a los problemas relacionados con ofrecer servicios óptimos a sus clientes; esto es particularmente importante en industrias como la hotelera, donde se busca “enamorar” a los huéspedes al brindarles opciones de acuerdo a sus intereses y presupuesto. Si bien en México existen herramientas computacionales como Softcalli, orientadas a apoyar la operación diaria en los hoteles, en la mayoría de los casos carecen de módulos de Inteligencia de Negocio que exploten todo el potencial de conocimiento que se encuentra almacenado ya en sus bases.

Por otro lado, herramientas específicas de Minería de Datos, como la suite de Inteligencia de Negocios de SQL Server son caras y no todos los hoteles cuentan con el presupuesto para aplicarlas, lo que los coloca en una posición de desventaja frente a grandes cadenas hoteleras; alternativas como Weka proveen una opción viable a este problema, pero presentan la problemática de trabajar en una plataforma distinta a la que utilizan sistemas como Softcalli. En este punto entran tecnologías como los Servicios Web que apoyan en la interoperabilidad de plataformas heterogéneas. El presente trabajo mostró una arquitectura que considera todo lo anterior para proporcionar un módulo de Inteligencia de Negocio que apoye a los hoteleros a identificar mejor las necesidades de sus huéspedes.

Como trabajo a futuro se tiene la implementación de la arquitectura presentada en este documento, desarrollando los componentes indicados para construir el módulo de inteligencia de negocios del Sistema Softcalli, donde se utilizará por una parte C# y

SQL Server y por otra Weka, Java y Glassfish. Además, se refinarán los elementos complementarios del modelo de arquitectura 4+1 vistas (casos de uso, lógica, despliegue y procesos).

**Agradecimientos.** Al Consejo Nacional de Ciencia y Tecnología, CONACyT, por el apoyo económico otorgado para la realización de esta investigación.

## Referencias

1. S. M. a. H. Ingram: *Staying Away from Home, The Business of Hotels*. New York, Elsevier Ltd, pp. 3–12 (2000)
2. Secretaría de Turismo (SECTUR). Available: <http://www.datatur.sectur.gob.mx/SitePages/versionesRAT.aspx> (2015)
3. Gang Li, Rob Law, Huy Quan Vu, Jia Rong, Xinyuan (Roy) Zhao: Identifying Emerging Hotel Preference Using Emerging Pattern Mining Technique. *Tourism Management*, Vol. 46, No. 1, pp. 311–321 (2015)
4. Borges, L. C., Marques, V. M., Bernardino, J.: Comparison of data mining techniques and tools for data classification. In: *Proceedings of the International C\* Conference on Computer Science and Software Engineering (C3S2E '13)*, Vol. 34, No. 1, pp. 113–116 (2013)
5. Ghosha, S., Biswas, S., Sarkar, D., Pratim Sarkar, P.: A novel neuro-fuzzy classification technique for data mining. *Egyptian Informatics Journal*, Vol. 15, No. 3, pp. 129–147 (2014)
6. Femina Bahari, T., Sudheep Elayidom, M.: An Efficient CRM-Data Mining Framework for the Prediction of Customer Behaviour. *Procedia Computer Science*, Vol. 46, No. 1, pp. 725–731 (2015)
7. Abbasi Khalifelum, Z., Soleimani Gharehchopogh, F.: Comparison and evaluation of data mining techniques with algorithmic models in software cost estimation. *Procedia Technology*, Vol. 1, No. 1, pp. 65–71 (2012)
10. Han, J., Kamber, M., Pei, J.: *Data Mining: Concepts and Techniques*. Waltham, USA, Elsevier Inc. (2012)
11. Ahumada Tello, A., Perusquia Velasco, J. M. A.: Business intelligence: Strategy for competitiveness development in technology-based firms. *Contaduría y Administración*, Vol. 61, No. 1, pp. 127–158 (2016)
12. Winda, R. T.: Extraction, Transformation, and Loading (ETL) Module for Hotspot Spatial Data Warehouse Using Geokettle. *Procedia Environmental Sciences*, Vol. 33, No. 1, pp. 626–634 (2016)

## Prototipo de un oxímetro de pulso con ESP8266 Wi-Fi

Gabriel Contreras Mota<sup>1</sup>, Rafael Lemuz López<sup>1</sup>, Carlos Guillén Galván<sup>2</sup>,  
Blanca Bermúdez Juárez<sup>1</sup>

<sup>1</sup> BUAP, Facultad de Ciencias de la Computación,  
Puebla, México

<sup>2</sup> BUAP, Facultad de Ciencias Físico Matemáticas,  
Puebla, México

gabriel.contreras.217@gmail.com, rafalemuz@cs.buap.mx, cguillen@cfm.buap.mx,  
bbj.solarium.cs.buap.mx

**Resumen.** En este trabajo se describe el diseño y construcción de un oxímetro de pulso con conexión WIFI. El dispositivo se conecta a un servidor que muestra la saturación de oxígeno y la frecuencia cardíaca. La información se visualiza en una aplicación web. El dispositivo se basa en el microcontrolador ESP8266 que establece la conexión inalámbrica y el envío de la información al servidor. En el artículo se describe el diseño del dispositivo y se muestran algunos resultados experimentales.

**Palabras clave:** Oxímetro de pulso, WiFi, microcontrolador ESP8266, aplicación web.

## Prototype of a Pulse Oximeter with ESP8266 Wi-Fi

**Abstract.** This paper describes the design and construction of a pulse oximeter with WIFI connection. This device connects to a server that shows oxygen saturation and heart rate. The information is displayed on a web application. The device is based on the ESP8266 microcontroller which establishes the wireless connection and sends the information to the server. In this article, the design of this device is described and some experimental results are shown.

**Keywords:** Pulse oximeter, WiFi, ESP8266 microcontroller, web application.

### 1. Introducción

El oxímetro de pulso es un dispositivo médico utilizado en el monitoreo no invasivo de la frecuencia cardíaca y el porcentaje de saturación de oxígeno (SpO<sub>2</sub>) en la sangre. La Oximetría de Pulso ha evolucionado en buena medida porque los dispositivos frecuentemente reducen su tamaño considerablemente permitiendo el desarrollo de

equipos portátiles e incrementando la distribución y su uso en el monitoreo del oxígeno en la sangre.

El oxígeno es uno de los elementos indispensables para la vida, el porcentaje normal en el cuerpo varía de 95% a 100% [1], teniendo algunas variaciones dependiendo de la altura sobre el nivel del mar en la que se encuentre, bajos niveles de oxigenación en la sangre generan una condición llamada Hipoxemia. La Hipoxemia puede ser provocada por diferentes causas, entre las más frecuentes se encuentran las enfermedades pulmonares o problemas cardíacos, una atención rápida es crucial para la salud del paciente, el uso del Oxímetro de Pulso proporciona una medición en tiempo real del estado del paciente facilitando su tratamiento.

La Oximetría es ampliamente utilizada en las intervenciones quirúrgicas, específicamente en el control de la anestesia, en el diagnóstico y tratamiento de enfermedades de obstrucción pulmonar crónica (EPOC), el control de la gravedad asma, así como Infecciones respiratorias como la influenza y la neumonía. En México desde el 2009 después del brote epidemiológico de influenza mundial, los casos de influenza han estado presente año con año, en el periodo de 2015-2016 se confirmaron 870 casos [2], el uso de la Oximetría de Pulso proporciona un método en el diagnóstico de los casos que se requiera la hospitalización del paciente debido a su bajo nivel de saturación de oxígeno en la sangre.

El internet de las cosas (IoT) ha intensificado su desarrollo en el último año, dispositivos cotidianos están ahora conectados a internet pudiendo ser controlados de manera remota. Esta misma tecnología puede ser aplicada a la medicina, el monitoreo en tiempo real del paciente y su entorno, así como su seguimiento posterior al alta hospitalaria traerían un beneficio en la salud del paciente.

Consultado el estado del arte sobre el tema encontramos el trabajo presentado por el Centro de Bioingeniería de Colombia, demuestra que la sustitución de elementos analógicos disminuye en gran medida el consumo eléctrico, el cambio de filtros analógicos a digitales y el uso de PWM en el control de los diodos emisores, su sistema se basa en tarjeta y sondas prediseñadas para uso médico [3].

La propuesta presentada por el grupo de investigación de la Universidad de Colombia implementa un sistema de pulsioximetría y capnografía en un microcontrolador PSoC, el cual utiliza filtros digitales para el corte de frecuencia, así mismo mantiene una conexión inalámbrica tipo TCP/IP a un dispositivo Android el cual se encarga del almacenamiento y visualización [4]. La investigación realizada por Instituto Politécnico Nacional de México destaca la utilidad de implementar la conexión inalámbrica, así mismo el uso de la comunicación TCP/IP. La implementación del estándar HL7 en la comunicación del dispositivo con el servidor garantiza la compatibilidad de los datos con otros dispositivos de medición, es sistemas está basado en una PIC con un módulo WiFi, la transmisión de datos se centra en el sistema hospitalario [5].

El objetivo de este trabajo es el desarrollo de un prototipo de un Oxímetro de Pulso y aplicación web que solucione el problema del monitoreo de pacientes a distancia, se busca que el dispositivo obtenga la señal cardíaca y de la saturación de oxígeno de la sangre del paciente, el instrumento por sí mismo será capaz de establecer una conexión a un servidor remoto al cual enviara la señal obtenida, posteriormente la aplicación

calculara el nivel de oxígeno y el pulso cardiaco mostrándolas en pantalla al igual que la curva pletismográfica; se emplean las tecnologías existentes, utilizando la menor cantidad de componentes.

## 2. Componentes y construcción

La oximetría de pulso se basa en la estimación de la saturación de oxígeno en la sangre por medio de la espectrofotometría que es la medición de la cantidad de energía radiante que puede absorber un elemento, en este caso la hemoglobina presente en la sangre. La estructura principal del dispositivo se basa en el esquema que se muestra en la Fig. 1, se maneja cuatro etapas para resolver el problema.

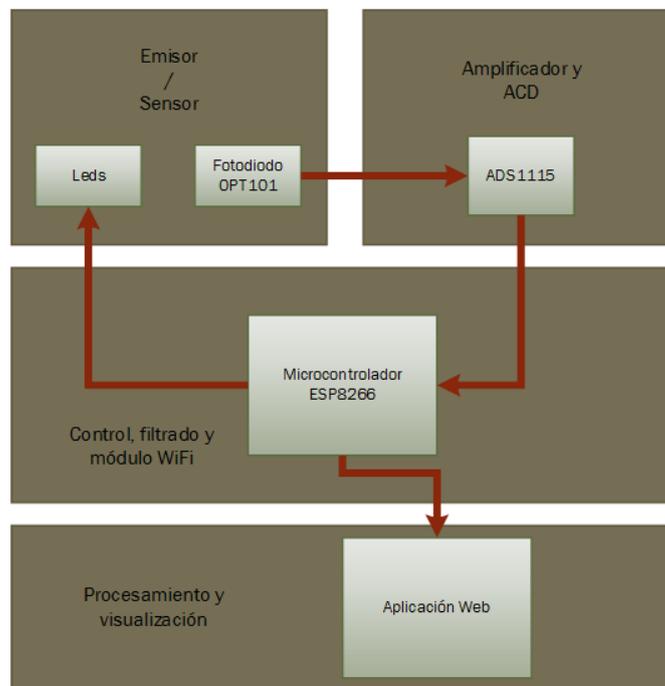


Fig. 1. Esquema de la absorción de la luz.

### 2.1. Emisor de luz y sensor

La hemoglobina es una proteína que se genera en la médula ósea, está presente en los glóbulos rojos y es la encargada del transporte de O<sub>2</sub> a cada rincón del cuerpo y regresando el CO<sub>2</sub> de la sangre a los pulmones para su expulsión. La hemoglobina tiene diferentes niveles de absorción de la luz al transportar O<sub>2</sub> y CO<sub>2</sub> permitiendo su medición a través de la diferencia que hay entre estos dos a diferentes longitudes de onda.

Oxihemoglobina se denomina a la hemoglobina que transporta O<sub>2</sub>, esta absorbe mayor cantidad de luz roja (640-660nm) y deja pasar la luz infrarroja (910-940nm). Por su parte, la desoxihemoglobina transporta CO<sub>2</sub> y absorbe la luz infrarroja, pero deja pasar la luz roja. En la Fig.2, se muestra la absorción de la oxihemoglobina y desoxihemoglobina a diferentes longitudes de onda, como se puede observar el intervalo de 640nm a 940nm es el segmento que presenta la mayor diferencia de absorción de luz.

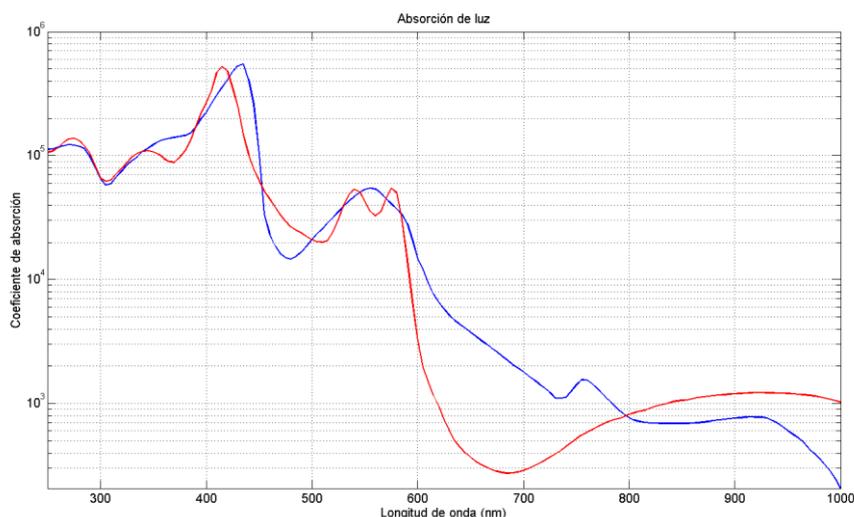


Fig. 2. Esquema de la absorción de la luz.

El funcionamiento del oxímetro de pulso se basa en lo anteriormente mencionado, utilizando dos emisores de luz, en este caso diodos lumínicos con longitudes de onda 660nm (rojo) y 940nm (infrarrojo), se sincronizan secuencialmente por medio del microcontrolador con un fotodetector.

El emisor consta de un led rojo con longitud de onda de 660nm y un led infrarrojo con longitud de onda de 940nm aproximadamente, ambos leds serán controlados por un pulso sincronizado con el fotodiodo, en la Fig. 3 vemos es esquema de la colocación de los diodos emisores y el fototransistor, en el cual las ondas de luz de cada longitud atraviesan el dedo y estas tendrán una variación de acuerdo a la hemoglobina.

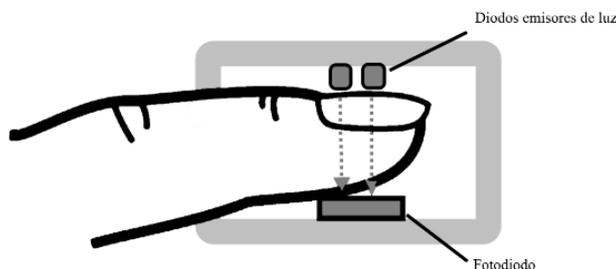


Fig. 3. Funcionamiento del sensor.

Utilizamos el componente OPT101, es un fotodiodo monolítico, tiene un rango de espectro promedio de 400 a 1100nm, además cuenta con un amplificador de transimpedancia, la Fig. 4 vemos el diagrama del fotodiodo, se utiliza la configuración típica proporcionada por el fabricante administrando un voltaje de 3.3volts [6].

### 2.2. Amplificador y ACD

La salida del fotodiodo es amplificada y convertida de analógico a digital por el componente ADS1115. Este convertidor digital(ADC) tiene una resolución de 16bits de los cuales solo 15bit serán utilizados y que un bit es utilizado por el signo, se manejan 32768 valores para la señal. Cuenta con un PGA que amplifica la ganancia del sensor para su conversión, este amplificador tiene 6 niveles de ganancia programable desde el dispositivo, la comunicación de este componente al microcontrolador se lleva a cabo por el bus de datos I2C, en la Fig. 4, se muestran los esquemas del componente [7].

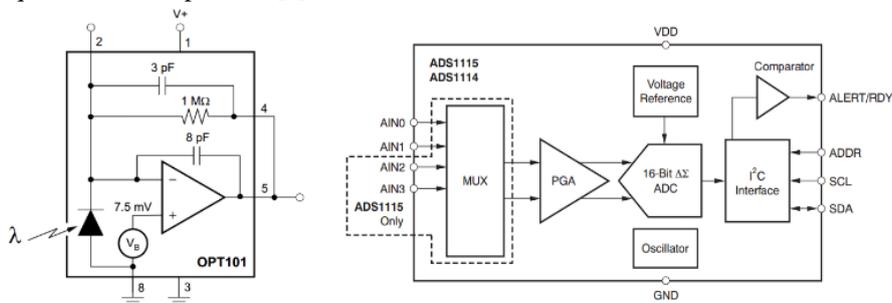


Fig. 4. Esquema del fotodiodo y del convertidor digital.

### 2.3. Control, filtrado y transmisor de la señal procesada

El microcontrolador en este caso el ESP8266 será el que se encargue de realizar las tres tareas, controlar los diodos emisores, filtrar la señal recibida por el convertidor digital y finalmente transmitir la señal resultante al servidor de la aplicación web. El microcontrolador ESP8266 es un chip de bajo costo de la marca Espressif, es un módulo Wi-Fi que puede realizar conexiones TCP/IP, en la Fig. 5 se observa el dispositivo, así mismo el esquemático de las entradas y salidas del dispositivo; principalmente el dispositivo fue diseñado para trabajar junto a un microcontrolador [8], posteriormente la empresa libero un SDK que le permitía al dispositivo ser programado; entre sus características principales son:

- CPU Tesilca Xtensa LX106 de 32-bit de 160MHz
- 96k de DRAM,64k de IRAM
- 4mb de memoria Flash
- Stack TCP/IP
- 100 metros de alcance
- Protocolos 802.11 b/g/n

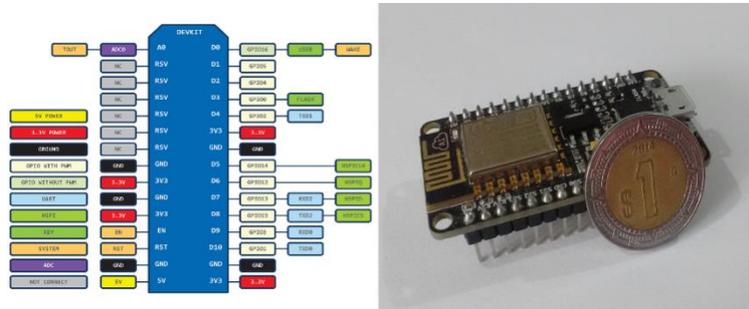


Fig. 5. Diagrama de conexión ESP8266-12.

El elemento a destacar de este microcontrolador es la arquitectura a 32bits, que permite la manipulación de datos de gran tamaño. El dispositivo puede ser programado de dos formas: interpretado, en este caso se manejan los lenguajes LUA y MicroPython y Compilado a partir de la API de Arduino que se basa en un lenguaje simplificado de C/C++ con su compilador y procesamiento basado en Java. Para mayor velocidad se utiliza el compilador nativo Gcc.

Para obtener la medición de la hemoglobina el microcontrolador controla por medio de pulsos en encendido y apagado de los diodos emisores, estos tendrán un tiempo de encendido de 50us seguidos de un desfase de 450us en el cual el sensor podrá obtener la medición.

Una vez obtenidos los datos de las mediciones de las dos ondas de luz, se filtra las señales por medio de un filtro digital, utilizamos un filtro pasa-banda IIR, que tiene una carga computacional menor, la frecuencia de corte se establece entre 0.5Hz a 5Hz. Su función es limitar la frecuencia al rango de las mediciones, en este caso las pulsaciones del corazón, así, el rango va de 30 a 300ppm, señales mayores a 5 Hz significarán ruido en la señal, principalmente generado por el movimiento del paciente, en la Fig. 6 podemos observar las dos diferentes señales en RAW y su posterior filtrado. Este filtro se rige por la formula siguiente:

$$H(z) = \frac{\sum_{k=0}^N b_k z^{-k}}{1 + \sum_{k=1}^M a_k z^{-k}} \quad (1)$$

Obtenida la señal digital es filtrada eliminando ruido y picos por el movimiento del paciente, se procede al envío de esta señal al servidor, el dispositivo se conecta a una red inalámbrica con conexión a internet, una vez establecida la conexión a internet se conecta al servidor listo para proceder al envío.

#### 2.4. Procesamiento y visualización

Para que el monitoreo del paciente pueda ser seguido en cualquier parte sin la necesidad que este se traslade, la aplicación web será el medio por el cual se procesara y visualizara esta información. La base principal será el servidor el cual se creó con Node js, el oxímetro desarrollado a partir del microcontrolador ESP8266 servirá a la vez como cliente con el servidor en el intercambio de los datos a través del protocolo

HTTP. Teniendo los datos de la medición del paciente la aplicación procesara para obtener la saturación de oxígeno y la frecuencia cardiaca, además dejara la señal intacta para su visualización. Debido a la variación de la oxihemoglobina y desoxihemoglobina con respecto a la absorción de la luz, se puede calcular el porcentaje de saturación de oxígeno conociendo la cantidad relativa de cada longitud de onda a través siguiente fórmula:

$$\% \text{ SaturaciondeOxigeno} = \frac{\text{Oxihemoglobina}}{\text{Oxihemoglobina} + \text{Desoxiheoglobina}} * 100 \quad (2)$$

Al tomar estas mediciones ambas señales pasan a través de la piel, tejido, hueso, etc. A esto se le llama componente estática (DC) ya que no pulsan en comparación a la sangre arterial conocida como la parte dinámica (AC), en la Fig. 6, se observa el esquema de las diferentes capas del dedo.

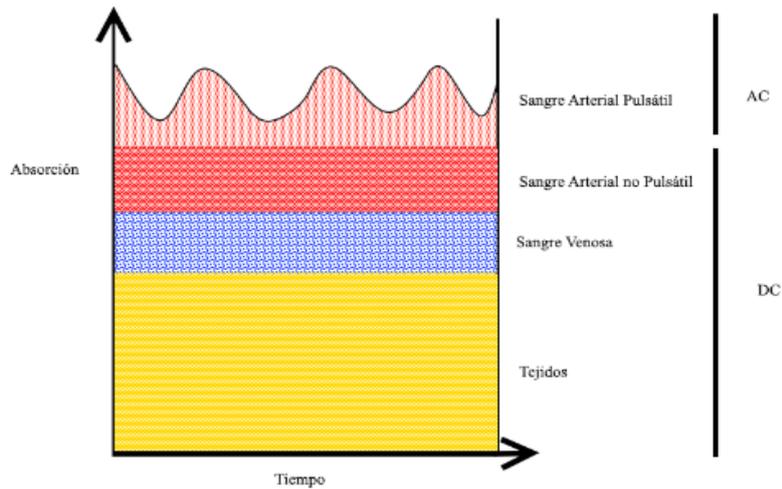


Fig. 6. Absorción de la luz de los componentes del dedo.

Debido a esto para obtener una medición correcta la saturación de oxígeno se calcula el cociente entre el valor de cada muestra con la siguiente formula:

$$R = \frac{\frac{ACrojo}{DCrojo}}{\frac{ACinfrarrojo}{DCinfrerrojo}} \quad (2)$$

Teniendo las dos señales en los rangos previamente establecidos, aplicamos la ecuación 2, obtenemos la razón, se establece el porcentaje a través de la curva generada mediante un proceso de calibración. En cuanto al pulso cardiaco solo restaría aplicar la trasformada rápida de Fourier en base a la frecuencia de salida.

El valor de R se encuentra ligado al porcentaje de SpO2, una vez calibrado el equipo, estos datos de comparación son guardados en el microcontrolador para su uso. El monitoreo en tiempo real se lleva a cabo accediendo a la página, una vez abierta esta

generará peticiones al servidor, en pantalla se muestra los niveles de oxígeno en la sangre, así como la frecuencia cardiaca, en la parte posterior se presenta la onda pletismográfica de la medición.

### 3. Resultados

El dispositivo se montó en una tarjeta de pruebas, la alimentación del dispositivo se realiza por medio de 5volts, el microcontrolador en la tarjeta que está montada lleva consigo un regulador de tensión a 3.3volt el cual alimenta a los demás dispositivos. En la Fig. 8 se puede ver los componentes ya funcionando.

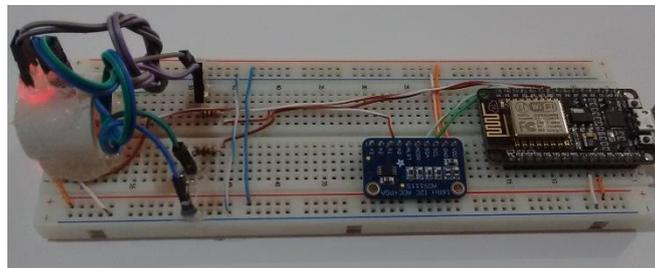


Fig. 7. Prototipo oxímetro de pulso.

Al obtener la señal del paciente se procedió a verificar el resultado final del filtrado, lo que se busca es eliminar ruido generado por el mismo dispositivo, por la luz ambiental y por los movimientos que el paciente llegara a producir, las mediciones fueron enviadas por medio de puerto serial a un programa de Matlab el cual grafica las señales, se compararon se las señales con y sin filtro como se observa en la Fig. 8.

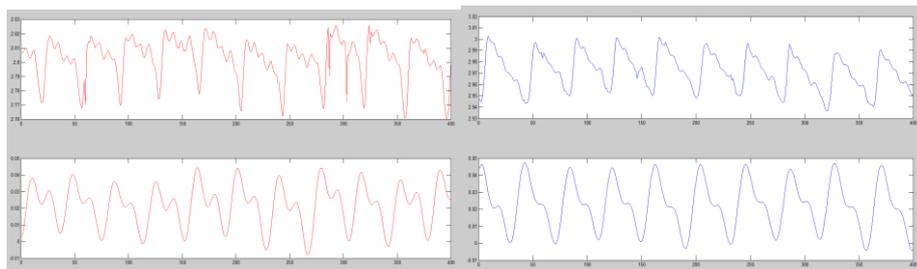
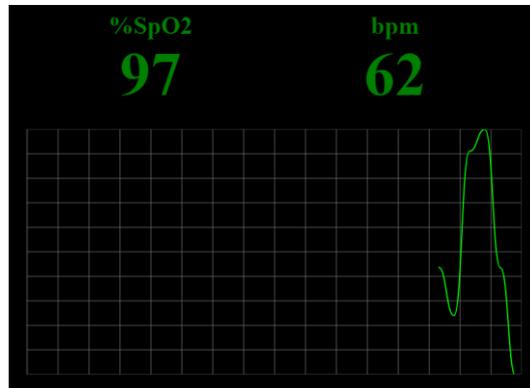


Fig. 8. Señal emisor IR (color azul) y señal emisor rojo (color rojo), la gráfica superior corresponde a la captura sin procesar y la inferior después de aplicar el filtrado.

La conexión del dispositivo a la red inalámbrica local y al servidor de la aplicación web se mantuvo estable en las pruebas realizadas con mediciones reales a periodos de 6 horas continuas por 7 días, los datos recibidos por el servidor fueron almacenados en archivos JSON, al identificar la dirección MAC del dispositivo se asigna al paciente, los archivos guardan la fecha y hora de la medición.



**Fig. 9.** Monitoreo en tiempo real.

En la Fig. 9 se muestra la captura de la visualización de la aplicación web, esta misma está configurada a modo responsivo por lo que el usuario podrá visualizarla en cualquier dispositivo.

#### 4. Conclusión

En este documento se presentó el prototipo de un Oxímetro de Pulso basado en el microcontrolador ESP8266 con conexión WiFi, el cual permite monitorear a distancia de la saturación de oxígeno y el pulso cardíaco, se logró mantener un monitoreo continuo, además de guardar un registro del mismo para su posterior análisis. En un futuro se busca la implementación de este sistema a otros signos vitales.

**Agradecimientos.** Los autores agradecemos el apoyo financiero del proyecto de red PRODEP, modelación matemática y computacional en medicina y dinámica poblacional.

#### Referencias

1. Mejía, H., Mejía, M.: Oximetría de pulso. Revista de la Sociedad Boliviana de Pediatría, Vol. 51, No. 2 (2012)
2. Cuitláhuac, M.: Actualización de la situación de influenza durante la temporada 2015-2016. Boletín Epidemiológico, Sistema Nacional de Vigilancia Epidemiológica, México, Vol. 33, No. 6, pp. 4-9 (2016)
3. Pérez, R. P., Pérez, J., Betancur, M. J., Bustamante, J.: Método alternativo del control de la intensidad lumínica para pulsioximetría en tiempo real. Ing. Biomed., Vol. 9, No. 17 (2015)
4. Gómez, C. A., Velasco, J.: Sistema de pulsioximetría y capnografía para dispositivos móviles Android. Ing. Biomed., Vol. 8, No. 15 (2014)
5. Ochoa, V., Sandoval, R.: Design of pulse oximeter with Wifi connectivity and interoperability with standard HL7 and IEE 1107.10404:2008. Systemics, Cybernetic And Informatics, Vol. 12, No. 7 (2014)

*Gabriel Contreras Mota, Rafael Lemuz López, Carlos Guillén Galván, Blanca Bermúdez Juárez*

6. Texas Instruments Incorporated: OPT101 Monolithic Photodiode and Single-Supply Trans-impedance Amplifier (2015)
7. Texas Instruments Incorporated: Ultra-Small, Low-Power, 16-Bit Analog-to Digital Converter with Internal Reference ADS1115 (2009)
8. Espressif System: ESP8266EX Datasheet (2016)

## **Aplicación multimedia para el entrenamiento en la certificación TOEFL mediante reconocimiento de voz**

Miguel Hernández Ramos, Rafael Lemuz López

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

miguellhdez@gmail.com, rlemuz@cs.buap.mx

**Resumen.** En este trabajo se presenta una aplicación para la práctica del inglés orientada a mejorar el desempeño de los alumnos en los exámenes de certificación TOEFL y motivar la práctica de hablar el idioma. Un aspecto esencial de la aplicación es que todas las respuestas a las preguntas de las diversas secciones son habladas. En el trabajo se describen, el diseño de la aplicación y las tecnologías de la información y comunicación utilizadas para su desarrollo, que van desde la base de datos, las APIs de reconocimiento y síntesis de voz, hasta los aspectos del diseño de la interacción humano computadora.

**Palabras clave:** TOEFL, entrenamiento, reconocimiento de voz, tecnologías Web.

### **Multimedia Application for Training in TOEFL Certification Using Speech Recognition**

**Abstract.** In this work we present an application to the practice of English geared to improving the performance of students in the certification exams TOEFL and motivate the practice of speak the language. An essential aspect of the application is that the answers to all the questions of the various sections are spoken. In the work described, the design of the application and the information and communication technologies used for its development, ranging from the database, the APIs speech recognition and synthesis, until those aspects of the design of the human computer interaction.

**Keywords:** TOEFL, training, speech recognition, Web technologies.

#### **1. Introducción**

El dominio del idioma inglés es importante porque permite mejorar la comunicación en una sociedad globalizada. También permite compartir experiencias y conocimientos

a través de los medios de comunicación tradicionales y emergentes basados en la Web. En especial en México es fundamental porque muchas de las relaciones laborales con los principales socios económicos son en inglés. Además, varias universidades actualmente tienen como requisito de titulación la acreditación de una lengua extranjera [1], y el idioma inglés, para muchos estudiantes, es la primera opción. Sin embargo, los métodos de enseñanza en general no incorporan las tecnologías de la información y la comunicación (TIC), y cuando lo hacen estas no están alineadas con los objetivos fundamentales de quienes toman un curso: aprender a hablar el idioma y acreditar los exámenes de certificación.

Lo anterior es la motivación fundamental de la aplicación multimedia que se propone, el objetivo es favorecer el desarrollo de ciertas habilidades como la percepción, la pronunciación, la habilidad auditiva y en particular se hace énfasis en que el alumno desarrolle su habilidad del habla.

En el 2015 el idioma inglés fue el principal idioma internacional, hoy en día se está convirtiendo en una habilidad básica para toda la fuerza laboral global [2]. Estudios recientes [3], muestran que Latinoamérica es una región con nivel bajo en cuanto al dominio de inglés, por primera vez, sólo tres países latinoamericanos quedaron en la banda del nivel más bajo. México se encuentra en la sexta posición de 14 países de América Latina y en la posición 40 de 70 países a nivel mundial.

Este trabajo analiza las soluciones existentes para el dominio del idioma inglés, en su mayoría aquellas que se sustentan de medios electrónicos para transmitir el conocimiento y determina en que puntos o partes del trabajo debe ponerse más dedicación para llenar los huecos que otros trabajos dejaron. La aplicación incorpora material de entrenamiento para la gramática, la pronunciación, la percepción y la lectura en cada uno de sus niveles para que el estudiante pueda desarrollar tareas mínimamente académicas.

## **2. Marco referencial**

El marco utilizado para este proyecto involucra temas de tecnologías web y APIs (Application Programming Interface) que pueden ser usadas en temas educativos.

### **2.1. Certificaciones de inglés**

Sirven para acreditar, de manera precisa, que la persona puede seguir un programa formativo en este idioma o que tiene las habilidades lingüísticas necesarias para desempeñar una función laboral completa o parcialmente en esta lengua. La elección de un certificado u otro depende del objetivo del interesado.

#### **2.4.1. TOEFL (Test of English as a Foreign Language)**

La Prueba del inglés como Lengua Extranjera fue creada por ETS, (Educational Testing Service), una organización con sede en los EU que se dedica al desarrollo de pruebas profesionales [6]. Es una prueba que evalúa el nivel de inglés americano de las personas que no son hablantes nativos de esta lengua. Evalúa la habilidad de una

persona al leer, escribir, hablar y comprender el inglés que se utiliza en la preparatoria o la universidad.

## **2.5. Tecnología de reconocimiento de voz**

El objetivo de la tecnología de reconocimiento de voz es hacer que la comunicación de los humanos con las máquinas principalmente sea más eficiente y fácil. Estas tecnologías de voz incluyen varias tecnologías: Speech analysis, Speech synthesis, Speech recognition y Speaker recognition.

## **2.6. Sintetizadores de voz**

También conocida como Síntesis del Habla, es la tarea de generar voz artificial partiendo de un texto. Un sistema Text To Speech (Texto a Conversación), también llamado TTS, es una aplicación que convierte un texto escrito a audio, permitiendo a un invidente escuchar un documento, un email o incluso llamar a un número de teléfono y poder escuchar una información que previamente alguien ha escrito.

## **2.7. Kits de desarrollo de software con reconocimiento de voz**

### **2.7.1. SAPI**

Microsoft ha desarrollado un set de herramientas para tareas de reconocimiento de voz desde las primeras versiones del sistema operativo. Este se incluye en la actualidad motores de reconocimiento y de síntesis de voz en varios idiomas que incluyen inglés, francés, chino, español, alemán, ruso, japonés entre otros. Desde la versión 5.0 de SAPI incluida en Windows 2000, se limita el acceso de la aplicación a los motores de reconocimiento y de síntesis de voz.

### **2.7.2. Java Speech**

La plataforma de desarrollo Java no tiene un motor de reconocimiento de voz, pero sí brinda al programador la facilidad para escribir aplicaciones de voz con un conjunto de clases e interfaces que modelan esta tarea. Este conjunto de clases e interfaces se las conoce como el JSAPI o Java Speech API. Estas clases fueron concebidas a manera de envoltura, ya que internamente en el entorno administrado de Java, estas clases no tienen implementación alguna. Es tarea del sistema operativo host a través de objetos nativos el proveer de la implementación de estas tareas de reconocimiento de voz. Java conecta a los objetos y a las interfaces de JSAPI con la implementación nativa a través de JNI14.

## **3. Soluciones existentes**

Se presentan sitios y/o aplicaciones web que al día de hoy son usadas por quienes están interesados en aprender o fortalecer sus conocimientos en el idioma inglés.

- Duolingo. Página Web y aplicación móvil basada en niveles que simulan un juego donde ganar puntos y avanzar de nivel es un reto, tiene contenido para casi cualquier situación de la vida diaria, pero carece de orientación para fines específicos.
- Interlingua. Aplicación móvil que presenta animaciones de tipo conversación. Está enfocada a la pronunciación.
- Saber inglés. Página de gran ayuda si se tiene conocimiento del idioma, enfocada principalmente a la gramática.
- Learn English. Página ideal para comenzar a aprender inglés. Lo mejor de este sitio es su sección de vocabulario ya que tiene todo tipo de palabras, conceptos y expresiones útiles para una comunicación básica en inglés.
- English Page. Está basada en el inglés americano, muchos cursos tradicionales la utilizan como apoyo adicional para sus estudiantes. Tiene cientos de ejercicios para cada una de las lecciones que ofrece.

#### 4. Metodología

La metodología aplicada para alcanzar el objetivo de este trabajo fue IDEAL que consta de las siguientes cinco fases (Bibiana D. Rossi, 2001):

1. Identificar la tarea mediante el plan de requisitos.
2. Desarrollo de prototipos. La concepción de la solución e implementación.
3. Ejecución de la construcción del sistema. Integración con otros sistemas, y la aceptación por el usuario del final.
4. Actuación para conseguir el mantenimiento perfectivo. Definir el mantenimiento del sistema global, de las bases de conocimientos.
5. Lograr la transferencia tecnológica mediante entrenamiento en sesiones de tutoría entre los diseñadores y los usuarios finales, (García Martínez, R y Britos P. 2004. Ingeniería de Sistemas Expertos. Editorial Nueva).

#### 5. Desarrollo

##### 5.1. Propuesta de solución

Como primera consideración para desarrollar la aplicación, de ahora en adelante *Angloparlante*, fue dar extras a las soluciones existentes. Las siguientes son características que la hacen diferente:

- Es una *aplicación para el entrenamiento* para un examen de certificación TOEFL, no enseña gramática ni vocabulario en inglés, es como su nombre lo indica aplicación para entrenar a las personas a fin de fortalecer los conocimientos que se tienen en el idioma inglés.
- Las respuestas a las evaluaciones para cada sesión de autoaprendizaje son habladas haciendo uso de un micrófono de computadora.
- Incluye actividades de la vida cotidiana del ámbito académico. Cuenta con 4 secciones: Reading, Listening, Writing y Speaking para las que tiene un apartado

de *learn* y otro de *practice*. La mayor parte de las actividades están dentro de Speaking debido a la naturaleza de la aplicación, el uso del reconocimiento de voz, y además porque se considera que es la sección dónde se ponen a prueba las otras 3 secciones.

- El aprendizaje puede ser guiado por un docente o de forma automática. Si se elige la forma guiada la respuesta es enviada vía mensaje, dentro de la misma aplicación, al docente y la calificación es asignada manualmente. En cambio, si se elige la forma automática, la aplicación recibe la entrada por voz, la convierte a texto y con el algoritmo del punto 5.2.3 se recibe la calificación, en caso de ser aprobatoria el alumno podrá realizar la siguiente actividad de evaluación, en otro caso podrá reintentar máximo 5 veces.
- Angloparlante conserva el último nivel alcanzado por el estudiante a fin de direccionarlo justo donde se quedó.
- A manera en que el estudiante avanza la dificultad es mayor, se usan palabras más difíciles de pronunciar, las respuestas son más extensas o cambian de contexto, pero siempre son frases de un entorno académico que es lo que evalúa TOEFL IBT.
- Los docentes intervienen únicamente cuando los alumnos eligen ser evaluados por ellos, y es en ese momento cuando orientan a los alumnos.
- Al realizar una evaluación, el estudiante, puede elegir entre ser evaluado por el docente o por la aplicación independientemente de cómo ha sido evaluado antes.
- Contiene tips para cada sección, estos son aleatorios y puede verse uno u otro cada vez que el estudiante ingresa a una sección.

## 5.2. Diseño e implementación del sistema

### 5.2.1. Modelo de casos de uso

El modelo de casos de uso describe a Angloparlante en términos de sus distintas formas de utilización, cada una de las cuales se conoce como un caso de uso. El estudiante y el docente son actores, es decir, entidades distintas a los usuarios, en el sentido de que éstos son las personas reales que utilizarán el sistema, mientras que los actores representan cierta función que una persona real realiza. En la terminología orientada a objetos se considera al actor una clase de usuario mientras que los usuarios se consideran como objetos o instancias de esa clase. Incluso, una misma persona puede aparecer como diversas instancias de diferentes actores.



Fig. 1. Diagrama de delimitación del sistema.

El diagrama de delimitación del sistema con nuestros actores se muestra en la figura 1, y tiene las características mencionadas.

Se puede notar que para interactuar con la aplicación ambos deben iniciar sesión y como consecuencia estar registrados. Actualmente la inscripción es manual debido a que no se cuenta con un administrador.

### 5.2.2. Diseño de la base de datos

La base de datos para Anglolarlante tiene la colección de datos necesaria para el correcto funcionamiento de la aplicación. Para realizarla, y una vez abstraído el problema, se determinó *el propósito de la base de datos* el cuál es disponer en todo momento de la información relacionada con la aplicación. El siguiente paso fue *determinar las tablas necesarias*, luego se *determinaron los campos necesarios para cada tabla*, el cuarto paso consistió en *determinar las relaciones* y finalmente se *perfeccionó el diseño* buscando errores usando datos de prueba para poder realizar consultas en modo consola y con esto determinar campos sobrantes o tablas innecesarias para disminuir la redundancia e inconsistencia de los datos.

La figura 3 muestra la estructura de la base de datos que se obtuvo al seguir los 5 pasos anteriores. En la figura se aprecian las tablas, los campos, la relación y la cardinalidad entre las tablas.

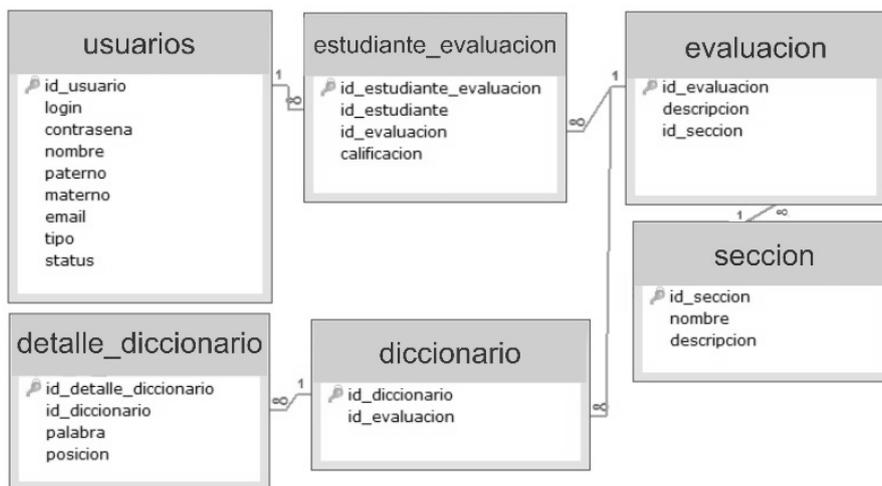


Fig. 3. Estructura gráfica de las tablas, la cardinalidad y relaciones de las tablas.

### 5.2.3. Técnicas y tecnologías empleadas

Respecto al reconocimiento de voz, se utilizó la modalidad de reconocimiento por medio de comandos debido a que disponemos de las palabras relacionadas a cada una de las respuestas, y además es una colección limitada dentro de la base de datos.

El reconocimiento de voz se hace con Web Speech API, una API diseñada por la comunidad de la W3C e impulsada por Google y openStream para la síntesis y el análisis del habla en aplicaciones Web [9].

En el diseño de la aplicación se consideraron aspectos de interacción humano computadora: un diseño sencillo y atractivo y responsivo para el usuario, con la capacidad de ocultar el menú lateral, fig. 4, para tener disponible la pantalla completa para las actividades.

El siguiente algoritmo realiza la evaluación automática, recibe como entrada dos arreglos, uno formado por las palabras reconocidas por Web Speech API y el otro con los datos asociados a la evaluación en curso, estos datos se recuperan previamente mediante una petición AJAX.

```
function evalua(arr_estudiante, arr_conocimiento){
  var arr_acertadas = [];
  var arr_errores = [];
  var resultado = [];
  for(var i=0;i<arr_conocimiento;i++){
    for(var j=0;j<arr_estudiante;j++){
      if(arr_conocimiento[i] == arr_estudiante[j]){
        arr_acertadas.push(arr_estudiante[j]);
      }
      else{
        arr_errores.push(arr_estudiante[j]);
      }
    }
  }
  resultado['pAciertos'] = arr_acertadas;
  resultado['pErrores'] = arr_errores;
  return resultado;
}
```

Posterior a la ejecución del algoritmo, escrito en JavaScript, lo que resta es mostrar el resultado al estudiante, Si el arreglo de errores está vacío y el arreglo de palabras acertadas coincide término a término con el arreglo de la petición la respuesta será correcta y su calificación aprobatoria, en caso contrario, si el 50% o más de las palabras de la petición no coinciden debe repetirse el ejercicio completo, en cambio si los errores son mínimos se le indica en que partes falló y tendrá la opción de corregir sólo eso. Respecto a los intervalos de calificación, si la respuesta llega en el primer intento la calificación es de 10, en el segundo intento (total o parcial) 9, para el tercer intento corresponde 8, para el cuarto 7, y para el quinto 6.

#### 5.2.4. Aplicación

La ventana principal de la aplicación tiene accesos directos a su último nivel alcanzado, sus calificaciones y un enlace para acceder a sus mensajes, esta ventana principal tiene un menú lateral con las secciones, cada una con su apartado de Learn y

Practice tal como se aprecia en la figura 4. Con la navegación lateral el estudiante puede ir a una u otra sección, y dentro de ella elegir si quiere aprender o practicar.

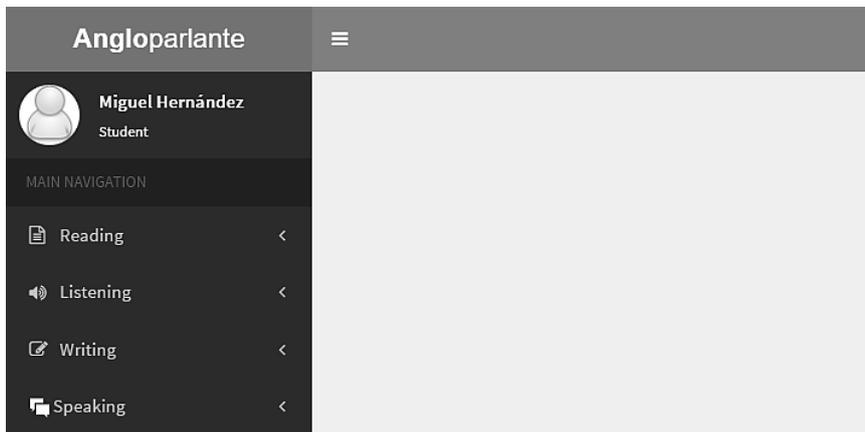


Fig. 4. Navegación lateral del estudiante.

La figura 5 muestra un ejemplo de evaluación. A la derecha está el texto que el usuario le dirá a Angloparlante. Para activar el reconocimiento el estudiante debe presionar el icono del micrófono y posteriormente comenzar a hablar.

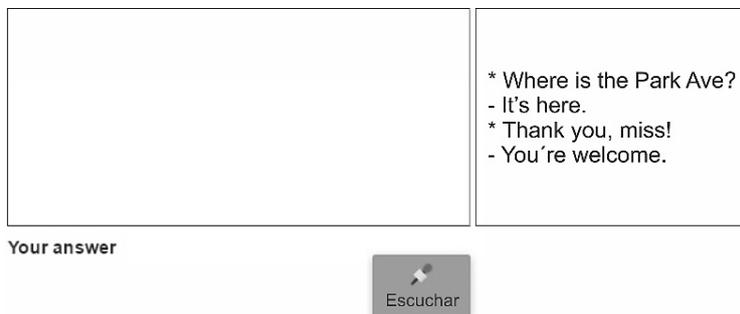


Fig. 5. Actividad de evaluación con reconocimiento de voz.

Para el caso del docente, sus actividades son de seguimiento al aprendizaje del estudiante como se aprecia en la figura 6.

Nombre	Lección superada	Regresar nivel	Activo	Suspender	Activar
Miguel Hernández Ramos	7	<	SI	×	✓
Violeta Elvia Mexicano Salas	0	<	SI	×	✓
Miguel H	0	<	SI	×	✓

Fig. 6. Ventana principal del docente.

## **6. Conclusiones**

En este trabajo presentamos un sistema de software para la práctica hablada del idioma inglés. En el trabajo se describe el diseño de la aplicación desde el diseño de la base de datos hasta el diseño de la interfaz gráfica para la interacción humano computadora. En un trabajo futuro se hará un estudio de usabilidad en un grupo piloto de estudiantes de nivel básico y medio para evaluar el sistema propuesto.

## **Referencias**

1. Criollo, R.: Titulación: Lengua Extranjera. Facultaddelenguas.com, available: <http://www.facultaddelenguas.com/caale/titulacion> (2016)
2. Ef.com.mx: EF Índice inglés Proficiency - inglés, Economía y Calidad de Vida. Available: <http://www.ef.com.mx/epi/insights/english-economics-and-quality-of-life> (2016)
3. Ef.com.mx: EF EPI - inglés e Innovación. Available: <http://www.ef.com.mx/epi/insights/english-and-innovation> (2016)
4. Castaño, C.: El rol del profesor en la transición de la enseñanza presencial al aprendizaje online (2003)
5. Marqués, P.: Multimedia educativo: clasificación, funciones, ventajas e inconvenientes (2004)
6. ETS Home. Ets.org, Available: <https://www.ets.org> (2016)
7. Cambridge Institute: Academia de inglés que imparte cursos de diferentes lenguas para que mejores tu nivel de idiomas - Cambridge Institute. CambridgeInstitute, Available: <http://www.cambridgeinstitute.net> (2016)
8. Salazar, J. A.: Reconocimiento de comandos de voz para control de las funciones multimedia dentro del hogar (2012)
9. Adorf, J.: Web Speech API. KTH Royal Institute of Technology (2013)



## Diseño de una ontología para el proceso de evaluación de las asignaturas técnico-científicas del Instituto Tecnológico de Orizaba

I. Colohua Cruz<sup>1</sup>, L. A. Reyes Hernández<sup>1</sup>, G. Hernández Chan<sup>2</sup>,  
J. L. Sánchez-Cervantes<sup>3</sup>

<sup>1</sup> División de Estudios de Posgrado e Investigación,  
Instituto Tecnológico de Orizaba,  
Orizaba, Veracruz,  
México

<sup>2</sup> Dirección de Ingeniería Eléctrica y Electrónica,  
Instituto Tecnológico de Mérida,  
Mérida, Yucatán,  
México

<sup>3</sup> Catedrático CONACYT comisionado al Instituto Tecnológico de Orizaba,  
Orizaba, Veracruz,  
México

icc90@outlook.com, l.a.reyes.h@gmail.com, gandhi.hernandez@itmerida.mx,  
jsanchezc@ito-depi.edu.mx

**Resumen.** Este artículo presenta el diseño de una ontología para apoyar al proceso de evaluación de las asignaturas técnico-científicas (Cálculo Diferencial, Física, Química, entre otras) que se imparten a nivel licenciatura en el Instituto Tecnológico de Orizaba (ITO). El E-learning es una de las estrategias actuales que facilita el proceso de evaluación mediante el uso de entornos Web o plataformas digitales, los cuales comparten estrategias pedagógicas similares a las habituales aplicadas en cursos con modalidad presencial. Las preguntas que se aplican en la evaluación de asignaturas técnico-científicas por lo general tienen un resultado único, en su mayoría de forma numérica, y por lo tanto es conveniente generar respuestas que actúen como distractores al momento de aplicar un Quiz al alumno. La ontología que se presenta pretende especificar el dominio adecuado para la evaluación de las asignaturas técnico-científicas en el ITO. Para su desarrollo se siguió la metodología de ontologías Methontology y su codificación se realizó con Protégé.

**Keywords:** Methontology, Protégé, quiz, ontología, distractores asignaturas técnico-científicas.

## **Design of an Ontology for the Evaluation Process in the Technical-Scientific Courses at Instituto Tecnológico de Orizaba**

**Abstract.** This paper presents the design of an ontology to support the assessment process in the technical-scientific subjects (Differential Calculus, Physical, Chemical, among others) which are taught in the firsts years of bachelor's degree at Instituto Tecnológico de Orizaba (ITO). E-learning is one of the current strategies that facilitates the process of evaluation through the use of Web environments or digital platforms, which share pedagogical strategies similar to the usually applied in classroom courses. The questions that are applied in the evaluation of technical-scientific subjects usually have a single result, mostly of the time represented by a number, so it is convenient to generate answers that act as distractors when applying a quiz to a student. The ontology presented here pretends to specify the appropriate domain for the evaluation of the technical-scientific subjects at ITO. It was developed following the Methontology method and codified with Protégé.

**Keywords:** Methontology, Protégé, quiz, Ontology, distractors, technical-scientific courses.

### **1. Introducción**

En la actualidad, los procesos de evaluación para las asignaturas técnico-científicas, por ejemplo Álgebra, Calculo, Física, Ecuaciones Diferenciales, Métodos Numéricos, entre otras que se imparten a nivel licenciatura en el Instituto Tecnológico de Orizaba (ITO) y en la mayoría de Instituciones de Educación Superior (IES), en donde se imparten programas de ingeniería, son laboriosos y extensos. Se ha observado que al aplicar los algoritmos Shuffle Option/Question (barajeo en preguntas y respuestas) en los exámenes de la materia de cálculo diferencial se cumple con los requisitos de las estrategias pedagógicas adecuadas para una evaluación, sin embargo, los entornos Web que aplican LMS (Learning Management System, Sistema de gestión de aprendizaje) como, por ejemplo, las plataformas Moodle y Chamilo, permiten la generación un Quiz pero no cumplen con las estrategias pedagógicas recomendadas [1]. No obstante, estas y otras plataformas Web enfocadas en la educación han tenido gran éxito puesto que cada día son más los usuarios que utilizan estas y otras herramientas en línea para promover y sacar provecho del E-learning (aprendizaje electrónico), el cual engloba a las aplicaciones y servicios que facilitan el proceso de enseñanza-aprendizaje, y que además se basan en el uso de las Tecnologías de la Información y la Comunicación (TIC) [2]. Además de implementar herramientas en línea para los procesos de evaluación, es necesario considerar técnicas para presentar un Quiz, el cual es una pequeña evaluación. Para ello, se recomienda que la evaluación contenga reactivos de tipo opción múltiple, que es una técnica que consiste en plantear una pregunta o problema, al cual se le denomina reactivo, que consta de un enunciado y una serie de respuestas, llamadas opciones. Entre estas respuestas hay una sola respuesta correcta,

llamada solución, y otras incorrectas, conocidas como distractores. El número de respuestas opcionales en cada reactivo depende de la probabilidad que se acepte de que el alumno conteste bien al azar [3].

Los entornos Web enfocados a la gestión del conocimiento educativo desempeñan un papel importante para el E-learning, el cual surge de la necesidad de la educación y el aprovechamiento de la tecnología, en especial las TICs, para este fin. Además, el E-Learning es uno de los pilares fundamentales para el auto aprendizaje y retroalimentación de algún tipo de conocimiento. Actualmente tiene un gran potencial en el mercado del software. En este sentido, es importante destacar que para el proceso de evaluación en línea no existen instrumentos que permitan medir los logros alcanzados por los estudiantes y se califiquen como “buenos o malos” ni tampoco hay ningún instrumento que cubra todas las necesidades de la evaluación, solo existen herramientas de apoyo para recopilar y administrar la información [4].

Por otra parte, el uso de las ontologías es común en áreas o dominios como los negocios, finanzas, Internet, medicina e industria y los dominios son usados para formalizar capitales intelectuales, integrar fuentes de datos, representar y organizar conocimiento [5]. También son empleados por el E-learning porque permiten visualizar de manera general los elementos necesarios en el dominio del mismo.

Por otra parte, para la elaboración de ontologías existen diversas metodologías, un ejemplo de ello es Methontology, la cual consta de 11 tareas que permiten llevar a cabo la elaboración de una ontología de forma correcta. Esta metodología permite construir ontologías en el nivel de conocimientos, y tiene sus raíces en las actividades como la gestión/administración, soporte y técnicas identificadas por el proceso de desarrollo de software propuesto por la organización IEEE [6].

Para fines prácticos es necesario que la ontología quede en un archivo con extensión .owl para ser implementada en algún sistema, y para algunos casos, es recomendable diseñar la ontología en una arquitectura modular, la cual se basa en la definición de una ontología de nivel superior que definirá las relaciones que deben aplicarse a un conjunto de sub-ontologías para cada sub-dominio [7].

Para la construcción de ontologías, una de las herramientas más conocidas es Protégé. Se trata de un software de código abierto que permite construir una ontología mediante su interfaz que ayuda al desarrollador en el proceso de construcción de la misma. Este software está disponible en su página oficial [8].

Los LMS están diseñados para apoyar al proceso de enseñanza-aprendizaje de manera virtual mediante un conjunto de herramientas que permiten la interacción y colaboración entre los actores del proceso (estudiantes, profesores, contenido), y tienen una orientación hacia el alumno, el profesor o para el proceso de aprendizaje. Los LMS que utilizan ontologías se dividen internamente en cuatro capas: interfaz (para el estudiante y/o profesor), servicios, contenidos (conocimientos o recursos) y actividades [9, 10].

El presente artículo presenta el diseño e implementación de una ontología para la administración y diseño de reactivos de evaluación de las asignaturas técnico-científicas impartidas en el Instituto Tecnológico de Orizaba. Para la codificación de la ontología se utilizó Protégé y se siguió la metodología dictada por Methontology.

## 2. Metodología

### 2.1. Elementos de una ontología

Los componentes de una ontología varían de acuerdo al dominio de interés y a las necesidades de dicho dominio. Por lo general, se encuentran los siguientes siete elementos [5].

1. **Clases:** Describen los conceptos del dominio y se organizan en taxonomías y por lo general se les aplican mecanismos de herencia.
2. **Relaciones:** Representan las interacciones entre los conceptos del dominio, por lo general contienen relaciones binarias. El primer elemento de la relación es el dominio y el segundo el rango.
3. **Funciones:** Son un tipo concreto de relación, donde se identifica un elemento mediante el cálculo a partir de otros elementos que se encuentren en la ontología.
4. **Instancias:** Representan objetos determinados de un concepto.
5. **Taxonomía:** Conjunto de conceptos organizados jerárquicamente.
6. **Axiomas:** Modelan sentencias que son siempre ciertas y permiten, junto con la herencia, inferir conocimiento que no esté indicado explícitamente en la taxonomía de conceptos.
7. **Propiedades (slots).** Son las características o atributos que describen a los conceptos.

### 2.2. Tareas de methontology

El diseño de la ontología que aquí se presenta está basado en las tareas definidas en la metodología Methontology, esto debido a que su completitud permite satisfacer los requerimientos propios del E-learning, por ejemplo motivar al alumno para procesar la información mientras se trabaja con los materiales de aprendizaje y que los medios de comunicación faciliten el procesamiento de la información para que el alumno obtenga el conocimiento requerido [11]. A continuación se describen las once tareas de conceptualización de la ontología.

#### — Tarea 1: Glosario de términos

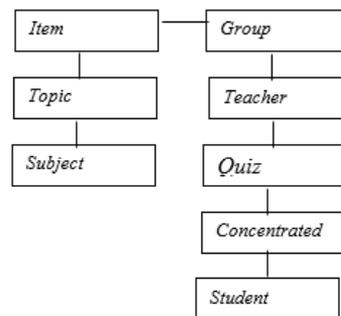
Este glosario está constituido por los términos de interés del dominio E-learning determinados para el proceso de evaluación de las materias técnico-científicas del ITO. La Tabla 1 muestra los elementos (conceptos y atributos) más importantes. Cabe mencionar que hay términos que son entendidos como metadatos del estándar de Especificación de Interoperabilidad de Preguntas y Pruebas de IMS (*IMS Question & Test Interoperability Specification*) que menciona el intercambio de temas, contenidos y resultados de datos entre herramientas de edición, bancos de reactivos, plataformas de aprendizaje y de evaluación [12]. No se detectaron sinónimos para los conceptos.

**Tabla 1.** Extracto del glosario de términos.

Nombre	Tipo	Descripción
<i>Image</i>	Atributo	Es un elemento que representa una imagen o una gráfica para el <i>Item</i> .
<i>Text</i>	Atributo	Es el enunciado que forma parte del <i>Item</i> .
<i>Distractor</i>	Atributo	Son aquellas respuestas que presentan errores de pensamiento que ocurren con frecuencia y que parecen verdaderas.
<i>Correct Answer</i>	Atributo	Es la respuesta correcta a un <i>Item</i> .
<i>Result</i>	Atributo	Es el puntaje obtenido del <i>Quiz</i> .
<i>Item</i>	Concepto	Es una pregunta estructurada que forma parte de un <i>Quiz</i> .
<i>Subject</i>	Concepto	Es la asignatura que imparte un maestro.
<i>Topic</i>	Concepto	Los temas que están diseñados con base en los contenidos de las asignaturas.
<i>Student</i>	Concepto	Persona que presenta un <i>Quiz</i> .
<i>Quiz</i>	Concepto	Es el conjunto de <i>Item</i> 's para ser evaluados.
<i>Concentrated</i>	Concepto	Es el conjunto de resultados de varios <i>Quiz</i> que presenta el <i>Student</i> .
<i>Group</i>	Concepto	Es el conjunto de <i>Student</i> 's de una asignatura que presentan un <i>Quiz</i> .
<i>Teacher</i>	Concepto	Es la persona responsable de impartir una materia ( <i>Subject</i> ) y elaborar un <i>Quiz</i> .

### — Tarea 2: Taxonomía de conceptos

La Fig. 1 muestra la jerarquía de los conceptos que fue definida a partir del glosario de términos de acuerdo al dominio de E-learning.



**Fig. 1.** Taxonomía de conceptos.

### — Tarea 3: Diagrama de relaciones binarias

La Fig. 2 muestra el diagrama que establece los tipos de relaciones entre los conceptos antes definidos. Se omiten los atributos de los mismos.

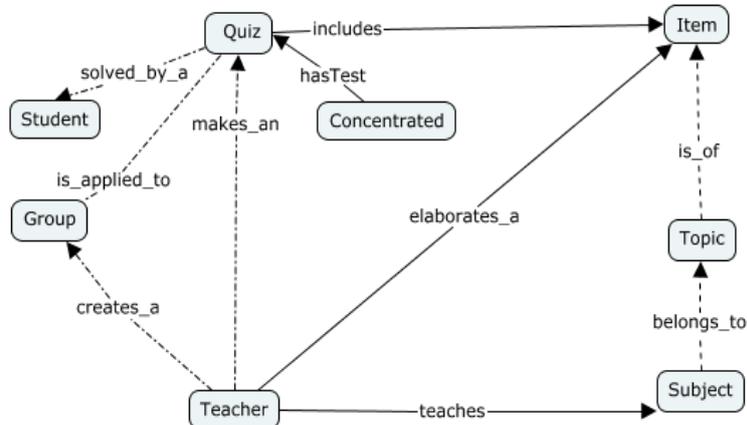


Fig. 2. Diagrama de relaciones binarias.

— Tarea 4: Construir el diccionario de conceptos

En la Tabla 2 se muestran los conceptos que se determinaron como los más relevantes en el proceso de E-learning. Estos conceptos se consideran los de mayor importancia debido a que, en la dinámica correspondiente al proceso de elaboración de evaluaciones, los actores principales son los profesores, ya que son ellos quienes definen y agregan reactivos a los exámenes, sin embargo, esto conlleva otras actividades anteriores, por ejemplo, crear los grupos y crear los exámenes con base en los contenidos de las asignaturas que hayan impartido. Por otra parte, el Quiz es por sí mismo, un elemento fundamental, ya que se trata del medio o la herramienta mediante la cual se logra la evaluación.

Tabla 2. Diccionario de conceptos.

Concepto	Relaciones	
Teacher	creates_a -> Group makes_an-> Quiz	elaborates_a-> Item teaches->Subject
Quiz	is_applied_to-> Group includes->Item	solved_by_a->Student

— Tarea 5: Detallar relaciones binarias

Para cada relación se especifica el nombre, conceptos fuente (A) y destino (B), cardinalidad del dominio de E-learning. Véase Tabla 3.

Tabla 3. Detalles de las relaciones binarias.

Nombre	Concepto A	Concepto B	Cardinalidad
teaches	Teacher	Subject	1: 1...*
creates_a	Teacher	Group	1: 1...*

<i>makes_an</i>	<i>Teacher</i>	<i>Quiz</i>	1: 1...*
<i>elaborates_a</i>	<i>Teacher</i>	<i>Item</i>	1: 1...*
<i>includes</i>	<i>Quiz</i>	<i>Item</i>	1 : *
<i>is_applied_to</i>	<i>Quiz</i>	<i>Group</i>	1: 1...*
<i>solved_by_a</i>	<i>Quiz</i>	<i>Student</i>	1...*: 1...*
<i>is_of</i>	<i>Item</i>	<i>Topic</i>	1: 1
<i>belongs_to</i>	<i>Topic</i>	<i>Subject</i>	1...*: 1
<i>hasTest</i>	<i>Concentrated</i>	<i>Quiz</i>	1...*: 1

#### — Tarea 6: Definir los atributos de instancias

Aquí se definen algunos atributos de instancias. Cabe mencionar que todas las clases tienen por lo menos una instancia, a continuación, en la Tabla 4 se muestran los más relevantes. *NombreStudent* representa un atributo de *Student*, el cual es el nombre completo del alumno, *IdEstudent* representa la matrícula del estudiante, los demás atributos que pertenecen a *Item* esta descritos en la Tabla 1.

**Tabla 4.** Atributos de instancia de las clases Student e Item.

Atributo	Concepto	Tipo de Valor	Rango	Cardinalidad
<i>NombreStudent</i>	<i>Student</i>	<i>String</i>	ilimitado	1
<i>IdEstudent</i>	<i>Student</i>	<i>String</i>	ilimitado	1
<i>Correct Answer</i>	<i>Item</i>	<i>int</i>	números reales	1
<i>Distractor</i>	<i>Item</i>	<i>int</i>	números reales	3
<i>Image</i>	<i>Item</i>		ilimitado	0: 1
<i>Text</i>	<i>Item</i>	<i>String</i>	ilimitado	1

#### — Tarea 7: Detallar atributos de clase

La Tabla 5 muestra los atributos de clase, los cuales representan características generales de un concepto, es decir, todas las instancias del concepto tendrán el mismo valor para ese atributo.

**Tabla 5.** Atributos de clase Quiz.

Atributo	Concepto	Tipo de Valor	Rango de Valores
Mínimo de <i>Item</i> 's por <i>Quiz</i>	<i>Quiz</i>	<i>Item</i>	5

#### — Tarea 8: Detallar constantes

Para cada constante se especifica el nombre, tipo de valor, valor y unidad de medida. Esta tarea no aplica para el dominio del E-learning.

#### — Tarea 9: Detallar axiomas formales

Los axiomas describen restricciones del modelo, así como las características que debe cumplir algún concepto. En la Tabla 6 se muestra el nombre, descripción, expresión lógica, conceptos o relaciones, y variables de los axiomas que intervienen.

**Tabla 6.** Axiomas formales.

Nombre	Descripción	Expresión Lógica	Conceptos/ Relaciones	Variables
<i>NoAnswer</i>	Un distractor es una respuesta que no es correcta	NO [ EXISTE( <i>Item</i> (?x) Y ?x. <i>Correct Answer</i> =false; ) ]	<i>Item</i> <i>Correct Answer</i>	?x
<i>Answer</i>	De una pregunta solo hay una respuesta correcta	EXISTE( <i>Item</i> (?x) Y ?x. <i>Correct Answer</i> =true ; )	<i>Item</i> <i>Correct Answer</i>	?x
<i>Question</i>	De un tema puede haber varias preguntas	N/A	<i>Topic</i> <i>Item</i>	N/A
<i>Assessment</i>	Una evaluación se compone de una serie de preguntas	N/A	<i>Quiz</i> <i>Item</i> <i>Distractor</i>	N/A

— **Tarea 10: Definir reglas**

En esta tarea se definen las reglas necesarias para la ontología de E-learning, se definen en la Tabla 7, recordando que para cada regla, se especifica el nombre, descripción, expresión y conceptos.

**Tabla 7.** Reglas.

Nombre	Descripción	Expresión	Conceptos
Barajar opciones	Se barajan las opciones siempre y cuando el atributo no se haya seleccionado.	Si [ <i>Quiz</i> (?x) Y <i>Item</i> (?y) Y ?x <i>include</i> ?y ] ENTONCES [ chose(?x, ?y) shuffe (?x, ?y)]	<i>Quiz</i> <i>Item</i>
Mostrar <i>Item</i> 's más frecuentes.	Muestra los <i>Item</i> 's utilizados con mayor frecuencia.	Si [ <i>Quiz</i> (?x) Y <i>Item</i> (?y)Y chose (?x, ?y) ] ENTONCES ?y.masUsu=?y. masUsu + 1	<i>Quiz</i> <i>Item</i>
Mostrar <i>Item</i> 's menos frecuentes.	Muestra los <i>Item</i> 's utilizados con menor frecuencia.	Si [ <i>Quiz</i> (?x) Y <i>Item</i> (?y)Y chose (?x, ?y) ] ENTONCES ?y.menUsu=?y.menUsu + 1	<i>Quiz</i> <i>Item</i>

— **Tarea 11: Definir instancias**

En esta tarea se definen las instancias de conceptos de la ontología, cada instancia cuenta con nombre, nombre del concepto al que pertenece y los valores de sus atributos de instancia, si se conocen. Esta tarea no aplica por el momento hasta la creación de la ontología del dominio de E-learning.

### 3. Resultados y discusión

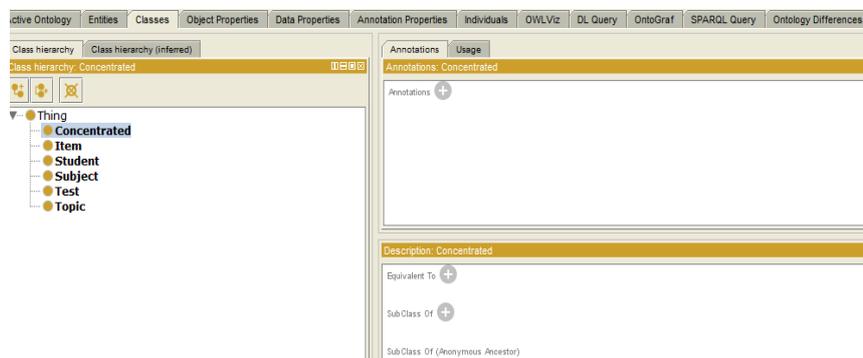
El proceso de evaluación en línea es una estrategia que permite saber el grado de aprendizaje de un alumno en un tema específico. Un buen Quiz cuenta con distractores que cumplen con la parte pedagógica de un modelo educativo, para ello se analizaron diversas formas de generar un distractor. Específicamente en el área de las matemáticas, por ejemplo en álgebra, se ha observado que uno de los errores más comunes que cometen los estudiantes es no respetar el orden de los operadores o de los paréntesis cuando tratan de resolver alguna ecuación, por lo que una forma de crear distractores es cambiando el orden de los operadores o de los paréntesis.

Siguiendo la metodología propuesta por Methontology, lo que se obtuvo al final de las 11 tareas descritas anteriormente fue la identificación de los conceptos principales de E-learning y la relación entre ellos, los cuales fueron tomados como base para la obtención de un diseño conceptual de una ontología que sirve de soporte para el proceso de creación de evaluaciones de asignaturas técnico-científicas. Finalmente, con base en el modelo obtenido, se procedió a la codificación de la ontología utilizando el software Protégé.

**Tabla 8.** Codificación en Protégé.

Concepto	Relaciones	
<i>Teacher</i>	<i>creates_a -&gt; Group</i> <i>makes_an-&gt; Quiz</i>	<i>elaborates_a-&gt; Item</i> <i>teaches-&gt;Subject</i>
<i>Quiz</i>	<i>is_applied_to-&gt; Group</i> <i>includes-&gt;Item</i>	<i>solved_by_a-&gt;Student</i>

El siguiente paso fue la codificación de la ontología con la herramienta Protégé, porque permite generar código de un lenguaje ontológico. A continuación se muestra la descripción de las pantallas principales donde se visualiza la creación de la misma. La Fig. 3 muestra el panel de clases, la cual, permite definir las, establecer la jerarquía y restricciones. En la Fig. 4 se visualiza la creación de las relaciones binarias.



**Fig. 3.** Creación de clases del dominio de E-learning en Protégé.

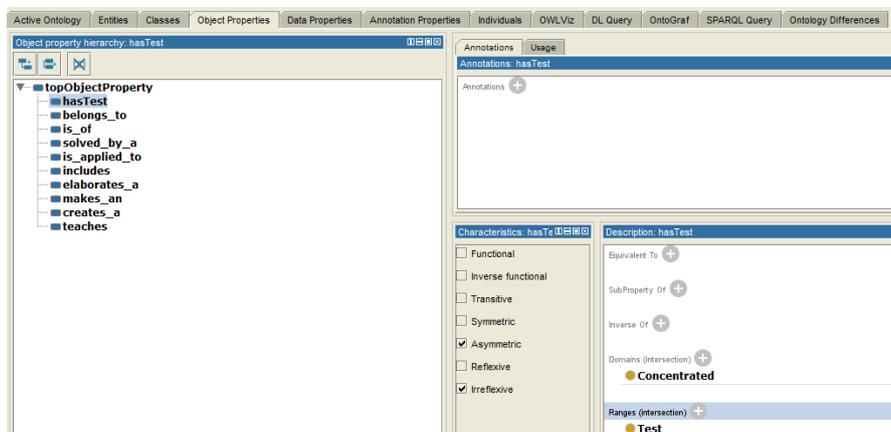


Fig. 4. Relaciones binarias del dominio E-learning.

Para la comprobación del diseño de la ontología se agregaron algunos individuos (instancias) para comprobar que las relaciones y las clases estén correctamente diseñadas. En la Fig. 5 se observan algunos los individuos de los conceptos de *Teacher*, *Topic*, *Student* y *Subject* con sus respectivos atributos de clase que se les proporcionaron para realizar algunas consultas basadas en SPARQL.

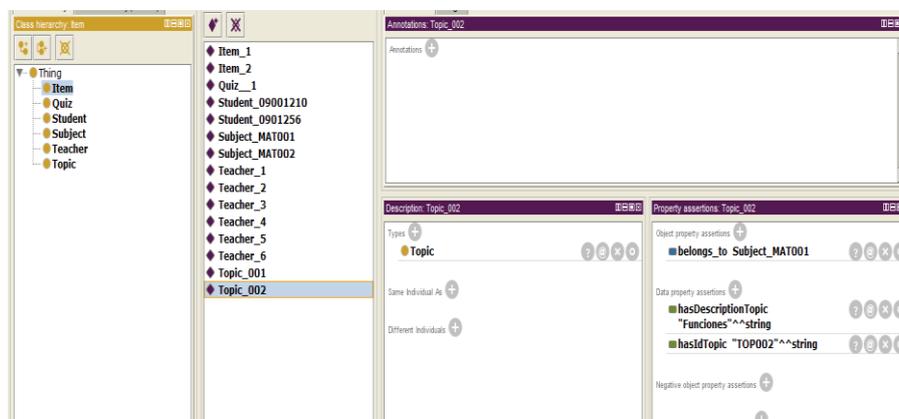


Fig. 5. Instancias agregadas a la ontología del dominio E-learning.

El lenguaje SPARQL (*Protocol and RDF Query Language*) está diseñado para elaborar consultas sobre datos representados para RDF (*Resource Description Framework*), el cual contiene funciones de consultas requeridas y opcionales para ser utilizadas al momento de expresar los resultados a través de diversas fuentes de datos [13]. A continuación, en la Fig. 6, Fig. 7 y Fig. 8 se muestran los resultados de algunas consultas que se elaboraron sobre la ontología y su respectivo código escrito en SPARQL. Es importante mencionar que hemos omitido los prefijos en las consultas basadas en SPARQL para optimizar espacio en el presente artículo.

a) Buscar el id y nombre de los profesores.

```
SELECT ?id ?name
WHERE {
  ?x rdf:type itoTest:Teacher.
  ?x itoTest:hasIdTeacher ?id.
  ?x itoTest:hasNameTeacher ?name
}
```

b) Buscar el id y descripción de las asignaturas técnico-científicas.

```
SELECT ?id ?description
WHERE {
  ?x rdf:type itoTest:Subject.
  ?x itoTest:hasIdSubject ?id.
  ?x itoTest:hasDescriptionSubject ?description
}
```

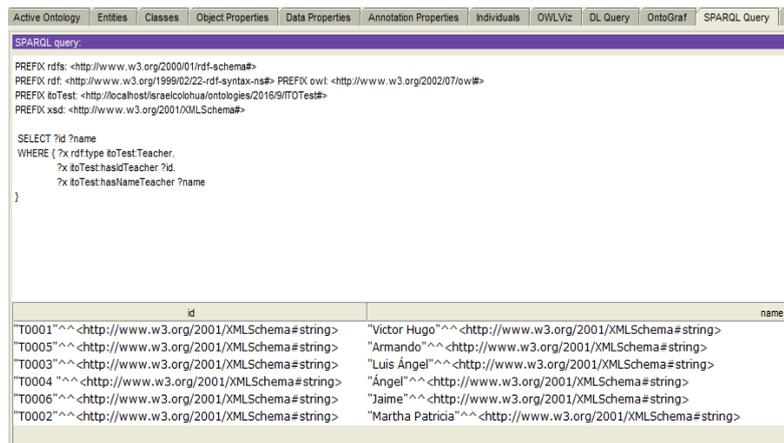


Fig. 6. Resultado de la consulta sobre los profesores.

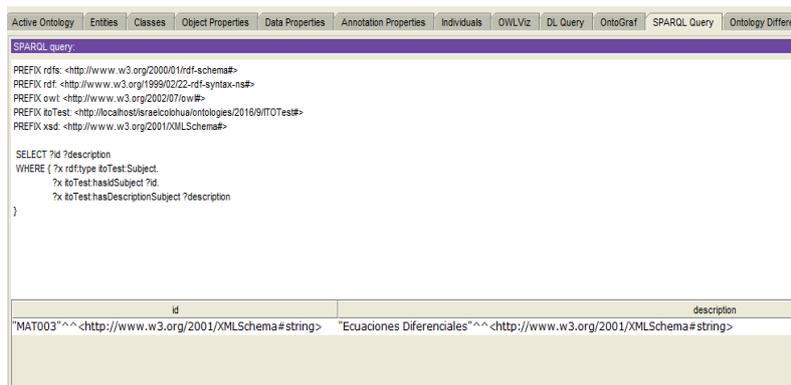


Fig. 7. Resultado de la consulta en las asignaturas.

- c) Buscar la descripción y respuesta correcta de los ítem's disponibles

```
SELECT (str (?des) as ?decripcion) ?rc
WHERE {?x rdf:type itoTest:Item.
       ?x itoTest:hasTextItem ?des.
       ?x itoTest:hasCorrectAnswer ?rc}
```

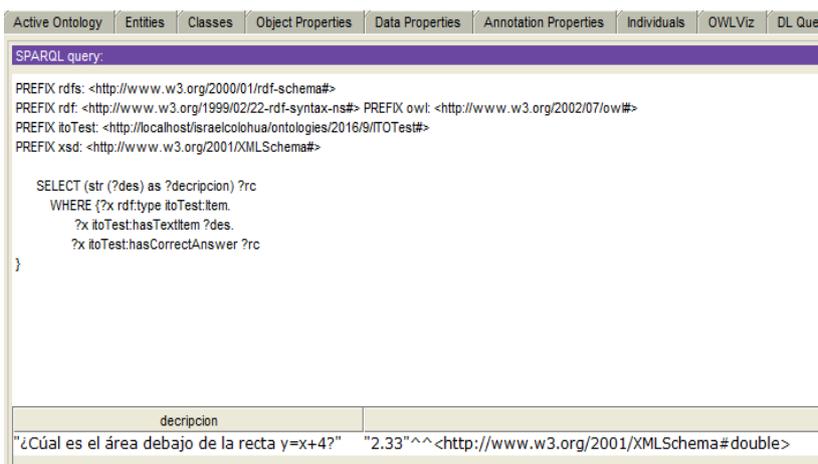


Fig. 8. Resultado de la consulta en los ítem's.

#### 4. Conclusiones

En la actualidad, los contenidos, materiales y cursos que son impartidos en línea son ofrecidos al estudiante a través de plataformas de E-learning, que comúnmente son aplicaciones Web que aplican herramientas y servicios propios de Internet, y que además ofrecen un espacio para configurar y organizar todo el material o recurso para el aprendizaje y lograr que el alumno adquiera conocimiento. A pesar de que existen diversas plataformas con las características antes mencionadas, una aplicación dotada de Web semántica es una de las mejores alternativas para implementar el aprendizaje en línea, ya que el uso de las ontologías tiene un gran potencial para este dominio debido a que una de sus principales características es que infieren conocimiento a partir de las definiciones dadas, es por eso que el desarrollo de la ontología presentada con base a los principios del E-learning facilita el proceso de evaluación de las asignaturas técnico-científicas en el ITO.

#### 5. Trabajo a futuro

Tomando en cuenta la importancia que tiene la educación como uno de los pilares fundamentales para el desarrollo de los países, como trabajo a futuro se pretende desarrollar una aplicación Web en la cual se implemente la ontología presentada en este

trabajo. De igual manera, se pretende que dicha aplicación haga uso de los algoritmos de barajeo de preguntas y respuestas, con la particularidad de que, para la generación de respuestas, se definan y apliquen métodos para la generación de distractores, esto con base en los investigación que se ha realizado para este trabajo acerca de los errores más comunes que se comenten en evaluaciones de matemáticas. Lo que se buscará con esto es aportar facilidad a la elaboración de *quiz* (exámenes) correspondientes a materias del área de matemáticas (técnico-científicas), a partir de lo que la Web semántica, en específico las ontologías, puedan beneficiar a la educación.

## Referencias

1. Antonio, L. R. D.: Entorno Web para la generación de reactivos en la evaluación de aplicación de derivadas. Tesis de Maestría, Departamento de posgrado e investigación, Instituto Tecnológico de Orizaba, Orizaba, Ver, Méx. (2014)
2. Baelo, A. B.: El E-learning, una respuesta educativa a las demandas de las sociedades del siglo xxi. España, Universidad de León (2009)
3. González, O. M.: Evaluación de opción múltiple v.s. evaluación tradicional, Un estudio de caso en ingeniería. Revista Virtual, Universidad Católica del Norte (2014)
4. Tabares, J. J., Jiménez, J. A.: Ontología para el proceso evaluativo en la educación superior. Revista Virtual, Universidad Católica del Norte, 42, pp. 68–79 (2014)
5. Ramos, E., Núñez, H.: Ontologías: componentes, metodologías, lenguajes, herramientas y aplicaciones. Caracas (2007)
6. Corcho, O., Fernández, M., Gómez, A., López, A.: Construcción de ontologías legales con la metodología METHONTOLOGY y la herramienta WebODE (2005)
7. Rodriguez, A., Hernandez, G., Colomo, R., Gomez, J. M., Garcia, A., Alor, G., Valencia, R.: Towards an Ontology to Support Semantics Enabled Diagnostic Decision Support Systems. Current Bioinformatics 7, No. 3, pp. 234–45 (2012)
8. Protégé: Protege.stanford.edu (2016)
9. Menéndez, V., Prieto, M., Zapata, A.: Sistemas de Gestión Integral de Objetos de Aprendizaje. Facultad de Matemáticas de la Universidad Autónoma de Yucatán (2010)
10. Clarenc, C. A.: Analizamos 19 plataformas de E-learning. Disponible en: [www.congresoelearning.org](http://www.congresoelearning.org), Grupo GEIPITE, Congreso Virtual Mundial de e - Learning (2013)
11. Tavangarian, D., Leypold, M. E., Nölting, K., Röser, M.: Is e-Learning the Solution for Individual Learning? University of Rostock, Germany (2004)
12. Bacon, D.: IMS Question and Test Interoperability. MSOR Connections, 3(3) (2003)
13. W3C.: w3.org. Disponible en: <https://www.w3.org/TR/sparql11-query> (2016)



## Un algoritmo para detectar la polaridad de opiniones en los dominios de laptops y restaurantes

Karen L. Vazquez, Mireya Tovar, Darnes Vilariño, Beatriz Beltrán

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

karnlet@gmail.com, {mtovar, darnes, bbeltran}@cs.buap.mx

**Resumen.** En este artículo se presenta un estudio para la detección de Polaridad en un conjunto de opiniones de usuarios emitidos hacia Restaurantes en el idioma español e inglés, así como para opiniones escritas a valoraciones de Laptops en el idioma inglés. Como solución propuesta se emplea Máquina de Soporte Vectorial con validación cruzada. Los mejores resultados se presentaron para el Dominio de Restaurantes en español con un 74% de precisión.

**Palabras clave:** Minería de opiniones, aprendizaje supervisado, procesamiento de lenguaje natural.

## An Algorithm to Detect the Polarity of Opinions in the Domains of Laptops and Restaurants

**Abstract.** In this article, we present a study for the detection of polarity over a dataset of opinions of users over Restaurants in the Spanish and English language, as well as, opinions about Laptops in the English language. As a proposed solution, we use Vector Support Machine with cross validation. The best results were presented for the Spanish Restaurant Domain with 74% accuracy.

**Keywords:** Opinion mining, supervised learning, natural language processing.

### 1. Introducción

El fácil acceso que tenemos a Internet y a su vez, a las grandes cantidades de información que se producen en la Web, la Inteligencia Artificial y más específicamente, el Procesamiento de Lenguaje Natural (PLN) proporcionan mecanismos de extracción de información. La información que se encuentra en Internet se presenta en la mayoría de los casos de manera no estructurada, un ejemplo de ello son las redes sociales, fuente de acceso a opiniones, productos o servicios que la sociedad genera a diario en estos sitios, esta información puede ser una fuente para la aplicación del PLN, que se encarga de la detección automática de los sentimientos

expresados en los textos y su clasificación según la polaridad que tienen, es el área de análisis de sentimientos, también llamada minería de opiniones.

En este documento se muestra la propuesta de solución para el análisis de sentimientos en opiniones expuestas por usuarios a través de las redes sociales, se revisan dos dominios de opiniones, Restaurantes y Laptops, el primero de ellos en el idioma español e inglés y el último únicamente en el idioma inglés. El objetivo es detectar la polaridad de cada opinión, es decir, clasificar la polaridad de la reseña dada por el usuario, siendo esta positiva, negativa, neutra o conflicto. Para el logro de este objetivo, se propone el uso de un clasificador, el seleccionado en esta investigación es Máquina de Soporte Vectorial utilizando validación cruzada para la prueba de datos, los resultados obtenidos muestran un desempeño adecuado con el algoritmo propuesto.

La distribución de este artículo se describe a continuación, en la Sección 2 se presentan los trabajos relacionados con esta investigación, en la Sección 3 se presenta el algoritmo o modelo propuesto, en la Sección 4.1 se presenta la información de los datos utilizados, y los resultados logrados en la Sección 4.2. Finalmente, las conclusiones se presentan en la Sección 5.

## **2. Trabajo relacionados**

El monitoreo de opiniones ha sido estudiado desde hace algunos años a través del Procesamiento del Lenguaje Natural, sin embargo, los estudios e investigaciones realizadas se han logrado en la mayoría de los casos en idiomas diferentes al español, los sistemas para monitoreo de opiniones realizados en esos idiomas y principalmente en el idioma inglés, han tenido resultados favorables. A continuación se describen algunos trabajos relacionados con esta investigación.

En [1] se describe el sistema de minería de opinión llamado “sentie”, que pretende determinar la polaridad del sentimiento expresado sobre algún aspecto concreto de una entidad. Como sistema de clasificación se utiliza MALLETT, y como característica utiliza las palabras textuales y los lemas. Este sistema participó en la tarea 12 de SemEval-2015 obteniendo un 79% de precisión en el dominio de laptops, para la determinación de la polaridad de sentimientos de un texto dado.

En el trabajo de investigación desarrollado por [2] se presenta una contribución a la tarea 5 de SemEval 2016, que se enfocaron en los lenguajes de inglés y francés para el dominio de restaurantes. Su sistema se basa en modelos compuestos, que mezcla rasgos lingüísticos con algoritmos de aprendizaje automático. De acuerdo con los resultados observados, los autores obtuvieron un 88% de precisión para el dominio de restaurantes (idioma inglés), y un 78% de precisión para la determinación de la polaridad en el dominio de restaurantes (lengua francesa).

En [3], los autores describen el sistema que utilizan en la tarea 5 de SemEval 2016. Su sistema se basa en aprendizaje automático supervisado, utilizando un clasificador de máxima entropía, campo aleatorio condicional, y un gran número de características tales como vectores globales, la asignación de Dirichlet latente, bolsa de palabras, iconos gestuales y otros. Se obtuvieron resultados muy competitivos en la tarea 5 de SemEval-2016.

En el trabajo de [4], se propone un esquema de ponderación supervisada en base a dos factores: la importancia de un término en un documento (ITD) y la importancia de

un término para expresar el sentimiento (ITS). Los resultados experimentales muestran que el método produce la mejor precisión en dos de tres conjuntos de datos.

En el estudio [5] se propuso la minería de opiniones basada en Ontologías, enfocado a textos de opiniones provenientes de usuarios con escritura en idioma inglés, proporcionando un nuevo método para el análisis de los sentimientos basados en el análisis vectorial.

Para los estudios realizados en [6], se propuso un método automático para clasificar la polaridad de las opiniones en consumidores de productos de una empresa. El algoritmo se basa en el uso de ontologías para encontrar todas aquellas opiniones realizadas a sus productos incluyendo dentro del análisis de opiniones una fusión de ontologías. Los autores concluyen que conforme aumenta el número de términos de cada ontología, no se nota una mejora en los métodos utilizando ontologías de alto nivel respecto a los métodos directos.

En el trabajo desarrollado en [7] se presentan dos modelos para descubrir la polaridad de mensajes en redes sociales, en particular extraídos del Twitter.

El primer modelo extrae las características léxico-sintácticas de cada tweet. El segundo modelo obtiene las características de cada tweet basándose en la centralidad de grafos.

En el trabajo de [8] clasifican los sentimientos en tweets, de acuerdo a su contenido en positivo, negativo o neutro; donde proponen mejoras a las dependencias objetivo de la clasificación de sentimientos en twitter con la incorporación de características dependientes de objetivos mediante tres pasos, en el primero; se hace una clasificación subjetiva para decidir si hay subjetividad en el tweet o es neutro, en el segundo paso, para los elementos subjetivos se le da su polaridad (positivo o negativo) y finalmente se realiza una optimización basada en grafos usando tweets que están relacionados. Para los dos primeros pasos hacen uso de máquinas de soporte vectorial (SVM-Light). Con las características dependientes de objetivos obtienen una precisión del 85.6%.

Los autores de [9] crean un corpus para los experimentos, añadieron juicios de polaridad contextuales a las anotaciones existentes en el Corpus de Opinión multi-perspectiva Pregunta-Respuesta (MPQA, por sus siglas en inglés), mediante un esquema de anotación. Para los experimentos usaron un lexicón con pistas subjetivas, que fueron agrupadas de acuerdo a su confianza y subjetividad. Posteriormente se expandió el lexicón mediante un diccionario y un tesoro. El algoritmo les permite identificar la polaridad contextual automáticamente para un amplio conjunto de expresiones de sentimientos. El mejor resultado obtenido considera 28 características y logran una precisión de 75.9%.

En Bakliwal et al. [10], construyen un conjunto de datos de tweets políticos. Cada tweet de este conjunto esta anotado como positivo, negativo o neutro. Se incluyeron tweets de sarcasmo. En sus experimentos omitieron el sarcasmo y lograron una precisión de casi el 59% con operaciones simples de búsqueda léxica.

En [11] presentan un enfoque híbrido para determinar el sentimiento de cada tweet. Realizan un pre-procesamiento analizando abreviaturas, lematización, eliminación de palabras cerradas, etc. Probaron seis conjuntos de datos de Twitter, y logran un 83.3% de medida  $F_1$ , una precisión de 85.7% y un recall de 82.2%.

En la siguiente sección se describe el algoritmo utilizado para el análisis de sentimientos.

### 3. Algoritmo propuesto

A continuación se describe el algoritmo propuesto para la determinación de la polaridad de las opiniones proporcionadas en la Tarea 5 de SemEval 2016 [12].

En el algoritmo se inicia con un pre-procesado de datos, en esta fase se realiza una limpieza de la información como es la eliminación de signos de puntuación, eliminación de palabras cerradas, etc. En la segunda fase se realiza la extracción de características utilizando la biblioteca *sci-kit-learn* que es una herramienta de aprendizaje automático en Python [13], [14]<sup>1</sup>. En la tercera fase se utiliza un sistema de clasificación utilizando Máquinas de Soporte Vectorial proporcionado por la herramienta de aprendizaje automático *sci-kit-learn*. En la Fig. 1 se muestra gráficamente, el algoritmo propuesto.

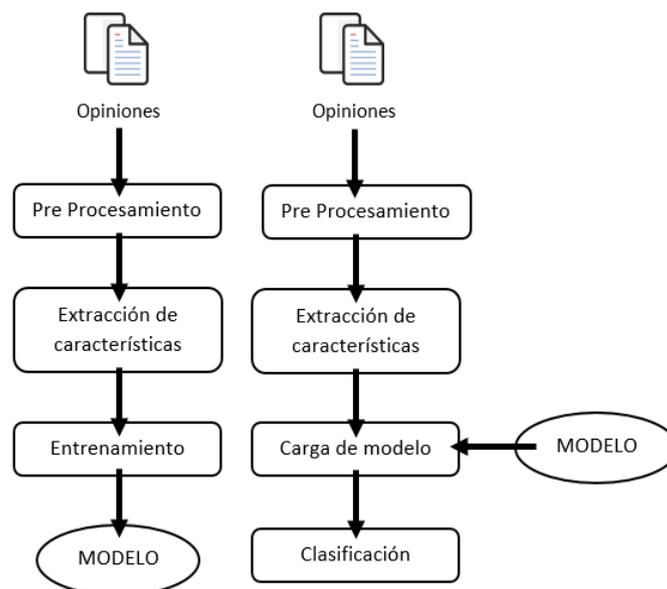


Fig. 1. Algoritmo propuesto.

- Pre procesamiento:
  - Extracción de opiniones del documento en formato XML: Filtrar únicamente opiniones del documento
  - Limpieza de opiniones: Eliminar palabras vacías o cerradas, signos de puntuación, acentos y caracteres aislados.
  - Tokenización: Tokenizar opiniones por palabra.
  - *Stemming*: Proceso heurístico que corta el final de las palabras y con frecuencia incluye la eliminación de los afijos derivativos: coches = coche, fuimos = fui.

<sup>1</sup> <http://scikit-learn.org/>

- Filtrar Categorías: Filtrar opiniones de los datos de entrenamientos por polaridad como: POSITIVO, NEGATIVO, NEUTRAL y CONFLICTO.
- Extracción de características:
  - Frecuencia y Frecuencia Inversa del Documento para un Término (*TF-IDF*): Este método indica qué tan relevante es la palabra con respecto al documento seleccionado y al corpus en general. Permite también calificar a los documentos del corpus con base en estas palabras clave, es decir, si las palabras tiene mayor peso, entonces el documento está más relacionado con ellas que uno con las mismas palabras pero con menor peso. La herramienta *sci-kit-learn* de Python proporciona varios *vectorizers*<sup>2</sup> de traducción de los documentos de entrada en vectores de características: *TfidfVectorizer*<sup>3</sup> es un vectorizer que usa TF-IDF como esquema de ponderación.
  - Sistema de clasificación:
    - En esta se utiliza el método de Máquina de Soporte Vectorial (SVM) que es una técnica de aprendizaje automático. En este caso se utiliza SVC que se basa en libsvm.

## 4. Resultados

En este apartado se describen los datos utilizados para el análisis de opiniones y los resultados correspondientes a cada dominio, con el algoritmo propuesto.

### 4.1. Conjunto de datos

Los datos utilizados son los datos de entrenamiento proporcionados por SemEval-2016 para la subtarea 2 de la tarea 5. Los dominios considerados son el de Restaurantes para el idioma español e inglés, y el dominio de Laptops para el idioma inglés. En la Tabla 1 se muestra el total de opiniones por dominio, idioma y polaridad.

**Tabla 1.** Total de opiniones por dominio, idioma y polaridad.

Dominio	Positivo	Negativo	Neutral	Conflicto
Restaurantes (español)	1,519	443	101	58
Restaurantes (inglés)	1,012	327	55	41
Laptops (inglés)	1,210	707	123	41

### 4.2. Resultados experimentales

Siguiendo el algoritmo, presentado en la sección 3, se realiza una evaluación de la polaridad detectada con el sistema de clasificación utilizando como característica TF-IDF. En las Tablas 2, 3 y 4 se presentan el total de opiniones evaluadas por Polaridad,

<sup>2</sup> Función utilizada en scikit-learn <http://scikit-learn.org/>.

<sup>3</sup> Convierte una colección de documentos sin formato a una matriz de características TF-IDF.

el resultado de *Precisión*, *Recall* y  $F_1$  por dominio, así como los datos de prueba por polaridad. Cabe mencionar que la realización de la prueba de datos se hizo con validación cruzada<sup>4</sup>, en la cual se toma cierto porcentaje para entrenar los datos y el porcentaje restante para la prueba. En este caso se considera el 16% como prueba de datos en todos los dominios y un promedio de 10 repeticiones. Es por eso que en detalle, en la Tabla 2 se muestran los resultados del dominio de Laptops en inglés y en la última columna se desglosa la división de opiniones evaluadas por polaridad, así mismo se presenta que el total de opiniones evaluadas fueron 304, lo mismo para la Tabla 2 pero en el caso de Restaurantes en idioma inglés, presentando un total de prueba de 222 opiniones evaluadas y por último en la Tabla 3, donde se observa que el total de pruebas para Restaurantes en idioma español fueron 343 opiniones.

**Tabla 2.** Resultados del algoritmo para el dominio de Laptops del idioma inglés.

Polaridad	Precisión	Recuerdo	$F_1$	Total de prueba
Conflicto	0.00	0.00	0.00	11
Negativo	0.64	0.63	0.63	62
Neutral	0.00	0.00	0.00	25
Positivo	0.78	0.92	0.84	206
Total	0.66	0.75	0.70	304

**Tabla 3.** Resultados del dominio de Restaurantes del idioma inglés.

Polaridad	Precisión	Recuerdo	$F_1$	Total de prueba
Conflicto	0.00	0.00	0.00	5
Negativo	0.67	0.33	0.44	43
Neutral	0.00	0.00	0.00	9
Positivo	0.79	0.96	0.87	165
Total	0.72	0.78	0.73	222

**Tabla 4.** Resultados del dominio de Restaurantes del idioma en español.

Polaridad	Precisión	Recuerdo	$F_1$	Total de prueba
Conflicto	0.00	0.00	0.00	14
Negativo	0.67	0.47	0.55	60
Neutral	0.00	0.00	0.00	11
Positivo	0.82	0.96	0.89	258
Total	0.74	0.80	0.76	343

## 5. Conclusiones

En este artículo se presenta un algoritmo de clasificación automática que permite identificar la polaridad de las opiniones proporcionadas en la tarea 5 de SemEval 2016. Los dominios considerados en las pruebas son de Restaurantes en español e inglés, y el dominio de Laptops para el idioma inglés. En base a los resultados obtenidos se observa que el método de SVM logra obtener resultados satisfactorios al obtener más del 66% de precisión en los tres dominios con la característica de TF-IDF. El algoritmo

<sup>4</sup> Valoración optimista que emplea los propios datos de entrenamiento para evaluar el modelo.

propuesto, no utiliza información adicional para enriquecer las opiniones proporcionadas por Semeval-2016, motivo por el cual el algoritmo no logra clasificar adecuadamente las opiniones de conflicto y neutral. Como trabajo a futuro se tiene considerado el enriquecimiento de los datos de entrenamiento para lograr mejorar la clasificación de las opiniones que tienen poca información. Así como el uso de otros métodos de clasificación supervisada y su prueba con los datos de entrenamiento y pruebas proporcionadas por Semeval-2016.

**Agradecimientos.** Esta investigación es parcialmente apoyada por el proyecto PRODEP-SEP ID 00570 (EXB-792) DSA/103.5/15/10854, por el proyecto ID 00570 VIEP-BUAP. Apoyado por el Fondo Sectorial de Investigación para la Educación, proyecto Conacyt 257357.

## Referencias

1. Saias, J.: Sentiue: Target and aspect based sentiment analysis in semeval-2015 task 12. In: Proceedings of the 9th International Workshop on Semantic Evaluation, Denver, Colorado, Association for Computational Linguistics, pp. 767–771 (2015)
2. Brun, C., Perez, J., Roux, C.: Xrce at semeval-2016 task 5: Feedbacked ensemble modeling on syntactico-semantic knowledge for aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation, San Diego, California, Association for Computational Linguistics, pp. 282–286 (2016)
3. Hercig, T., Brychcín, T., Svoboda, L., Konkol, M.: Uwb at semeval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation, San Diego, California, Association for Computational Linguistics, pp. 354–361 (2016)
4. Deng, Z. H., Luo, K. H., Yu, H. L.: A study of supervised term weighting scheme for sentiment analysis. *Expert Systems with Applications*, 41, pp. 3506–3513 (2014)
5. Peñalver, I., Garcia, F., Valencia, R., Rodríguez, M. A., Moreno, V., Fraga, A., Sánchez, J. L.: Feature-based opinion mining through ontologies. *Expert Systems with Applications*, 41, pp. 5995–6008 (2014)
6. Balaguer, E. V., Rosso, P., Locoro, A., Mascardi, V.: Análisis de opiniones con ontologías. *Polibits*, 41, pp. 29–36 (2010)
7. Sanzón, Y. M., Vilariño, D., Somodevilla, M. J., Zepeda, C., Tovar, M.: Modelos para detectar la polaridad de los mensajes en redes sociales. *Research in Computing Science*, 99, pp. 29–42 (2015)
8. Jiang, L., Yu, M., Zhou, M., Liu, X., Zhao, T.: Target-dependent twitter sentiment classification. In: The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, Portland, Oregon, USA, pp. 151–160 (2011)
9. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase level sentiment analysis. In: HLT/EMNLP 2005, Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, Vancouver, British Columbia, Canada (2005)
10. Bakliwal, A., Foster, J., van der Puil, J., O'Brien, R., Tounsi, L., Hughes, M.: Sentiment analysis of political tweets: Towards an accurate classifier. In: Proceedings of the Workshop on Language in Social Media, Atlanta, Georgia, Association for Computational Linguistics, pp. 49–58 (2013)
11. Khan, F. H., Bashir, S., Qamar, U.: Tom: Twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems*, 57, pp. 245–257 (2014)

12. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez, S. M., Eryigit, G.: Semeval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation, San Diego, California, Association for Computational Linguistics, pp. 19–30 (2016)
13. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, pp. 2825–2830 (2011)
14. Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., Varoquaux, G.: API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pp. 108–122 (2013)

## **Avances en el desarrollo de un clasificador de imágenes termográficas de planta del pie diabético basado en una red neuronal de retropropagación**

José Federico Ramírez Cruz, Edmundo Bonilla Huerta, Lauro Reyes Cocolletzi,  
José Crispín Hernández Hernández

Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala,  
México

**Resumen.** Actualmente el incremento de enfermedades crónicas (diabetes) y el diagnóstico preventivo en el campo de la medicina nos llevan a usar la tecnología para prevenir complicaciones a largo plazo o incluso a predecir con un rango de confiabilidad el porcentaje de probabilidad a ser propensos a estas patologías. En la literatura médica se tiene una relación entre la temperatura de algunas zonas de las plantas del pie en relación a el flujo sanguíneo y el grado de azúcar en este, por lo que en este trabajo se presenta una propuesta para la implementación de una red neuronal (tipo retropropagación) que será capaz de realizar la clasificación de patrones de imágenes digitalizadas en base a la variación térmica e identificar patrones para la clasificación de angiosomas planares de individuos sanos y posibles individuos diabéticos no diagnosticados (a través de anomalías en la lectura de la imagen térmica). La red neuronal de retropropagación implementada tiene la característica que pre-clasifica a través del almacenamiento de los pesos de los patrones de entrenamiento, para evaluar la mayor relación (a través de la correlación de Pearson) existente entre las distintas imágenes evaluadas lo que conlleva a un mayor grado de confiabilidad en la clasificación.

**Palabras clave:** Correlación de Pearson, red neuronal de retropropagación, angiosomas planares.

## **Advances in the Development of a Thermographic Image Classifier of Diabetic Foot Plant Based on Backpropagation Neural Network**

**Abstract.** Currently the increase in chronic sickness (diabetes) and preventive diagnosis in the medical field lead us to use technology to prevent long-term complications or even to predict with a range of reliability the percentage of probability to be prone to these pathologies. In the medical literature there is a relationship between the temperature of some areas of the plant in relation to the blood flow and the degree of sugar in the foot, so in this paper we present a proposal for the implementation of a neural network (Type backpropagation) that will be able to perform the classification of patterns of digitized images based on

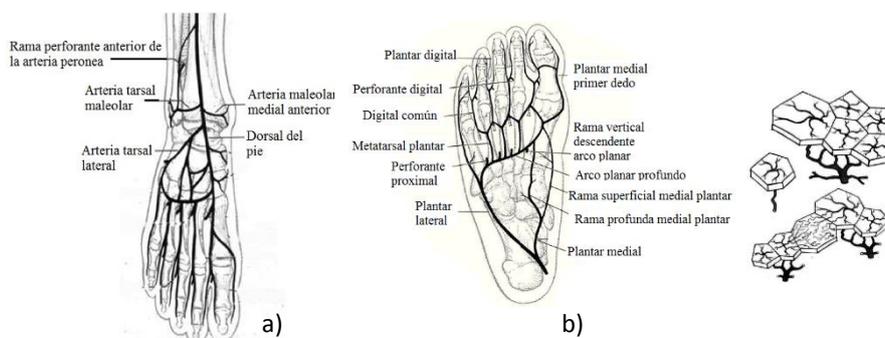
the thermal variation and to identify patterns for the classification of plantar angiosomes of healthy individuals and possible undiagnosed diabetic individuals (through abnormalities in the reading of the Thermal imaging). The implemented backpropagation neural network has the characteristic that pre-classifies through the storage of the weights of the training patterns, to evaluate the greater relation (through the Pearson correlation) existing between the different evaluated images which leads to a higher degree of reliability in the classification.

**Keywords:** Pearson correlation, backpropagation neural network, planar angiosomes.

## 1. Introducción

El pie tiene importantes funciones como la sustentación del peso corporal, la postura bípeda, el equilibrio y la locomoción, condiciones indispensables para la mayoría de las actividades del ser humano. La vitalidad de sus estructuras depende de su sistema arterial, cuyo elemento central es el arco plantar profundo, del que se originan las arterias metatarsianas plantares, fundamentales para la irrigación de los dedos y de la región plantar profunda. En el paciente sano el flujo arterial al pie es redundante, pero no es así en el paciente isquémico [R. Fernández-Samos Gutiérrez, 2012].

La teoría del angiosoma explica las variaciones que existen en los aportes de sangre arterial a la piel y estructuras adyacentes entre las diferentes regiones del cuerpo, proporciona la base para interpretar muchos procesos fisiológicos y patológicos [R. Fernández-Samos Gutiérrez, 2012].



**Fig. 1.** a) Arterias presentes en el pie, b) arterias planta del pie con sus colaterales, c) esquema de angiosoma.

El objetivo de este trabajo es presentar avances en la cuantificación de anomalías presentes en la imagen térmica de las plantas de los pies a través de una termografía (la termografía arroja características de la irrigación de la sangre que a su vez influye en la temperatura alrededor de la arteria por la teoría del angiosoma) que nos permita evaluar y cuantificar defectos congénitos o patologías para la prevención sin tener que realizar un análisis invasivo sobre el paciente [Neville et al. 2009].

## **2. Estado del arte**

Algunos puntos importantes en la toma y análisis de las imágenes (muestreo) se retoman en base a trabajos previos existentes, estos se consideran a continuación:

Las imágenes térmicas se representan en colores RGB que a su vez se ligan a una escala en grados centígrados ( $^{\circ}\text{C}$ ). A través de imágenes térmicas de ambos pies, se pueden identificar diferencias entre áreas correspondientes del pie derecho y el pie izquierdo con el fin de aportar información cuantitativa acerca de la distribución de la temperatura plantar, por ejemplo la información se obtiene mediante la segmentación de cada imagen térmica en cuatro regiones del angiosoma planar y a través de una clasificación de los píxeles se obtienen regiones diferentes que cubre un área planar [H. Peregrina et al. 2013].

Investigaciones con termografía relacionan la existencia de una asociación positiva entre una alta diferencia asimetría térmica de los angiosomas planares contra el riesgo de padecer diabetes tipo 2 o ulceraciones que conlleven a problemas de amputación [Neves et al. 2015].

En pies diabéticos la ocurrencia de una ulcera es frecuentemente asociada con la hipertermia. La hipertermia se define como una temperatura superior a  $2.2^{\circ}\text{C}$  en una región dada de uno de los pies comparada con la temperatura de la misma región del pie contralateral observable en la termografía [L. Vilcahuaman et al. 2014].

Cabe mencionar que en el estudio realizado por K. Ammer encontró que las deformidades presentes en los dedos (callosidades o varices) no influyen directamente o en mayor medida en la variación de la temperatura censada en los pies con la termografía o con los riesgos en la ulceración de los pies, aunque se debe tomar en consideración estas deformidades para considerar en otras patologías [K. Ammer et al. 2001].

Para pacientes diabéticos diagnosticados el mayor riesgo corresponde la ulceración de los pies, posteriormente infecciones vasculares y culminar con amputación de miembros. La termografía es un medio de bajo costo para determinar complicaciones vasculares debidas al riesgo de ulceraciones en la planta del pie. Esto se lleva en base a la relación ente la osteomielitis correlacionada por la imagen infrarroja y el estudio clínico que valide los resultados obtenidos con termografía [J.R.Harding et al. 1998].

## **3. Metodología**

Se toman imágenes térmicas (formato jpg) a través de una cámara Fluke serie 13ABC a una muestra aleatoria doble (dos fotogramas térmicos por cada sujeto) de 20 individuos en un rango de edad de 20 – 23 años estas imágenes son procesadas digitalmente y son la base para el entrenamiento y pruebas de la programación del clasificador a través de una red neuronal de retropropagación por medio de mecanismos de computo.

Para la adquisición de las imágenes térmicas se propone dar suficientemente tiempo al sujeto de prueba el equilibrar la temperatura ambiente con la temperatura de sus pies (un mínimo de 15 minutos) en un espacio donde la temperatura sea de  $23^{\circ}\text{C} \pm 0.3$  durante la toma de la imagen. Colocar al individuo en posición horizontal y aislar con un fondo negro los pies para prevenir radiación externa que afecte la captura de la

imagen. La cámara infrarroja se puede colocar a 1.1 m de distancia de los pies del sujeto pero en general depende de la resolución (en pixeles) de la cámara con la que se cuenta [L. Vilcahuaman et al. 2014].

La red de retropropagación en una red neuronal supervisada tiene limitantes inherentes como lo es el tiempo de aprendizaje basado en el número de iteraciones (épocas) para su entrenamiento en las salidas deseadas así como variaciones en la salida, pero es bastante buena para clasificar pues su lentitud es compensada con el grado de fiabilidad en el momento de clasificar, esto principalmente, debido a la supervisión de pequeñas cambios de los pesos hasta llegar al entrenamiento y aprendizaje deseado.

Inicialmente se plantea digitalizar la imagen e implementar como entrada a la red neuronal cada pixel y su información de forma individual, esto es, si la imagen termografía tiene una resolución de 750 x 450 pixeles el número de pixeles a procesar es de 337,500 pixeles en tres capas RGB (337,500 de profundidad 3) pero debido a la alta demanda de computo necesario se descarta esta opción y se utilizan regiones térmicas identificables (alrededor de unos cientos de pixeles a procesar como se presenta más adelante) en las muestra obtenidas de las imágenes digitalizadas.

Estas regiones son las zonas frías, calientes y templadas en el área del pie, obtenidas por la cámara termográfica.

La obtención de las regiones frías, calientes y templadas se realiza primero separando el fondo de la forma (la forma es la imagen termográfica del pie), esto se lleva a cabo realizando una transformación de los planos de imagen RGB al formato estándar de imagen YIQ, así se obtienen los pixeles que aportan la luminancia (Y = luminancia ) que pertenecen a la forma y nunca al fondo (esto debido a que la toma de muestra se realiza en un ambiente controlado).

En base a la posición de los pixeles obtenidos de la forma se generan seis cuadrantes como se muestra en la figura 2, para delimitar el uso de pixeles a las regiones que contienen la información de interés.

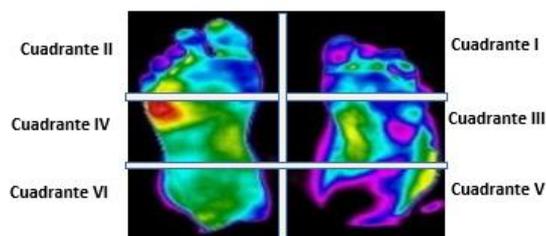


Fig. 2. Ilustración de los cuadrantes definidos para analizar las regiones de interés.

Conociendo la posición de los pixeles que pertenecen a la forma se procede a obtener los valores de la posición de estos pixeles en los tres planos por separado cada uno (valor de la posición del pixel en el plano R, en el plano G, en el plano B). Esta información se almacena en 3 matrices y se evalúa el valor mayor de la posición de los 3 planos y se asigna como el valor predominante y así obtener las áreas templadas, frías y calientes de la imagen termográfica.

Las entradas hacia la red neuronal constan básicamente de la información obtenida de las áreas frías, templadas y calientes en los seis cuadrantes en los que se divide la imagen, así, cada cuadrante tendrá puntos de interés representados por la cantidad de áreas con tonos rojos, verdes y azules. Esto representa la distribución de la temperatura en la imagen y esta información es la base para evaluar el grado de similaridad (para este análisis se usan las dos regiones calientes, templadas y frías de mayor área en cada cuadrante, si se desea se pueden utilizar mayor número de áreas en cada cuadrante).

Se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado en las neuronas de salida con la salida que se desea obtener y se calcula un valor de error para cada neurona de salida. A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida [Freeman et al. 1991].

Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan obtenido un error que describa su aportación relativa al error total. Basándose en el valor del error dado, se reajustan los pesos de conexión de cada neurona, de manera que se almacena estos pesos para el patrón aprendido y en la siguiente vez que se presente el mismo patrón o similar se reconocerá inmediatamente, para el caso de que sea más de un patrón se almacenaran los pesos aprendidos para el siguiente patrón, entonces con respecto a la salida que esté más cercana a la deseada de los pesos de los patrones obtenidos se clasificara de forma más rápida.

Se propone el almacenar los pesos de entrenamiento obtenido para cada uno de los patrones a clasificar para obtener una categorización optimizada en la relación de un patrón nuevo de entrada.

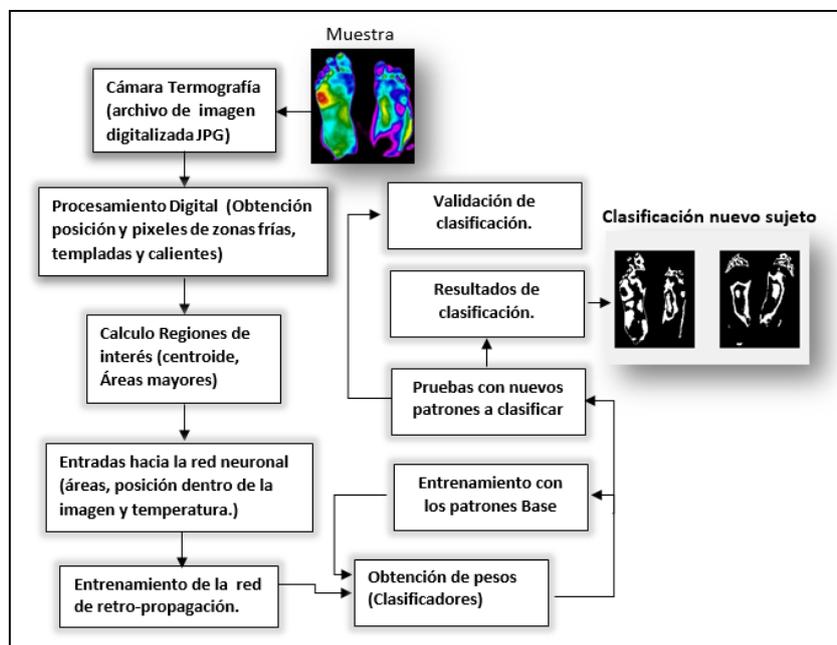


Fig. 3. Diagrama a bloques de operación del algoritmo.

### 3.1. Clasificación y distribución de las regiones de interés en la termografía

**Imágenes digitalizadas.** Las imágenes digitalizadas tienen 5 patrones base deseadas con la que se entrenó a la red neuronal (salida deseada donde el error tienda a cero), las demás imágenes que sean introducidas a la red neuronal deben clasificarse de acuerdo a la similitud que tengan con respecto a estos patrones base.

Estos patrones se toman como base porque representan características muy distintas unas de otras del conjunto de muestreo obtenido de los sujetos de prueba y sirven para representar las anomalías que se podrán encontrar en pacientes por diagnosticar o ya diagnosticados.

Con los 5 patrones base se entrena a la red neuronal para la clasificación de 5 características definidas en la distribución térmica de la planta de los pies.

Estas características se definen como un rango normal y cuatro irregularidades presentes en las muestras de los sujetos para clasificar, la descripción de las categorías se presentan a continuación.

**Patrón 1.** - Al tomar una muestra con la cámara termográfica lo ideal es que se presente simetría en ambas plantas de los pies, como en ocasiones esto no es posible el clasificador también considera este caso, es decir, si algún pie se clasifica en este patrón quiere decir que se debe realizar otra toma de muestra del paciente pues la termografía carece de la simetría mínima requerida para clasificar.

**Patrón 2.**-Simetría mínima en umbral, analiza las regiones de los falanges y la zona angular de la planta del pie, establece patrón sano.

**Patrón 3.**-Pie plano, característica de deformación congénita tratable y candidato a estudios clínicos para establecer propensión a la diabetes. Para el patrón 3, el pie plano influye en encontrar las características de los angiosomas y la probabilidad de presentar un cuadro clínico de diabetes por lo que se recomendaría realizar un estudio de la sangre.

**Patrón 4.**-Anormalidad 1 en la distribución de la temperatura debido a problemas en los angiosomas, pie de sujeto con posible padecimiento diabético.

**Patrón 5.**-Anormalidad 2 en la distribución de la temperatura de los angiosomas debido a problemas de circulación de sangre, posible cuadro diabético.

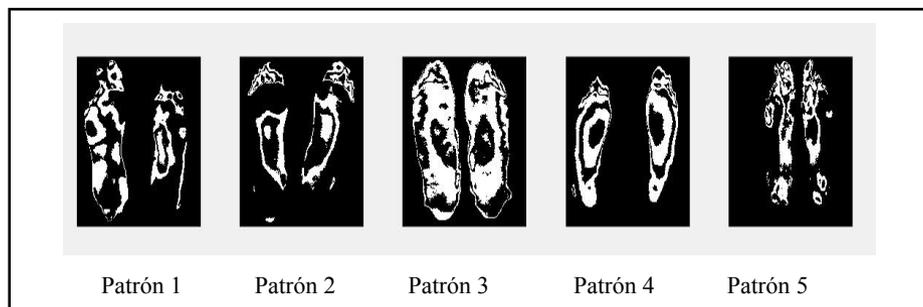


Fig. 4. Patrones base para la clasificación de la salida de la red neuronal.

La binarización de las imágenes se emplea para resaltar en la salida y desplegar resultados obtenidos de la clasificación y solo es de carácter de interface visual al usuario.

La obtención de las áreas o regiones de interés (total de pixeles) que conforman las regiones térmicas distribuidas en la planta del pie se obtienen una vez realizado el procesamiento digital de la imagen, estas áreas y su posición en la imagen definen las entradas de la red neuronal.

Se establecen los valores de entrada con respecto a el área de los pixeles encontrados en los seis cuadrantes utilizados por ejemplo, Área 1=1500 pixeles, Área 2=2058 pixeles, Área 3 = 3156 pixeles, Área 4 = 1204 pixeles, Área 5= 3560 pixeles y Área Total = sumatoria de todas las áreas anteriores. Se divide a cada una de las regiones por un factor de 100000 o cercano al número de pixeles totales que forman a la imagen para obtener una representación proporcional de los pixeles que conforman a las distintas regiones en valores positivos < 1. Esto se realiza por que la red neuronal utiliza datos continuos para el cálculo de los pesos.

Entonces, con esto tendremos una representación de los pixeles que conforman a la imagen y no a cada pixel de manera individual, al final un conjunto de pixeles en una imagen presentan características similares y esto es lo que se obtiene al detectar y procesar las regiones de interés.

Para este caso de estudio se trabaja con 12 regiones para cada imagen y la sumatoria total de estas regiones. Al final la suma de las regiones indica el total de pixeles que se han detectado y que son relevantes de la imagen.

La salida obtenida se clasifica de acuerdo a la correlación de cada salida obtenida con respecto a las salidas aprendidas y clasificadas utilizando el coeficiente de correlación de Pearson, por lo que la red neuronal proporciona la información y clasifica pero el coeficiente de correlación de Pearson interpreta esa información y valida el resultado.

#### 4. Pruebas y resultados obtenidos

Se introducen las imágenes de prueba para observar el funcionamiento del clasificador.

**Tabla 1.** Parámetros de salida de la red neuronal durante el entrenamiento de los patrones base para la clasificación.

Patrones de entrenamiento	Salida red neuronal	
	Patrones binarizada (deseada)	Salida entrenamiento (obtenido)
#1	[1 0 0 1]	[0.999 -0.0231 -0.00269 -0.0394 0.99]
#2	[0 0 1 0]	[-0.000105 0.0191 0.9985 0.0452 -0.0001]
#3	[1 0 1 1]	[0.9991 0.00048 1 0.99893 0.00020]
#4	[0 1 0 1]	[-0.9290 0.99897 -0.8998 -0.3172 0.99904]
#5	[1 1 1 1]	[0.99896 0.99999 0.99904 0.03998 0.99999]

La red neuronal procesa la información dada y clasifica de acuerdo a la mayor similitud que encuentre en los puntos calientes, templados y fríos distribuidos en la

imagen. Se clasifica como uno de los cinco patrones con base en los pesos obtenidos en el entrenamiento de la red neuronal.

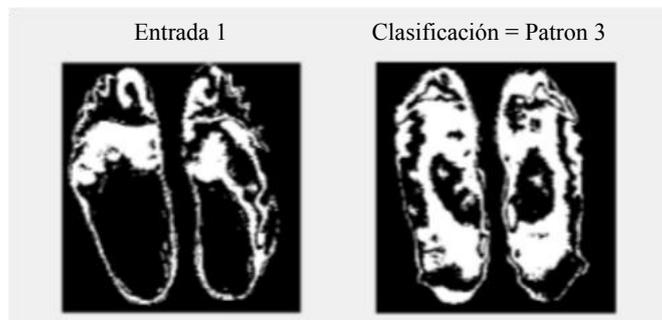
La tabla 1, muestra los valores de trabajo e información puntual de los patrones de entrenamiento de la red neuronal como son la discretización en parámetros binarios de las salidas deseadas.

La tabla 2 a su vez muestra resultados arrojados para 5 patrones de prueba que clasifico la red neuronal, el error y el factor de correlación que definió a que patrón corresponde para su clasificación.

**Tabla 2.** Resultados de 5 patrones de prueba y el error.

Patrones de prueba	Salida entrenamiento	Correlacion de Pearson	Error (comparacion salida deseada- salida obtenida)
#6	[-0.00010 -1.9112e-06 0.99859 -4.5245e-06 -0.00011 ]	0.867	[0.00010 1.9112e-06 0.00141 4.5245e-06 0.00010753]
#7	[0.99932 1 0.99936 -0.00130 0.99999 ]	0.8856	[0.00068 0 0.00064 0.0013029 1e-05]
#8	[-0.00169 0.98841 0.00048 0.00028 0.98895 ]	0.875	[0.00170 0.01159 0.00048 0.00028 0.01105]
#9	[0.99839 -2.9323e-05 -5.0367e-05 -8.5986e-05 0.99839 ]	0.908	[0.00161 2.9323e-05 5.0367e-05 8.5986e-05 0.00161]
#10	[0.94341 -0.00077 0.99998 0.95196 -0.13268 ]	0.9423	[0.05659 0.00077 2e-05 0.04804 0.13268 ]

La respuesta de salida de clasificación obtenida con el clasificador y lo observado de forma visual cualitativa arroja resultados con variación como se muestra en la figura 5, aunque debe de considerarse la región de interés que está tomando el clasificador corresponde a la distribución térmica de las zonas calientes, templadas y frías.



**Fig. 5.** Resultados arrojados por el clasificador, se clasifican como pertenecientes al patrón 3.

El resumen de los resultados de clasificación de las muestras obtenidas de los sujetos de prueba se muestra en la tabla 3.

Cabe resaltar que se introducen por cada sujeto dos muestras y el resultado esperado es que replique la clasificación del patrón de la muestra 1 con la muestra 2, sin embargo en algunos casos no sucede esto. Esto se debe a variaciones entre una muestra y otra en algunos casos sutiles y en otros casos no tanto de ahí que se realiza la validación de los resultados obtenidos a través la validación cruzada de n-iteraciones, obteniendo resultados de una tasa de error de 0.88 %, lo que significa que tiene aún muchas áreas de oportunidad para la mejora del clasificador, pues idealmente se necesita obtener una

tasa de error menor al 0.05% para denotar que el clasificador discretiza con un porcentaje muy alto de validación.

**Tabla 3.** Resultados de clasificación de los sujetos de prueba, en la muestra 1 y la muestra 2.

Sujeto de prueba	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muestra 1 (clasificación arrojada por la red neuronal)	2	3	2	1	1	1	1	2	3	2	4	5	5	4	3	2	1	1	2	5
Muestra 2 (clasificación arrojada por la red neuronal)	2	2	1	4	3	1	1	1	3	3	3	5	1	4	2	2	1	1	3	5
Desviación estandar muestra 1	1,4327																			
Desviación estandar muestra 2	1,3534																			

Se agrega solo como referencia la desviación estándar de los resultados de la distribución de las muestras en cuanto a cuál sería la dispersión de la clasificación de los sujetos de prueba, esto está limitado pues el rango de cambio no es muy grande ya que solo se puede clasificar dentro de cinco patrones posibles.

## 5. Conclusiones

La estabilización de una red neuronal de retropropagación necesita de ciertas características adicionales e inherentes a ella para clasificar de forma adecuada, que pueden ser lógica difusa, Bayes Ingenuo, Control clásico o estadística.

Lo importante es fusionar las redes neuronales con otras metodologías para obtener un resultado optimizado.

Para nuestro caso el coeficiente de correlación de Pearson utiliza las salidas y los 5 patrones de entrenamiento para llevar a cabo la clasificación de las salidas, es decir, se soporta la red neuronal con un análisis estadístico.

Además, se implementa el almacenamiento de pesos con base a los patrones a clasificar por lo que se puede ser una ventaja en algunos casos el entrenar exclusivamente para un patrón y en algunos otros casos una desventaja pero esto dependerá de la aplicación a realizar.

## 6. Trabajo a futuro

Continuar con el desarrollo del clasificador tomando en cuenta los tres planos de colores RGB y la relación con la variación térmica ( $^{\circ}$  C), tratar digitalmente la información a través de técnicas como ecualización de histograma, filtrado digital entre otras técnicas para definir las áreas de interés para obtener mejores resultados a través de un pre-procesamiento de la imagen más robusto.

Un avance mayor sería que agrupara patrones nuevos incrementando automáticamente las opciones de clasificación, es decir, no solo se contara con cinco

patrones posibles a clasificar si no que automáticamente aumentara el número de patrones según se encontraran nuevas características en la distribución de las zonas térmicas del pie.

Realizar cálculos estadísticos de acuerdo a la distribución de las áreas de interés y que afectan en mayor medida al diagnóstico del paciente y no solo de forma cualitativa.

Obtener imágenes termografías de pacientes diagnosticados con diabetes positivo y evaluar el clasificador con estos sujetos de prueba.

Hacer más robusto el clasificador para la detección de ulceraciones y no solo tomar en consideración el procesamiento de las variaciones de temperatura en las plantas de los pies.

Evaluar la posibilidad de tomar toda la información de la imagen termográfica (todos los píxeles) con técnicas robustas como el cómputo paralelo o el uso de GPU's.

## Referencias

1. Peregrina, H., Morales, L. A.: Thermal image processing for quantitative determination of temperature variations in plantar angiosomes. In: IEEE International Instrumentation and Measurement Technology Conference (2013)
2. Neves, E. B., Almeida, J., Rosa, C., Vilaça, J., Reis, V. M., Mendes, R.: Anthropometric Profile and Diabetic Foot Risk: A Cross-Sectional Study Using Thermography. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2015)
3. Vilcahuaman, L., Harba, R., Canals, R., Zequera, M., Wilches, C., Arista, M. T., Torres, L., Arbañil, H.: Detection of Diabetic Foot Hyperthermia by Infrared Imaging. In: Engineering in Medicine and Biology Society 36th Annual International Conference of the IEEE (2014)
4. Ammer, K., Melnizky, P., Rathkolb, O., Ring, E. F.: Thermal Imaging Of Skin Changes On The Feet Of Type II Diabetics. Engineering in Medicine and Biology Society In: Proceedings of the 23rd Annual International Conference of the IEEE (2001)
5. Harding, J. R., Wertheim, D. F., Williams, R. J., Melhuish, J. M., Banerjee, D., Harding, K. G.: Infrared Imaging in Diabetic Foot Ulceration. Engineering in Medicine and Biology Society. In: Proceedings of the 20th Annual International Conference of the IEEE (1998)
6. Neville, R., Attinger, E. C.: Revascularización de angiosomas específicos en isquemia crítica de la extremidad: ¿Es importante la arteria diana? Anales de cirugía Vascul, revista internacional de Cirugía vascular y Terapéutica Endovascular (2009)
7. Fernández, R.: El modelo angiosoma en la estrategia de revascularización de la isquemia crítica. Servicio de Angiología y Cirugía Vascul, Complejo Asistencial Universitario de León, España (2012)
8. Freeman, J. A., Skapura, M. D.: Neural Networks: Algorithms, Applications, and Programming Techniques. California, U.S., Addison-Wesley Publishing Company (1991)
9. Mitchell, T. M., Ryszard, S., Michalski, J. G.: Machine Learning. New York, US, McGraw Hill (1996)

## **Evaluación de una ontología para el intercambio de información materno-infantil entre instituciones de salud en México**

E. V. López Hernández<sup>1</sup>, Beatriz A. Olivares Zepahua<sup>1</sup>, G. S. Hernández Chan<sup>2</sup>,  
José L. Sánchez Cervantes<sup>2</sup>

<sup>1</sup> Instituto Tecnológico de Orizaba, División de Estudios de Posgrado e Investigación,  
Orizaba, Veracruz, México

<sup>2</sup> Instituto Tecnológico de Mérida, Dirección de Ingeniería Eléctrica y Electrónica,  
Yucatán, México

erwinl@acm.org, bolivares@ito-depi.edu.mx, gandhi.hernandez@itmerida.mx,  
cervantes82@msn.com

**Resumen.** El proceso de evaluación de una ontología es una tarea complicada, ya que hay pocas metodologías o herramientas que permitan realizar esta actividad de manera automática y fundamentada. Existen diversos enfoques para realizar la evaluación de ontologías dependiendo del tipo específico y sus propósitos. En este artículo se describe el proceso de evaluación de la calidad de la ontología para el intercambio de información materno-infantil mediante la implementación de la aplicación Web OOPS!, la cual a través de un catálogo de errores comunes permite detectar problemas de diseño y modelado de ontologías.

**Palabras clave:** Ontología, calidad, evaluación, OOPS!

## **Evaluation of an Ontology for the Exchange of Maternal and Child Information Among Health Institutions in Mexico**

**Abstract:** The process of evaluating ontology is a complicated task, since there are few methodologies or tools that allow this activity to be performed automatically and grounded. There are several approaches to perform the evaluation of ontologies depending on their specific type and purpose. This article describes the process of evaluating the quality of the ontology for the exchange of maternal and child information through the implementation of the OOPS! Web application, which through a catalog of common errors allows to detect design and modeling problems of ontologies.

**Keywords:** Ontology, quality, evaluation, OOPS!

## **1. Introducción**

El presente trabajo presenta el proceso de diseño, elaboración y evaluación de una ontología orientada al intercambio de información materno-infantil, es decir, de mujeres y niños en periodo perinatal y neonatal en hospitales de México.

El problema que se ha observado gira entorno a la atención tanto de mujeres embarazadas y recién nacidos que son pacientes de hospitales o centros de salud especializados en esta área de la medicina. El problema radica sobre todo en cuanto al manejo de información y el intercambio de la misma. Esto es, se ha notado que existe una diferencia en la forma en que los hospitales almacenan y procesan la información. Esta diferencia, además de retrasar el proceso de atención, y complicar el intercambio de la misma, provoca problemas como falta e inconsistencia de información.

Cabe señalar, que si bien la información relacionada con expedientes médicos es importante, en el caso particular de la atención materno-infantil lo es aún más, ya que implica un vínculo o relación entre la información de la madre y del hijo, relación que por ninguna razón debe perderse o verse afectada por cuestiones de seguridad.

Dada esta situación, la solución propuesta es el diseño de una ontología que, en primera instancia, permita, de alguna forma, estandarizar los diferentes formatos de la información que se maneja en los distintos hospitales, y, como consecuencia de ello, facilite su intercambio. Todo ello con el objetivo de apoyar a las instituciones de salud a disminuir las tasas de mortalidad materno-infantil en México.

En el proceso de elaboración de la ontología se aplicó la metodología Methontology, la cual proporciona actividades, guías y un ciclo de vida que permiten llevar a cabo el desarrollo de la ontología. Sin embargo, Methontology no describe la manera de evaluar la calidad de lo obtenido. Durante el subproceso de validación y evaluación de la ontología es importante contar con herramientas formales que permitan determinar si la ontología construida cumple con los requerimientos específicos, las necesidades, expectativas o si es posible reusar el conocimiento sobre un dominio determinado. Para este proceso se ha utilizado la herramienta OOPS!

## **2. Diseño de ontología para el intercambio de información materno-infantil**

Las instituciones públicas de salud en México funcionan bajo el esquema de derechohabencia, es decir, una persona debe registrarse ante la institución para tener derecho a recibir atención médico-hospitalaria de dicha institución. Sin embargo, todas las instituciones de salud, públicas o privadas, están obligadas a atender los casos de emergencia que se presenten, con independencia de la derechohabencia del paciente; por supuesto, contar con el expediente clínico del paciente facilita el brindar una atención de calidad ya que permite conocer tanto el historial como la situación actual del paciente más allá de los síntomas de la emergencia. Esto es particularmente importante en el caso de las mujeres embarazadas y los recién nacidos, pues la falta de información relevante en el momento de la atención deriva fácilmente en la muerte. Pese a la importancia de contar con el expediente clínico en el momento de la atención, no es fácil obtenerlo cuando el paciente es derechohabiente de una institución distinta

a la que brinda la atención, debido a que, aun en el caso de expedientes electrónicos, es común que algunos términos, sobre todo administrativos, difieran entre instituciones.

Por lo anterior, se planteó el diseño de una ontología para el intercambio de información materno-infantil entre los sistemas de instituciones de salud en México. La ontología considera requerimientos legales vigentes en México aplicables a la atención materno-infantil y al intercambio de expedientes médicos tales como Norma Oficial Mexicana NOM-004-SSA3-2012 del expediente clínico, NOM-024-SSA3-2012 Intercambio de información en salud, entre otros. Además se implementó el estándar internacional *Health Level Seven (HL7)* para minimizar las incompatibilidades entre sistemas de información en salud y la reutilización parcial de la solución semántica SNOMED-CT la cual proporciona una manera estandarizada de representar frases clínicas capturadas por los especialistas.

En la figura 1 se muestran los conceptos sobresalientes obtenidos en el desarrollo de la ontología, se observa que el concepto Persona es el concepto principal de la ontología, ya que cuenta con relaciones con los demás conceptos. Además se visualiza que los otros conceptos no cuentan con relaciones entre ellos, esto se debe a que la ontología está diseñada para consultar información del expediente clínico de mujeres embarazadas y de recién nacido.

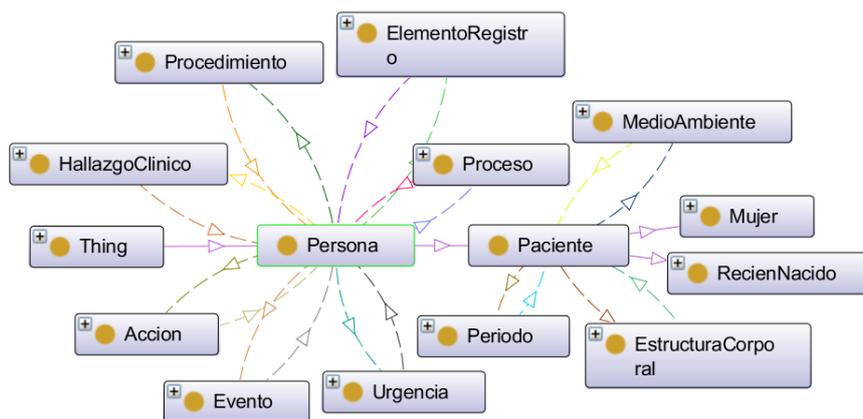


Fig. 1. Diagrama de relaciones binarias.

Para comprobar que la ontología cumple con el objetivo de su desarrollo, se considera contar con una aplicación cliente que envíe la solicitud de información acerca de un paciente en particular y una aplicación proveedor que dé respuesta a la solicitud. En la figura 2 se muestran las tecnologías y herramientas necesarias para que sistemas de salud en México logren compartir expedientes clínicos de mujeres embarazadas y recién nacidos. Cabe hacer notar que el proceso de las aplicaciones soportaría expedientes de todo tipo de pacientes, sin embargo la ontología, por el momento, sólo considera el mapeo de información específicamente relacionada con salud materno-infantil.

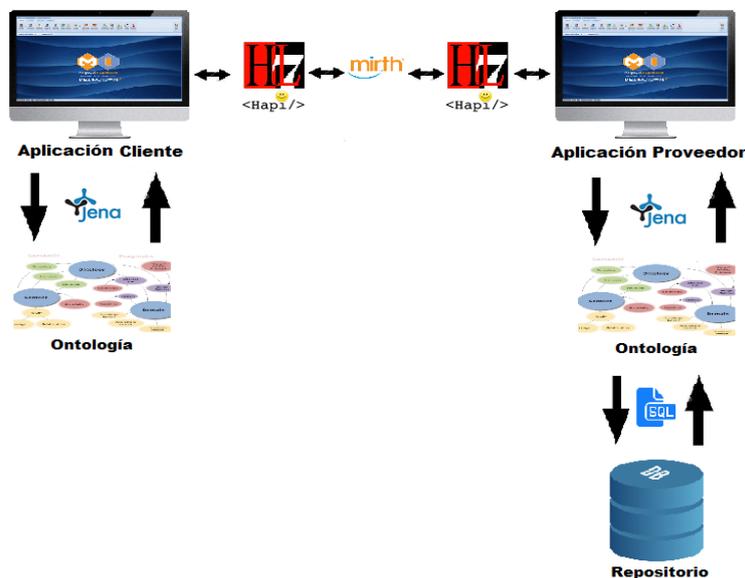


Fig. 2. Esquema de la funcionalidad de la ontología.

Las tecnologías necesarias para el intercambio de información indicadas en la figura anterior son:

- HL7 es un conjunto de estándares que apoyan la práctica clínica y la gestión, prestación y evaluación de los servicios de salud, que se reconocen como los más utilizados en el mundo [1].
- Mirth Connect es un motor de integración de software libre basado en Java, desarrollado específicamente para integraciones sanitarias y que provee de manera nativa el estándar de mensajería HL7 [2].
- Jena es una API Java que se utiliza para construir aplicaciones de Web Semántica y *Linked Data*. Trabaja con modelos RDFS y OWL (*Web Ontology Lenguaje*) para agregar semántica adicional a los datos RDF [3].
- HAPI es una biblioteca y una aplicación enriquecida para editar, validar y transmitir mensajes HL7 [4].

La funcionalidad de la aplicación cliente es la siguiente:

1. Primero es necesario configurar la aplicación cliente, que consiste en crear los individuos de la ontología que representen la unidad médica que hará las solicitudes y la(s) unidad(es) médica(s) que le proveerá(n) la información de pacientes así como el canal de comunicación entre ellas para el intercambio de información.
2. Captura de información necesaria para identificar al paciente desde el cliente.
3. Consulta de la ontología, haciendo uso de Jena, para extraer los elementos necesarios para la construcción del mensaje HL7.
4. Construcción de mensaje HL7

5. Envío del mensaje, a través de *Mirth Connect*, utilizando HAPI, hacia el canal descrito en la ontología.

6. Recepción del mensaje de respuesta HL7, interpretando sus campos con ayuda de la ontología para finalmente presentar la información en la pantalla del cliente.

Por otro lado, la funcionalidad de la aplicación proveedor es:

1. Al igual que en la aplicación cliente, es necesario realizar una configuración creando individuos en la ontología que representen la unidad médica que escucha el mensaje, el o los canales de comunicación por donde recibe mensajes, la información del repositorio de la unidad médica (tablas y campos de la base de datos) y conceptos de la ontología que está dispuesto a compartir de acuerdo a la legislación vigente.

2. Recepción de mensaje HL7 procesando los elementos que lo forman.

3. Consulta de la ontología para conocer la equivalencia entre la solicitud HL7 y los elementos específicos del repositorio que serán accedidos.

4. Construcción de consulta SQL para obtener la información del repositorio.

5. Construcción de mensaje HL7 de respuesta obteniendo la equivalencia con ayuda de la ontología.

6. Envío de la respuesta.

### **3. Proceso de evaluación de ontologías**

Barchini y Álvarez presentan en [5] varias definiciones de calidad relacionadas:

“La calidad es un conjunto de propiedades y características de un producto o servicio que le confieren su aptitud para satisfacer necesidades explícitas o implícitas

“La calidad de software es el grado con el que un sistema, componente o proceso cumple los requerimientos específicos y las necesidades o expectativas del cliente o usuario”.

“La calidad de una ontología es el grado con el que la misma puede ser localizada y recuperada, responde a la estructura invariante del dominio, con un mínimo compromiso ontológico, cumple con los requisitos específicos y puede usarse o reusarse de manera efectiva”.

Para realizar la evaluación de ontologías y asegurar la calidad de éstas existen diversos enfoques dependiendo del tipo específico y propósitos de la ontología a evaluar. A continuación se mencionan algunas clasificaciones propuestas por diversos autores.

Barchini y Álvarez [5] también proponen cuatro dimensiones que permiten definir de manera operativa la calidad de una ontología.

1) Descriptiva: referente a brindar información sobre características intrínsecas e identificadoras de la ontología, tal como los desarrolladores de la ontología, los principales destinatarios, los expertos del dominio y los usuarios finales.

2) Estructural: grado en que la ontología especifica de manera explícita, formal y consensuada los conceptos (entidades, atributos, procesos), definiciones, interrelaciones y restricciones del dominio. Dicho de otra manera, corresponde a la sintaxis y semántica formal de la ontología.

3) Funcional: capacidad de la ontología para proporcionar funciones que satisfagan las necesidades específicas. Permite determinar el grado de concordancia con los requisitos y cuestiones competentes para las que se diseñó.

4) Operacional: capacidad de la ontología para usarse, comunicarse, interactuar e integrarse entre agentes software y/o personas. Esta dimensión refleja la capacidad de uso y reúso que ofrece la ontología.

Gangemi [6] define tres tipos principales de evaluación:

1) Evaluación funcional: se centra en verificar que la ontología cumpla con su objetivo.

2) Evaluación de usabilidad: se ocupa de metadatos y anotaciones.

3) Evaluación estructural: se centra en las propiedades estructurales de la ontología como grafo.

En [7] se menciona que la evaluación de una ontología es mejor si se realiza por niveles en lugar de tratar de evaluarla directamente en conjunto. Los niveles de evaluación se han definido de diversas maneras sobresaliendo los siguientes:

- Léxico de vocabulario o capa de datos: se centra en el vocabulario utilizado para representar o identificar conceptos. La evaluación en este nivel tiende a implicar comparaciones con diversas fuentes de datos relativos al dominio del problema.

- Jerarquía o taxonomía: una ontología incluye típicamente una estructura jerárquica con relaciones tipo “es-un” entre conceptos. Aunque también se definen otro tipo de relaciones entre conceptos, la relación tipo “es-un” es importante y el foco de los esfuerzos específicos de evaluación de este nivel.

- Otras relaciones semánticas: dado que la ontología es posible que contenga otras relaciones además de las tipo “es-un”, estas relaciones se evalúan por separado. Esto normalmente incluye medidas como la precisión y la recuperación.

- Nivel de aplicación o contexto: es posible que una ontología forme parte de una colección mayor de ontologías, además de referenciar o ser referenciada por diferentes definiciones dentro de dicho conjunto. En este caso, es importante tener en cuenta este contexto para su evaluación. Otra forma de contexto es la aplicación en la que la ontología se va a utilizar; la evaluación analiza cómo los resultados de la aplicación se ven afectados por el uso de la ontología.

- Nivel sintáctico: la evaluación en este nivel es de particular interés para las ontologías que se han construido, en su mayoría, de forma manual. La ontología está descrita en un lenguaje formal particular y debe coincidir con los requisitos sintácticos de ese lenguaje.

- Estructura, arquitectura y diseño: esta evaluación es de gran interés en ontologías construidas manualmente, donde busca que la ontología cumpla con ciertos principios y criterios de diseño predefinidos. Este tipo de evaluación suele proceder de forma completamente manual.

De acuerdo a [8] la mayoría de los enfoques de evaluación caen en una de las siguientes categorías:

- Basado en la evaluación Estándar Oro: consiste en medir la similitud entre ontologías; existen dos maneras de comparar ontologías: 1) Léxica, que evalúa la similitud entre los léxicos (conjunto de etiquetas que detonan conceptos) y 2)

Conceptual, donde se comparan las estructuras taxonómicas y las relaciones en las ontologías.

- Basados en el uso de la ontología en una aplicación: donde se evalúa la ontología creando una aplicación que permita comprobar que dicha ontología cumpla con el objetivo para el cual se desarrolló.
- Basados en comparaciones con fuentes de datos: la ontología se evalúa comparando su contenido con fuentes de datos sobre un mismo dominio.
- Basados en metodologías y normas: la evaluación se lleva a cabo por especialistas que tratan de evaluar la eficacia de la ontología, tomando en cuenta un conjunto de criterios predefinidos, normas y requisitos entre otros elementos.

#### 4. Evaluación de la ontología utilizando OOPS!

Las aplicaciones que se describen en el apartado anterior permiten verificar la ontología de manera funcional, es decir, validar que cumpla con el objetivo principal para el que se creó, sin embargo no verifican otras condiciones de calidad.

OOPS! es una aplicación Web, independiente de cualquier entorno de desarrollo de ontologías, para la detección de errores de modelado. Esta herramienta está diseñada para ayudar a los desarrolladores de ontologías durante la actividad de validación, misma que se divide en Diagnóstico y Reparación. Actualmente, OOPS! proporciona mecanismos para detectar errores por lo que se ubica en la etapa de Diagnóstico [9].

Es importante mencionar que no todos los problemas encontrados tienen el mismo nivel de importancia, OOPS! maneja tres niveles de resultados.

1. **Crítico:** Es crucial corregir el problema, de lo contrario podría afectar la consistencia de la ontología, el razonamiento, la aplicabilidad, entre otros.
2. **Importante:** Aunque el problema no es crítico para la funcionalidad de la ontología, es importante corregir este tipo de problema.
3. **Menor:** En realidad no es un problema serio, pero la corrección de este tipo de defecto permite que la ontología sea aún más estable.

##### 4.1. Revisión de la ontología en busca de errores

En la figura 3 se muestra el código fuente introducido en la aplicación para su evaluación.

```
<!-- http://www.semanticweb.org/ee/ontologies/2016/4/untitled-ontology-67#UrgenciaCalificada -->
<Class rdf:about="http://www.semanticweb.org/ee/ontologies/2016/4/untitled-ontology-67#UrgenciaCalificada">
  <rdfs:label rdf:datatype="xsd:string">UrgenciaCalificada</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/ee/ontologies/2016/4/untitled-ontology-67#Urgencia"/>
  <disjointWith rdf:resource="http://www.semanticweb.org/ee/ontologies/2016/4/untitled-ontology-67#UrgenciaNoCalificada"/>
  <rdfs:comment rdf:datatype="xsd:string">Problema de salud, habitualmente de presentación súbita, que se atendió en el área de urgencias, que pone en riesgo la vida, órgano, tejido o función del paciente y que, por lo tanto, requiere de una atención médica inmediata.</rdfs:comment>
</Class>
```

Fig. 3. Código fuente de la ontología.

La aplicación analizó la ontología utilizando Jena, se revisó en busca de problemas de acuerdo a las políticas de OOPS! Durante esta fase de exploración, se detectaron los elementos que intervienen en la ontología y sus posibles errores; además, se generaron advertencias con respecto a la sintaxis de RDF y algunas sugerencias de modelado.

## 4.2. Resultado de la evaluación

El resultado de la evaluación de la ontología se muestra en la figura 4 donde se visualizan cuatro problemas hallados, de los cuales tres son de nivel Menor y uno de nivel Importante. A continuación se describe cada uno de los problemas encontrados.

[Expand All] | [Collapse All]

Results for P08: Missing annotations.	169 cases   Minor 🟡
Results for P13: Inverse relationships not explicitly declared.	10 cases   Minor 🟡
Results for P22: Using different naming conventions in the ontology.	ontology*   Minor 🟡
Results for P41: No license declared.	ontology*   Important 🔴

Fig. 4. Resultados de la evaluación.

1. Falta de anotaciones: este problema tiene que ver con elementos de la ontología que no cuentan con propiedades de anotación que brinden mayor información al usuario (por ejemplo `rdfs:label`, `lemon:LexicalEntry`, `skos:prefLabel` o `skos:altLabel`).

2. Falta de relaciones inversas: este problema aparece cuando alguna relación (a excepción de aquellas que son simétricas, `owl:SymmetricProperty`) no tiene una relación inversa (`owl:inverseOf`) definida dentro de la ontología. Las relaciones inversas son importantes porque facilitan las inferencias y consultas.

3. Uso de diferentes nomenclaturas en la ontología: si bien no existe un estándar de nomenclatura, se sugiere utilizar *CamelCase* o separación mediante delimitadores (“-” o “\_”), además se prefiere utilizar la misma nomenclatura para las mismas cosas y no una mezcla de nomenclaturas.

4. Licencia no declarada: se refiere a la omisión de información acerca de la licencia que se aplica a la ontología.

## 4.3. Fase de mantenimiento (Depuración de errores)

De acuerdo con los resultados de la evaluación de la ontología se procedió a resolver los problemas encontrados. En el primer caso se agregaron anotaciones a cada uno de los elementos de la ontología que no contaban con etiquetado y/o comentario. En la figura 5 se observa que al concepto “AtencionMedica” se le agregaron anotaciones de etiquetado y de comentario.

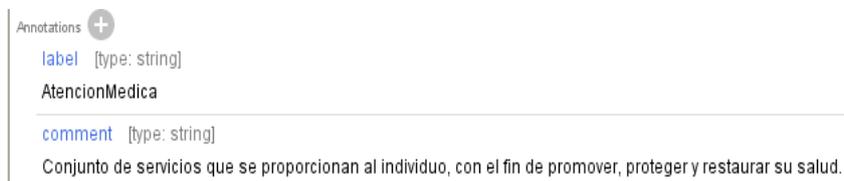


Fig. 5. Anotaciones en Protégé.

La tabla 1 contiene las relaciones inversas que se crearon para la ontología.

Tabla 1. Relaciones inversas agregadas a la ontología.

Concepto	Relaciones
<b>HallazgoClinico</b>	Hallazgo_Persona->Persona
<b>Procedimiento</b>	Procedimiento_Persona->Persona
<b>Proceso</b>	Proceso_Persona->Persona
<b>Urgencia</b>	Urgencia_Persona->Persona
<b>Accion</b>	Accion_Persona->Persona
<b>Evento</b>	Evento_Persona->Persona
<b>ElementoRegistro</b>	Elemento_Persona->Persona
<b>MedioAmbiente</b>	albergaA->Paciente

Se tomó en cuenta la recomendación de la herramienta OOPS! y se utilizó el delimitador “\_” para definir las relaciones binarias, para los conceptos se decidió colocar letras en mayúscula al inicio de cada concepto y para la separación de palabras. En la figura 6 se muestra la implementación de estas recomendaciones.



Fig. 6. Implementación de nomenclaturas para nombrar elementos de la ontología.

Por último se agregó información referente al tipo de licencia, versión previa, última versión y comentario a la ontología como se muestra en la figura 7.

<b>license</b>
Libre de: Compartir - copiar y redistribuir el material en cualquier medio o formato Adaptar - remezcla, transformar y construir sobre el material para cualquier propósito, incluso en el comercio.
<b>versionInfo</b> [type: string]
3.1
<b>priorVersion</b> [type: string]
3.0
<b>comment</b> [type: string]
Ontología que permite el intercambio de información materno-infantil entre instituciones de salud en México.

Fig. 7. Información de la licencia, versiones y comentario de la ontología.

## 5. Conclusiones y trabajo futuro

Es sumamente importante contar con herramientas o métodos que permitan la identificación de problemas en el diseño y/o planteamiento de una ontología. Existen muchas formas de evaluación, como las mencionadas en la introducción de este trabajo, muchas de ellas se centran en evaluaciones cuantitativas, sin embargo se encontraron pocas referencias activas a herramientas que realicen evaluación cualitativa, por tal motivo se optó por utilizar la aplicación OOPS! para la detección de errores debido a su naturaleza cualitativa.

OOPS! identifica los errores más comunes cometidos en el planteamiento y desarrollo de una ontología, por lo que ayuda a identificar y solucionar problemas de calidad de una manera más sencilla; los resultados obtenidos en el proceso de evaluación con OOPS! permitieron corregir deficiencias que podrían afectar el correcto funcionamiento de la ontología planteada.

Como trabajo a futuro se implementarán las aplicaciones descritas en el apartado 2 Ontología para el intercambio de información materno-infantil de tal manera que se realice la evaluación funcional de la ontología planteada; como parte de las pruebas se buscará la colaboración de algunas instituciones de salud en México para realizar las validaciones correspondientes.

**Agradecimientos.** El autor agradece al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo en el financiamiento del proyecto de Tesis y al Tecnológico Nacional de México por haberlo aceptado y ser parte de él.

## Referencias

1. Sitio Oficial Health Level Seven International. Disponible en: <http://www.hl7.org>
2. Sitio Oficial Mirth. Disponible en: <http://www.mirthcorp.com/products/mirth-connect>
3. Sitio Oficial Jena. Disponible en: <https://jena.apache.org/documentation/ontology>
4. Sitio Oficial HAPI. Disponible en: <http://hl7api.sourceforge.net>

5. Barchini, G. E., Álvarez, M. M.: Dimensiones e indicadores de la calidad de una ontología. Revista Avances en Sistemas e Informática, Disponible en: <http://www.redalyc.org/articulo.oa?id=133115523004> (2010)
6. Gangemi, A.: A theoretical framework for ontology evaluation and validation. SWAP, Vol. 166 (2005)
7. Brank, J., Grobelnik, M., Mladenić, D.: A survey of ontology evaluation techniques. In: Proceedings of the conference on data mining and data warehouses (2005)
8. Lu, Q.: OntoKBEval: a support tool for OWL ontology evaluation. Diss. Concordia University (2006)
9. Poveda, M., Suárez, M. C., Gómez, A.: Validating ontologies with Oops! In: International Conference on Knowledge Engineering and Knowledge Management, Springer Berlin Heidelberg (2012)



## Identificación de lengua de señas mediante técnicas de procesamiento de imágenes

Manuel Martínez Gutiérrez<sup>1</sup>, José Rafael Rojano-Cáceres<sup>2</sup>,  
Ismael Everardo Bárcenas Patiño<sup>1</sup>, Fredy Juárez Pérez<sup>3</sup>

<sup>1</sup> Universidad Veracruzana, Maestría en Sistemas Interactivos Centrados en el Usuario,  
Xalapa, México

<sup>2</sup> Universidad Veracruzana, Facultad de Estadística e Informática,  
Xalapa, México

<sup>3</sup> Instituto Tecnológico Superior de Álamo Temapache,  
Facultad de Ciencias de la Computación,  
Tuxpan, México

cheleazar1805@gmail.com, rrojano@uv.mx, iebarcenaspa@conacyt.mx,  
fredyjuarezp@gmail.com

**Resumen.** En este trabajo se aborda el procesamiento de imágenes obtenidas a partir de una cámara digital con el objeto de caracterizarlas empleando técnicas básicas de visión computacional. Se puede definir como técnicas básicas a las transformaciones en el dominio de la frecuencia. La aplicación de diversas técnicas tiene como objeto la identificación de señas que provienen de la Lengua de Señas Mexicana, en particular de señas del tipo estático como es el caso de un subconjunto del alfabeto. Como muestra del enfoque planteado se presenta el desarrollo de un prototipo que sirva de apoyo para que una persona que realiza una seña pueda saber si está correctamente realizada.

**Palabras clave:** Lengua de señas, procesamiento digital, inclusión social.

## Identification of Sign Language Using Image Processing Techniques

**Abstract.** This work deals with the processing of images obtained from a digital camera in order to characterize them using basic computer vision techniques. The transformations in the frequency domain can be defined as basic techniques. The application of different techniques has as an objective the identification of signs that come from the Mexican Sign Language, in particular signs of the static type as one subset of the alphabet. As a sample of the approach, the development of a prototype is presented in order to help a person in the correct realization of signs.

**Keywords:** Sign language, digital processing, social inclusion.

## **1. Introducción**

En la vida cotidiana, las computadoras manejan una gran cantidad de datos provenientes de diferentes fuentes y formatos como pueden ser sensores, bases de datos, redes sociales, textos, etc. Aunado a este proceso, nos encontramos ante la necesidad que tienen las personas de utilizar diferentes dispositivos de comunicación que enriquezcan y faciliten la Interacción Humano Computadora (IHC). Como resultado se tiene la necesidad desarrollar técnicas computacionales que permitan la búsqueda de patrones o datos característicos en imágenes, ondas de audio o pulsos eléctricos, entre otros, para llevar a cabo tareas que solo hasta el momento puede realizar mejor el ser humano [1].

De esta forma, con la intención de mejorar tanto la Experiencia de Usuario (EU) así como la facilidad de interacción con las computadoras se vienen proponiendo diversos enfoques de interacción natural [2]–[6] entre los cuales se considera tanto el tratamiento digital de imágenes junto con la adquisición desde diversas fuentes de datos como puede ser un sensor como Kinect.

En este trabajo se plantea una propuesta para la identificación de señas derivadas de la Lengua de Señas Mexicana empleando algoritmos básicos para el procesamiento de la imagen tal como lo son las transformaciones espaciales (transformación en escalas de grises, binarización, el uso de umbrales, etc.). El objetivo propuesto tras emplear diferentes técnicas para el procesamiento de imágenes es el de desarrollar un prototipo a través del cual una persona con o sin entrenamiento en Lengua de Señas (LS) pueda validar si lleva a cabo de forma correcta la seña. A este proceso de acuerdo con Cruz-Aldrete [7] se le denomina “señar” otros términos también son el de “signar”. Aunado a lo anterior es necesario plantear que la Lengua de Señas es específica para cada país, en el caso de México se le conoce como Lengua de Señas Mexicana. Por otro lado la LS se define como aquel lenguaje viso-espacial en donde se genera comunicación empleando las manos y sus posiciones en relación con el cuerpo y la mirada intencional [8].

Así la estructura de este trabajo considera una introducción, descripción sobre las técnicas de procesamiento de imágenes empleadas, el diseño de un prototipo, identificación de la seña, conclusiones y referencias.

## **2. Técnicas de procesamiento de imágenes**

En esta sección se presentan y explican brevemente las diferentes técnicas de procesamiento de imágenes empleadas para el desarrollo del prototipo que reconoce un conjunto de Lengua de Señas estáticas. A continuación se explica cada una de ellas.

### **2.1. Escala de grises**

La primera técnica a explicar consiste en la aplicación de un filtrado de transformación en escala de grises. Para transformar una imagen a escala de grises mientras que esta proviene del modelo RGB, se puede realizar de diversas maneras, en este caso se realiza un promedio de los valores RGB de entrada y se realiza una asignación de dicho valor promediado igual en cada uno de los píxeles de la imagen.

La función que representa esta transformación se muestra en (1). Donde los valores RGB transformados pueden ser ya sea tomando el promedio de los valores R, o G, B. En este caso la transformación en escala de grises sirve como etapa previa de la binarización, ya que al estar todos los píxeles en dicha escala facilita la siguiente etapa del proceso:

$$(r', g', b') = \begin{cases} (r, r, r), & o \\ (g, g, g), & o \\ (b, b, b). \end{cases} \quad (1)$$

## 2.2. Negativo

Otro filtro básico es el de la transformación a valores negativos. El filtro negativo consiste en realizar una resta del valor máximo de la imagen (255) menos el valor RGB del pixel, esto por cada pixel de la imagen. La transformación en negativo se define a partir de un rango de valores de intensidad comprendido entre  $[0, L - 1]$  y mediante la aplicación de la ecuación (2) el resultado es la inversión de los valores dado el máximo del rango de entrada  $L - 1$  menos el pixel de entrada  $r$ :

$$s = L - 1 - r. \quad (2)$$

## 2.3. Threshold

El filtro threshold o mejor conocido como filtro binario, permite transformar una imagen a binario o blanco y negro. Este filtro necesita que el usuario ingrese un valor mayor a 0 y menor a 255, con este valor le permite al filtro catalogar los colores para saber si cambiarlo por el color blanco o el color negro. Se utiliza en este caso para remarcar la mano con el color negro y el fondo quede de color blanco:

$$s = \begin{cases} 0 & \text{si } r \leq c, \\ L - 1 & \text{en otro caso.} \end{cases} \quad (3)$$

## 2.4. Recorte binario

Se definió una función denominada recorte binario cuyo propósito es la extracción de la mano, misma que se encuentra en color negro. De esta forma el recorte binario recorre la imagen buscando los puntos en color negro estableciendo los puntos mínimos y máximos que conforman la imagen para descartar el resto del fondo.

## 2.5. Zoom

Uno de los filtros más importantes de la aplicación consiste en la aplicación de un aumento digital, así como su correspondiente decremento digital. El aumento digital consiste en incrementar el tamaño de la imagen dependiendo de un valor que introduce el usuario, a partir de ahí se debe realizar una ecuación para separar los píxeles y rellenarlos con valores de algún vecino. El decremento digital es parecido al aumento,

pero en su función inversa. Se utiliza para proporcionar una altura a la imagen que el usuario ingresa, en este caso se maneja el tamaño de la imagen a 150 píxeles.

## 2.6. Rotación

La función para rotación se emplea como un mecanismo de prueba de tolerancia al identificar una imagen. En este caso se consideró una tolerancia de  $\pm 5$  grados. La función de rotación es tal y como se describe en la ecuación (4) en donde  $R$  se define como en (5):

$$x' = Rx + t, \quad (4)$$

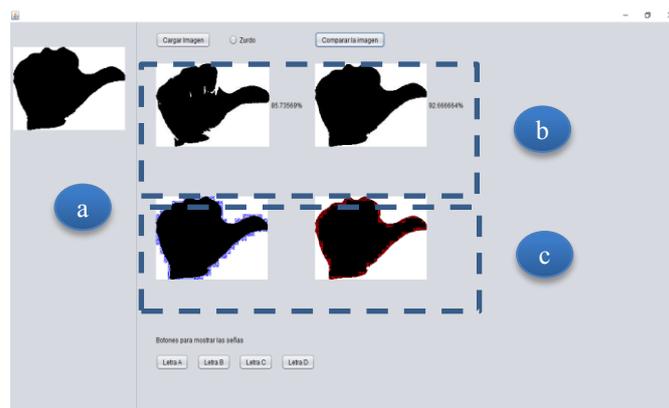
$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}. \quad (5)$$

## 2.7. Closing y opening

El filtro opening permite agrandar espacios pequeños en las imágenes, esto lo logra aplicando el filtro erosión y luego dilatación. El filtro closing a diferencia del opening se centra en cerrar los pequeños agujeros que se pueden considerar como ruido en una imagen, esto lo logra aplicando el filtro de dilatar y luego el de erosionar. Así la aplicación de estos filtros permite eliminar el ruido de la imagen, así como el ruido que se genera al aplicar diversos filtros para el tratamiento.

## 3. Prototipo desarrollado

El desarrollo del prototipo se realizó en el lenguaje Java, con una interfaz gráfica que permitiera ver la funcionalidad del sistema y que mediante su empleo el usuario pudiera conocer si al realizar una determinada seña se realizaba de forma correcta, veáse Fig. 1.



**Fig. 1.** Pantalla del prototipo. a) imagen de usuario, b) imágenes identificadas como correctas, c) errores de comparación.

Para esta versión del prototipo las imágenes a verificar pueden ser tomadas a partir de una cámara del celular o de una cámara web y posteriormente se salvan a un archivo jpg o png que deberá ser introducido al sistema para su validación. La aplicación consta únicamente de una sola ventana que se encuentra dividida en cuatro funcionalidades descritas en detalle en la siguiente sección.

### 3.1. Casos de uso

Las funcionalidades que puede realizar el usuario en el sistema, se limitan a cuatro actividades que se denotan en el diagrama de caso de usos mostrado en la Fig. 2. El usuario puede realizar tres de esas actividades sin ningún requisito, pero para comparar la imagen es necesario que primero cargue una imagen. A continuación se explica brevemente cada caso de uso.

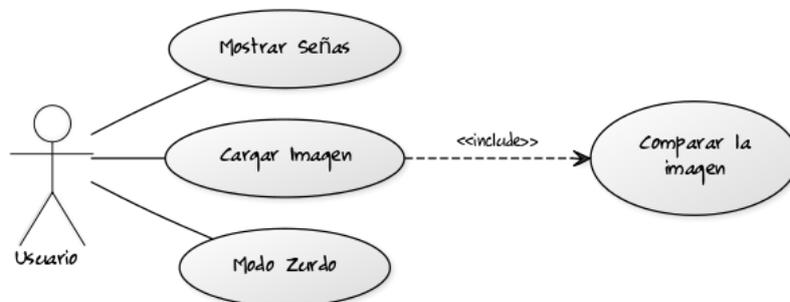


Fig. 2. Diagrama de casos de uso para el prototipo desarrollado.

**Mostrar señas.** El propósito de esta función es mostrar las letras que corresponden a una determinada seña. En este caso para prueba el conjunto de señas es reducido y se limita a cuatro letras del alfabeto dactilológico. Así cuando el usuario desea conocer cómo se seña la letra “a” entonces el sistema mostrará la letra “a” en el área b de Fig. 1. Así mismo, dentro del área b, se muestran resultados asociados al porcentaje de reconocimiento de la seña.

**Cargar imagen.** La funcionalidad de cargar imagen, como su nombre lo indica, permite al usuario subir al sistema una imagen de su mano realizando una seña. Para su correcto funcionamiento, se espera que la mano se vea claramente y el fondo sea oscuro. Al realizar el proceso de carga también se realizará la llamada a los diferentes pasos del procesamiento de imágenes descritos en la sección anterior. En concreto al realizar la carga se efectúa la binarización de la imagen: escala de grises, negativo y threshold. Una vez que tiene la imagen en binario se aplican los filtros morfológicos opening y closing para disminuir el ruido de la imagen. Al resultado del paso anterior se le aplica una función que recorta la imagen para obtener la imagen con la mano únicamente. El último paso en esta función es redimensionar la imagen a un tamaño esperado por la aplicación, en este caso es una altura de 150cm. para esto se utiliza el filtro zoom in y zoom out.

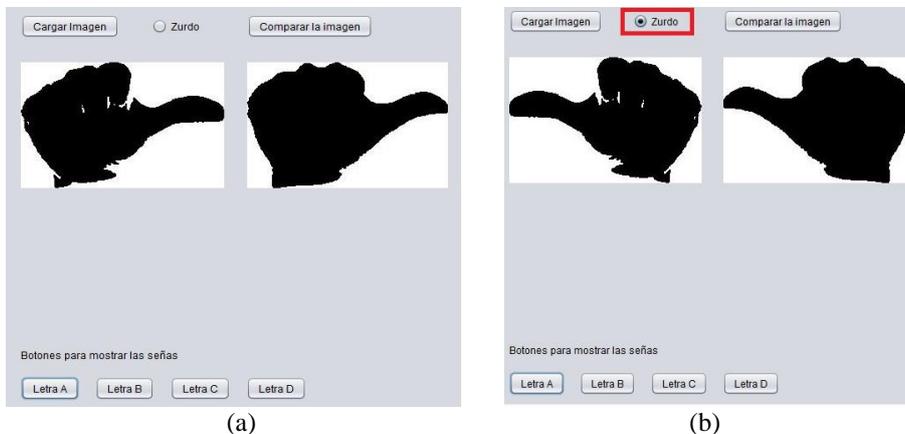


Fig. 3. Ventana de la aplicación (a) mostrando la letra A, (b) mostrando la letra A en zurdo.

**Modo zurdo.** El modo zurdo como su nombre lo indica es una función que al ser activada aplica una reflexión horizontal a las imágenes. La activación del modo zurdo tiene como propósito reconocer aquellas imágenes que hayan sido grabadas por personas zurdas.

**Comparar la imagen.** Finalmente, la funcionalidad de comparar la imagen realiza un cálculo que consiste en determinar el porcentaje de similitud entre un conjunto de señas previamente registradas en el sistema con relación a la imagen que carga el usuario.

En este caso al realizar la comparación de las imágenes se mostrará en la aplicación un contorno en tono rojo para las zonas que debería ocupar la mano y azul para las zonas que no debería ocupar la mano. Véase Fig. 4.

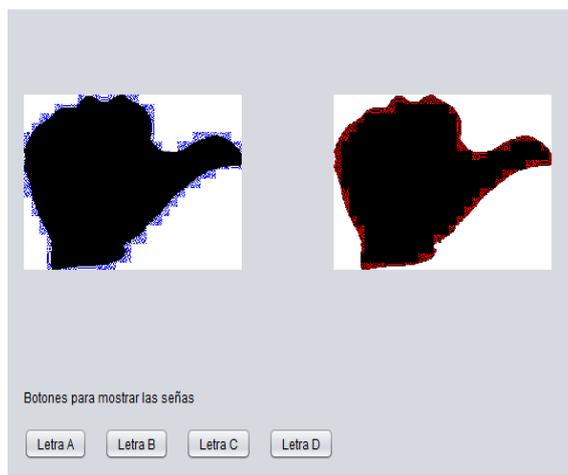


Fig. 4. Área de la aplicación que muestra los resultados de la comparación de la imagen.

#### 4. Identificación de la seña

El proceso de identificación de la seña se basa en la comparación de dos imágenes, la primera se adquiere a través de la cámara por parte del usuario y la segunda se encuentra previamente almacenada en el sistema. A partir de las dos imágenes, el sistema recorre pixel por pixel comparando el valor de tonalidad de estos. En el caso de que coinciden los valores binarizados el sistema aumentará un contador, en caso contrario el sistema identificará cual era el resultado esperado tomando como base la imagen propia del sistema. Así, dependiendo del valor esperado el sistema guarda un registro del exceso o la falta de la coincidencia. De forma que se indicará con un color rojo aquellos pixeles que representan lo que debería haber ocupado o bien en azul representa lo que no debe ocupar, es decir ambos son un tipo de error natural pues es muy complejo que la mano siempre ocupe la misma posición aun cuando lo haga la misma persona. Una vez que el sistema termina de recorrer los pixeles, realiza la media de aciertos entre el número de pixeles recorridos para determinar el porcentaje de precisión. El proceso anterior se debe repetir para comparar la imagen del usuario con todas las imágenes que tiene el sistema y a partir de la media más alta se deduce la seña que realizó el usuario. De esta forma, la imagen de entrada se compara contra todas las imágenes que se encuentran en el sistema dado que no existe conocimiento de cuál podría ser la letra a identificar. Así mismo en el sistema existen variaciones de la misma letra considerando aspectos como el de rotación o bien la flexión de los dedos que como se expresó son naturalmente cambios de la postura al grabar la seña. De esta forma el algoritmo comparará contra todas las variaciones con el propósito de tolerar cierta imprecisión derivada de la generación de la seña por parte del usuario.

```
Función compara (imagen A, imagen B)
  doble porcentaje=0
  entero contador=0
  entero numeroPixel=0
  Desde entero I=0, hasta I == ancho(A)
    Desde entero J=0, hasta J == alto(A)
      entero tonalidadA = tonalidad(I,J) de A
      entero tonalidadB = tonalidad(I,J) de B
      Si tonalidadA == tonalidadB
        Contador = contador + 1
        numeroPixel= numeroPixel + 1
    Fin desde
  Fin desde
  Porcentaje = contador/numeroPixel
  Porcentaje = porcentaje*100
  regresa porcentaje
Fin función
```

Fig. 5. Pseudo-código de la función compara imagen.

## **5. Pruebas**

Para corroborar que el algoritmo planteado en la Fig. 5 era funcional para la identificación de la seña se procedió a realizar dos tipos de pruebas, la primera toma como parámetros la imagen de la cual ya se sabe que letra es, la segunda prueba hace una evaluación considerando el resto de letras.

En la primera prueba, que funcionó como experimento de control, se ingresó al sistema la imagen previamente identificada esperando que al compararla con su correspondiente se tuviese una precisión del 100%. Sin embargo, se encontró que el algoritmo solo daba precisiones superiores al 90% sin llegar al cien por cien. Tras un análisis se observó que la función implementada en Java generaba un pequeño ruido, pero este se trató mediante la aplicación de un filtro de threshold con lo cual dichos errores desaparecieron produciendo una certidumbre del 100%.

Para el segundo tipo de prueba, se usaron diferentes imágenes considerando una ejecución normal del sistema. Es decir, se capturaron nuevas señas y se procedió a ingresarlas al sistema para su identificación contra las que tenía registradas. Para la verificación se capturaron diversas imágenes aplicando rotación, misma que fue considerable al ser de aproximadamente de entre 10 y 40 grados. De esta forma se ingresaron grupos de cinco imágenes por cada letra, mismas que fueron realizadas por dos personas, con lo cual nos daba como punto de comparación y reconocimiento un total de veinte imágenes de prueba.

Como resultado de la identificación de las señas se encontró que el sistema reconoció correctamente dieciséis de las veinte imágenes. Entre las imágenes que no identificó se encuentran la letra a y la letra d que correspondían a la grabación de la seña realizada por una persona no registrada en el sistema, mientras que para la misma persona aun cuando se aplicaron las rotaciones y las variaciones naturales generó una precisión del 100%. Cabe hacer algunas precisiones, la variación de la rotación se debe el valor mayor a que la persona que ingresó la imagen tenía escasos conocimientos sobre Lengua de Señas por lo cual la emisión de la seña tenía un gran error, mientras que la persona con la cual tuvo mejor precisión era un usuario que conoce la LS y por tanto producía menores variaciones al momento de ingresar la seña.

## **6. Conclusión**

En esta investigación se abordó el procesamiento de imágenes mediante la aplicación de técnicas esenciales como las descritas en la sección 2, así mismo se implementó un procedimiento que permitiera identificar la seña realizada teniendo como fuente la imagen procesada. En este trabajo se optó por emplear técnicas derivadas de la visión por computadora para demostrar su potencial en el reconocimiento de señas provenientes de la Lengua de Señas Mexicana, en particular de señas estáticas, es decir aquellas que no requieren movimiento continuo para su ejecución. De acuerdo a las pruebas realizadas se encontró que el porcentaje de reconocimiento fue superior al 75%. De esta forma se considera que el prototipo cumple 1) con la función de procesar imágenes para su tratamiento digital, y 2) identificar las imágenes mediante un algoritmo sencillo como el descrito en la Fig. 5, al final del artículo.

Finalmente, como trabajo futuro se considera ampliar la base de imágenes empleadas para prueba, es decir en primera instancia proveer todo el conjunto de letras del abecedario y posteriormente realizar una prueba con diversas señas estáticas de la lengua de señas. De esta forma se provee un medio para que personas puedan por un lado practicar el uso de la Lengua de Señas, y por otro lado se fomenta la inclusión de personas con discapacidad auditiva al generar una actitud sensibilizadora en las personas oyentes a través del aprendizaje de dicha lengua.

## **Referencias**

1. Rich, E., Knight, K.: Artificial intelligence. McGraw-Hill, New (1991)
2. Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., Presti, P.: American sign language recognition with the kinect. In: Proceedings of the 13th international conference on multimodal interfaces, pp. 279–286 (2011)
3. Liang, R. H., Ouhyoung, M.: A real-time continuous gesture recognition system for sign language. In: Automatic Face and Gesture Recognition, Third IEEE International Conference on, pp. 558–567 (1998)
4. Sturman, D. J., Zeltzer, D.: A survey of glove-based input. IEEE Comput. Graph. Appl., Vol. 14, No. 1, pp. 30–39 (1994)
5. Starner, T., Weaver, J., Pentland, A.: Real-time american sign language recognition using desk and wearable computer based video. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 20, No. 12, pp. 1371–1375 (1998)
6. Gavrilu, D. M.: The visual analysis of human movement: A survey. Comput. Vis. image Underst., Vol. 73, No. 1, pp. 82–98 (1999)
7. Cruz, M.: Gramática de la lengua de señas mexicana. Universidad del Valle, Colombia (2008)
8. D. O. de la Federación: Ley General para la Inclusión de las Personas con Discapacidad. D. Of. la Fed. (2011)



## Breve revisión sobre el reconocimiento de la Lengua de Señas

Héctor Eduardo Sánchez Barrera<sup>1</sup>, José Rafael Rojano-Cáceres<sup>2</sup>,  
Guillermo MoleroCastillo<sup>1</sup>, Etelvina Archundia Sierra<sup>3</sup>

<sup>1</sup> Universidad Veracruzana, Maestría en Sistemas Interactivos Centrados en el Usuario,  
Xalapa, México

<sup>2</sup> Universidad Veracruzana, Facultad de Estadística e Informática,  
Xalapa, México

<sup>3</sup> Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

zs15019631@estudiantes.uv.mx, rrojano@uv.mx, ggmolero@conacyt.mx,  
etelvina@cs.buap.mx

**Resumen.** La lengua de señas involucra diversos factores como la posición de la mano, la velocidad con la que se realiza la seña e inclusive los gestos de la cara, por lo que interpretar el significado de una seña es un trabajo complejo. Debido a ello desde el punto de vista de la computación se ha venido trabajando bajo diferentes propuestas e investigaciones las cuales involucran el uso de diversas tecnologías y técnicas que tratan de resolver esta problemática. En el presente trabajo se lleva a cabo una breve revisión de las diferentes investigaciones que están trabajando con el problema de la clasificación de las señas.

**Palabras clave:** Lengua de señas, revisión, aprendizaje de máquina.

## Brief Review on Recognition of the Sign Language

**Abstract.** Sign language involves various factors such as the position of the hand, the speed with which the sign is made and even the gestures of the face, so interpreting the meaning of a sign is a complex task. Thus, under the point of view of computing, we have been working with different proposals and researches which involve the use of different technologies and techniques that try to solve this problem. In the present work we offer a brief review about the different researches working with the problem of the classification of the signs.

**Keywords:** Sign language, review, machine learning.

## **1. Introducción**

Los estudios sobre la Lengua de Señas (LS) han venido a cobrar gran relevancia entre la comunidad científica, por un lado, derivado de la promulgación de leyes que reconocen a dicha lengua en el contexto cultural de una nación. Por ejemplo, en el caso de México la promulgación de la Ley General para la Inclusión de las Personas con Discapacidad [1]. Por otro lado, debido a la complejidad inherente de la interpretación de las señas, así como que las señas no son universales, sino que forman parte de la cultura y vivencia de cualquier país, existe entonces el planteamiento para su tratamiento mediante medios computacionales. Por ello, se vienen realizados diversos estudios científicos relacionados al estudio de la LS que van desde los aspectos lingüísticos [2] hasta su tratamiento computacional [3]–[6], los cuales incluyen el empleo de Modelos Ocultos de Markov (MOM), Redes Neuronales (RN), entre otros para la clasificación de datos derivados de electromiogramas. Así mismo se puede encontrar la aplicación de técnicas como las Máquinas de Soporte Vectorial (MSV) para la clasificación de datos derivados de sensores de profundidad como el Kinect o Leap Motion. Así mismo otros trabajos plantean la necesidad de sensibilizar a la población oyente a través de la formación y la capacitación mediante medios computacionales [7], o bien fortalecer su comprensión y humanización mediante el uso de avatares [8].

De esta forma, considerando lo planteado por Trujillo-Romero & Caballero-Morales [9] para el ser humano resulta de vital importancia el poder comunicarse, ya que es a través de la comunicación que socializamos, trabajamos y sencillamente sobrevivimos.

Por lo cual, dado que en el día a día la forma de comunicación típica imperante es el habla oral y en segundo término estaría la escrita, se presente un grave problema social para la comunidad de personas sordas ya que su interacción con los oyentes encuentra una grave barrera de comunicación debido a que la lengua de comunicación es completamente diferente.

Para proveer por tanto una descripción de las implicaciones de la LS se retoma lo que para Sun, Zhang, Bao & Xu [3] es dicha lengua, un tipo de lenguaje visual, que consiste en una secuencia de gestos gramaticalmente estructurados. A su vez Cruz-Aldrete [2] lo describe como una lengua visogestual, en donde se utiliza la vista, manos, gestos faciales y el espacio circundante para interactuar con otros.

En el presente artículo se aborda una revisión no extensiva de las investigaciones que se han realizado en los últimos ocho años con base al reconocimiento de Lengua de Señas. De esta forma la estructura del artículo es: trabajos realizados, discusión, conclusión y referencias.

## **2. Trabajos realizados**

Como punto de partida del estudio, se consideró realizar una revisión sobre los artículos existentes en el área de Ciencias de la Computación que se han publicado bajo el término de “Lengua de Señas”. Como resultado se encontró que existen 2,169 artículos partiendo desde 1982 a la fecha, considerando la búsqueda en el título, el resumen y las palabras clave.

Con base en el censo se seleccionaron algunos términos clave que fueran significativos a la investigación de los autores, con lo cual se encontró que el 71.8 % de los artículos hacen referencia al término “Lengua de Señas” entre sus términos clave. Otro dato relevante encontrado en la BD Scopus es el grupo relativo a la Lengua de Señas tratada y que corresponde a cinco lenguas principales americana, hindú, china, árabe y japonesa.

En la Fig. 1 se aprecia la distribución relativa a las lenguas tratadas en problemas de reconocimiento de Lengua de Señas, siendo la lengua americana la más abordada y la lengua china con el menor número de investigaciones, por otro lado, se puede observar que no existen trabajos que aborden la lengua señas mexicana por lo que resulta un área de interés para la investigación.

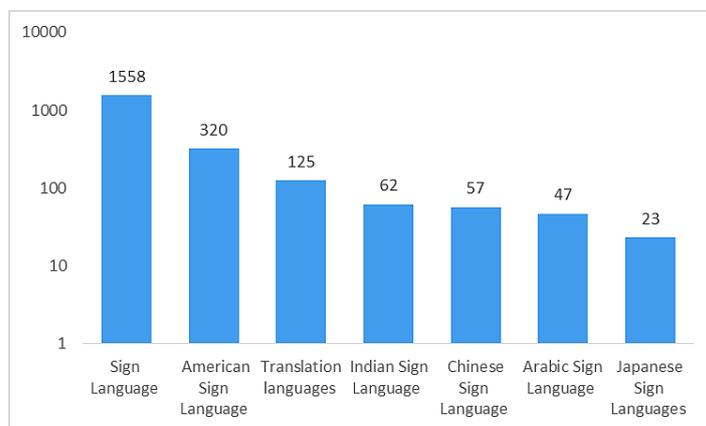


Fig. 1. Total de lenguas tratadas en el área de Ciencias de la Computación.

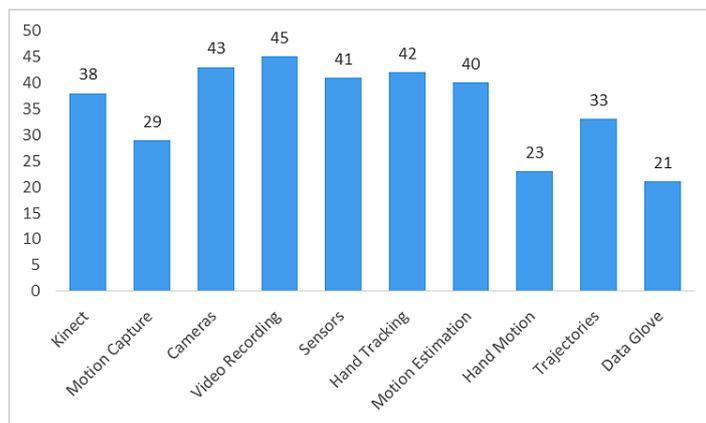


Fig. 2. Distribución de artículos que indican como términos clave al dispositivo.

En un segundo término se exploró las técnicas computacionales mencionadas, así como los dispositivos para su adquisición. En relación a los dispositivos se encontró entre las palabras clave a las cámaras, al Kinect, a los guantes y al uso de técnicas de

captura de movimiento, entre otros términos relacionados. En algunos casos cabe decir que los autores decidieron combinar los términos cuando el término principal aparecía en los diversos conteos. Entonces para el caso de los dispositivos se tiene una distribución como la presentada en la Fig. 2. Tal como se puede observar los dispositivos más utilizados son las cámaras y el Kinect, mientras que las técnicas más abordadas son las grabaciones de video y el seguimiento de las manos, lo que da lugar a explorar nuevas técnicas para el reconocimiento de la seña.

En relación a las técnicas computacionales los autores agruparon en tres grupos principales las técnicas: a) aquellas derivadas de la visión por computadora, b) aquellas derivadas de la minería de datos y c) aquellas que hacen uso de técnicas basadas en avatares. En la Tabla 1 se presenta un resumen de las palabras clave relacionadas.

**Tabla 1.** Términos clave con mayor relevancia al área especificada.

Técnicas de Avatares		Técnicas de Visión		Técnicas de Minería de Datos	
Animation	192	Computer Vision	199	Hidden Markov Models	206
Virtual Reality	94	Image Processing	126	Neural Networks	148
Three Dimensional Computer Graphics	62	Image Segmentation	94	Markov Processes	99
Computer Graphics	44	Signal Processing	93	Support Vector Machines	73
Signing Avatars	34	Image Analysis	45	Principal Component Analysis	48
Avatar	26	Video Signal Processing	35	Recognition Accuracy	47
Virtual Character	23	Color	27	Recognition Rates	47
		Video Sequences	23	Learning Algorithms	24
				Machine Learning	22
				Decision Trees	20

De esta forma en las secciones siguientes se abordan brevemente aspectos puntuales relacionados con el contenido de tales artículos para brindar una descripción general de los mismos.

## 2.1. Técnicas con avatares

Al respecto del trabajo relacionado con avatares se hacen referencia aquellas investigaciones que buscan mediante el uso de diseños tridimensionales humanizados presentar interfaces que describan la interacción mediante Lengua de Señas para el usuario. Un ejemplo de esta línea de trabajo es el propuesto por Bouzid y Jemni [8] que pretende crear un avatar virtual que interprete de manera automática la notación de la Lengua de Señas. Esta investigación utiliza un lenguaje de marcas de escritura de signos SWML (SignWriting Markup Language) el cual es un formato basado en XML, que almacena, indexa y procesa la notación de símbolos escritos. Este documento SWML es utilizado por un módulo de renderizado 3D que a su vez es usado por un agente virtual para convertirlo en una animación en 3D y así mostrar la animación de la seña.

En el trabajo descrito anteriormente se denota la importancia de las características de la seña, en la cual se toma en cuenta la configuración de la mano, ubicación,

dirección y orientación. Estas características son necesarias para poder reconstruir la seña por lo que se puede decir que también lo son para poderla clasificar adecuadamente.

## 2.2. Interfaces para captura de datos

Dentro de las investigaciones que hacen uso de dispositivos externos se tiene el trabajo de SignTutor [6] o tutor de señas, que busca enseñar la lengua de señas mediante un sistema que utiliza una cámara de video y dos guantes de diferente color: azul y amarillo. La aplicación cuenta con las opciones para aprender la lengua de señas y un apartado de entrenamiento. Inicialmente la investigación comenzó con un reconocimiento de 40 a 50 señas definidas, pero en la práctica se llegó a reconocer hasta 5,119 señas diferentes. En la investigación se utilizaron diferentes técnicas como: redes neuronales, redes bayesianas dinámicas, máquinas de estado finito o plantillas.

En esta investigación el primer paso consiste en la grabar un video con la realización de la seña, después se continua con la detección de la mano y el rostro, en la siguiente fase se analiza el movimiento, forma y posición de la mano, además del movimiento de la cabeza, ojo y boca, en la penúltima fase se clasifica la seña utilizando los modelos ocultos de Markov y finalmente en la última etapa se obtiene una respuesta favorable o desfavorable dependiendo de la información obtenida por el modelo.

También se puede citar la investigación que utiliza un electromiograma [4], el cual cuenta con 8 canales, para convertir el lenguaje de señas thai a su correspondiente lenguaje escrito. Su metodología se basa en cuatro etapas las cuales se describen a continuación:

1. La adquisición de datos por medio de 8 canales de señales alrededor de los músculos del antebrazo
2. Colección de datos, en cual el alfabeto thai fue grabado 200 veces, en donde 100 de los datos se utilizaron para el entrenamiento y 100 de ellos para pruebas.
3. Extracción de características: Se usa para seleccionar distintas características de la señal y para destacar los datos en el proceso de clasificación.
4. Algoritmo de clasificación del algoritmo, se encarga de determinar la calidad y rendimiento del sistema.

En este trabajo no se describen las características que toman en cuenta para la clasificación de la seña, pero si se puede observar de forma general el proceso que sigue para la identificación de la seña.

De manera similar en el sistema de traducción de lengua de señas chino [10, 11], se ha implementado el uso sistemas microelectromecánicos en el brazo y en los dedos de la mano para poder obtener un conjunto de datos, a los cuales se les realiza un preprocesamiento para extraer las características más relevantes, posteriormente los datos serán procesados por el algoritmo de árboles de decisión C4.5 para llevar a cabo la clasificación.

Del lado de la tecnología no intrusiva se encuentran LeapMotion y Kinect. La tecnología LeapMotion, es capaz de reconocer el movimiento de la mano en 3D, es una barra censorsa que detecta los gestos de la mano y la posición de los dedos en aplicaciones de software interactivo [5], mientras que la tecnología Kinect es utilizada

por la consola de videojuegos de Xbox 360 y Xbox One desarrolladas por la empresa Microsoft, principalmente utilizada para videojuegos aunque también se ha utilizado en diferentes áreas como la medicina y la arquitectura cada una con sus propios fines, en la medicina ayudando a los médicos a mejorar sus cirugías y en la arquitectura utilizada en el diseño de planos, otra utilidad que se le ha dado al uso de estas tecnologías, como ya se ha mencionado es en la clasificación de la lengua de señas en donde se han obtenido resultados favorables [3, 5, 9, 12].

Así mismo dentro de este tipo de tecnología existe trabajo que sigue la misma serie de pasos que los trabajos descritos anteriormente, en este trabajo se utiliza la lengua de señas india para pronosticar el texto [13], la diferencia radica en la tecnología, técnica y el conjunto de señas que utilizan, ya que en esta investigación hacen uso de la barra sensora LeapMotion para obtener los datos y de un algoritmo de random forest para llevar a cabo la clasificación. Igualmente, el uso del Kinect en el trabajo de Galicia et al. [14], es utilizado para obtener los datos de la seña tomando en cuenta la altura y profundidad de la posición de la mano y utiliza un random forest de árboles de decisión para clasificar 5 letras del alfabeto mexicano, teniendo como resultado una precisión del 76.19%.

### **2.3. Técnicas inteligentes**

En la clasificación de la seña, se encuentra una investigación que utiliza árboles de decisión, mediante el uso de la tecnología Leap Motion [15]. Lo que se pretende en esta investigación es que este reconocimiento se realice totalmente de manera dinámica y estática, es decir, que se pueda crear un modelo que permita clasificar de manera correcta las señas, en donde automáticamente aprenda nuevas para así, no tener que recurrir a un conjunto de datos estáticos para obtenerla. Cabe destacar este proyecto se destaca las características de la figura, dirección, posición y velocidad de la mano para crear el modelo de reconocimiento de las señas.

Al igual que en el trabajo anterior el trabajo sobre el sistema de señas para la lengua indonesia [16] toma en cuenta la mano que se utiliza, así como la forma de la mano, la posición de partida, y la dirección. Cada uno de estos componentes tiene que combinarse en una información única que representa cada palabra en una oración.

Por otro lado, en [3] se hace uso mediante el uso del Kinect y una máquina de soporte vectorial (SVM, por sus siglas en inglés Support Vector Machine) la clasificación de un conjunto de señas correspondientes a palabras, e incluso también algunas frases completas. En esta investigación se utilizaron histogramas de gradientes orientados y de flujo óptico, y se aplicó un modelo latente de SVM que implementaron para así mejorar el rendimiento del reconocimiento al nivel de palabra.

Así en dicha investigación se emplea el Kinect para grabar a la persona y mediante el uso de la SVM se identifica los momentos claves al realizar la seña. Lo importante en este proceso es la manera en que la SVM sabe cuándo inicia y termina una seña lo que le permite clasificar a un conjunto de ellas, y por consecuencia poder traducir frases completas.

Mientras que en el sistema de diálogo de la lengua de señas china [10] se hace uso de un guante con sensores en los que se utiliza dos técnicas de clasificación la primera consiste en utilizar árboles de decisión difusos, los cuales son utilizados para eliminar la gran cantidad de candidatos imposibles y segundo se utiliza los MOM utilizado como

extractor de características, el sistema es capaz de reconocer 5,113 señas, en las cuales se utilizó una muestra de 61,356 provenientes de 6 señantes, con lo que se obtuvo una precisión del 91.6% utilizando como método de validación la validación cruzada (cross validation).

Cada una de las investigaciones mencionadas anteriormente tienen resultados descritos del orden de muy favorables con respecto a la clasificación de las señas, sin embargo, también hay que indicar que sus experimentos se realizan bajo un contexto específico y controlado por lo que al ser presentados en un ambiente real puede que no se obtenga el mismo resultado.

### 3. Discusión

Las investigaciones descritas anteriormente involucran diferentes tecnologías, y conforme al avance del tiempo la tecnología empleada cambia, en 2009 se utilizó un lenguaje de marcas para definir las características de la seña, posteriormente se utilizaron cámaras y guantes, los cuales permitían identificar la mano y obtener una imagen de esta. El Kinect por otro lado viene a simplificar la detección de las partes del cuerpo humano, entre ellas la mano, pero a pesar de ello no es capaz de identificar los movimientos de los dedos de la mano.

Por otro lado, en el 2014 se trabajó con tecnologías enfocadas en el reconocimiento de los movimientos de la mano, una mediante el uso de impulsos eléctricos del brazo (Electromiograma), y el otro mediante el uso de sensores (LeapMotion), ambas tecnologías dejan fuera el reconocimiento facial y se orientan en la detección de los dedos de la mano dando así una mejor precisión para reconocer el movimiento, al dejar de fuera el reconocimiento facial se dejan fuera algunas señas, debido a que en algunas señas la expresión del rostro puede cambiar por completo el significado de esta.

En cuanto a la técnica utilizada varía en cada investigación, siendo la HMM las que se emplean en la mayoría de ocasiones, tal y como se muestra en la **Tabla 1**. Por otro lado, aquí se comentó brevemente el uso de las SVM. Así mismo entre otras de las técnicas más utilizadas se encuentran las Redes Neuronales que utilizan algoritmos de aprendizaje supervisado.

En las investigaciones se identifican señas estáticas y dinámicas siendo la diferencia entre ellas la forma de señado por parte de la persona que puede incluir movimiento o no, y en el caso de frases se consideran un conjunto de señas estáticas con una secuencia.

### 4. Conclusión

La descripción presentada en este artículo brinda un panorama general sobre la problemática en el reconocimiento de la seña, sin que la misma se exhaustiva. En el mismo se ha realizado un planteamiento sobre las diferentes tecnologías empleadas para el proceso de la adquisición de la imagen, así como las técnicas que más se han utilizado. Se pudo observar que en todos los casos descritos se reportan resultados favorables en el proceso de reconocimiento de la seña, pero vale la pena puntualizar que en todos los casos los experimentos se realizan bajo un contexto específico y

controlado por lo que al ser presentados en un ambiente real podrían no presentar las mismas condiciones de evaluación ya que hay que tomar en cuenta los aspectos de iluminación, ansiedad, y cualquier otra variación que afecte al normal desarrollo del ejercicio.

Un punto relevante a considerar en la revisión, y derivado del análisis de palabras clave se encontró que diversas lenguas se han venido explorando entre ellas la americana, la china, india, pero la lengua mexicana aún no tiene la misma presencia en artículos indexados en Scopus. Por lo cual es un punto de interés y atención para la comunidad multidisciplinaria.

Otro aspecto a considerar con relación a la experiencia de los autores, es que importante reconocer que el problema del reconocimiento de señas, por su propia naturaleza es amplio y puede ser abordado desde diferentes perspectivas. Por ello en la revisión realizada se observa que aún no existe un modelo o una guía que resuelva de forma eficiente el problema.

Finalmente se puede observar que a pesar de la tecnología y las técnicas utilizadas no se ha llegado a una solución universal, en concreto, solo se han obtenido soluciones parciales aplicados a contextos específicos, por lo cual se refuerza la necesidad de seguir investigando y proponiendo enfoques para apoyar al mejor tratamiento y reconocimiento de la lengua de señas.

## **Bibliografía**

1. D. O. de la Federación: Ley General para la Inclusión de las Personas con Discapacidad. D. Of. la Fed. (2011)
2. Cruz, M.: Gramática de la lengua de señas mexicana. Universidad del Valle, Colombia (2008)
3. Sun, C., Zhang, T., Bao, B. K., Xu, C.: Latent support vector machine for sign language recognition with Kinect. In: IEEE Int. Conf. Image Process., Vol. 6, No. 2, pp. 4190–4194 (2013)
4. Amatanon, V., Chanhang, S., Naiyanetr, P., Thongpang, S.: Sign language-Thai alphabet conversion based on Electromyogram (EMG). In: 7th Biomed. Eng. Int. Conf., pp. 1–4 (2014)
5. Khan, N., McClean, S., Zhang, S., Nugent, C.: Sign Language Recognition Using Leap Motion, A Support Vector Machine Approach. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Vol. 9454, pp. 50–59 (2015)
6. Benoit, A., Caplier, A.: SignTutor : An Interactive Language. *The Hand*, pp. 81–93 (2009)
7. Rojano, J. R., Morales, C., Rebolledo, G., Ortega, J. A., Muñoz, J.: Raise awareness in society about deafness: A proposal with Learning Objects and Scenarios. In: The Second International Conference on Higher Education Advances (2016)
8. Bouzid, Y., Jemni, M.: An avatar based approach for automatically interpreting a sign language notation. In: IEEE 13th Int. Conf. Adv. Learn. Technol. ICALT, pp. 92–94 (2013)
9. Trujillo, F., Caballero, S. O.: 3D data sensing for hand pose recognition. 23rd Int. Conf. Electron. Commun. Comput. CONIELECOMP, pp. 109–113 (2013)
10. Chen, Y., Gao, W., Fang, G., Yang, C., Wang, Z.: CSLDS : Chinese Sign Language Dialog System. pp. 1 (2010)

11. Fang, G., Gao, W., Zhao, D.: Large Vocabulary Sign Language Recognition Based on Fuzzy Decision Trees. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 34, No. 3, pp. 305–314 (2004)
12. Weichert, F., Bachmann, D., Rudak, B., Fisseler, D.: Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, Vol. 13, No. 5, pp. 6380–6393 (2013)
13. Chavan, P.: Indian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier (2016)
14. Galicia, R., Carranza, O., Jiménez, E. D, Rivera, G. E.: Mexican sign language recognition using movement sensor. pp. 573–578 (2015)
15. Jiang, M., Abulizi, A., You, X.: Decision-tree-based Algorithm for 3D Sign Classification Lab of Computational Linguistics. School of Humanities, Tsinghua Univ., Beijing 10084, China, Dept. of Math, Beijing University of Aeronautics and Astronautics, Beijing, 100191, P. R. Ch, pp. 1200–1204 (2014)
16. Rakun, E., Andriani, M., Wiprayoga, I. W., Danniswara, K., Tjandra, A.: Combining Depth Image and Skeleton Data from Kinect for Recognizing Words in the Sign System for Indonesian Language. *SIBI, Sistem Isyarat Bahasa Indonesia*, pp. 978–979 (2013)



## Sistema de planificación de trayectorias utilizando visión artificial

Jorge Cerezo-Sánchez<sup>1</sup>, Griselda Saldaña-González<sup>1</sup>, Mario M. Bustillo-Díaz<sup>2</sup>,  
Apolonio Ata-Pérez<sup>2</sup>

<sup>1</sup> Universidad Tecnológica de Puebla, División de Mecatrónica,  
Puebla, México

<sup>2</sup> Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México

{Jorge.cerezo, Griselda.saldana}@utpuebla.edu.mx, {bustillo, apolonio}@cs.buap.mx

**Resumen.** En este trabajo se presenta un sistema de visión artificial para el análisis de la trayectoria de un objeto, lo que permite resolver un laberinto robotizado analizando las rutas de un objeto móvil en el espacio. El prototipo consta de un laberinto que tiene solución al centro, de manera que independientemente de donde se coloque un objeto móvil, éste se reconoce y se mapea a un sistema coordenado. En función de la posición, un sistema de control de lazo cerrado se encarga de que el robot realice acciones que permiten la solución de dicho laberinto. El sistema cuenta con una interface de usuario realizada en LabVIEW desde donde el usuario puede proporcionar parámetros y visualizar el proceso. La PC se conecta con un PIC que recibe instrucciones, las interpreta y traduce posicionando los servomotores que permiten el movimiento del laberinto para su solución.

**Palabras clave:** Clasificación de imágenes, LabVIEW, sistema de visión, seguimiento de trayectoria, reconocimiento de patrones.

### Path Planning System Using Artificial Vision

**Abstract.** In this paper an artificial vision system for the analysis of the trajectory of an object is presented, which allows to solve a robotic labyrinth analyzing the routes of a moving object in the space. The prototype consists of a labyrinth with a unique solution at the center, so that regardless the position where a mobile object is placed, it is recognized and mapped to a coordinate system. Depending on the position, a closed loop control system ensures that the robot performs actions that allow the solution of the labyrinth. The system has a user interface made in LabVIEW from where the user can provide parameters and visualize the process. The PC connects with a PIC that receives instructions, interprets and translates them to position the servomotors that allow the movement of the labyrinth for its solution.

**Keywords:** Images classification, LabVIEW, vision system, path planning, pattern recognition.

## **1. Introducción**

Uno de los procesos más importantes en las aplicaciones de visión, es la localización de un objeto de interés dentro del campo de visión de la cámara, una tarea que se puede lograr mediante el reconocimiento de patrones [1-2].

El reconocimiento de patrones es un método para encontrar regiones en una imagen en escala de grises que coincidan con un patrón de referencia. A menudo se utiliza para localizar las posiciones del patrón, o rasgos característicos de un objeto en una imagen. Las posiciones pueden utilizarse para calcular longitud, ángulos, y otros parámetros. Como resultado, el reconocimiento de patrones se ha usado ampliamente en diversas aplicaciones tales como alineación, medición, e inspección [3-4]. El reconocimiento de patrones tiene una ventaja sobre el análisis de partículas o la detección de bordes debido a que la búsqueda del patrón no se basa en distinguir el brillo del objeto en comparación con el fondo de la imagen.

El objetivo de este trabajo consiste en el diseño e implementación de un sistema automático de reconocimiento de objetos por visión aplicado a la solución de un laberinto utilizando LabVIEW [5-6] y las herramientas de visión IMAQ. El toolbox de IMAQ ofrece un conjunto muy completo de funciones para adquisición y procesamiento digital de imágenes que mejoran la eficiencia de los proyectos y reducen el esfuerzo de programación de los usuarios obteniendo mejores resultados en menos tiempo. El sistema propuesto presenta la ventaja de realizar una inspección sin necesidad de entrar en contacto con el producto, alcanzando un alto nivel de flexibilidad y repetitividad.

Como primer punto en este trabajo, en la sección 2 se menciona la herramienta IMAQ de LabVIEW y las funciones que permiten realizar el reconocimiento de patrones, en la sección 3 se abordará una descripción del sistema desarrollado, incluyendo la parte de hardware y software del prototipo, a continuación en la sección 4, se discutirán brevemente algunas pruebas realizadas al sistema y los resultados obtenidos, para finalmente presentar las conclusiones en la sección 5.

## **2. Módulos de visión de LabVIEW**

LabVIEW soporta las técnicas de procesamiento de imágenes con la ayuda de bloques funcionales divididos en dos categorías, de bajo y alto nivel. El asistente de visión IMAQ es un módulo de alto nivel que permite realizar las funciones de visión de cinco dominios, procesamiento de imágenes, procesamiento de color, procesamiento en escala de grises, procesamiento binario, visión artificial e identificación. Cada una de estas funciones de procesamiento contiene otras funciones entre las cuales se incluyen histogramas, reconocimiento de patrones, detección de bordes, seguimiento de objetos, análisis de contornos, detección de forma, análisis de partículas, filtros, detección de textura, segmentación, entre otras.

El reconocimiento de patrones localiza rápidamente regiones en imágenes en escala de grises que corresponden a un patrón conocido. Este proceso puede implementarse en dos etapas. La primera es la de aprendizaje, mientras que el reconocimiento es la segunda. El valor de escala de grises se extrae del patrón proporcionado por el usuario

durante la etapa de aprendizaje. El algoritmo organiza y almacena la información en una forma que facilita la búsqueda en la imagen a inspeccionar.

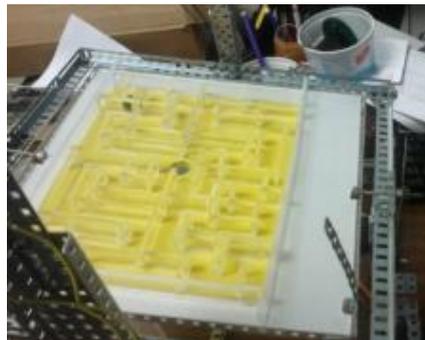
En la etapa de reconocimiento, el algoritmo extrae el valor de gris de la imagen a inspeccionar, y encuentra coincidencias localizando las regiones donde se observa la más alta correlación con la información aprendida del patrón.

Existen varios algoritmos implementados para realizar el reconocimiento de patrones, entre los que se incluye la correlación cruzada normalizada, el reconocimiento piramidal, el método de valor de gris y el método del gradiente.

En la implementación del sistema, se utilizó el método de Gray Scale Value Pyramid configurado con coincidencia para un solo objeto, y un score mínimo de 700.

### **3. Descripción del sistema**

El presente proyecto consiste en la implementación de un laberinto, el cual se soluciona empleando técnicas de visión. El sistema cuenta con una base cuadrada donde se ha implementado un laberinto de acrílico, con un ángulo de  $90^\circ$  en las esquinas y que tiene una única solución al centro. Esta base está montada sobre una estructura metálica que puede moverse en la dirección  $x$  e  $y$ , con referencia al plano que forma la base empleando servomotores que son controlados por microcontroladores PIC.



**Fig. 1.** Laberinto con solución al centro.

Para la implementación del sistema de visión se utiliza una cámara web ubicada en el plano  $z$ , calibrada por mediciones con respecto de las esquinas del laberinto. Dicha cámara se encarga de identificar una canica de color negro para determinar el siguiente movimiento que debe realizarse para hacer que la canica llegue a un orificio central donde se resuelve el laberinto.

#### **3.1. Desarrollo mecánico**

El laberinto se encuentra montado en una estructura metálica elaborada con varillas de aluminio, la cual cuenta con servomotores de 15 Kg. Conectados a flechas de 20 cm. éstas a su vez están conectadas a bases móviles. El movimiento de los motores en ambos ejes es sincronizado a través de tres microcontroladores de la familia 16F876A. Los microcontroladores están conectados en una arquitectura maestro-esclavo. El PIC

maestro recibe las instrucciones desde el sistema de visión y las interpreta, a su vez envía instrucciones a los PIC'S esclavos posicionando los servomotores en el ángulo deseado.

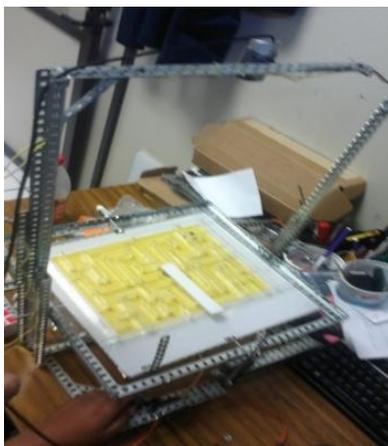


Fig. 2. Estructura mecánica del prototipo.



Fig. 3. Ubicación de los motores para posicionamiento del laberinto.

### 3.2. Desarrollo del software

El sistema de visión se implementó empleando el entorno de desarrollo de LabVIEW, debido a la facilidad con que permite utilizar diferentes dispositivos de adquisición de imágenes. Además permite desarrollar interfaces de usuario de forma intuitiva y los algoritmos de visión y reconocimiento de patrones pueden ser implementados con mayor facilidad que con otro tipo de programas.

La implementación del sistema inicia con la calibración y alineamiento de la cámara con respecto del plano, para evitar procesos de calibración demasiado pesados para el sistema, este procedimiento se logró identificando cuatro puntos en el laberinto y alienado la cámara de manera manual con cuatro regiones de interés que corresponden a las esquinas del laberinto, posteriormente se realiza la adquisición de una imagen en tres dimensiones. A continuación se identifica el objeto de interés (canica) y se separa del fondo de la imagen, para finalmente comparar sus características con las de un patrón previamente definido. Finalmente, el algoritmo permite obtener como resultado la posición de dicho objeto en el plano de la imagen.

Para ubicar el objeto de interés, se generó un patrón esférico con la imagen de la canica que se desea encontrar y se indicó al programa que la discriminación se hiciera por forma.

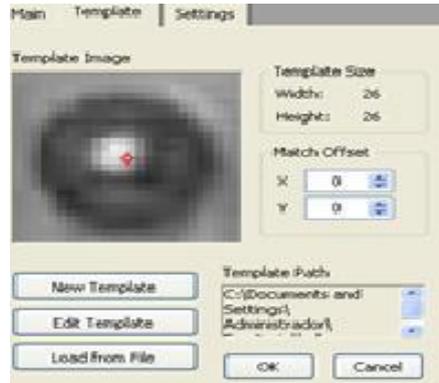


Fig. 4. Imagen del patrón.

Se realiza un filtro a la imagen adquirida para convertirla en escala de grises y a continuación se determinó una Región de Interés (ROI) que consiste en todas las áreas que no pertenecen al patrón y que corresponde con el área del laberinto.

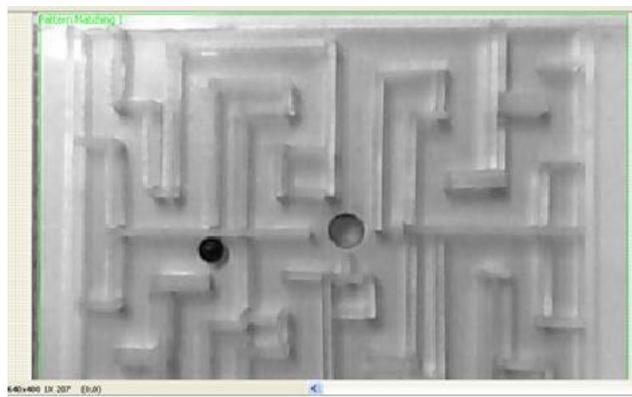


Fig. 5. Región de interés (ROI).

Una vez definido el patrón a seguir y eliminando las áreas que se requieren discriminar del mismo (Fig.6), el sistema de visión devuelve un clúster de información con los datos de origen en  $(x, y)$  que es la primera posición donde el sistema detecta la canica, y referencia  $(x, y)$  que es la posición actual del objeto, una vez conocida la posición de referencia y actual de la canica, se realiza un mapeo del laberinto obtenido con el sistema de visión en una imagen compuesta en formato JPG, y se dibuja sobre esta imagen una bola que corresponde a la posición de la canica real para verificar que las lecturas del sistema de visión corresponden a las medidas en el mundo real, actualizando ésta cada milisegundo con las nuevas posiciones de la canica.

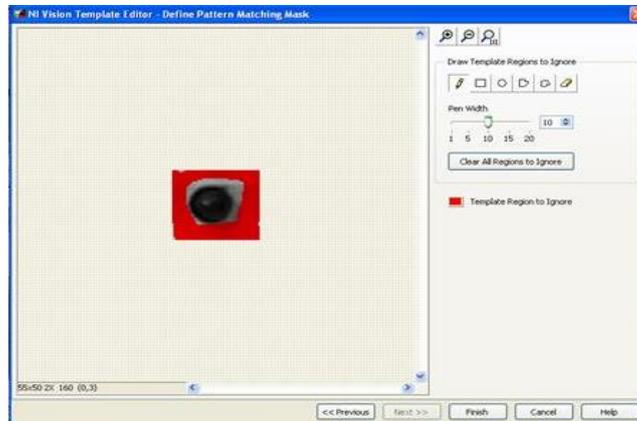


Fig. 6. Eliminación de las áreas no deseadas del patrón.

Una vez que se mapea la imagen y se obtiene la posición del objeto en el plano, se analizan las regiones cercanas con respecto a las esquinas del laberinto y se define un movimiento hacia adelante, atrás, izquierda o derecha según sea el caso para hacer que la canica se vaya aproximando a la solución del centro. Esta información se envía a los motores por medio de una cadena de mando en formato RS232, en la estructura mecánica, ya que corresponde con un movimiento de la flecha para posicionar el laberinto en  $x$  e  $y$ .

En la implementación del prototipo, se crea un sistema de control de lazo cerrado, que considera las regiones donde se encuentra el objeto, para realizar un movimiento, de forma tal que, independientemente de donde se coloque la canica el programa será capaz de dar solución a cualquier caso presentado.

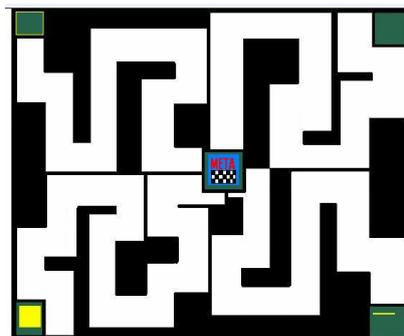


Fig. 7. Imagen del Laberinto en LabVIEW con la solución al centro.

### 3.3. Algoritmo de decisiones

Para que el sistema de visión pueda realizar un movimiento mecánico del laberinto se requiere de un sistema de toma de decisiones que sea capaz de reaccionar con un movimiento de la canica y trasladarla a un movimiento mecánico.

Las posiciones de la canica entran en un árbol de toma de decisiones, para poder determinar el movimiento de la bola por cada muestra proporcionada por la cámara, se realiza una máquina de estados para retroalimentar al sistema con las coordenadas de la bola y éstas son tomadas como parámetros para el árbol de toma de decisiones.

Una vez definido el movimiento que se tiene que realizar con el robot, este mediante un protocolo RS232, envía una cadena serial que corresponde al movimiento necesario para trasladar la canica al siguiente nodo de evaluación. La correspondencia en pixeles de la posición de la canica es mapeada hacia el dibujo del laberinto (Fig. 5) y es superpuesta sobre la imagen para corroborar la correspondencia de la posición física con lo reconocido por el sistema de visión.

La figura 8 muestra la máquina de estados que muestra el funcionamiento del sistema de seguimiento.

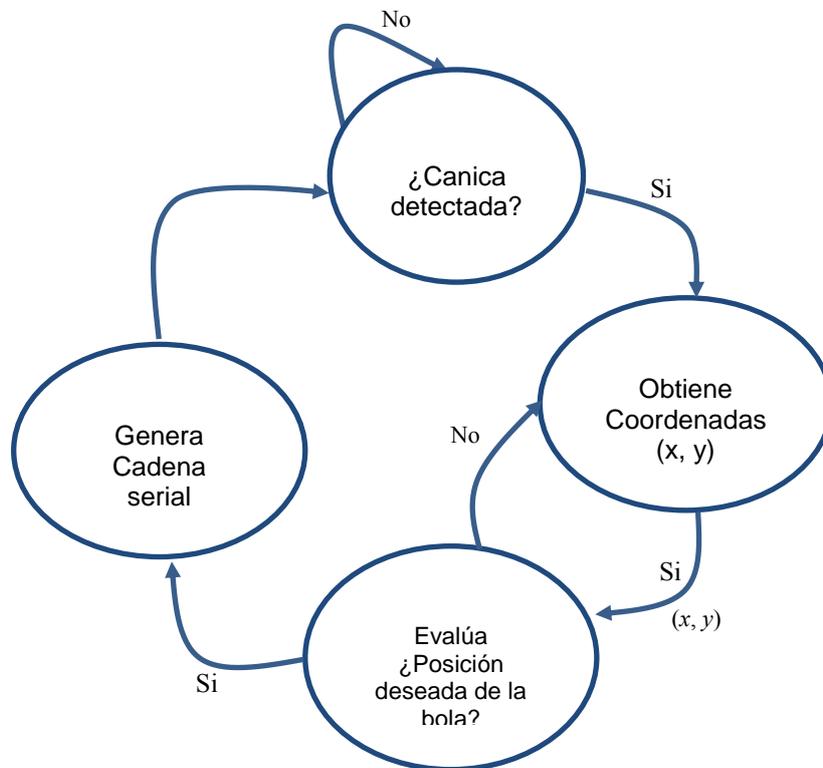


Fig. 8 Máquina de estados para el sistema de seguimiento.

Una vez definida la cadena serial que corresponda al caso para región de interés y recibida por un PIC maestro, este interpreta cada comando proveniente del sistema de visión para controlar dos microcontroladores esclavos que son los que físicamente modulan el PWM correspondiente para cada motor. Dando como resultado una cadena no determinística de movimientos en el plano físico X, Y y llevando a la canica a solucionar en el centro.

El diseño de esta máquina de estados permite resolver casos especiales, por ejemplo si la canica por alguna situación regresa o rebota en alguna pared, el diseño dinámico de la máquina de estados permite reorientar la canica a la siguiente posición.

#### **4. Pruebas y resultados**

Para comprobar la funcionalidad del Sistema, se realizaron múltiples pruebas colocando la canica dentro del laberinto desde diferentes puntos de inicio. Se observó que el movimiento de los motores fuera el adecuado para dar solución al laberinto. Desde la interface de usuario se observa que en todos los casos se podía identificar la canica y que el sistema de visión ejecuta de manera satisfactoria el algoritmo de seguimiento. El sistema es capaz de identificar y corregir de forma adecuada varios casos incluyendo cuando la canica por acciones mecánicas regresa a un punto anterior, el sistema de visión detecta y realiza el movimiento para reenviar la canica al punto siguiente.

#### **5. Conclusiones**

En este trabajo se presentó un prototipo funcional de un sistema de visión para seguimiento de objetos aplicado a la solución de un laberinto. El sistema se desarrolló exitosamente utilizando el toolbox IMAQ de LabVIEW, que permitió un desarrollo rápido y más simple en comparación con una programación que no es de tipo gráfica. Esta aplicación permitió desarrollar un Sistema de visión bastante robusto para el seguimiento de objetos, definición de trayectorias, y corrección de casos en tiempo real.

#### **Referencias**

1. Kwon, K S., Ready, S.: Practical Guide to Machine Vision Software: An Introduction with LabVIEW. N.J., USA, Wiley, pp. 279 (2014)
2. Fukunaga, K.: Introduction to Statistical Pattern Recognition. 2nd ed, USA, Academic Press, pp. 615, (1990)
3. Saldaña, G., Estévez, J., Gracios, C. A.: Sistema Automático de Clasificación Aplicado a la Industria Automotriz. Revista de Ciencia e Ingeniería del Instituto Tecnológico Superior de Coatzacoalcos, Vol.1, No.1, pp. 266–270 (2014)
4. Malamas, E. N.: A survey on industrial vision systems, applications and tools; Image and Vision Computing. Vol.21, No.2, pp. 171–188 (2003)
5. Travis, J., Kring, J.: LabVIEW for Everyone: Graphical Programming Made Easy and Fun. 3er Edition, New Jersey, Prentice Hall, pp. 960 (2006)
6. Lajara, J. R., Pelegrí, J. P.: LabVIEW: Entorno gráfico de programación. 2nd Edition, México, Alfaomega, pp. 477 (2013)

## **RENTOL: Un algoritmo de agrupamiento basado en K-means**

Erendira Rendon Lara, Itzel María Abundez Barrera

Instituto Tecnológico de Toluca, División de Estudios de Posgrado e Investigación,  
México

erendir@prodigy.net.mx erendonl@toluca.tecnm.mx,  
iabundezb@toluca.tecnm.mx

**Resumen.** Sin lugar a duda el algoritmo K-means es el más utilizado en la comunidad de aprendizaje no supervisado. Desafortunadamente es muy sensible a la selección de los centroides iniciales. Debido a ello, se han propuesto un gran número de métodos para la selección de los centros iniciales. En este artículo se presenta un algoritmo de agrupamiento que tiene como base al algoritmo K-means, en el cual se implementa un nuevo método para la selección de centros iniciales. Para evaluar los resultados se utilizaron índices internos de validación de agrupamiento. Los resultados del algoritmo propuesto fueron comparados con los algoritmos K-means y el Algoritmo K-means++. De acuerdo a las pruebas realizadas, RENTOL mejoró los resultados de K-means y K-means++.

**Palabras clave:** Agrupamiento K-means, índices de validación, algoritmo K++ means.

## **RENTOL: A Clustering Algorithm Based on K-means**

**Abstract.** K-means algorithm is the most widely used in the community unsupervised learning. Unfortunately it is very sensitive to the selection of initial centroids. For this reason, there are proposed many methods for selection of initial centers. This article presents a clustering algorithm, which is based on K-means algorithm, which a new method for selecting implemented centers initial. To evaluate the results internal clustering validation indices were. The results of the proposed algorithm were compared with the K-means algorithm and K-means ++ algorithms. According to tests, the results of RENTOL improved K-means and K-means ++.

**Keywords:** K-means clustering, validation indexes, K++ means, algorithm.

## 1. Introducción

Agrupamiento es una técnica de clasificación no supervisada, cuyo objetivo es encontrar o descubrir grupos en un conjunto de patrones u objetos, además de ser una de las tareas más importantes en el análisis de datos y en la minería de datos [8]. La técnica de agrupamiento ha sido ampliamente utilizada en la detección de anomalías, identificación de características sobresalientes de conjuntos de datos, etc., en diferentes áreas del conocimiento como: biología, antropología, medicina, estadística, matemáticas entre otras. Por lo cual se han desarrollado una gran diversidad de técnicas desde sus inicios en los años 50's.

Existe en la literatura diferentes clasificaciones de los algoritmos de agrupamiento [10], aunque la más general es proporcionada por Jain [5], donde los algoritmos son clasificados en dos categorías: jerárquicos y de partición. Los algoritmos jerárquicos trabajan interactivamente encontrando una secuencia anidada de particiones tomando en cuenta un criterio para unir o dividir grupos en función de una medida de similitud. Por otra parte los algoritmos de partición, encuentran particiones del conjunto de datos sin ninguna jerarquía. La mayoría de los algoritmos jerárquicos tienen un orden (complejidad) cuadrático [6], por lo que tienen problemas cuando trabajan con grandes volúmenes de datos, mientras que los algoritmos de partición tienen una complejidad menor.

Podemos encontrar en la literatura una gran cantidad de algoritmos de partición [10]. Pero uno de los más utilizados y referenciados es el algoritmo de agrupamiento K-Means. Esto se atribuye a varias razones que a continuación se enlistan:

- La facilidad para comprender su funcionamiento, así como su implementación.
- El algoritmo puede tomar diferentes caminos si se utilizan diferentes criterios como: distancia, método de selección de centroides iniciales o el criterio de terminación del algoritmo.
- El tiempo de complejidad es lineal

A pesar de las ventajas mencionadas anteriormente, no está exento de algunas desventajas que han sido reportadas ampliamente en la literatura [5], las más importantes se enlistan a continuación:

- Tiene dificultades para detectar grupos con formas no esféricas y de tamaño diferente. Cuando los grupos naturales son muy grandes existe una alta probabilidad que seleccionen, como centroides iniciales a puntos del mismo grupo natural, esto ocasionará que el algoritmo divida al grupo natural.
- Es muy sensible al ruido y a puntos atípicos. Si se llegará a elegir como centroide uno de estos puntos, los centroides resultantes pudieran no ser tan representativos como debieran serlo.
- Converge a un mínimo local de la función criterio, obteniendo una solución pobre.
- Es altamente sensible al método para seleccionar los centroides iniciales.

Una forma de contribuir a la eficiencia del algoritmo, es mejorar el método de selección de centros iniciales. Así en este trabajo se presenta un algoritmo de

agrupamiento basado en el algoritmo K- Means, que incluye un método diferente de seleccionar los centros iniciales.

El resto del artículo está organizado de la siguiente manera: en la sección 2 se describe el algoritmo K-Means mencionando algunas de sus ventajas y desventajas, en la sección 3 se describen brevemente los índices de validación internos utilizados en este trabajo. En la sección 4 se presentan algunos trabajos sobre métodos de inicialización de centros, en la sección 5 se describe el algoritmo propuesto, denominado RENTOL. En las secciones 6 y 7 se presentan los resultados obtenidos y la discusión de ellos respectivamente.

## 2. El algoritmo K-means

La mayoría de los algoritmos de partición tienen como base la optimización de una función criterio que denominaremos  $F$ , el valor de esta función depende de las particiones del conjunto de datos  $\{C_1, \dots, C_K\}$ , es decir:

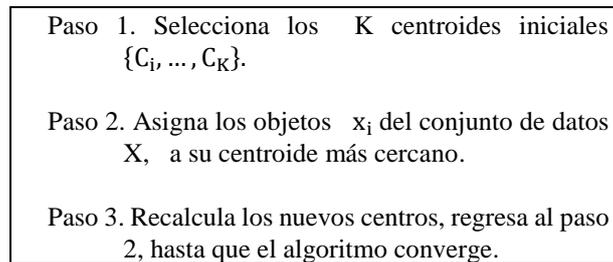
$$F: P_K(X) \rightarrow \mathbb{R}, \quad (1)$$

donde  $P_K(X)$  son las particiones del conjunto de datos  $X = \{x_1, \dots, x_n\}$  en  $K$  grupos no vacíos.  $x_i$  es un vector  $n$  – dimensional (objeto) del conjunto de datos  $X$ . El algoritmo K-Means converge a un mínimo local, utilizando la función criterio  $F$ , de la sumatoria de las distancias  $L^2$  entre cada objeto y su centroide más cercano. A este criterio normalmente se le denomina error cuadrático. Que puede ser expresado como sigue:

$$F(\{C_1, \dots, C_K\}) = \sum_{i=1}^K \sum_{j=1}^{p_i} \|x_{ij} - \bar{C}_i\|^2, \quad (2)$$

donde  $K$  es el número de grupos,  $p_i$  es el número de objetos del grupo  $i$ ,  $x_{ij}$  es el  $j$  – ésimo objeto del  $i$  – ésimo grupo y  $\bar{C}_i$  es el centroide del  $i$  – ésimo grupo el cual es calculado de la siguiente manera:

$$\bar{C}_i = \frac{1}{p_i} \sum_{j=1}^{p_i} x_{ij}, \quad i = 1, \dots, K. \quad (3)$$



**Fig. 1.** Pseudo-código del algoritmo K-means.

En la Figura 1, se presenta el pseudocódigo del algoritmo K-Means. El algoritmo inicia seleccionando o calculando los centroides iniciales, dependiendo del criterio de selección de centroides, posteriormente asigna los objetos a su centroide más cercano,

para después recalcular los nuevos centroides esto lo realiza hasta que el algoritmo converja (paso 3).

### 3. Índices validación de agrupamiento

Los índices de validación de agrupamiento son utilizados para evaluar los resultados de los algoritmos de agrupamiento. Existen tres enfoques para el estudio de esta validación: criterio externo, relativo e interno. El criterio interno evalúa los resultados de un algoritmo de agrupamiento considerando información propia del conjunto de datos (como por ejemplo la matriz de proximidad), sin considerar información adicional [7]. En este trabajo se utilizaron los siguientes índices internos para comparar los resultados obtenidos.

#### 3.1. Índice de validación S\_Dbw

El índice de validación S\_Dbw fue propuesto por M. Halkidi [7]. Este índice evalúa los resultados de un algoritmo de agrupamiento, en función de la densidad y separación de los grupos. Para lo cual mide la varianza intra-grupo e inter-grupo, representada por la Ec. 4:

$$S\_Dbw = Scatt + Dens\_bw. \quad (4)$$

#### 3.2. Índice de validación PS

El índice PS fue propuesto por ChienHsing Chou [3], el cual calcula un promedio de la distancia simétrica hacia los otros centros, con la Ec. 5:

$$PS(C) = \frac{1}{K} \sum_{i=1}^K \left[ \frac{1}{p_i} \sum_{j=1}^{p_i} \frac{d_c(x_i, \bar{C}_i)}{d_{\min}} \right], \quad (5)$$

donde  $d_c(x_i, \bar{C}_i) = d_s(x_i, \bar{C}_i) d_e(x_i, \bar{C}_i)$ ;  $d_e = (x_i, \bar{C}_i)$  es la distancia Euclidiana entre el punto  $x_i$  y el centroide  $\bar{C}_i$  y  $d_s$  es la distancia simétrica [13].

#### 3.3. Índice de validación CS

El índice CS fue propuesto por Chien-Hsing Chou [4]. Este índice evalúa la calidad del agrupamiento tomando en cuenta la densidad y tamaño de los grupos. Se define con la Ec. 6:

$$CS(C) = \frac{\frac{1}{K} \sum_{i=1}^K \left\{ \frac{1}{|C_i|} \sum_{x_j \in C_i} \max \{d(x_j, x_k)\} \right\}}{\frac{1}{K} \sum_{i=1}^K \left\{ \min_{y \in C, j \neq i} \{d(\bar{C}_i, \bar{C}_j)\} \right\}}, \quad (6)$$

donde  $\{C_1, \dots, C_K\}$  son los centroides del conjunto de datos  $X$ , de los grupos encontrados por el algoritmo de agrupamiento y  $d$  es una función de distancia, esta medida es una función de la relación de las sumas de la dispersión dentro del grupo y la separación entre grupos.

#### 4. Trabajos relacionados

M. Emre Celebi [13], presenta una comparación de métodos de inicialización del algoritmo K-Means de orden lineal con respecto al número de objetos: El método Forgy, MacQueen, maximin, Bradley y Frayyad, k-means++, greedy k-means++, Var-Part y PCA-Part. Esta comparación fue realizada con conjuntos de datos reales y sintéticos. Además la comparación se realizó en función de la calidad de agrupamiento y la velocidad de procesamiento, para lo cual se emplearon diferentes índices de validación internos y externos (calidad) el tiempo de CPU y número de iteraciones (velocidad). Concluyendo que los métodos de inicialización del algoritmo K-Means que dieron malos resultados fueron: Forgy, Macqueen y maximin.

El método PCA-Part [14], utilizó un método de agrupamiento jerárquico de tipo divisivo, para lo cual hicieron uso de PCA (Principal Component Analysis). El método inicia, considerando un solo grupo, que contiene a todos los objetos de la muestra en estudio y selecciona iterativamente el grupo con el mayor SSE (suma de cuadrados debida al error), el cual es seleccionado para ser dividido en dos grupos, considerando el hiperplano que pasa por el centro del grupo y que es ortogonal a la dirección del eigenvector de la matriz de covarianza. Este procedimiento se repite hasta que se encuentran los  $K$  grupos deseados. Así los centros de los grupos obtenidos son considerados como centroides iniciales del algoritmo K-Means.

El Algoritmo K-Means++, propuesto por Arthur [1], es una modificación del clásico K-means, esta diferencia radica en el método de selección de los centroides iniciales. El método de inicialización de centros, inicia seleccionando aleatoriamente el primer centro  $C_1$  del conjunto de datos  $X$ , después escoge el siguiente centro  $C_i = x' \in X$  con probabilidad  $\frac{D(x')^2}{\sum_{x \in X} D(x)^2}$  proporcional a  $D(x')^2$ , donde  $D(x')^2$  es la distancia más corta entre el punto  $x$  y los centros seleccionados previamente. El proceso de selección se repite hasta que  $C_i = K$ .

Onoda et al. [11], realizan un análisis exhaustivo del método K-means ++ y el método KKz [9], además proponen un nuevo método de selección de centroides, este método se basa en el análisis de componentes independientes (ICA), el cual consiste básicamente de dos pasos:

- 1.- Obtiene los  $K$  componentes independientes  $IC_1, \dots, IC_k$  del conjunto de datos  $X$ .
- 2.- Selecciona los  $K$  centros iniciales  $C_i$  ( $i = 1, \dots, k$ ) de  $C_i = x' \in X$  con un valor mínimo de  $\frac{IC_i \cdot x'}{|IC_i||x'|}$ .

## 5. Método propuesto

Sea  $X = x_1, x_2, \dots, x_N$  un conjunto de objetos de tamaño  $N$ , donde  $x_i \in \mathbb{R}^n$ , es decir  $N$  objetos (vectores) con  $n$  características. El algoritmo RENTOL divide el conjunto  $X$  en  $K$  particiones,  $C = \{c_1, \dots, c_K\}$ , tal que  $\cup_{i=1}^K c_i = X$ ,  $c_i \cap c_j = \emptyset$  para  $1 \leq i \neq j \leq K$ . Cada una de las  $c_i$  particiones o grupos tendrán un centroide  $\bar{c}_j$ .

### 5.1. Algoritmo propuesto: RENTOL

El algoritmo propuesto requiere como parámetro de entrada el número de grupos  $K$  a formar, al igual que el algoritmo K-means. El algoritmo inicia calculando 2 centroides, estos centros son los objetos más alejados del conjunto de datos  $X$  (paso 2).

El algoritmo RENTOL requiere como entrada el número de grupos a formar,  $K$

Paso 1. Inicia considerando  $K' = 2$

Paso 2. Encuentra los dos objetos más alejados entre sí del conjunto de datos  $X$ , que serán los dos centros iniciales

Paso 3. Asigna cada  $x_i \in X$  para  $i = 1, \dots, N$  al centroide más cercano  $C = \{c_1, \dots, c_{K'}\}$ ,  $j = 1, \dots, K'$  obteniéndose los grupos  $C = \{c_1, \dots, c_{K'}\}$

Paso 4. Calcula los nuevos centros de los grupos  $C = \{c_1, \dots, c_{K'}\}$ , como,

$$\bar{c}_j = \frac{1}{|c_j|} \sum_{i=1}^{|c_j|} x_{ij}$$

Paso 5. Repite los pasos 3 y 4 utilizando los nuevos centroides (obtenidos en el paso anterior)  $C = \{\bar{c}_1, \dots, \bar{c}_{K'}\}$ , hasta que los objetos no cambien de grupo.

Paso 6. Si  $K' = K$  el algoritmo termina

Paso 7. Escoge el siguiente centro  $x_i \in X$ , que sea el más alejado de los centros  $C = \{\bar{c}_1, \dots, \bar{c}_{K'}\}$

Paso 8. Incrementa en uno el número de centros  $K' = K' + 1$

Paso 9. Regresa al paso 3 con  $K'$

Fig. 2. Pseudo-código del algoritmo RENTOL.

En el paso 3, los objetos del conjunto  $X$  son asignados a su centroide más cercano. En el paso 4 se recalculan los centros y se repite el paso 3 con los nuevos centros. Los pasos 3 y 4 se repiten hasta que los objetos no cambien de grupo. En el paso 6 se verifica si el número de grupos actual ( $K'$ ) es igual al número de grupos deseado  $K$ , si es así, entonces el algoritmo termina. En caso contrario se calcula el siguiente centroide (paso 7).

Este centro será el objeto más alejado de los centros previamente calculados, en el paso 8 se incrementa en uno el valor  $K'$ . En el paso 9 el algoritmo regresa al paso 3 con el valor actualizado de  $K'$ . En la Fig. 2 se presenta el pseudocódigo del algoritmo.

## 6. Resultados experimentales

### 6.1. Descripción de los conjuntos de datos

Para comprobar los resultados obtenidos por el algoritmo propuesto se utilizaron 27 conjuntos de datos sintéticos, con características bidimensionales estos conjuntos describen figuras geométricas (círculos, elipses, cuadros y rectángulos) de diferentes tamaños, también describen formas irregulares. Además se utilizaron conjuntos de datos reales que fueron obtenidos Irvine, CA: University of California, School of Information and Computer Science<sup>1</sup> (ver Tabla 1).

Table 1. Descripción de los conjuntos de datos.

Conjunto de datos	Número de Registros	Número de Características	Número de clases
Iris	150	4	3
Wine	178	13	3
Diabetes	768	8	2
Glass	214	9	7

### 6.2. Pruebas realizadas

Para cada una de la pruebas se ejecutaron, los algoritmos K-means, K-means++ y RENTOL, haciendo variaciones del valor de  $K$ , desde  $K = 2, \dots, 9$ . Se utilizó la distancia Euclidiana, además se midió la calidad de agrupamiento utilizando índices de validación internos: PS, S\_Dbw, CS.

### 6.3. Resultados

En las tablas 2 y 3 se muestran los resultados obtenidos de las pruebas realizadas, cuando se utilizaron las muestras con los algoritmos K-means, K-means++ y RENTOL.

<sup>1</sup> <http://ics.uci.edu/mllearn/MLRepository.html>.

En el último renglón aparece el porcentaje de aciertos por cada algoritmo de acuerdo al índice de validación.

**Table 2.** Resultados con datos reales.

	<b>K-Means</b>	<b>K++Means</b>	<b>Rentol</b>
Iris	2	5	3
Wine	2	2	3
Diabetes	2	2	2
Glass	2	8	9
Porcentaje	25%	25%	75%

Por otro lado también se realizaron pruebas con 27 conjuntos de datos sintéticos, donde se obtuvieron los porcentajes de eficiencia de acuerdo al índice PS, ver Tabla 3.

**Table 3.** Resumen de resultados con datos sintéticos.

<b>Algoritmo</b>	<b>PS (Promedio % )</b>
K-Means	74.00
K++Means	70.03
Rentol	88.88

Es importante resaltar que en las corridas de los experimentos se utilizaron los índices mencionados anteriormente, sin embargo solo se presenta los obtenidos por el índice PS ya que fueron los mejores. Aunque los porcentajes de los otros índices fueron menores se mantuvo la tendencia, es decir el algoritmo Rentol siempre obtuvo porcentajes mayores.

## 7. Conclusiones

En este artículo se propone un algoritmo de agrupamiento llamado RENTOL. Los resultados fueron comparados con el algoritmo K-means y K-means++, para medir la eficiencia se utilizaron los índices de validación de agrupamiento PS, CS y S\_Dbw. En todas las pruebas se utilizó la distancia Euclidiana. Los experimentos se realizaron utilizando datos sintéticos y datos reales.

Los resultados obtenidos, revelan que el algoritmo propuesto obtuvo mejores resultados que los algoritmos K-Means con inicialización aleatoria y K-means++ de acuerdo a los índices de validación PS, CS y S\_Dbw. Además elimina la dependencia de la selección de centros iniciales. Aunque el algoritmo que se propone tiene una complejidad mayor al K-means. Sin embargo en un trabajo próximo se trabajará en eliminar esta desventaja.

En el futuro se realizarán experimentos para la segmentación de imágenes médicas e imágenes de un microscopio electrónico.

## Referencias

1. Arthur, D., Vassilvitskii, S., K-means++: The advances of careful seeding. In: Proc. of the 18th annual ACM-SIAM symposium on discrete algorithms, pp.1027–1035 (2007)
2. Celebi, M., Emre, K., Hassan, A., Vela, A.: A Comparative study of efficient initialization methods for k-means clustering algorithm. Expert System with Applications, 40, pp. 200–210 (2013)
3. Chow, C. H., Su, M. C., Lai, E.: Symmetry as a new measure for Cluster Validity. In: 2th WSEAS Int. Conf. Scientific Computation and Soft Computing, Crete, Greece, pp. 209–213 (2002)
4. Chow, C. H., Su, M. C., Lai, E.: A new Validity Measure for Clusters with Different Densities. Pattern Anal. Applications, 7, pp. 2005–2020 (2004)
5. Jain, A. K.: Data clustering: 50 years beyond k-means. Pattern Recognition Letters, 31(8), pp. 651–666 (2010)
6. Jain, A. K., Murty, M. N., Flynn, P. J.: Data Clustering: A Review. ACM Computing Surveys, pp. 651–666 (1999)
7. Halkidi, M., Vazirgiannis, M.: Quality scheme assessment in the clustering process. In: Proc. PKDD (Principles and Practice of Knowledge in databases), Lyon, France, Lecture Notes in Artificial Intelligence, Springer-Verlag, Vol. 1910, pp. 265–279 (2000)
8. Kantardzic, M.: Data mining: Concepts, Models, Methods, and Algorithms. Wiley Inter-Science (2001)
9. Katsavounidis, I., Kuo, C. J., Zhang, Z. J., Zhang, Z.: A new initialization technique for generalized Lloyd iteration. IEEE Signal Processing Letters, Vol. 1, No. 10, pp. 144–146 (1994)
10. Kaufman, L., Rousseeuw, P. J.: Finding groups in data. Wiley Inter-Science (1990)
11. Onada, T., Sakai, M., Yamada, S.: Careful seeding method base on independent components analysis for k-means clustering. Journal of Emerging Technologies in Web Intelligence, 4(1), pp. 51–59 (2012)
12. Peña, J. M., Larranaga, P.: An empirical comparison of four initialization methods for K-means. Pattern Recognition Letters, 20(10), pp. 1027–1040 (1999)
13. Su, M. C., Chow, C. H.: A Modified Version of the K-Means Algorithm with a Distance Based on Cluster Symmetry. IEEE Trans. Pattern Anal. and Machine Intelligence, Vol. 23, No. 6, pp. 674–680 (2001)
14. Su, T., Dy, J. G.: In search of deterministic methods for initializing K-means and Gaussian mixture clustering. Intelligent Data Analysis, 11(4), pp. 319–338 (2007)



# Multi-sensory Tracking of Elders in Outdoor Environments on Ambient Assisted Living

Javier Jiménez Alemán

Fluminense Federal University, Niterói,  
Brazil

[jjimenezaleman@ic.uff.br](mailto:jjimenezaleman@ic.uff.br)

**Abstract.** Ambient Assisted Living (AAL) is an emergent area that provides useful mechanisms that allows tracking elders through opportunistic sensing, for example, using mobile devices. In that sense, it is necessary a permanent attention to these people by caregivers and relatives and a growing necessity to create mechanism to support this task. In our proposal, we have a network of caregivers that work together to ensure the wellbeing of elders in their daily activities. This paper aims to model the interaction process of different actors with SafeRoute, an AAL system that pretends monitoring elders in their day-to-day daily living activities in outdoor environments. SafeRoute merges sensor data provided by different sensors built in mobile devices to provide alert mechanisms for caregivers.

**Keywords:** Interaction model, AAL, sensors, mobile application.

## 1 Introduction

The increase in life in Brazil and all over the world expectancy brings with an increase in the prevalence of health problem, such as physical limitations (balance problems), perceptual problems (vision and hearing problems) and cognitive problems (loss of memory, difficult to make parallel tasks). Events such as falls tend to occur frequently when persons are older. That fact implies an increase of the permanent attention to these people by his caregivers and relatives and a growing necessity to create mechanism to support this task [1].

Ambient Intelligence (AmI) refers to a vision in which people are empowered by an electronic environment that is sensitive and responsive to their needs, and is aware of their presence. Its target is improving quality of life by creating the desired atmosphere through intelligent and inter-connected systems. Inside AmI, Ambient Assisted Living (AAL) can be defined as the use of information and communication technologies (ICT) in a person's daily living and working environment to enable them to stay active longer, remain socially connected and live independently into old age [2]. These technologies can supply security for people with a wide range of physical and cognitive challenges, although the focus is often on supporting elders. AAL uses sensing technology

embedded in objects, or in the environment to promote health and enhance wellbeing. Furthermore, AAL technologies also allow best communication between elders with his relatives and friends [3].

Concerning AAL, Smart Homes concerns the methods, concepts, electronic systems, products and services that assist elders or people with disabilities in their daily lives in a controlled environment. However, the sense of security changes dramatically when elders are in an uncontrolled environment where the risk to accidents is multiplied. It is very important to increase the safety of this people, to reduce the response time in emergency situations (fall, loss). Nowadays, smartphones not only work like communication devices, but also are equipped with several sensors like accelerometer, GPS system and camera. All these sensors make possible a wide range of applications like the assistance to elders and people with disabilities.

Attending that, was created the SafeRoute project, a system able to assist elders with activities related to their day-to-day activities in outdoor environments that use geo-localization technologies built-in mobiles devices. The purpose of this work is to present the SafeRoute projects, who aim to monitor elders who follow predefined routes and to notify to their caregiver in case of emergencies (losses and falls). We considered the inclusion of actors and other elements in our model that contribute to improve the monitoring process of elders and to reduce response time in case of emergencies.

## **2 Related Work**

In the case of AAL systems, it is necessary to provide to elders the simplest way to interact with the systems, to minimize as much as possible the error product of this interaction. It is recommended limiting the possible options, dialogues has to be linear and parallel tasks must to be avoided. For example, Kopal [4] is a system, which provides a feedback mechanism based on speech, which helps people with dementia and their caregivers. The caregiver may use the own mobile for ask for information to the elder through the own mobile. OutCare [5] is an example of an AAL application that attempts to help people with Alzheimer's and their caregivers. In this case, the authors implemented a multiagent system (MAS) that monitoring outdoor daily routines for people with Alzheimer, sending alerts to their caregivers in case of deviation via SMS, e-mail, and voice.

Another system that tries to help people with mental diseases is AlzNav from Fraunhofer Portugal [6]. This system intends to help elder and people in general who are in early stages of dementia. This open source project proposes an interface that guides the user through a compass indicating the right direction to go on the trip. Furthermore, it provides a location system that allows caregiver to send a SMS with the location for the people who need care.

Another system that attempts help people with cognitive disabilities in their day-to-day journey is [7]. In this case, it consists in a system location to assist caregivers to be aware of current location of these disabled people. The system is divided in three main parts: one application for the person with cognitive impairment, and two apps for your caregiver (mobile and web server).

The author proposes a decision algorithm for creating alerts whenever necessary and used augmented reality to provide help to user through few options. The user can choose a fixed destination (generic location, such as shopping, work or home) according to the type of destinations stored in the database.

Authors of [8] presented a sensor-based map navigation approach in very early stages for users who due to disabilities or lack of technical knowledge. Their vision is that artificial companion can be achieved based on the integration of a variety of information sources which includes both user profiles and real-time sensor data. They described the challenges and potential solutions related to this vision and provide a number of scenarios that illustrate its utility. The main contributions of this paper is to provide an AAL system in the context of AAL that allow monitoring elders in outdoors, using a set of sensors available in smartphones and other information sources.

We also proposed an interaction model of a hybrid context that combine traditional elements of an AAL environment with new human actors and functionalities in our system, finding to create a safety network around the city. This new social actors will act as sensors and actuators, to detect emergencies and reduce the response time in this cases.

### **3 Interaction Model**

In this section, we define our interaction model, considering several elements and actors. In the studied context, our objective is to monitor the daily routines of elders in outdoors, and acting according the diagnostic product of the mentioned status. There are different people involved in the care for the elders. Relatives and health attendants will be monitoring constantly the elder status, doctors will be alerting in case of and medical emergency and standworkers will be receiving information about the status of elders in the proximity of their workplaces and will help elders in case of emergency. It is possible to infer a diagnostic of the elder status, for example, if elder falls, his health condition or his motor activity (walking, standing, etc.).

We defined different kinds of actions that will act over the correspondent actuator, and according the previous diagnostic, we divide actions in communication actions and human help actions. Other important aspect that we realized is concerning the different information sources that our system can consume to diagnose the elder status. We taking in consideration not only the sensor observations of sensor built-in the smartphones, but also other external information (maps, environmental information located in the web, profile information of elder), and we also included ambient devices and wearable devices, but focused mainly in smartphones as main sensing tool. We consider smartphones as other sensor inside our Ambient Intelligence system.

Summarizing, relatives and health assistants will form a safety network that will be monitoring the elder status in their daily activities; also, doctors would receive alerts in case of severe emergencies. In addition, finding to reduce the risk of accidents, we include an innovative tracking system formed by stand workers who work in these newsstands that often are geographically distributed through out cities in Brazil, these

people will be aware of the elders circulating in the vicinity of their workplace, to provide quick help in case of emergency.(Figure 1).



Fig. 1. Interaction Model.

## 4 Safe Route System

In order to develop a computational tool that would resolve our interaction model and to help elders in outdoor environments was developed, a system called SafeRoute. We describe in detail the developed system: we present the architectural design that guided the development of the SafeRoute System and we presented their components.

### 4.1 Architectural Design

The SafeRoute system is a solution to the interaction model previously presented. The proposed approach consists in a system to help caregivers in the care of elders who are doing their daily activities outdoor. The system is composed by three main components: the Android application CareofMe, the SafeRoute web system and a display panel to visualize the elder general status. These components work in combination to merge data from the different sensors embedded in the smartphone as other sources of information, finding to visualize the elder state. Figure 2 presents the architecture of the SafeRoute system, composed by the CareofMe Android application and the SafeRoute web system. The display panel is not detailed in this representation since it only show the elder status provided by the SafeRoute web system.

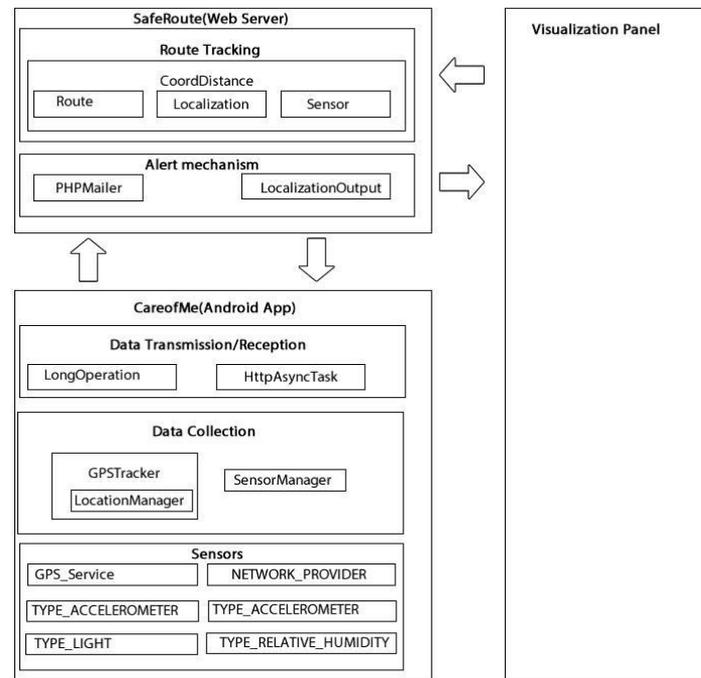


Fig. 2. SafeRoute system architecture.

## 4.2 Android Application CareofMe

The Android application CareofMe uses a combination of sensors technologies available in smartphones to identify possible risk situations for elders. Regarding the elder localization functionality, the GPS sensor integrated into smartphones provide the elder location with an error margin between 3 to 7.8 meters. However, the battery life consumption of the smartphone when this sensor is usually considered very high, for this reason, an alternative for the elders localization was considered, in this case, through Internet (Wi-Fi). GPSTracker is the class in charge for managing the elder location data through the LocationManager class belonging to the GoogleMap Android v2 API. The accelerometer sensor embedded in the smartphone was used to develop the functionality related to the elder's falls identifications. The SensorManager class is responsible to man-age the accelerations axes values, and to identify falls when abrupt changes in these values are identified.

CareofMe allows choosing a route or a safety net. Once the elder start his movement, the application allows identifying his exact geographical location (Figure 3). In the interface implementation, it was considered typical usability requirements for elders, who may have vision problems or memory problems. For this reason, the application interface has been designed trying to minimize the elder's interaction. In addition, the letters and buttons size was designed trying to facilitate its visualization.

Regarding the feedbacks at risk situations for the elders, these are some of the alarms that the application can provide:

- **Movement away of a predetermined route or safety network:** If a movement away of the allowed area is detected, an alert action will be issued to the elder (Figure 4). In addition, another alert will inform to the elder how to return to the route or secure network that he is following.
- **Fall detection:** If a fall for an elder is detected, the application will ask to the elder about their condition (Figure 5). If a help request is confirmed by the elder, the application will automatically inform to the caregiver. The application also allows to the elderly inform that he is right, finding to minimize the amount of false alarms, but if a reasonable time elapses between the fall detection and the response of the elder, an alert is issued for the caregiver.



Fig. 3. Route and safety networks monitoring in CareofMe.

### 4.3 Web SafeRoute System

The decision to develop a web platform emerged as the most natural choice; since according to the proposed model, it is necessary facilitate the access to data provided by the different sources and their subsequent fusion. The web system SafeRoute was designed as a tool to provide constant monitoring of the elder state and have the responsibility for the routes and predefined secured networks management Figure 6 shows the edition interface of a safety network.

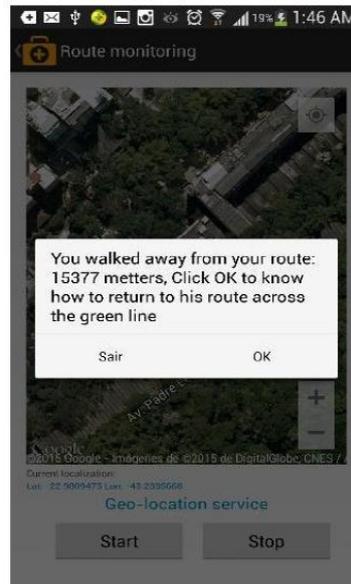


Fig. 4. Movement away feedback issued by CareofMe.

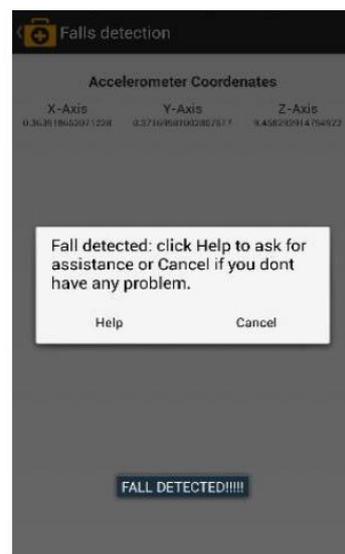


Fig. 5. Fall detection feedback by CareofMe.

In addition, the web system will process the alerts to caregivers in case of a dangerous situation inferred for the elder status. The monitoring feature uses the `CoordDistance` class to calculate the distance between the elder's location in real time (Localization) and the locations of predefined routes or network (Route). This distance is calculated by the `CoordDistance` class following the Haversiana formula between two geographical points. The proposed approach considers distance between the elder's

geographical location, the route points and the lines connecting them. A PHPMailer class is responsible for sending alerts to caregivers about the elder's situation through an e-mail with the elder's current position.



**Fig. 6.** Safety networks in the SafeRoute Web system.

#### 4.4 Control Panel

In order to create a mechanism to improve the monitoring process of the elders' activities in outdoors was created a display panel to visualize the elder status. In this way, caregivers who are monitoring the elders activity will have an overview of the elder's geographical location, as well as their status (movement away of the route or safety net, falling, etc.) and of the environmental variables.

The information presented in the panel is the result of the feedback of the SafeRoute web system, which carried out all the reasoning and inference of the elder state.

## 5 Evaluation

To evaluate our work we carried out a user-centered test with three elders. Our test consisted in that every elder followed a predefined route (Figure 7) inside our safety network. We simulate falls (in a controlled environment) and intentional departures from the safety network. Our objective with this test was to compare localization accuracy using localization sensors provided by the Arduino platform with the geographic localization obtained by sensor embedded in the smartphone and used by the CareofMe application (GPS, Wi-Fi).

Results showed in Figure 8 and Figure 9 showed a better accuracy for locations obtained by the CareofMe application in comparison with location sensors of Arduino. However, the results showed the necessity to implement calibration functionalities and



We showed that it is possible helping elders better and to reduce the time response in case of emergency in this context with the inclusion of new social actors as the stand workers, who acts like sensors and actuators in our model. With our model, we described a hybrid context, when caregivers and embedded technologies work together to create a safety network for care elders in outdoors environments and how is it possible that using our system, the response time in case of emergencies could be reduced.



**Fig. 9.** Tracked route using Arduino sensors.

In addition, we showed the results of an initial user-center test that had the objective to compare the localization accuracy between our proposed localization method and an alternative method provided by the Arduino platform. We showed that with the combination of different localization methods our system obtain a better accuracy that an alternative method that only use a simple localization sensor.

However, we considered indispensable to make a more extensive experimental phase. We aim to evaluate prototype through a technology evaluation consisting in some post-test questionnaires to measure the satisfaction level of users. We also plan to develop some new intelligent functionalities that respond automatically to user and to the environment behavior to allow a faster response in case of emergency, such as the inclusion of other available sensors and other information sources (environmental sensors, location sensor, accelerometer, camera, etc.), for example, we pretend manipulate the camera for automatically taking photos of users in case of emergency.

## References

1. Nealon, J., Moreno, A.: Applications of Software Agent Technology in the Health Care Domain. Whitstein Series in Software Agent Technologies and Autonomic Computing (2003)
2. Active Assisted Living Programme: <http://www.aal-europe.eu> (2015)
3. Roussaki, I.: Hybrid context modeling: A location-based scheme using ontologies. In: 4th Annual IEEE International Conference on Pervasive Computing and Communications Workshop (2006)

4. Fudickar, S., Schnor, B.: KopAL: A mobile orientation system for dementia patients. *Intelligent Interactive Assistance and Mobile Multimedia Computing*, 53, pp. 109–118 (2009)
5. Wan, J.: Orange alerts: Lessons from an outdoor case study. In: 5th International Conference on Pervasive Computing Technologies for Healthcare, *Pervasive Health*, pp. 446–451 (2011)
6. Alznav: <http://alznav.projects.fraunhofer.pt> (2015)
7. Ramos, J.: Geo-localization System for People with Cognitive Disabilities. In: 11th International Conference on Practical Applications of Agents and Multi-Agent Systems, pp. 59–66 (2014)
8. Alirezaie, M., Klugl, F., Lout, A.: *Knowing Without Telling: Integrating Sensing and Mapping for Creating an Artificial Companion* (2016)



Impreso en los Talleres Gráficos  
de la Dirección de Publicaciones  
del Instituto Politécnico Nacional  
Tresguerras 27, Centro Histórico, México, D.F.  
noviembre de 2016  
Printing 500 / Edición 500 ejemplares

