

# RISCE Revista Internacional de Sistemas Computacionales y Electrónicos

Marzo 2011  
Número 2, Volumen 3, Año 3



75  
*Años*  
INSTITUTO POLITÉCNICO NACIONAL  
1936-2011



**RISCE** Revista Internacional de Sistemas Computacionales y Electrónicos; es una publicación bimestral del Instituto Politécnico Nacional, Av. Luis Enrique Erro S/N, unidad “Profesional Adolfo López Mateos”, Del. Gustavo A. Madero, C.P. 07738, México D.F. a través de la Escuela Superior de Computo; Av. Juan de Dios Bátiz S/N esquina Miguel Othón de Mendizábal. “Unidad Profesional Adolfo López Mateos”. Col. Lindavista C.P. 07738, México, D. F. tel. 57296000 ext. 52000. Certificado de reserva de Derechos al uso Exclusivo del título No. 04-2008-062613190500-203, ISSN en trámite. Los artículos son responsabilidad exclusiva del autor y no reflejan necesariamente el criterio de la institución, a menos que se especifique lo contrario. Se autoriza la reproducción total o parcial, siempre y cuando se cite explícitamente la fuente.

La revista se especializa en el área de los sistemas computacionales y electrónicos; tanto en el desarrollo, como en la investigación en:

Ciencias de la Computación

Cómputo educativo

Cómputo Móvil

Comunicaciones

Disciplinas Emergentes

Electrónica

Física Electrónica

Ingeniería de Cómputo

Ingeniería de Software

Innovación Tecnológica

Inteligencia artificial

Matemática computacional

Procesamiento de señales

Robótica y cibernética

Sistemas de Información

Tecnologías de la Información

### **Distribución**

La revista cuenta con 300 ejemplares que se distribuyen en:

Europa, Asia y América Hispana; mediante CD ROM y correo electrónico

## Directorio



### INSTITUTO POLITÉCNICO NACIONAL

**DRA. YOLOXÓCHITL BUSTAMANTE DÍEZ**  
DIRECTORA GENERAL

**ING. JUAN MANUEL CANTÚ ALVAREZ**  
SECRETARIO GENERAL

**DR. EFREN PARADA ARIAS**  
SECRETARIO ACADEMICO

**DR. JAIME ALVAREZ GALLEGOS**  
SECRETARIO DE INVESTIGACIÓN Y POSGRADO

**ING. ERNESTO MERCADO ESCUTIA**  
SECRETARIO DE SERVICIOS EDUCATIVOS

**ING. OSCAR JORGE SÚCHIL VILLEGAS**  
SECRETARIO DE EXTENSIÓN E INTEGRACION SOCIAL

**M. EN C. FERNANDO ARELLANO CALDERON**  
SECRETARIO DE GESTION ESTRATEGICA

**C.P. ROBERTO ALVAREZ ARGUELLES**  
SECRETARIO DE ADMINISTRACION

**LIC. JUDITH CLAUDIA RODRIGUEZ ZUÑIGA**  
DEFENSORA DE DERECHOS POLITECNICOS



### ESCUELA SUPERIOR DE CÓMPUTO

**ING. APOLINAR FRANCISCO CRUZ LÁZARO**  
DIRECTOR

**DR. FLAVIO ARTURO SÁNCHEZ GARFIAS**  
SUBDIRECTOR ACADÉMICO

**DR. JESÚS YALJÁ MONTIEL PÉREZ**  
JEFE DE LA SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

**LIC. ARACELI LOYOLA ESPINOSA**  
SUBDIRECTORA DE SERVICIOS EDUCATIVOS E INTEGRACIÓN SOCIAL

**M. EN C. JUAN VERA ROMERO**  
SUBDIRECTOR ADMINISTRATIVO

**DR. EDUARDO BUSTOS FARIAS**  
EDITOR DE RISCE

## **Miembros del comité Revisor**

(Todo el comité técnico está formado por doctores en ciencias o su equivalente)

Francisca Losavio de Ordaz (Venezuela) (Universidad Central de Venezuela)

Alfredo Matteo (Venezuela) (Universidad Central de Venezuela)

Emmanuel F. Moya Anica (México)

Edgardo Manuel Felipe Riverón (Cuba) (México) (CIC)

Luis Enrique Palafox Maestre (México)

Eduardo F. Caicedo Bravo (Colombia)

Hilda Ángela Larrondo (Argentina)

Guillermo Leopoldo Kemper Vásquez (Perú)

Elizabeth León Guzmán (Colombia)

María Cecilia Rivera (Chile)

Satu Elisa Schaeffer (Finlandia) (UANL)

Rafael Canetti (Uruguay)

Javier Echaiz (Argentina)

Pablo Belzarena (Uruguay)

Carlos Beltrán González (Italia) (Università di Genova)

Elena Fabiola Ruiz Ledesma (México)

Jonatan Gómez (Colombia)

Armando De Giusti (Argentina)

Juan José Torres Manríquez (México)

Jesús Yaljá Montiel Pérez (México)

Luis Alfonso Villa Vargas (México)

Marco Antonio Ramírez Salinas (México)

Félix Moreno González (España) (UPM)

Salvador Godoy Calderón (México) (CIC)

José Luis López-Bonilla (México) (IPN ESIME ZAC)

Lorena Chavarría Báez (México)

Miguel Santiago Suárez Castañón (México)

## INDICE

<b>Predicción de las Distancias de Cobertura para los Perfiles de Ráfaga en Redes</b>	
<b>WiMAX.....</b>	<b>6</b>
<b>Construcción y Diseño de un Clúster Tipo HPC Virtualizado .....</b>	<b>11</b>
<b>Aplicación de Técnicas de Minería de Datos al Dominio de un Ámbito Laboral .....</b>	<b>16</b>
<b>Ambiente de Entretenimiento para Cómputo Forense, Basado en Casos .....</b>	<b>25</b>
<b>Stimulation of three Biological Systems Using Low Level Laser Radiation .....</b>	<b>30</b>
<b>Instrucciones para los autores .....</b>	<b>33</b>

# Predicción de las Distancias de Cobertura para los Perfiles de Ráfaga en Redes WiMAX

Juan Angel Alcaráz Chávez, Héctor Manuel Trejo Sánchez, José Ernesto Rojas Lima

Academia de Comunicaciones, Departamento de Ingeniería en Comunicaciones y Electrónica (ICE)  
Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME) Unidad Zacatenco – Instituto Politécnico Nacional (IPN)  
Unidad Profesional “Adolfo López Mateos”, Av. Instituto Politécnico Nacional S/N, Col. Lindavista, C.P. 07738.  
México D.F., México  
errojas@ipn.mx

**Resumen** — Este artículo presenta una predicción de las distancias de cobertura de los diferentes esquemas de modulación y codificación (perfiles de ráfaga) empleados en las redes WiMAX basadas en el estándar IEEE 802.16-d, las cuales son de suma importancia para poder establecer un mecanismo de adaptación al enlace, evaluando 4 tipos de terreno en un ambiente de una sola celda.

**Keywords**— IEEE 802.16-d; Modulación y Codificación Adaptable; Adaptación al Enlace; Perfiles de Ráfaga; WiMAX.

## I. INTRODUCCIÓN

En los últimos años, la creciente demanda de servicios de altas tasas de datos, aplicaciones multimedia, alta calidad de la información, ha generado nuevos retos para el acceso de banda ancha. Debido a esta razón las comunicaciones inalámbricas están ampliando sus campos de acción, permitiendo obtener funcionamientos comparables a las soluciones que emplean cables, pero con menores costos de infraestructura e implementación de las redes.

Bajo este argumento, surge la Interoperabilidad Mundial para el Acceso por Microondas (WiMAX: Worldwide Interoperability for Microwave Access), basada en el estándar IEEE 802.16, la cual es una tecnología de última milla que puede ser empleada como una alternativa viable y atractiva para satisfacer estas demandas. El estándar IEEE 802.16 está encauzado a las Redes Inalámbricas de Área Metropolitana (WMAN: Wireless Metropolitan Area Network), operando principalmente en dos capas del modelo OSI (OSI: Open System Interconnection), la capa Física (PHY: Physical) y la capa de Control de Acceso al Medio (MAC: Medium Access Control) [1] [2].

Las versiones de la familia de estándares IEEE 802.16 son las siguientes: La versión original del estándar IEEE 802.16 define los protocolos de la capa MAC y la capa PHY para frecuencias de 10-66 GHz operando en condiciones de Línea de Vista (LOS: Line Of Sight). La versión IEEE 802.16a, se enfocó en bandas de más baja frecuencia en el rango de 2 a 11 GHz, lo cual permitió operar en condiciones Sin Línea de Vista (NLOS: Non-Line Of Sight), en la capa PHY adicionales esquemas como el Multiplexaje por División Ortogonal de Frecuencia (OFDM: Orthogonal Frequency Division

Multiplexing) y en la capa MAC se incluye el soporte a la técnica de Acceso Múltiple por División de Frecuencias Ortogonales (OFDMA: Orthogonal Frequency Division Multiple Access) [2]. La versión IEEE 802.16-2004 denominada también como WiMAX fijo, incluye Sistemas de Antenas Avanzadas (AAS: Advanced Antenna System), utiliza Modulación y Codificación Adaptable (AMC: Adaptive Modulation and Coding). La versión IEEE 802.16e-2005, incorpora la funcionalidad que permite movilidad, portabilidad, además de soportar la transferencia automática del enlace “handover” o “handoff”. La versión IEEE 802.16f define una Base para la Gestión de Información (MIB: Management Information Base) para la capa PHY y MAC, la cual consiste en un Sistema de Gestión de Red (NMS: Network Management System) que se encarga de gestionar nodos y proporcionar un flujo de servicios para dicha base de datos [3]. La versión IEEE 802.16j, soporta una arquitectura de Relevos Móviles por Saltos-Múltiples (MMR: Mobile Multi-Hop Relay), agregando Estaciones de Relevos (RS: Relay Stations), esto nos ayuda a mejorar la QoS, extender y mejorar la capacidad de la red, de igual manera describe una técnica llamada “túneleo” para soportar la agregación de RS, para administrar o transportar el tráfico agregado [4]. El estándar IEEE 802.16m está previsto terminarse en este año 2010 y es candidato para la interface de radio del estándar para las Telecomunicaciones Móviles Internacionales- Avanzadas (IMT-Advanced: International Mobile Telecommunications-Advanced), el cual define los requerimientos establecidos por la Unión Internacional de Telecomunicaciones (ITU: International Telecommunication Union) para 4G brindando una mayor cobertura en entornos de alta interferencia, optimiza la interoperabilidad tecnologías tales como sistemas 3G, Wi-Fi y Bluetooth [5] [6]. Este trabajo está enfocado en el estándar IEEE 802.16-d, definido para el acceso inalámbrico de banda ancha fijo.

El artículo está estructurado de la siguiente manera: en la sección 2, mostramos la revisión de las técnicas de modulación y codificación en el estándar IEEE 802.16. En la sección 3, presentamos el modelo de evaluación utilizado. En la sección 4, mostramos los resultados obtenidos y finalmente en la sección 5, se presentan las conclusiones y se plantea parte del trabajo futuro.

## II. REVISIÓN DE LAS TÉCNICAS DE MODULACIÓN Y CODIFICACIÓN EN EL ESTÁNDAR IEEE 802.16

Los sistemas inalámbricos basados en el estándar IEEE 802.16, utilizan esquemas de modulación y codificación con el objetivo de transmitir altas tasas de datos cuando la calidad del canal es buena y transmitir una tasa de datos más baja cuando la calidad del canal es pobre, a lo cual se le denomina adaptación al enlace [2] [7].

A la combinación de los esquemas de modulación y codificación, se les denomina perfiles de ráfaga (Burst Profiles) [2]. En la tabla 1, observamos los diferentes esquemas de modulación y codificación soportados por WiMAX, tanto para el enlace de subida, como para enlace de bajada.

TABLA I. ESQUEMAS DE MODULACIÓN Y CODIFICACIÓN SOPORTADOS POR EL ESTÁNDAR IEEE 802.16

	Enlace de Bajada	Enlace de Subida
<b>Modulación</b>	BPSK, QPSK, 16-QAM, 64-QAM	BPSK, QPSK, 16-QAM; 64-QAM es opcional.
<b>Codificación</b>	Obligatorios: Códigos Convolucionales a $R_c = 1/2, 2/3, 3/4, 5/6$ Opcionales: Códigos convolucionales, Turbo, a $R_c = 1/2, 2/3, 3/4, 5/6$ ; Códigos de repetición a $R_c = 1/2, 1/3, 1/6$ , LDPC, Códigos Reed-Solomon para OFDM-PHY.	Obligatorios: Códigos Convolucionales a $R_c = 1/2, 2/3, 3/4, 5/6$ Opcionales: Códigos Convolucionales y Turbo a $R_c = 1/2, 2/3, 3/4, 5/6$ ; códigos de repetición a $1/2, 1/3, 1/6$ , LDPC.

En el caso de IEEE 802.16-d, se definen dentro de su capa PHY, 7 perfiles de ráfaga, mostrados en la tabla II, dichos perfiles están en función de la tasa de datos, potencia mínima requerida ó sensibilidad  $S_{min}$ , la Relación Señal a Interferencia más Ruido (SINR: Signal to Interference plus Noise Ratio) y la tasa de datos [8] [9].

TABLA II. PERFILES DE RÁFAGA EN EL ESTÁNDAR IEEE 802.16-d

Perfil de ráfaga	Modulación	Razón de codificación $R_c$	$S_{min}$ [dBm]	Tasa de datos [Mbps]
1	BPSK	1/2	-83.05	6.91
2	QPSK	1/2	-80.05	13.82
3	QPSK	3/4	-78.25	20.74
4	16-QAM	1/2	-73.05	27.65
5	16-QAM	3/4	-71.25	41.47
6	64-QAM	2/3	-66.25	55.30
7	64-QAM	3/4	-65.05	62.21

La elección de los diferentes perfiles de ráfaga de acuerdo a las condiciones del canal, tiene como finalidad proporcionar una tasa de datos promedio óptima, esto se logra utilizando técnicas de Modulación y Codificación Adaptable.

Cabe mencionar que dentro de la capa física del estándar IEEE 802.16-d, no se encuentran definidos los mecanismos de adaptación al enlace, por lo que la forma e implementación de estos, se dejan a criterio del operador del sistema [8].

## III. MODELO DE EVALUACIÓN

### A. Consideraciones Generales

Propusimos una celda circular con un radio de 1 km, en cuyo centro se localiza una antena omnidireccional, montada en una torre con una altura  $h_{tx} = 30$  m, la antena opera en condiciones NLOS, el ancho de banda es de 7 MHz y la frecuencia de operación es de 3.5 GHz [10].

### B. Modelo de propagación

Empleamos los modelos SUI (Stanford University Interim) IEEE 802.16, para evaluar de forma precisa las características de propagación en el medio inalámbrico, los cuales ha sido aprobados para frecuencias por debajo de los 11 GHz [11].

Las pérdidas debido a la trayectoria de propagación de los modelos SUI, están dadas por la siguiente ecuación [10]:

$$L_p = A + 10\gamma \log\left(\frac{d}{d_0}\right) + X_r + X_h + X \quad \text{para } d > d_0 \quad (1)$$

Donde  $d$  es la distancia entre el transmisor y el receptor en metros,  $d_0$  es una distancia de referencia igual a 100 m, el factor  $A$  representa las pérdidas en el espacio libre a la distancia  $d_0$  y  $\gamma$  es el exponente de las pérdidas por trayectoria el cual se obtiene a partir de la ecuación (2):

$$\gamma = a - bh_{tx} + c/h_{tx} \quad (2)$$

Donde  $h_{tx}$  es la altura de la antena de la estación base (BS: Base Station) sobre la superficie de la tierra, en metros ( $10 < h_{tx} < 80$ ).

Los términos  $X_r$  y  $X_h$  son factores de corrección de frecuencia y la altura de la antena sobre el suelo, definidas como:

$$X_r = 6.0 \log\left(\frac{f}{2000}\right) \quad (3)$$

$$X_h = -10.8 \log\left(\frac{h_m}{2.0}\right) \quad \text{Para terrenos tipo A y B} \quad (4)$$

$$X_h = -20.0 - \log\left(\frac{h_m}{2.0}\right) \quad \text{Para terreno tipo C} \quad (5)$$

Donde  $f$  es la frecuencia en MHz y  $h_m$  es la altura de la antena del usuario sobre el suelo, entre 2 y 10 m. El factor  $X$  representa los desvanecimientos por obscurecimientos causados por los obstáculos en la trayectoria de propagación, los cuales no serán considerados en esta evaluación.

Las constantes  $a$ ,  $b$  y  $c$  se eligen de acuerdo al tipo de terreno, los cuales se clasifican como: A, B, C, donde el terreno tipo A se refiere a un terreno montañoso con densidad de árboles grande, el terreno tipo B se considera un terreno que es plano en su mayoría, con densidad de árboles ligera, y el terreno C corresponde a un terreno principalmente plano con densidad de árboles ligera.

Asimismo para este análisis, consideramos un cuarto terreno, el cual corresponde al de la Zona Norte de la Ciudad de México, donde sus principales características se muestran en la tabla III, [12].

TABLA III. CARACTERÍSTICAS PARA LOS DIFERENTES TIPOS DE TERRENOS

Parámetros	Terreno A	Terreno B	Terreno C	Terreno Ciudad de México
a	4.6	4.0	3.6	$\gamma = 4.1667$
b	0.0075	0.0065	0.005	
c	12.6	17.1	20	

### C. Cálculo del enlace

El objetivo principal del cálculo del enlace, es determinar los niveles de potencia que deben alcanzar tanto el transmisor como el receptor para poder establecer una comunicación aceptable. En la tabla IV, se muestran los parámetros considerados en el cálculo del enlace [12].

TABLA IV. PARÁMETROS CONSIDERADOS PARA EL CÁLCULO DEL ENLACE

Parámetros	Enlace de bajada	Notas
Número de antenas transmisoras ( $N_{ATx}$ )	2	Ganancia por diversidad en la transmisión $G_{DTx} = 10\log_{10}(N_{ATx})$
Ganancia de la antena transmitora ( $G_{ATx}$ )	10 dBi	Dependerá del fabricante: Andrew Model (DB910SN-W)
Pérdidas debido a los efectos no lineales del amplificador de potencia ( $L_{BACKOFF}$ )	0 dB	Se considera un amplificador altamente lineal
Pérdidas del transmisor ( $L_{Tx}$ )	3 dB	Debido a los conectores y cables
Ancho de banda del canal (BW)	7 MHz	IEEE 802.16 soporta varios anchos de banda del canal
Nivel del ruido del receptor (N)	-154.54 dB @ 7 MHz	$-174+10\log_{10}(BW)$
Figura de ruido en el receptor ( $N_f$ )	8 dB	Depende del equipo
Ganancia por diversidad Macroscópica ( $G_{DM}$ )	0 dB	Si se utiliza por diversidad de selección
Sensibilidad del receptor ( $S_{min}$ )	-83.05 dBm	Para una modulación BPSK 1/2, Tasa de datos = 6.91 Mbps
Ganancia de la antena receptora $G_{Rx}$	0 dBi	
Pérdidas por trayectorias por penetración de construcciones	0 dB	Se asume un escenario al aire libre
Radio de la celda	1 km	Valor propuesto para el rango de cobertura
Pérdidas de propagación ( $L_p$ )	127.2249 dB	Usando los parámetros de propagación ajustado al modelo de un terreno con irregularidades.
Potencia a la salida del amplificador ( $P_{Tx}$ )	34.17 dBm (2.6138 W)	Se calcula en función del radio y otros parámetros.

La PRx en dBm a una distancia d entre el transmisor y el receptor se define como:

$$P_{Rx} = P_{Tx} + \sum G - \sum L - L_p \quad (6)$$

Donde  $P_{Tx}$  es la potencia a la salida del amplificador en dBm,  $\sum G$  es la suma de todas las ganancias en dB,  $\sum L$  es la suma de todas las pérdidas en dB, excepto las pérdidas por trayectoria y  $L_p$  son las pérdidas por la trayectoria de propagación dadas en dB a una distancia d.

En la ecuación (6), requerimos que  $P_{Rx} \geq S_{min}$  en el borde de la celda, para garantizar un buen funcionamiento del sistema. Si igualamos  $P_{Rx} = S_{min}$ , entonces la  $P_{Tx}$  se puede determinar por la ecuación:

$$P_{Tx} = S_{min} - \sum G + \sum L + L_p \quad (7)$$

La distancia máxima para cada perfil de ráfaga, se puede determinar a partir de la ecuación (7), sustituyendo  $L_p$  de la ecuación (1):

$$P_{Tx} = S_{min} - \sum G + \sum L + \left( A + 10\gamma \log_{10} \frac{d}{d_0} + X_r + X_h \right) \quad (8)$$

En la Figura 1, se representan las zonas de cobertura de los perfiles de ráfaga empleados en WiMAX.

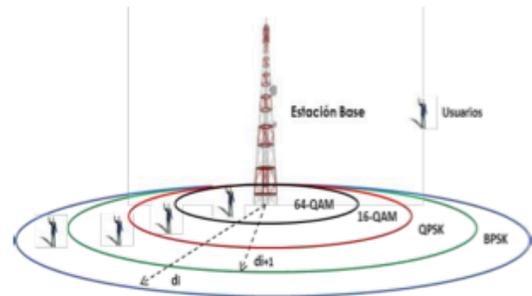


Figura 1. Distancias de cobertura de los perfiles de ráfaga en WiMAX, para una sola celda.

Para determinar las distancias de cobertura de los perfiles de ráfaga empleados en WiMAX, despejamos d de la ecuación (8), obteniendo así:

$$d_i = d_0 10^{\frac{P_{Tx} - S_{min-i} + \sum G - \sum L - A - X_r - X_h}{10\gamma}} \quad (9)$$

Donde  $S_{min-i}$  es la sensibilidad para cada perfil de ráfaga, con  $i = 1, 2, 3, 4, 5, 6, 7$ .

## IV. RESULTADOS

De la Fig. 2, inicialmente comprobamos que las pérdidas por trayectoria aumentan en relación a la distancia y a las características del terreno, en base a lo anterior notamos que el terreno tipo A presenta mayores pérdidas debido a su alta irregularidad y densidad de obstáculos, mientras que el terreno tipo C muestra menores pérdidas, ya que es un terreno más plano y con una menor densidad de obstáculos.

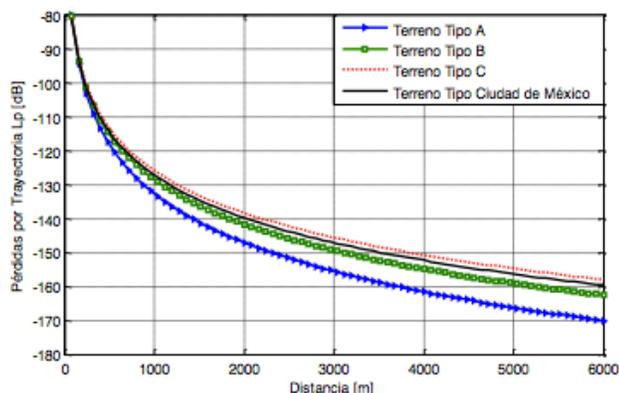


Figura 2. Pérdidas por trayectoria para los 4 tipos de terreno, basados en el modelo de propagación SUI IEEE 802.16.

En la tabla V, se muestran los resultados numéricos de la  $P_{Tx}$  requerida en los 4 tipos de terreno fijando el radio de la celda a 1 km, sin considerar el margen de desvanecimientos. A partir de los resultados apreciamos que en el terreno tipo A, se requiere una mayor  $P_{Tx}$  para poder cubrir un radio de 1 km, esto es provocado por sus grandes pérdidas por trayectoria mostradas, mientras que el terreno tipo C requiere una menor  $P_{Tx}$  para cubrir el mismo radio de 1 km, debido a que las pérdidas por trayectoria presentadas, son menores que las del terreno Tipo A.

TABLA V. VALORES DE LA POTENCIA DE TRANSMISIÓN REQUERIDA PARA CADA TIPO DE TERRENO @ 1 KM

Potencia de Transmisión requerida	Tipo de Terreno			
	A	B	C	Ciudad de México
PTx en [dBm]	39.6711	35.4711	32.8877	34.1727
PTx en [W]	9.2706	3.5246	1.9443	2.6138

A partir de la ecuación (11), determinamos las distancias de cobertura alcanzadas para cada perfil de ráfaga en el terreno tipo Ciudad de México, donde analizando la Fig. 3, notamos que el perfil de ráfaga 64-QAM 3/4 cubre una mayor distancia de todos los perfiles de ráfaga contendientes, posteriormente en este orden de cobertura, tenemos a los perfiles QPSK 3/4, BPSK 1/2, 16-QAM 3/4, QPSK 1/2, 16-QAM 1/2 y finalmente el perfil 64-QAM 2/3 presenta la menor distancia de cobertura.

En la tabla VI, se muestran valores numéricos de las distancias de cobertura alcanzadas por cada perfil de ráfaga en los terrenos tipo A y Ciudad de México, donde observamos una gran similitud de los resultados obtenidos, teniendo una variación promedio de 3.0 m, entre ambos terrenos. Realizando una comparación con la Fig.3, notamos que hay una correspondencia entre los resultados obtenidos.

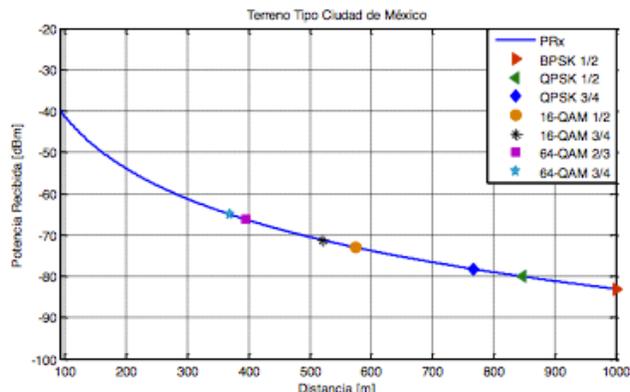


Figura 3. Potencia recibida y perfiles de ráfaga para el terreno tipo Ciudad de México.

TABLA VI. DISTANCIAS ENTRE LOS PERFILES DE RÁFAGA PARA LOS TERRENOS TIPO A Y CIUDAD DE MÉXICO

Perfil de Ráfaga	Distancia entre perfiles de ráfaga para el terreno tipo A [m]	Distancia entre perfiles de ráfaga para el terreno tipo Ciudad de México [m]
BPSK 1/2	865.8360 - 1000	847.2299 - 1000
QPSK 1/2	794.1390 - 865.8360	767.0111 - 847.2299
QPSK 3/4	618.6574 - 794.1390	575.4434 - 767.0111
16-QAM 1/2	567.4285 - 618.6574	520.9583 - 575.4434
16-QAM 3/4	446.3092 - 567.4285	395.1882 - 520.9583
64-QAM 2/3	421.3178 - 446.3092	369.8317 - 395.1882
64-QAM 3/4	100 - 421.3178	92 - 369.8317

## V. CONCLUSIONES

En este trabajo determinamos las distancias de cobertura alcanzadas por los perfiles de ráfaga utilizados en las redes inalámbricas WiMAX basadas en el estándar IEEE 802.16-d. Empleando los modelos de propagación SUI IEEE 802.16, determinamos las pérdidas por trayectoria en los 4 terrenos evaluados, tomando como punto de comparación, el resultado más extremo y poder así establecer una comparación con el terreno tipo Ciudad de México. Con las ecuaciones de la  $P_{Tx}$  y  $P_{Rx}$  comprobamos que el terreno tipo A, requiere una mayor  $P_{Tx}$ , debido a que presenta mayores pérdidas por trayectoria. Mediante el uso de los valores obtenidos del cálculo del enlace y las ecuaciones de  $P_{Tx}$  procedimos a realizar la estimación de las distancias de cobertura de cada perfil de ráfaga, para posteriormente establecer un mecanismo de adaptación al enlace. En base a lo anterior cabe mencionar que se tiene contemplado como trabajo futuro, realizar la evaluación de algoritmos de adaptación al enlace, mediante un modelo de simulación.

## VI. REFERENCIAS

- [1] Georgios S. Paschos, Petteri Mannersalo, and Thomas Michael Bohnert, "Cell Capacity for IEEE 802.16 Coverage Extension", IEEE Communications Society CCNC 2008.
- [2] Jeffrey G. Andrews, Arunabha Ghosh, Rias Muhamed, Fundamentals of WiMAX Understanding Broadband Wireless Networking, Prentice Hall, Massachusetts United States, February 2007.
- [3] Loutfi Nuaymi, WiMAX Technology for Broadband Wireless Access, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2007.
- [4] Kwang-Cheng Chen, J. Roberto B. de Marca, Mobile WiMAX, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2008.
- [5] Kamran Etemad, Intel Corporation, "Overview of Mobile WiMax Technology and Evolution", IEEE Communications Magazine, Vol. 0163-6804/08, October 2008.
- [6] Qinghua Li and Xintian Eddie Lin, Jianzhong (Charlie) Zhang and Wonil Roh, "Advancement of MIMO Technology in WiMAX: From IEEE 802.16d/e/j to 802.16m", IEEE Communications Magazine, Vol. 0163-6804/09, June 2009.
- [7] Chadi Tarhini, Tijani Chahed, "On capacity of OFDMA-based IEEE 802.16 WiMAX including Adaptive Modulation and Coding (AMC) and inter-cell interference".
- [8] Part 16: Air Interface for Fixed Broadband Wireless Access Systems, IEEE Std 802.16-2004, LAN/MAN Standards Committee of the IEEE Computer Society and IEEE Microwave Theory and Techniques Society, October 2004.
- [9] Ralf Pabst, Jan Ellenbeck Marc Schinnenburg, Christian Hoymann, "System Level Performance of Cellular WiMAX IEEE 802.16 with SDMA-enhanced Medium Access", IEEE WCNC 2007.
- [10] Harry r. Anderson, Fixed Broadband Wireless System Design, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2003.
- [11] Marcos D. Katz, Frank H.P. Fitzek, WiMAX Evolution: Emerging Technologies and Applications, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2009.
- [12] José Ernesto Rojas Lima, Jaime Pedro Abarca Reyna, Ricardo Morales Ballesteros, " Cell Coverage Issues On WiMAX Networks For Fixed Broadband Wireless Access", 2nd International Conference on Wireless Communications in Underground and Confined Areas, Val-d'Or Québec Canada, August 25 to 27-2008.

# CONSTRUCCIÓN Y DISEÑO DE UN CLÚSTER TIPO HPC VIRTUALIZADO

Juan Alberto Antonio Velazquez,  
IPN-CIDETEC, México D.F.  
juanjonatan@yahoo.com.mx

Jesús Antonio Alvarez Cedillo,  
Juan Carlos Herrera Lozada  
IPN-CIDETEC, México D.F.  
{jaalvarez, jlozada}@ipn.mx

**Abstract—** Virtualization has an important role in computer security, and most companies use it to protect the information using virtual systems that communicate with physical systems. Companies also see advantages as saving space and lower energy use which in turn is included in saving a lot of money. In this work we define the different parallelization techniques that exist and how a cluster works Mosix in implementing programs.

You can define virtualization techniques used to handle different operating systems. Virtualizer software used was Virtual-box can be installed on different platforms and accept the installation of several operating systems including to Opensuse.

**Keywords:** *Virtualization, Mosix, clusters, Virtualbox*

## Introducción

Actualmente, en lo que se refiere a la capacidad de procesamiento, existen varias alternativas en el mercado, pero muchas de esas alternativas resultan en costos muy altos, tal es el caso de los mainframes y servidores. El utilizar hoy en día tecnologías muy caras resulta para las pequeñas empresas, tecnológicos, universidades y cualquier entidad que no tenga recursos, muy difícil el acceso a este tipo de tecnologías en México. En lo que se refiere a las universidades públicas y tecnológicas centralizadas y descentralizadas, la investigación en las áreas de matemáticas, químicas, médicas, etc. han creado una rama que necesita del supercómputo para su desarrollo y esta rama es la Bioinformática, la cual se encarga de estudiar las ciencias antes mencionadas.

Para el estudio o la investigación de la Bioinformática [1], es necesario tener sistemas computacionales (con múltiples procesadores), capaces de poder resolver problemas en tiempos muy cortos, algo que una computadora simple no puede realizar. Hay muchas tecnologías que actualmente trabajan en el ámbito del procesamiento a gran escala, muchas de éstas incluyen arreglos de discos, procesadores, utilizan memoria a gran escala lo cual ayudan a realizar operaciones en tiempos relativamente cortos. Una de estas tecnologías son los servidores y mainframes, que algunas empresas como: IBM, Sun Microsystems, Hewlett Packard y DELL entre otros trabajan en el área del procesamiento a gran escala. Pero esta tecnología está al alcance de medianas y grandes empresas, las cuales tienden a utilizar este tipo de tecnologías por su alto

costo. Actualmente existe otra alternativa para las pequeñas empresas, universidades y departamentos de investigación: los clusters [4].

## I. CONCEPTOS BASICOS.

En la actualidad, las empresas empiezan a reconocer la importancia de la virtualización, sobre todo por el valor que toma en las mismas para solventar esquemas de seguridad, espacio y sobre todo el ahorro de energía eléctrica y de dinero.

Existe poca información de clusters implementados virtualmente. En la universidad de Barcelona se ha trabajado con el Grid Cloud con ayuda de sistemas virtualizadores como Xen y el sistema para crear Cloud como Open Nébula [3]. En lo referente a Mosix y su trabajo en clusters virtuales, los creadores de Mosix, demuestran que por medio de máquinas virtuales se pueda crear un Grid de interconexión de clusters virtuales, cada uno trabajando en un solo equipo y administrado por dos Workstation [4].

### I.1 Cluster Virtual

Al igual que el cluster físico la creación del cluster virtual implica tener 4 nodos conectados por medio de hardware virtual, como tarjetas de red virtual, procesadores, discos duros, cantidad de memoria, USB, etc.

En Virtual Box se puede escoger la marca de tarjeta de red y también el tipo de conexión interna para comunicar los hosts entre sí.

Algunas marcas que maneja la máquina virtual de Virtual Box [5] son: Intel PRO/1000 MT Desktop, PC net-Fast III, entre otras (ver Figura 4.2).

Aquí no existe un dispositivo central llamado switch, sino que internamente la conexión virtualizada toma el lugar de la red física, pero sin dispositivos ni cables de por medio.

A su vez los discos duros y las memorias virtuales toman un papel importante antes de instalar un sistema operativo, ya que estos dispositivos pueden tomar el tamaño de almacenamiento requerido para las funciones del sistema

operativo. Por ejemplo, en lo referente a la memoria RAM, podemos tomar diferentes tamaños de memoria que pueden ser desde los 4Mb hasta los 2048 Mb. Los discos duros se pueden configurar en la máquina virtual y su capacidad puede ser modificada por el Administrador del sistema virtual, desde 4 Mb hasta los 2 Tb de espacio.

Ya teniendo los nodos configurados con el tipo de red, se elige el sistema operativo Opensuse 11.1 y se escoge el disco duro virtual, que anteriormente se había configurado su tamaño. La instalación y configuración del sistema operativo Opensuse, se puede hacer en un disco montado en la unidad física de la computadora o bien mediante una imagen de disco con extensión \*.nrg o \*.iso.

## I. 2 Diferencias entre el cluster virtual y cluster físico

Las diferencias entre el cluster virtual y el cluster físico son notables a simple vista, pero en lo referente a las velocidades del procesador, como a las capacidades del disco duro y la memoria RAM dependen en las máquinas físicas de su plataforma, que pueden ser de 32 y 64 bits. La máquina virtual depende de estas plataformas para su funcionamiento. La ventaja de las plataformas es que cada máquina virtual puede acoplarse a sus características físicas, haciendo que una máquina virtual de 32 bits pueda funcionar en una máquina física de 64 bits.

Los clusters físicos cuentan con dispositivos que se encuentran físicamente instalados en una placa base y sus características imponen una velocidad de procesamiento y de acceso a la memoria RAM, además de que el disco duro puede almacenar la información que se requiera a una velocidad de lectura/escritura que depende de la placa base. Todos estos dispositivos de hardware trabajan conjuntamente con el microprocesador, con lo cual al trabajar a la velocidad de procesamiento, todos los dispositivos son compatibles con el procesador y la placa madre si es un sistema a 64 bits.

## II. METODOLOGIA.

El presente trabajo se basó en la investigación documental, estableciendo la base teórica del problema, investigando y documentando las técnicas de Virtualización, así como los diferentes tipos de clusters junto a la programación paralela para la solución de problemas.

El virtualizador a utilizar en el proyecto es Virtual Box y se van a crear 4 máquinas virtuales llamados nodo0, nodo1, nodo2 y nodo3; configurándose la red interna para la comunicación entre ellos, además de almacenamiento y acceso remoto. De igual forma se instaló el Mosix, para la conexión entre los nodos del cluster.

El trabajo consiste en realizar un cluster físico. El cual cuenta con cuatro equipos de cómputo conectados a un medio de comunicación físico por medio de un switch compartiendo recursos. Estos equipos de cómputo que se facilitan, tienen un procesador Pentium 4 a 1.5 GHz, con un disco duro de 80 Gb

de almacenamiento. Cada equipo de cómputo cuenta con tarjetas de red que se instalan en el bus PCI. Una de las tarjetas trabaja a una velocidad de 10/100 Mbps; la otra tarjeta trabaja a una velocidad de 1 Gbps y es la que se ocupa para la transmisión de la información.

Uno de los cuatro nodos (ver Figura 1), fue elegido para ser el nodo maestro. Este nodo maestro trabaja para la comunicación con los demás nodos con tarjetas de red veloces. La tarjetas de red que utiliza, una es de 10/100 Mbps y se conecta a internet en la red de la institución y la otra tarjeta de red, es de 10/100/1000 Mbps y es la que conecta a los nodos esclavos; los nodos esclavos se conectan por medio de la tarjeta de 10/100/1000 Mbps, con lo cual hace que la red entre los nodos sea eficiente y rápida.

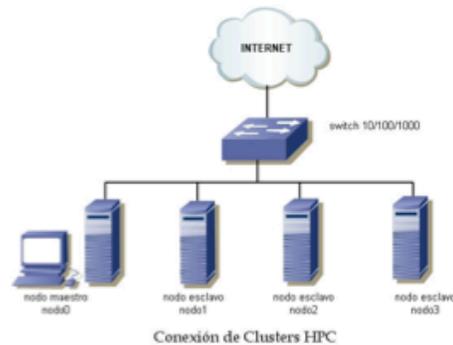


Figura 1 Esquema del cluster con 3 nodos y un nodo maestro.

El medio de comunicación es cable UTP categoría 5e, que conecta los nodos al switch. Se ha elegido el sistema operativo Linux [6], en su distribución Opensuse en la versión 11.1 ya que el parche Mosix es compatible con esta distribución. Esta instalación se hizo para cada uno de los nodos, los cuales se configuran en red dentro del sistema, asignando las direcciones de red. El cluster virtual es configurado de la misma manera que el cluster físico, pero la diferencia de éste con el sistema físico, es que depende del hardware virtual con el que se cuenta dentro de la máquina virtual, la cual es Virtual Box en su versión 3.0.2 de la empresa Sun Microsystems. Igual que en los nodos físicos, en la máquina virtual se instalaron y configuraron cuatro nodos virtuales con el sistema operativo Linux en su distribución Opensuse 11.1 [7], configurada en una red interna en la cual los nodos son capaces de comunicarse vía red compartiendo recursos de software y hardware.

En el cluster físico se han implementado 4 computadoras con características similares y estas son:

- 1.- Cuatro nodos interconectados con un enlace con un SWITCH Enterasys a 1000Mbps.
- 2.- Procesadores CELERON CORE DÚO, 2 Gb de memoria RAM, ya cada uno con un disco duro de 160 Gb.
- 3.- Sistema operativo Linux Opensuse versión 11.1 de Novel.

4.- El nodo maestro (nodo0), cuenta con 2 tarjetas de red eth0 y eth1, en donde eth1 tiene salida a internet por medio del numero IP 148.204.67.192. Este nodo incluye los servicios de servidor web Apache2 y Openssh para poder acceder al cluster desde internet.

5.- El middleware de conexión a nivel de kernel MOSIX [8].

### III. RESULTADOS

Para desarrollar las pruebas, fue necesario hacer pruebas de balanceo y de carga en cada uno de los nodos conectados en red y que serían ejecutados por Mosix. Para esto se buscaron los algoritmos que por su complejidad matemática responden a un tiempo de retardo, estos algoritmos son los algoritmos de factorial, un algoritmo con 3 bucles anidados (Stress), algoritmo que calcula el Pi por el método de Leibniz. Se realizan en cada uno de los algoritmos las pruebas que se realizaron fueron en base a muestras numéricas que en promedio de diez corridas por muestra, arrojaron diferentes tiempos en cada uno de los equipos en los que fueron ejecutados como: en una computadora (secuencial), el cluster físico y en el cluster virtual. Además de su análisis de Speedup, para ver la ganancia en tiempos comparados entre los tiempos secuenciales y los clusters físicos y virtuales.

Para el algoritmo factorial, las muestras que se tomaron en cuenta (Tabla 1), para desarrollar el algoritmo de factorial, fueron de acuerdo al alcance de los tipos de datos que el lenguaje c puede alcanzar.

Tabla 1 Muestras, tiempos calculados secuencialmente, en cluster físico y cluster virtual con el algoritmo del Factorial.

CALCULOS DE TIEMPOS EN EL ALGORITMO DE FACTORIAL					
Muestras	T. secuencial	T. Cluster físico	T. Cluster virtual	T. Speedup físico	T. Speedup virtual
5	0.002	0.001	0.001	1.00	1.00
8	0.002	0.001	0.001	1.00	1.00
20	0.003	0.001	0.002	2.00	0.50
27	0.003	0.002	0.002	0.50	0.50
33	0.003	0.002	0.002	0.50	0.50
40	0.003	0.002	0.003	0.50	0.00
50	0.003	0.002	0.003	0.50	0.00
58	0.004	0.003	0.003	0.34	0.34
62	0.004	0.003	0.003	0.34	0.34
70	0.004	0.003	0.003	0.34	0.34
89	0.01	0.003	0.004	2.34	1.51
105	0.03	0.003	0.004	9.03	6.53
115	0.031	0.004	0.005	6.78	5.23
145	0.038	0.005	0.006	6.64	5.37
163	0.043	0.011	0.014	2.95	2.11

Utilizando las muestras del 5 al 163, se pueden notar los resultados de tiempos que secuencialmente se desarrollaron en un solo nodo, sin ninguna conexión o software que lo paralelizara, se puede ver mediante la gráfica su desempeño (Figura 2).



Figura 2 Gráfica de comparación entre tiempos secuencial, cluster físico y el speedup aplicado al algoritmo de Factorial.

El algoritmo de Stress, es un algoritmo que dada su complejidad computacional  $O(n^3)$ , al ejecutarlo en un solo nodo puede tener un retardo de ejecución, debido a sus tres ciclos anidados que utiliza este algoritmo, al utilizar una muestra cada vez mayor, su complejidad computacional aumenta considerablemente.

Las muestras se escogieron de un valor de 25 a 100 (Tabla 2), debido a que en el algoritmo mientras el valor de n sea más alto más sobrecarga a la computadora, más el análisis para escoger entre ejecución secuencial o paralela.

Tabla 2 Muestras, tiempos calculados secuencialmente en cluster físico y cluster virtual con el algoritmo de Stress.

CALCULOS DE TIEMPOS EN EL ALGORITMO DE STRESS					
Muestras	T. Secuencial	T. Cluster físico	T. Cluster virtual	T. Speedup Físico	T. Speedup Virtual
25	0.004	0.002	0.002	1.004	1.004
31	0.014	0.003	0.003	3.680666667	3.680666667
39	0.039	0.003	0.004	8.04	8.789
46	0.078	0.003	0.004	15.07	15.876
50	0.08	0.007	0.007	9.023	9.678
52	0.098	0.007	0.008	10.45	11.348
68	0.185	0.01	0.013	13.54	13.41576923
80	0.245	0.162	0.171	0.757345679	0.677748538
91	0.337	0.175	0.182	1.262714286	1.188648352
100	0.396	0.293	0.31	0.747535836	0.673419355

El speedup que se muestra en esta comparación de tiempos puede verse en la gráfica (Figura 3), el tiempo obtenido con la fórmula del speedup nos arroja los tiempos de las mejoras del análisis paralelo, contrario al análisis secuencial que arroja el algoritmo de Stress.

En la tabla 3, se muestran los tiempos que dieron como resultado de la ejecución del algoritmo de PI de Gregory-Leibniz, ya sea secuencialmente, en el cluster físico, cluster virtual y los diferentes tiempos de la fórmula de Speedup.

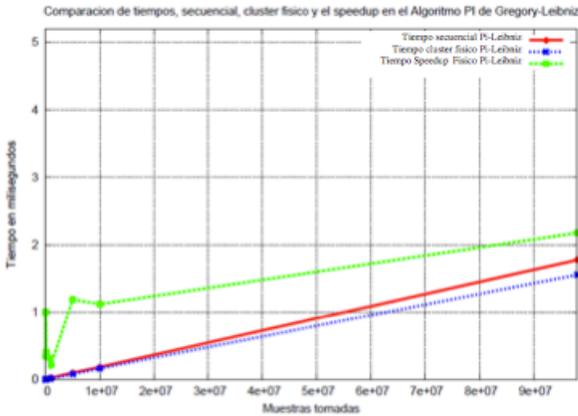


Figura 3. Gráfica de comparación entre tiempos secuencial, cluster virtual y el speedup aplicado al algoritmo de Stress.

Tabla 3 Muestras, tiempos calculados secuencialmente en cluster físico y cluster virtual con el algoritmo PI Gregory-Leibniz.

CALCULOS DE TIEMPOS EN EL ALGORITMO DE PI GREGORY-LEIBNIZ					
Muestras	T. Secuencial	T. Cluster físico	T. Cluster virtual	T. Speedup Físico	T. Speedup Virtual
1500	0.002	0.001	0.001	1.002	1.002
5800	0.002	0.001	0.002	1.002	0.002
7000	0.002	0.002	0.002	0.002	0.002
8500	0.004	0.002	0.002	1.004	1.004
22000	0.004	0.002	0.003	1.004	0.337333333
91000	0.004	0.003	0.003	0.337333333	0.337333333
147000	0.004	0.003	0.004	0.337333333	0.004
789000	0.016	0.013	0.015	0.246769231	0.082666667
1000000	0.025	0.016	0.017	0.5875	0.495588235
5000000	0.1	0.08	0.088	0.35	0.236363636
10000000	0.185	0.164	0.17	0.31804878	0.273235294
98000000	1.775	1.555	1.675	1.9164791	1.834701493

La fórmula del speedup aplicable para comparar los tiempos entre el cluster físico y el nodo individual que calcula secuencialmente el cluster virtualizado. Al aplicar esta fórmula se puede ver (Figura 4) que la mejora resulta suficiente para escoger al cluster paralelo sobre la ejecución de programas en el nodo secuencial.

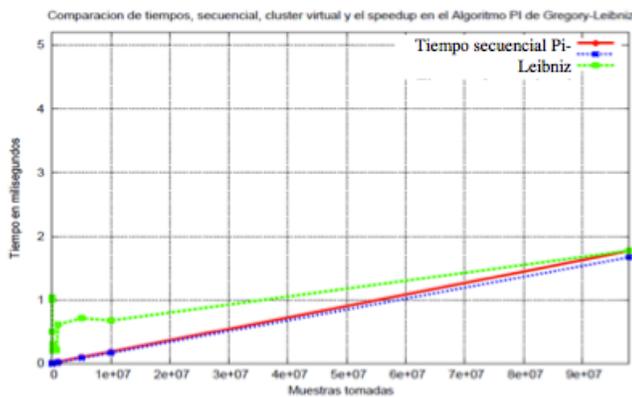


Figura 4 Gráfica de comparación entre tiempos secuencial, cluster virtual y el speedup aplicado al algoritmo de Pi de Gregory-Leibniz.

## I. CONCLUSIONES

Se puede ver tanto en el cluster virtual como en el cluster físico de una buena conexión en red de alta velocidad y de una

buena capacidad de memoria RAM para trabajar óptimamente. Las pruebas de desempeño que se realizaron en cada uno de los cluster nos arrojaron resultados favorables en lo relacionado al tiempo de ejecución. En el cluster físico en comparación al cluster virtual, el desempeño obtenido es ligeramente mayor al que se realizó en el cluster virtual, debido al consumo de memoria de todas las máquinas virtuales, que en la realidad el tamaño de memoria no se compara al de las máquinas físicas. Cabe mencionar que las empresas que utilizan máquinas virtuales, utilizan servidores como anfitriones con una buena capacidad de memoria y un procesador potente para realizar las operaciones.

En la prueba de desempeño se observó el comportamiento de cada uno de los algoritmos ejecutados tanto física como paralelamente. El algoritmo que se ejecuto más lento fue el recursivo, como: el cálculo del factorial, los cuales en el estudio de su complejidad computacional el resultado dado es  $O(n!)$ , con lo cual nos demuestra el retardo en la ejecución. En todos los casos en los que se hicieron las pruebas de desempeño tanto física como virtualmente, el algoritmo con un mejor desempeño fue el algoritmo por el método de Gregory-Leibniz.

El desempeño virtual en algunas de las ejecuciones obtuvo un tiempo elevado que no se esperaba, debido a que la ejecución tardó más de lo que se hace en el cluster físico.

Mosix no es solamente un programa que forma parte de un sistema operativo, también es un programa que trabaja silenciosamente y que cuando se está ejecutando algún programa en paralelo la carga de trabajo lo migra de nodo en nodo buscando un nodo disponible para su ejecución. Tanto en el cluster físico como en el cluster virtual se pueden añadir más nodos, con esto se puede hacer más potente el trabajo del cluster, sin embargo la desventaja de Mosix es que trabaja con licencias. En este trabajo se utilizó una licencia gratuita para conectar hasta 6 nodos, con lo cual hace que el desempeño en algunas ocasiones no sea el esperado. Si se quiere expandir a más de 6 nodos (lo cual es recomendable), se tiene que comprar una licencia según el número de nodos que se quieren anexar al cluster. Con esto el desempeño de un trabajo ejecutado en paralelo se espera sea eficiente y en menor tiempo.

De la elaboración del proyecto en donde se ha creado el cluster virtual, se concluye que cumple con las expectativas esperadas, como lo es la construcción del mismo y que al ser comparado con el cluster físico, los tiempos de ejecución fueron óptimos. En comparación con otros clusters que se han realizado con Mosix de forma física [26], se puede decir que el cluster virtual realiza las operaciones de balanceo y alto desempeño en forma óptima. En lo relacionado a la virtualización, no se encontraron trabajos acerca de clusters virtuales, debido a que aún se están desarrollando aplicaciones y sus beneficios [9].

## II. REFERENCIAS

[1] Lahoz-Beltrá Rafael., Bioinformática: simulación, vida artificial e inteligencia artificial., Edit. Ediciones Díaz de Santos., año 2004., 574 pp.

[4] D. Sloan, Joseph., High performance Linux clusters with OSCAR, Rocks, openMosix, and MPI., 2a. ed. Edit. O'Reilly Media, Inc., 2005 - 350 pp.

[3] Pages Montanera Enric, Gestion sostenible de clústers de recursos virtuales, memoria de proyecto de ing, 83 págs, Universidad Autonoma de Barcelona, 2009.

[4] Barak Amnon, The Mosix Organizational Grid A white Papper, Department of Computer Science The Hebrew University of Jerusalem, Israel, August 2005.

[5] Becker Dirk., VirtualBox: Installation, Anwendung, Praxis., Edit. Galileo Press., año 2009., 321 pp.

[6] Bookman Charles., Building And Maintaining Linux Clusters., Edit. Sams Publishing, año 2003, 265 pp.

[7] M. Surhone Lambert, T. Timpledon Miriam, F. Marseken Susan., Opensuse: Operating System, Linux Kernel, OpenSUSE Project, Novell, SUSE Linux Distributions., Edit. Betascript Publishers., año 2009., 108 pp.

[8] Barak A., The evolution of the MOSIX Multi-computer UNIX System., Edit. Hebrew University of Jerusalem. Dept. of Computer Science., Univ., 1989.

[9] S. Weygant, Peter., Clusters for high availability, 2a. ed., Edit. Hewlett-packard Profesional., 2003, 345pp.

# Aplicación de Técnicas de Minería de Datos al Dominio de un Ámbito Laboral

Fernando Vázquez Torres  
Instituto Politécnico Nacional  
UPIICSA – SEPI- Informática  
Distrito Federal, México  
[fvazquez\\_t@hotmail.com](mailto:fvazquez_t@hotmail.com)

Pilar Gómez Miranda  
Instituto Politécnico Nacional  
UPIICSA – Depto. Computación  
Distrito Federal, México  
[pgomez84@hotmail.com](mailto:pgomez84@hotmail.com)

**Resumen**—El objetivo es presentar la aplicación de técnicas de minería de datos en el descubrimiento de conocimiento en medidas seriales como una forma fácil y óptima para analizar medidas seriadas muy cortas y repetidas con factor de bloqueo en una aplicación al dominio del ámbito laboral, donde se obtiene conocimiento nuevo al aplicar esta metodología híbrida (técnicas de minería de datos: estadística e inteligencia artificial), a partir de la cual obtenemos una posible solución al problema de conocimiento, sobre todo cuando aparentemente no hay atributos relevantes

**Keywords**—descubrimiento de conocimiento; dominiolaboral, minería de datos.

## INTRODUCCIÓN

Muy a menudo, es posible que muchos datos obtenidos, mientras se supervisa un cierto proceso, como resultado de medidas seriadas muy cortas y repetidas durante el tiempo que dura un proceso. Parece que las características de estas medidas no constituyen un serio problema cuando se analizan con técnicas de series de tiempo clásicas, sin embargo, surge la siguiente pregunta: ¿Qué pasará cuando el número de medidas es muy pequeñas?

Con frecuencia, en una situación como esta, mucha información adicional está disponible de individuos y eventos involucrados en el proceso. Sin embargo, esta información podría no ser medidas seriadas y repetidas, muy estrechamente relacionada con lo que se lleva a cabo en el proceso. En efecto, los individuos a menudo constituyen un factor de bloqueo que actúan sobre las medidas seriadas, lo que sugiere las siguientes preguntas: ¿Cómo es posible aprovechar la ventaja de esta información adicional? ¿Cómo podría esta información adicional ser

manejada en relación con las medidas repetidas de serie, si dicha información no se obtienen mediante la medición de las características de los individuos que participan en el proceso? Para responder a estas preguntas, proponemos una metodología híbrida basada en técnicas de minería de datos. Dicha metodología permite el descubrimiento de conocimiento en los ámbitos donde las medidas seriadas muy cortas y repetidas presentan un factor de bloque –compuesto de individuos– y donde se define el factor de bloque [1] al factor (objeto o variable) que tiene un efecto sobre la respuesta, que aunque no es directamente de interés, se debe considerar en el experimento para obtener comparaciones homogéneas en los grupos de observaciones donde dicho factor se mantiene constante. Para evitar la influencia del factor de bloque que conforma el individuo sobre las medidas seriadas, se determinó realizar la diferencia entre las medidas seriadas del evento actual y las medidas seriadas del evento anterior (o viceversa). Así, se puede medir el efecto *per se* de un evento dado sobre la variable de interés, independientemente de las características del individuo. Estos datos, sólo toman en cuenta el incremento o decremento de la variable Y debido a la ocurrencia del evento, comparando lo sucedido antes y después de la ocurrencia del mismo.

En el caso de estudio, realizaremos la aplicación de la metodología propuesta al análisis de medidas seriadas muy cortas y repetidas con factor de bloque, presentes en un Dominio poco Estructurado (DPE) del dominio

del Ámbito Laboral para obtener conocimiento de la evolución del Programa de Becas de Capacitación para Trabajadores Desempleados (PROBECAT), así como de su efectividad a través del análisis de la información relacionada con: la colocación de la gente capacitada, características de los municipios donde se dio la capacitación y las necesidades del sector productivo. Los resultados obtenidos en este caso de estudio son una aplicación de esta metodología “Caracterización e Interpretación Automática de Descripciones Conceptuales” (CIADEC) [2] y la metodología para el “Descubrimiento de Conocimiento en Medidas Seriadadas Cortas y Repetidas” (KDSM) [3] que permitieron conocer el efecto que cada curso impartido aporta al municipio donde actúa el PROBECAT y se pueda incidir en el funcionamiento del programa mismo de manera más oportuna y eficaz. Además, muestran la resolución del problema que surge cuando no se obtienen variables caracterizadoras—de los municipios—al aplicar la metodología KDSM, pues con las técnicas hasta ese momento usadas no se obtenía ninguna descripción significativa de los municipios. Por lo que en esta parte damos a conocer la propuesta de solución y los resultados obtenidos.

### *Metodología CIADEC*

El sistema CIADEC implementa el modelo “Caracterización e Interpretación Automática de Descripciones Conceptuales en Dominios poco Estructurados usando variables cuantitativas” es un sistema híbrido (inteligencia artificial, estadística y lógica difusa) que surge de la necesidad de automatizar la caracterización e interpretación de clases en dominios poco estructurados [4-5] previamente particionados. Mediante la automatización de la metodología formal, denominada “Generación Automática de Reglas Difusas en Dominios poco Estructurados con variables cuantitativas [2]”, se pretende reducir el tiempo para la caracterización e interpretación de descripciones conceptuales usando variables cuantitativas, dando agilidad tanto a las actividades asociadas al análisis de

datos como a la obtención de información relevante que posteriormente sea útil en la gestión y toma de decisiones en esa clase de dominios. Además, la automatización de esta metodología ofrece un conjunto de funcionalidades que permiten:

- Construir un sistema de reglas.
- Visualizar funciones de pertenencia de una variable; **Error! No se pueden crear objetos modificando códigos de campo.** a las distintas clases.
- Caracterizar las diferentes situaciones que se presentan en una clasificación previa validada por el experto.
- Evaluar un conjunto de objetos nuevos de acuerdo a las reglas generadas.
- Interpretar las distintas situaciones que un individuo tiene respecto a las variables relevantes que lo definen.
- Validar la calidad de las clases resultantes respecto a su utilidad.

*En resumen, CIADEC permite determinar qué variables caracterizan, estadísticamente hablando, a las clases. Es fundamental la contribución de CIADEC, ya que mejora la actuación del experto, ahorrando su tiempo de análisis, pues sólo trabaja variables relevantes y caracterizadores [4, 5].*

### *Metodología KDSM*

La utilización de la informática para monitorizar un proceso, por ejemplo, el seguimiento a un paciente que esta sometido a una terapia; ofrece una gran cantidad de información tanto del proceso como de los actores (individuos y los eventos que intervienen sobre los individuos) en el proceso. Con frecuencia encontramos que muchos datos, que provienen de monitorizar un proceso, son resultado de medidas seriadas en el tiempo de duración de dicho proceso. Además, éste proceso suele repetirse tantas veces sea necesario para lograr el objetivo esperado.

Aparentemente, estas dos características de las medidas seriadas y repetidas en el tiempo, no ofrecen mucho problema para ser analizadas con técnicas

clásicas de series de tiempo. Sin embargo, ¿qué sucede cuando la cantidad de medidas es sumamente pequeña?. Por otra parte, en situaciones así, se cuenta con una gran cantidad de información adicional sobre los actores del proceso y el proceso mismo, que no son medidas seriadas, pero que sí guardan una estrecha relación con lo que sucede en el proceso. Además, frecuentemente los actores conforman un factor de bloque sobre las medidas seriadas; entonces, ¿cómo aprovechar esta información adicional?, ¿cómo se trabajaría dicha información, en relación a las medidas seriadas y repetidas, si ésta no se conforma de medidas sino de características de los actores en el proceso?.

Para dar respuesta a estas interrogantes hemos establecido la metodología KDSM [3] para el descubrimiento de conocimiento en dominios donde se presentan medidas seriadas muy cortas y repetidas con factor de bloque (conformado por los individuos) y donde se cuenta con información adicional de los actores en el proceso como lo son sus variables características.

La metodología KDSM cumple a grandes rasgos con 3 tareas principales:

- La caracterización de la estructura de los individuos sobre la primera toma de medidas, es decir, el establecimiento de las condiciones iniciales;
- El análisis del efecto de cada evento aislado eliminando el factor de bloque; y
- La identificación de las características relevantes de los eventos, la descripción de su estructura y su interpretación.

### **DESCRIPCIÓN DE LOS DATOS DEL ÁMBITO LABORAL**

El conjunto de matrices de datos contiene información sobre las variables característica de 68 municipios del Estado de Chihuahua en México (matriz X), medidas seriadas relativas a la cantidad de personas colocadas (ocupadas) en un puesto de trabajo matriz Y) y la variables características de los cursos de capacitación (matriz Z).

#### **Descripción de los datos de la matriz X**

Esta matriz contiene 26 variables (ni, mun, pt, h, m, p12, pea, pead, almas, alfem, anmas, anfer, ama, ac, as, ami, tem, tvh, ovp, pov, trc, ext12, tmax, tmin, tmed, pma) de tipo numérico.

#### **Descripción de las variables en la matriz Y**

Esta matriz contiene siete variables (ec, ic1, ic2, ic3, ic4, ic5, ic6) de tipo numérico

#### **Descripción de las variables en la matriz Z**

Esta matriz contiene diez variables numéricas (no.aut, u.op, prog., t.i., i.h., i.m., t.e., e.h., e.m., t.c.) y seis variables tipo cadena (ec, cre, esp, cen\_cap, localidad, municipio)

### **APLICACIÓN DE LAS TÉCNICAS DE MINERÍA DE DATOS AL DOMINIO DEL ÁMBITO LABORAL**

Para el análisis de los datos del Gobierno de Chihuahua, se aplicó la metodología KDSM lográndose sus tres tareas principales:

- La caracterización de la estructura conformada por los municipios sobre el primer índice de contratación, es decir, el establecimiento de las condiciones iniciales;
- El análisis del efecto de cada curso aislado eliminando el factor de bloque que conforman los municipios; y
- La identificación de las características relevantes de los cursos, la descripción de su estructura y su interpretación.

#### **Caracterización de la estructura de los municipios para el establecimiento de las condiciones iniciales**

Al realizar los pasos de la metodología KDSM § III, obtuvimos como resultado una base de conocimiento conformada por reglas que describen la estructura de los municipios en relación al primer Índice de Contratación (IC).

A continuación, Figura 1, podemos ver el árbol jerárquico obtenido al realizar la clasificación de la

matriz de basales  $Y_0$ . Es decir la clasificación de las primeras medidas seriadas del IC. La Figura 1 le sugirió al experto que el corte más conveniente era en 3 clases:  
 Clase C1: Municipios de Meoqui, Ocampo, Camargo, Rosario, Delicias, Guadalupe y Calvo, Aquiles Serdan, Valle de Zaragoza y Parral.

Clase C2: Municipios de Madera, Santa Isabel y Ojinaga.

Clase C3: Municipios de Cuahémoc, San Francisco del Oro, Balleza, Jiménez y Allende.

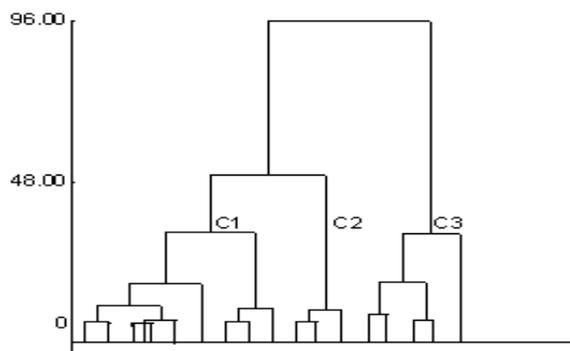


Figura 1. Estructura de los municipios

Para iniciar la interpretación de dichas clases se analizó la Figura 2 donde se puede visualizar la caracterización del patrón de curva típico de cada clase (curva media de cada clase); además de la tendencia general de las clases y la variabilidad entre ellas.

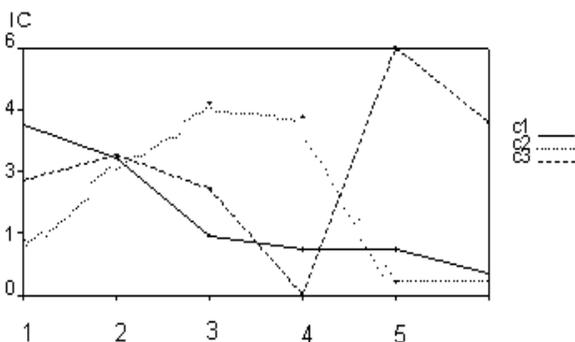


Figura 2. Medidas del IC en 3 clases

Podemos ver que la clase C1 (de 9 municipios) presenta variables que le distinguen especialmente porque se encuentran dentro de rangos donde los límites mínimos y máximos son los más pequeños y mayores de las tres clases. Como ejemplo se citan algunos de las

variables que son de especial interés para el experto: extensión territorial oscila desde 335 km<sup>2</sup> a 16,066 km<sup>2</sup> (Figura 3), población económicamente activa (PEA) varía entre 702 hasta 44,416 personas, PEA desocupada va desde 5 a 473 personas y las unidades de empresas manufactureras de 0 a 560 empresas. Además, se puede ver en la Figura 2 (línea continua) que la contratación más elevada se realiza en la primera y segunda mediciones.

En la clase C2 (tres municipios) se presentan valores que en general se sitúan más cerca del punto medio de los rangos citados anteriormente. Donde la extensión territorial varía desde 1,040 km<sup>2</sup> a 9,500 km<sup>2</sup> (Figura 3), población económicamente activa (PEA) varía entre 1375 hasta 9935 personas, PEA desocupada va desde 18 a 235 personas y las unidades de empresas manufactureras de 9 a 110 empresas. Además, se puede ver en la Figura 2 (línea de puntos) que la contratación se realiza de forma más equilibrada distribuyendo su grosor entre la segunda y la cuarta mediciones.

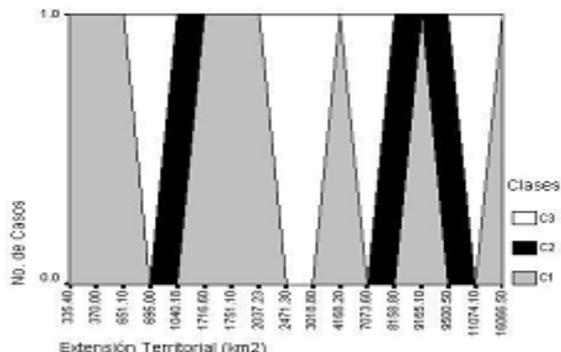


Figura 3. Extensión territorial por clase

Finalmente en la clase C3 (cinco municipios) se tienen variables con valores dispersos en rangos menores a la clase C1. La extensión territorial varía desde 695 km<sup>2</sup> a 1,1074 km<sup>2</sup> (figura 3), población económicamente activa (PEA) va desde 1,909 hasta 42,671 personas, PEA desocupada de 18 a 532 personas y las unidades de empresas manufactureras de 16 a 452 empresas. Además, se puede ver en la Figura 2 (línea discontinua) que el comportamiento muestra que en la primera y segunda mediciones se colocan en promedio

de 2 a 3 personas pero se aprecia que en la quinta medición hay un incremento considerable.

### Proceso de caracterización usando la metodología de CIADEC

En lo habitual cuando los individuos (en nuestro caso municipios) de un dominio complejo y real son descritos por variables cuantitativas, no es común encontrar aquellas que caractericen a las distintas clases de una partición dada (o de referencia); así para realizar su análisis se debe relajar este concepto al de variables parcialmente caracterizadoras [50], esto es, variables que son compartidos por otras clases y que en un sistema de reglas, se representan con un grado de pertenencia difuso a cada una de las clases de la partición de referencia, es decir, reglas difusas (en el sentido de certeza a una clase). Esto plantea un problema serio en la determinación de las variables caracterizadores y en consecuencia en la caracterización y calidad de las clases resultantes de una clasificación de referencia, considerando la calidad de una clasificación, desde un punto de vista subjetivo como la utilidad o significado que las clases resultantes puedan tener para el experto, ya que no existe un criterio objetivo que determine esta calidad.

Como una aproximación al proceso de caracterización y en consecuencia a la obtención de una clasificación “útil” o de “calidad” para los propósitos del estudio se propone realizar los siguientes pasos:

- 1) una estadística descriptiva que nos proporcione información preliminar sobre la variabilidad de las mediciones, descripción grafica de los boxplot para identificar las variables caracterizadores y algunos otros parámetros útiles si los hubiera.
- 2) la inclusión del conocimiento a priori del experto para obtener las restricciones semánticas (reglas) sobre las clases resultantes de la partición que faciliten el significado de las clases.
- 3) la aplicación de *CIADEC* para la obtención del sistema de reglas que proporcionen las características relevantes de éstas para

- 4) determinar la calidad de las clases en términos del “significado” o “utilidad” de éstas.
- 5) estas clases a la vista y análisis del experto decidir si la estructura descubierta es útil, sino repetir el proceso; considerando otra clasificación donde se puede o no incluir nuevas restricciones semánticas, nuevo conocimiento del experto o bien combinando variables en forma de reglas difusas que permitan obtener una nueva estructura de forma que ésta tenga significado para el objetivo del estudio. Si la clasificación es útil entonces continuamos con el
- 6) proceso de interpretación de resultados y la estructura descubierta en los datos puede usarse como nuevo conocimiento para la toma de decisiones o para continuar con
- 7) la segunda tarea de *KDSM*

En resumen las variables de *ET*, población total, *PEA* y *PEA* desocupada, total de escuelas y unidades económicas del sector manufactura son importantes, de acuerdo a la experiencia del experto, debido a que la impartición de cursos y su contenido están sumamente relacionados con la población sin empleo y el territorio que éste ocupa. En la Figura 4 se pueden apreciar los boxplots de las variables de interés para lograr el objetivo del experto.

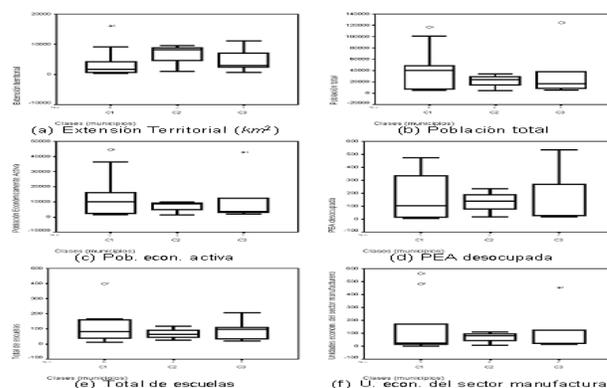


Figura 4. Boxplots de las variables de interés para el experto

## ANÁLISIS DE LOS RESULTADOS

A partir de estas variables se aplicó *CIADDEC* para obtener el sistema de reglas que permitiera la caracterización de cada una de las clases. Sin embargo, no se obtuvieron variables caracterizadoras y por lo tanto, la clasificación obtenida no fue útil para los propósitos del estudio. Como una aproximación al problema de encontrar una clasificación “útil” se decidió combinar las variables más significativas para el experto: *ET* (extensión territorial) y *PEAD* (Pob. Econ. Act. Desoc.) e introducir conocimiento, a priori del experto, a través de reglas difusas que se utilizan como “sesgo” en la siguiente tarea de la metodología *KDSM*. Para la “fuzificación” de estas variables se tomo la experiencia del experto definiendo los rangos de valores de dichas variables en tres etiquetas lingüísticas: pequeño (p), mediano (m) y grande (g), así, se tiene:

Para la variable *ET*:

- hasta 1,000 m<sup>2</sup> la etiqueta pequeña (p),
- entre 1,000 y 10,000 m<sup>2</sup> la etiqueta mediana (m) y
- más de 10,000 m<sup>2</sup> la etiqueta grande (g)

Para la variable *PEAD*:

- hasta 30 personas la etiqueta pequeña (p)
- entre 30 y hasta 270 personas la etiqueta mediana (m)
- más de 270 personas la etiqueta grande (g)

Así, considerando *A* el conjunto de variables, se tiene que  $A = \{NI, \dots, ET, \dots, PEAD, \dots, PMA\}$  y el conjunto de reglas obtenidas son:

- i. Si ( $x_{iET} \leftarrow p \wedge x_{iPEAD} \leftarrow p$ ) → *CR*
- ii. Si ( $x_{iET} \leftarrow p \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow m$ ) → *CU*
- iii. Si ( $x_{iET} \leftarrow p \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow g$ ) → *CU*
- iv. Si ( $x_{iET} \leftarrow m \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow p$ ) → *CR*

- v. Si ( $x_{iET} \leftarrow m \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow m$ ) → *CD*
- vi. Si ( $x_{iET} \leftarrow m \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow g$ ) → *CU*
- vii. Si ( $x_{iET} \leftarrow g \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow p$ ) → *CR*
- viii. Si ( $x_{iET} \leftarrow g \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow m$ ) → *CT*
- ix. Si ( $x_{iET} \leftarrow g \wedge$  **¡Error! No se pueden crear objetos modificando códigos de campo.**  $\leftarrow g$ ) → *CU*

donde:  $x_{iET}$  es el valor de la variable *ET* (extensión territorial) para el *i*-ésimo municipio y  $x_{iPEAD}$  es el valor de la variable *PEAD* (población econ. activa desocupada) para el *i*-ésimo municipio.

Una vez obtenidas las reglas se ponen a consideración del experto para que el valore la representación que forman en la estructura y determine cuáles de ellas le son de utilidad.

### Análisis del efecto de cada curso

Una vez que se llevaron a cabo los pasos de la metodología *KDSM* correspondientes a la primera tarea y, que se obtuvieron las reglas que representan el conocimiento que es de utilidad para el objetivo del análisis, se procede a realizar la segunda tarea de *KDSM* que comprende los pasos correspondientes al análisis del efecto de cada curso aislado, a partir de las diferencias de los índices de contratación *IC*, eliminando el factor de bloque que conforman los municipios sobre los cursos.

El total de reglas fueron nueve, las cuales fueron evaluadas en el conjunto de objetos e indujeron una nueva clasificación conformada por las clases: *U, D, TyR*. Donde *U, D* y *T* reflejan el conocimiento

del experto y  $R$  conforma la clase residual o de los objetos que no se contemplan en dicho conocimiento.

En la Figura 5, podemos ver el árbol jerárquico obtenido al realizar la Clasificación Basada en Reglas [6] (*CIBR*) de la matriz de las diferencias entre los índices de contratación. Donde el experto determinó que el corte más conveniente era en 13 clases debido ya que 2 de ellas reflejan una situación de gran interés y utilidad para los objetivos de su estudio de donde logró resultados importantes y novedosos que se verán más adelante.

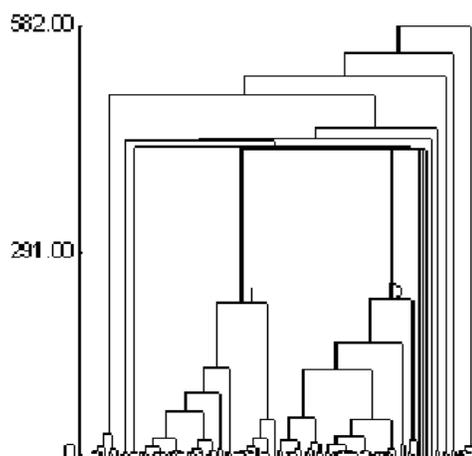


Figura 5. Estructura de los cursos

Clase *U*: *COC-03-02, CCA-15-14, CCU-01-00, CMA-01-00, CD-05-04, CCA-10-09, CCA-04-03, CMA-04-03, CCA-12-11, CMA-02-01, CCA-07-06, CCA-02-01, CCU-03-02, CCA-17-16, CCA-14-13, COC-04-03, CCA-08-07, COC-02-01, CCA-05-04, CD-03-02, CD-04-03, CCU-02-01, CD-01-00, CCU-04-03, CCA-11-10, CCA-09-08, CD-06-05, CCA-03-02, CCA-06-05, CCA-01-00, COC-01-00, CCA-16-15, CCA-13-12, CMA-03-02 y CD-02-01.*

5 cursos se impartieron en el municipio de Ocampo. Su especialidad fue corte y confección y su modalidad de autoempleo.

18 cursos en el municipio de Camargo, con especialidad de costura industrial y en modalidad mixta.

5 cursos en el municipio de Madera, con especialidad de corte y confección en modalidad de autoempleo.

5 cursos en Cuahémoc de corte y confección y en modalidad de autoempleo.

7 cursos en Delicias de corte y confección y de modalidad autoempleo.

Clase *D*: *CPA-19-18, CGC-02-01, CRO-02-01, CPA-13-12, CJI-02-01, CPA-16-15, CGC-01-00, CPA-05-04, CME-01-00, CPA-20-19, CPA-02-01, CME-04-03, CPA-06-05, CPA-01-00, CPA-17-16, CPA-10-09, CPA-04-03, CPA-11-10, CPA-08-07, CPA-14-13, CME-02-01, CRO-01-00, CPA-15-14, CPA-03-02, CPA-07-06, CJI-04-03, CPA-09-08, CPA-12-11, CJI-03-02, CPA-18-17, CME-03-02, CPA-21-20, CJI-05-04, CJI-06-05, CGC-03-02 y CJI-01-00.*

22 cursos en el municipio de Parral, con especialidad en costura industrial y modalidad de capacitación mixta.

4 cursos en Guadalupe y Calvo de corte y confección y de modalidad autoempleo.

5 cursos en Meoqui de corte y confección y de modalidad autoempleo.

3 cursos en Rosales de corte y confección y de modalidad autoempleo.

7 cursos en Jiménez de costura industrial y de modalidad capacitación mixta.

Residuales: *CSF-04-03, CB-04-03, CSF-01-00, COJ-01-00, COJ-04-03, CA-03-02, CSI-02-01, CVZ-03-02, CB-02-01, CA-02-01, CVZ-02-01, CA-04-03, CAS-01-00, CVZ-04-03, CAS-03-02, CVZ-01-00, COJ-03-02, CSF-02-01, CA-01-00, CAS-02-01, CSI-03-02, CB-01-00, CSF-03-02, CAS-04-03, CSI-01-00, CB-03-02 y COJ-02-01.*

El experto decidió omitir las clases residuales de su análisis debido a que no representan información de su interés. Por otra parte, representan sólo el 20 % de los datos que para este caso de estudio en particular son irrelevantes.

Para iniciar la interpretación de éstas clases se caracterizó el patrón de curva típico de cada una de las dos clases de interés para el experto (curva media de cada clase: Clase *U* y Clase *D*) Figura 6 para visualizar la tendencia general de las clases, así como la variabilidad entre ellas. Se descartaron las clases

residuales debido a que no tienen ningún significado para el experto. Se puede observar como ambas clases presentan una tendencia inversa entre ellas, donde la Clase *U* (línea continua) se encuentra entorno a valores negativos, con excepción de la cuarta medición, es decir que el grosor de la contratación se presenta en esta medida, y la Clase *D* (línea a puntos), entorno a valores positivos, que aunque discretos, indican que la contratación en general es estable a lo largo del tiempo de monitorización del PROBECAT.

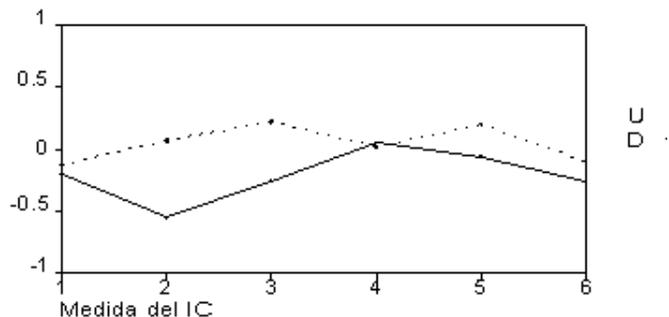


Figura 6. Medidas del IC en 2 clases

### CONCLUSIONES

En este trabajo se utilizó la metodología *KDSM* para el análisis de medidas seriadas muy cortas y repetidas y, la metodología de *CIADEC* enfocada al proceso de caracterización y en consecuencia a la obtención de una clasificación “útil” o de “calidad”, que nos permitan obtener la estructura semántica de forma que, ésta tenga significado para el objetivo de nuestro estudio; los datos fueron tomados de un conjunto de municipios, donde éstos actúan formado bloques con dichas medidas.

La metodología *CIADEC* fue diseñada para descubrir nuevo conocimiento en dominios donde se presentan este tipo de datos; obteniendo a la fecha resultados muy satisfactorios en una primera aplicación al ámbito medioambiental. Los trabajos de mejora de esta metodología continúan y su aplicación se ha extendido a otro ámbito: la capacitación laboral; donde también se obtuvieron resultados satisfactorios e importantes. Se puede concluir, en cuanto a la

metodología se refiere, que ésta integra técnicas habituales de estadística, inteligencia artificial y lógica difusa para dar una posible solución al problema cuando no se encuentran variables relevantes o caracterizadores y medidas seriadas como complemento de la metodología *KDSM* para caracterizar la matriz de basales  $Y_0$  en función de la matriz de datos  $X$ .

Por otra parte, el uso de reglas difusas como combinación de variables permitió la obtención de una clasificación “útil” para el experto en función del objetivo de estudio.

En cuanto a la aplicación de las metodologías antes mencionadas al ámbito laboral, se debe recordar que una de las funciones de la Secretaría de Trabajo y Previsión Social (STPS) a través de los Servicios Estatales de Empleo es dar seguimiento a los resultados obtenidos con el PROBECAT, midiendo su eficiencia mediante una relación costo/beneficio; es decir, la inversión económica con respecto a la gente ya capacitada y que se ha incorporado al sector laboral.

En un esfuerzo por mejorar esta labor se ejecutó un programa piloto para valorar la conveniencia de utilizar las metodologías antes mencionadas (*KDSM* y *CIADEC*) como herramientas auxiliares para el análisis de los aproximadamente 180 diferentes tipos de cursos en los 2427 municipios en México. El programa piloto consistió del análisis de un sólo tipo de curso (manufactura textil) que se impartió en un conjunto de 17 municipios del Estado de Chihuahua y de los resultados obtenidos descritos en el apartado 5.2.3 se obtuvieron las conclusiones que se detallan a continuación.

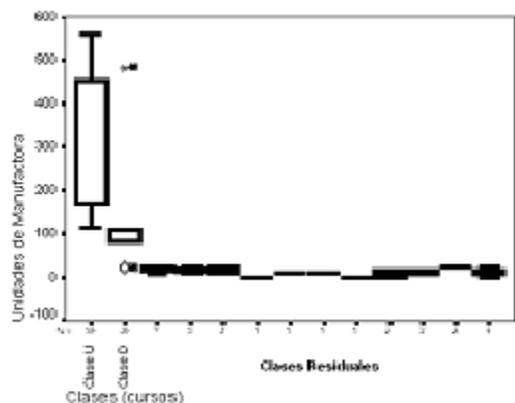


Figura 7. Proyección de UESM en las clases U y D

En base al boxplot de la Figura 7 que representa la distribución de las unidades económicas del sector manufactura (cantidad de empresas) en las clases U y D,. Es así, que basándose en este hecho la STPS puede optimizar el PROBECAT de forma que se aproveche al máximo el nicho de oportunidad que para el gobierno, sector empresarial y trabajadores desempleados representa la coordinación gobierno-empresa en la capacitación.

Por otra parte, conociendo la distribución de las empresas (figura 7), la modalidad del curso y las características de los municipios (boxplots de la figura 4 la STPS podrá realizar la planeación de la capacitación decidiendo entre las diferentes modalidades de los cursos, sus contenidos, y otros; de forma que respondan a las necesidades económicas de los municipios, así como las necesidades del sector empresarial y se maximice la relación costo/beneficio.

Finalmente, en base a variables relativas a la población (boxplots de la Figura 4(b), 4(c) y 4(d)), la STPS podrá actuar ante una situación en la cual el mercado ya no contrate la cantidad suficiente de egresados y que se pierda el equilibrio de la relación costo/beneficio.

### AGRADECIMIENTOS

Los autores agradecen el apoyo de las siguientes instituciones para la realización de este trabajo de investigación a: La Secretaría de Investigación y

Posgrado del IPN, La Secretaría Académica del IPN, La Comisión de Fomento a las Actividades Académicas del IPN (COFAA).

### REFERENCIAS

- [1] Peña, D. *Estadística, Modelos y Métodos. Modelos lineales y series temporales*, volumen II. Alianza, Madrid, segunda edición, 1989J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] Vázquez F. & Gibert K.. *Automatic generation of fuzzy rules in ill structures domains with numerical variables*. Research LSI-01-51-R, Technical University of Catalonia, Barcelona, Spain, <http://www.lsi.upc.es/dept/techreps/html/R01-51.html>. December 2001.K. Elissa, "Title of paper if known," unpublished.
- [3] Rodas J., Gibert K., & Rojo J. *Electroshock Effects Identification Using Classification Techniques*. Springer's Lecture Notes of Computer Science Series, Crespo, Maojo and Martin (Eds.):238-244, Second International Symposium, ISMDA 2001.
- [4] Vázquez F. & Gibert K. *Generación Automática de Reglas Difusas en Dominios Poco Estructurados con Variables Numéricas*. En Actas de la Conferencia de la Asociación Española para la Inteligencia Artificial, volumen 1, Págs. 143-152, España, CAEPIA 01. Nov 2001.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
- [5] Vázquez F. & Gibert K. *Implementation of the methodology "Automatic Characterization and Interpretation of Conceptual Descriptions in ill-Structured Domains*. Research LSI-02-28-R, Technical University of Catalonia, Barcelona. España, Enero 2002. <http://www.lsi.upc.es/dept/techreps/html/R02-28.html>.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [6] Gibert K. & Sonicki Z. *Classification Based on Rules and Medical Research*. In Rocco Curto, editor, VIII International Symposium on Applied Stochastic Models and Data Analysis, pages 181-186, ASMDA97, Italy, 1997.

## AMBIENTE DE ENTRENAMIENTO PARA CÓMPUTO FORENSE, BASADO EN CASOS.

Félix Saucedo Garnica

Centro de Innovación y Desarrollo Tecnológico en  
Cómputo. Instituto Politécnico Nacional.

México DF.

Email: [fsaucedog0200@ipn.mx](mailto:fsaucedog0200@ipn.mx)

### I. RESUMEN.

EL constante reporte de vulnerabilidades en sistemas de información, el aprovechamiento de fallas, bien sea humanas, procedimentales o tecnológicas, sobre infraestructuras de computación en el mundo, ofrecen un escenario perfecto para que se cultiven tendencias relacionadas con intrusos informáticos. Estos intrusos poseen diferentes motivaciones, alcances y estrategias que desconciertan a analistas, consultores y cuerpos especiales de investigaciones, pues sus modalidades de ataque y penetración de sistemas varían de un caso a otro. A pesar de esto, el cómputo forense nos ofrece un espacio de análisis y estudio que nos permite procurar una reflexión profunda sobre los hechos y las evidencias que se identifican en el lugar donde se llevaron a cabo las acciones catalogadas como criminales. En este momento es preciso establecer un nuevo conjunto de herramientas, estrategias y acciones que permitan descubrir en los medios informáticos, la evidencia digital que sustente y verifique las afirmaciones que sobre los hechos delictivos se han materializado en el caso bajo estudio. Este trabajo pretende mostrar una panorámica general de la informática forense, resaltando en primer lugar su importancia, sus objetivos y usos. Acto seguido se explica la necesidad de las herramientas informáticas de entrenamiento y de análisis de las evidencias digitales en los medios informáticos.

### II. INTRODUCCION.

La computación forense es una disciplina naciente que desde sus inicios ha establecido un reto tanto para los profesionales de ciencias de la computación y tecnología de información, como para los criminalistas tradicionales y la administración de la justicia en general. La necesidad de contar con un profesional que reconozca y actúe conforme a los requisitos de ley y siga los procedimientos básicos en criminalística, ahora en fenómenos o casos donde la informática y la tecnología se hacen presentes, es una ventana para repensar la práctica de las ciencias forenses en un entorno digital [7][1].

La seguridad de la información es muy reciente. Sus implicaciones van desde el análisis forense de un incidente común de pérdida o de ocultamiento de información en una computadora personal, hasta aspectos de defensa nacional. Esta disciplina está adquiriendo una gran importancia dentro del área de la información electrónica, esto debido al aumento del valor de la información y al uso que se le da a ésta, al desarrollo de nuevos espacios donde es usada, por ejemplo el internet, y al extenso uso de computadoras por parte de las compañías de negocios tradicionales como los bancos. Es por esto que cuando se realiza un crimen, muchas veces la información queda almacenada en forma digital. Sin embargo, existe un gran problema, debido a que los computadores guardan la información de información forma tal que no puede ser recolectada o usada como prueba utilizando medios comunes, se deben utilizar mecanismos diferentes a los tradicionales. Es de aquí que surge el estudio de la computación forense como una ciencia relativamente nueva [3][1].

### III. EVIDENCIA DIGITAL.

La evidencia digital es cualquier información que, sujeta a una intervención humana u otra semejante ha sido extraída de un medio informático [8]. En este sentido, la evidencia digital es un término utilizado de manera amplia para describir cualquier registro generado por o almacenado en un sistema computacional que puede ser utilizado como evidencia en un proceso legal.

En este sentido, el documento mencionado establece que la evidencia digital puede ser dividida en tres categorías, a saber:

1. Registros almacenado el en equipo de tecnología informática. Como ejemplos tenemos: correos electrónicos, archivos de aplicaciones, imágenes, etc.

2. Registros generados por los equipos de tecnología informática. Como ejemplos tenemos: registros de auditoría, registros de transacciones, registros de eventos, etc.
3. Registros que parcialmente han sido generados y almacenados en los equipos de tecnología informática. Como ejemplos tenemos: hojas de cálculo financieras, consultas especializadas en bases de datos, vistas parciales de datos, etc.

La evidencia digital es la materia prima para los investigadores donde la tecnología informática es parte fundamental del proceso. Sin embargo y considerando el ambiente tan cambiante y dinámico de las infraestructuras de computación y comunicaciones, es preciso detallar las características propias de dicha evidencia en este entorno. La evidencia digital, para aquellos que la identifican y analizan en la búsqueda de la verdad, posee, entre otros elementos que la hacen un constante desafío, las características siguientes:

1. Es volátil
2. Es anónima
3. Es duplicable
4. Es alterable y modificable
5. Es eliminable

Estas características nos advierten sobre el exigente labor que se requiere por parte de los especialistas en temas de informática forense, tanto en procedimientos como en técnicas y herramientas tecnológicas para obtener, custodiar, revisar, analizar y presentar la evidencia presente en una escena del delito. Por tanto es necesario mantener un conocimiento detallado de las normas y regulaciones legales asociadas con las pruebas y el derecho procesal, así como de las técnicas y los procesos que permitan mantener la confiabilidad de los datos recogidos, la integridad de los medios, el análisis detallado de los datos y la presentación idónea de los resultados.

#### IV. HERRAMIENTAS DE INFORMÁTICA FORENSE.

Las herramientas informáticas son la base esencial de los análisis de las evidencias digitales en los medios informáticos. Sin embargo, es preciso comentar que éstas requieren una formalidad adicional que permita validar tanto la confiabilidad de los resultados de la aplicación de las mismas, como la formación y el conocimiento del investigador que las utiliza. Estos dos elementos hacen del uso de las herramientas una constante reflexión y un cuestionamiento por parte de la comunidad científica y práctica de la informática forense en el mundo.

Dentro de las herramientas frecuentes utilizadas en procedimientos forenses en informática, detallamos algunas, las cuales son aplicaciones que tratan de cubrir todo el proceso en la investigación forense en informática. Tabla 1.

	Licencia	Control de integridad	Análisis	Administración del caso
Encase	Sí	Sí	Sí	Sí
Forensic Toolkit	Sí	Sí	Sí	Sí
Winhex	Sí	Sí	Sí	Sí
Sleuth	No	Sí	Sí	Sí

Tabla 1. Herramientas utilizadas en procedimientos forenses en informática.

La informática forense es un desafío interdisciplinario que requiere un estudio detallado de la tecnología, los procesos y los individuos que permitan la conformación de un cuerpo de conocimiento formal, científico y legal para el ejercicio de una disciplina que apoye el esclarecimiento de los hechos alrededor de los incidentes o los fraudes en las organizaciones. La informática forense es la manifestación natural del entorno digital y de la sociedad de la información para responder a la creciente ola de incidentes, fraudes y ofensas, todo esto en medios informáticos y a través de medios informáticos, con el fin de enviar un mensaje claro a los intrusos indicando que se está preparado para responder a sus ataques y se continúa aprendiendo para dar con la verdad de sus acciones.

Estas herramientas siguen el proceso operativo de la computación forense, que se representa en la figura 1.

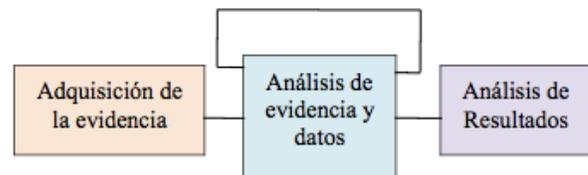


Figura 1. Modelo operativo de la computación forense [2].

#### V. EL INTRUSO.

Es importante conocer la mente del intruso y como éstos piensan, actúan y operan para desarrollar la pericia para predecir y analizar los posibles rastros y acciones que se desarrollen en medios tecnológicos.

Esto ofrece al administrador o al informático forense una ventaja estratégica y conceptual sobre el caso que revise, pues se pondrá en el lugar del infractor y tratará de ver como actuaría en un caso semejante. Así mismo, el estudio de otros ataques le dará mayores elementos de juicio para establecer patrones de análisis que ayuden a detallar lo ocurrido y así apoyar las investigaciones relacionadas con el caso.

#### A. INTRUSO INTERNO

Características psicológicas:

- Motivado por situación personal o laboral.
- Inestabilidad emocional.
- Socialmente hábil para recabar información y conocer a sus víctimas.

Características técnicas:

- Conocimiento detallado de fallas en procedimientos y regulaciones internas.
- Conocedor y estudioso de la operación de la organización y su modelo de procesos y controles.
- Conocedor de los mecanismos de seguridad y control.

#### B. INTRUSO INTERNO

Características psicológicas:

- Generalmente motivado por reto tecnológico y compensación económica
- Sensación de control y poder sobre un tercero
- Relaciones basadas en conocimiento y logros

Características técnicas:

- Conocedor y estudioso de las fallas tecnológicas de los sistemas objetivo
- Conocedor y usuario de técnicas de evasión de investigaciones
- Cuenta con un laboratorio de pruebas para verificar previamente sus acciones

Si bien el rol del intruso es importante y atractivo, no es posible conocer los detalles de los rastros, si el administrador, el profesional de tecnologías de información a cargo de la administración y control de las computadoras que posiblemente están involucradas, no se encuentra capacitado o no cuenta con las herramientas necesarias para analizar las evidencias digitales. En este rol, el administrador de la seguridad debe comprender los conceptos de protección y control de tecnologías de información, no solo para detallar las medidas tecnológicas de seguridad y control configuradas, sino para identificar y analizar las diferentes formas de alerta, detección, registro y monitoreo que la infraestructura tiene definidas para

prevenir algún tipo de incursión de autorizada.

En este sentido el administrador se enfrentará al reto de la inseguridad de la información y sus diferentes fuerzas, reconocerá las relaciones entre las tecnologías de protección y las fallas de seguridad informática, para afianzar su visión de intruso.

Existen diferentes consideraciones para concebir o para clasificar a los atacantes, así como las motivaciones que los mueven a actuar en una situación particular. Las tablas 2.1 y 2.2 muestran algunos atacantes y sus posibles motivaciones.

Motivaciones	Ciberterroristas	Phreakers	Script kiddies
Reto		X	
Ego		X	X
Espionaje			
Ideología	X		
Dinero		X	
Venganza	X		X

Tabla 2.1. Algunos tipos de atacantes y sus motivaciones. Obtenido de Furnell, S. 2002, p55.

Motivaciones	Crackers	Desarrollo de virus	Atacante interno
Reto		X	X
Ego		X	
Espionaje	X	X	X
Ideología			
Dinero	X	X	X
Venganza		X	X

Tabla 2.2. Algunos tipos de atacantes y sus motivaciones. Obtenido de Furnell, S. 2002, p55.

Furnell habla de atacantes con perfil de, entre otros términos, ciberterrorista, phreakers, script kiddies, crackers, desarrollador de virus. Cada uno de ellos se mueve por motivos diferentes que llevan una carga emocional que es importante analizar y no solamente, las consideraciones técnicas de sus acciones, que dicen del nivel de conocimiento del individuo

## VI. EL INFORMÁTICO FORENSE O ADMINISTRADOR DEL SISTEMA.

Este profesional será quien se encargue de la búsqueda de va verdad, en el análisis de la información residente en los dispositivos tecnológicos, en la construcción del caso con las evidencias digitales requeridas para esclarecer los móviles de los hechos, que se han podido presentar, bien sea en medios informáticos o electrónicos, o en

combinación de hechos físicos y tecnológicos. Las características que deben tener el informático forense o el administrador del sistema se presentan en la figura 2.



Figura 2. Características del administrador de sistema [7].

#### A. Dificultades que se le podrían presentar al Informático forense

El profesional de la computación forense enfrentara en sus labores algunas dificultades, las cuales podrán relacionarse con el personal, los procesos, escasa o ninguna documentación de los sistemas, etc.

A continuación se enumeran algunas de las dificultades a las cuales podrá enfrentarse [1]:

- No contar con los registros de auditoría. Esto puede suceder, porque el aplicativo no los tiene implementados o si los tiene, están desactivados (la entidad podría justificar que los registros están degradando la computadora).
- Registros incompletos o no claros de las pistas de auditoría. Esto ocurre porque solo se graban algunos campos para no cargar el sistema o no existen descripciones detalladas de los registros.
- No se realiza un buen levantamiento de información de la arquitectura del sistema y se dificulta determinar la forma y quién realizó la transacción fraudulenta.
- Poca habilidad en el manejo de las herramientas.
- Resistencia por parte de los funcionarios para suministrar información porque no les agrada ser investigados o porque podrían estar relacionados con el ilícito.

- Restricción de acceso a la información de la entidad. Si se cuenta con el conocimiento y las herramientas necesarias, los funcionarios de seguridad informática y/o auditoría de sistemas de la entidad podrían adelantar la investigación forense y no habría mayor dificultad en el acceso a la información; pero si se requiere que por la especialización del tema lo realice un tercero, éste investigador deberá trabajar de manera estrecha con las áreas de seguridad bancaria, jurídica y la auditoría de sistemas

## VII. CONCLUSIÓN

Utilizando herramientas de cómputo forense podemos obtener información relativa a cómo, cuándo y dónde se produjo un ataque informático, esto con la finalidad de reparar los daños o de averiguar la procedencia y los objetivos de dicho ataque. Sin embargo, la continua evolución de los sistemas de información permite que los intrusos redefinan sus métodos y técnicas de ataque, por lo que es necesario contar con herramientas y técnicas de análisis de evidencias digitales y así definir una metodología e implementar un plan de prevención, minimización de daños o de recuperación en caso de un ataque. Es necesario conocer las fortalezas y debilidades de las herramientas de computación forense a nivel comercial y de software libre para adaptar una aplicación que forme el perfil de administrador y entrene al informático forense y así mismo dé confiabilidad de los resultados.

## VIII. REFERENCIAS

- [1] Computación forense: una forma de obtener evidencias para combatir y prevenir delitos informáticos. Yuri Vladimir López Manrique. Guatemala, marzo de 2007.
- [2] Desarrollo de una interfaz gráfica multiplataforma para una herramienta de computación forense. Edward Andrés Corredor Rondón. Bogotá D.C. 2007
- [3] Informática forense: generalidades, aspectos técnicos y herramientas. Óscar López, Haver Amaya, Ricardo León, Beatriz Acosta. Universidad de Los Andes Bogotá, Colombia
- [4] Computational Forensic Techniques for Intellectual Property Protection. Jennifer L. Wong, Darko Kirovski, and Miodrag Potkonjak. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 23, no. 6, June 2004
- [5] Non-parametric statistical techniques for

computational forensic engineering. Jennifer lee Wong.

[6] Relative generic computational forensic techniques  
Jennifer l. wong and miodrag potkonjak. University  
of California, Los Angeles.

[7] Computación forense, Descubriendo los rastros  
informáticos. Jeimy Cano. Alfaomega 2009.

[8] Guidelines for the Management of IT Evidence.  
APEC Telecommunications and Information Working  
Group 29th Meeting | 21-26 March 2004 | Hong  
Kong, China

## Stimulation of three Biological Systems Using Low Level Laser Radiation

Alexandre Michtchenko, Mauricio Hernández Vizuet.

Instituto Politécnico Nacional, ESIME-SEPI, Zacatenco, C.P. 07738, México, D.F.

E-mail: [amichtchenko@ipn.mx](mailto:amichtchenko@ipn.mx), [mahhhevi@ipn.mx](mailto:mahhhevi@ipn.mx).

### Abstract

*The present work determines some parameters of light that are determinants in the activation or inhibition of the growth stimulus in a biological system. For this work intensity and dose, as important parameters, were considered. Semiconductor lasers with different wavelengths were used to irradiate the biological systems. Three biological systems were used to studied the stimulation effects when they were exposed to laser radiation. The three biological systems were Cucumis sativus L (cucumber) seeds, Triticum aestivum L (wheat) and human fibroblasts. The responses of the three biological systems were analyzed using statistical processes when intensity and dose were varied. The results obtained indicate that when the appropriate light parameters are used activation or inhibition can occur.*

**Keywords:** laser radiation, cucumber seeds, wheat seeds, Human Fibroblasts.

### 1. Introduction

Recently has appeared different articles that reported the use of microwaves and laser radiation to modify the behavior of biological systems [1-7]. Laser radiation has been applied to stimulate the growth of E. Coli cultures, the DNA replication and division, the activation of some genes and the activation of the growth of some seeds [1, 3-7] between others. It is important to mention that each biological system requires different parameters of radiation for stimulation.

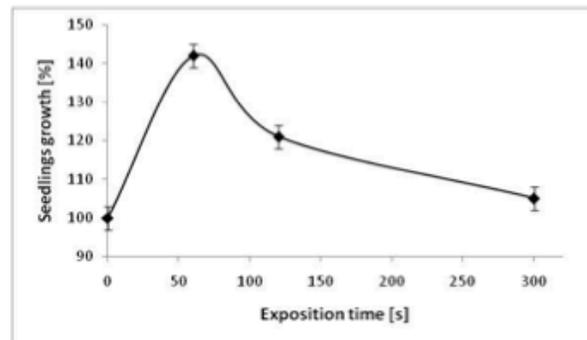
There are three factors that are determinant to obtain an early response in a biological system; these parameters are the wavelength (nm), the intensity (W/m<sup>2</sup>) and the dose (J/m<sup>2</sup>) [5, 8]. The wavelength determines the penetration of laser radiation into tissues while the intensity and the dose determine the stimulation effect and its magnitude [8]. From the above mentioned if appropriate parameters of laser radiation are applied, activation of the biological system can occur.

The present work investigated the behavior of three different biological systems when they were irradiated with different laser wavelengths. The three biological systems studied were cucumber seeds (*Cucumis sativus* L), wheat seeds (*Triticum aestivum* L) and human fibroblasts.

### 2. Stimulation of Cucumber seeds

The cucumber seeds were irradiated with a laser wavelength  $\lambda=650$  nm. The intensities of laser radiation were 2.5 mW/cm<sup>2</sup>, 5 mW/cm<sup>2</sup>, 10 mW/cm<sup>2</sup> and 20 mW/cm<sup>2</sup> and the exposition times were 60s, 120s and 300s.

In the case of stimulation of cucumber seeds with laser radiation, the highest effect registered was obtained for an intensity of 10mW/cm<sup>2</sup> and an exposition time of 60s (Figure 1). The increase with respect to control for this treatment had a value of 43.7%. No other intensities and exposition times presented important changes on the growth of cucumber seedlings.

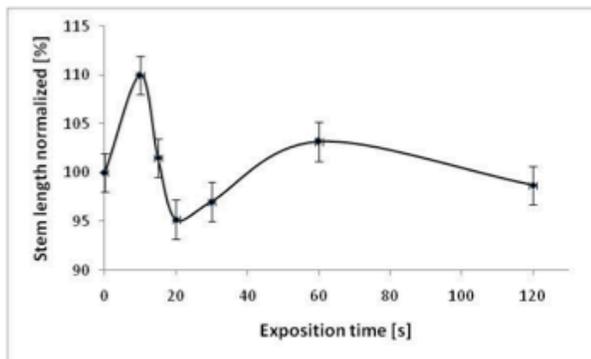


**Figure 1. Growth of cucumber seedlings for an intensity of 10mW/cm<sup>2</sup> and three different exposition times: 60s, 120s and 300s. The control seeds are represented as 0s.**

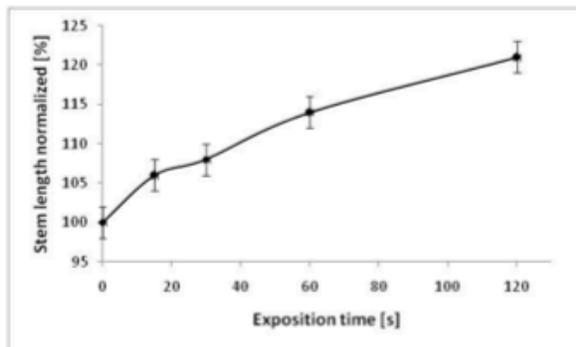
### 3. Stimulation of Wheat seeds

The wheat seeds were irradiated with four different laser wavelengths; 650 nm, 850 nm, 904 nm and 980 nm. In the case of the laser wavelength  $\lambda=650$  nm, the highest effect on the stem length was registered for an intensity of  $20\text{mW}/\text{cm}^2$  and an exposition time of 10s. The stem length increased 10.1% with respect to control (Figure 2).

For the case of a wavelength  $\lambda=850$  nm, the highest effect on the stem length was registered for an intensity of  $12.5\text{mW}/\text{cm}^2$  and an exposition time of 120s. The stem length increased 21.7% with their respective control (Figure 3).



**Figure 2. Stem length for wheat seeds treated with an intensity of  $20\text{mW}/\text{cm}^2$  and six different exposition times: 10s, 15s, 20s, 30s, 60s and 120s. The wavelength for this figure was  $\lambda=650$  nm.**

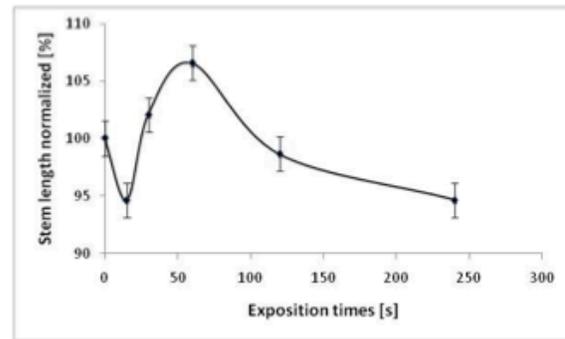


**Figure 3. Stem length for wheat seeds treated with an intensity of  $12.5\text{mW}/\text{cm}^2$  and four different exposition times: 15s, 30s, 60s and 120s. The wavelength for this figure was  $\lambda=850$  nm.**

In the case of a wavelength  $\lambda=904$  nm there were not important changes on the stem length for wheat

seeds treated with different intensities and exposition times.

For the case of a wavelength  $\lambda=980$  nm the highest effect of the growth of the stem length was registered for an intensity of  $60\text{mW}/\text{cm}^2$  and an exposition time of 60s. The stem length increased 8.43% with respect to control (Figure 4). No other intensities and exposition times registered important increases on the stem length.

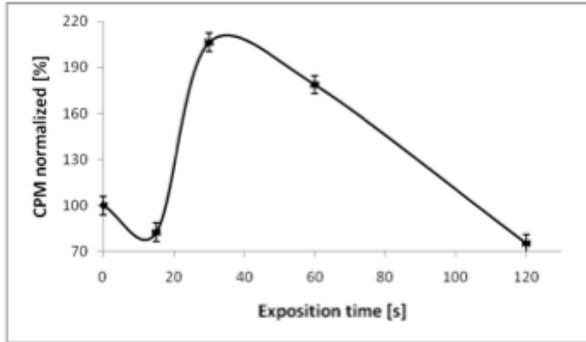


**Figure 4. Stem length for wheat seeds treated with an intensity of  $60\text{mW}/\text{cm}^2$  and five different exposition times: 15s, 30s, 60s, 120s and 240s. The wavelength for this figure was  $\lambda=980$  nm.**

### 4. Stimulation of Human Fibroblasts

The human fibroblasts were radiated using a laser wavelength  $\lambda=650$  nm to produce stimulation effects. The intensities of laser radiation were  $3.75\text{mW}/\text{cm}^2$ ,  $7.5\text{mW}/\text{cm}^2$ ,  $15\text{mW}/\text{cm}^2$  and  $30\text{mW}/\text{cm}^2$ ; the exposition times to laser radiation were 15s, 30s, 60s and 120s. To evaluate the stimulation effects on human fibroblasts the incorporation of  $[^3\text{H}]$ -thymidine was used. The results were obtained 24hr, 48hr and 72 hr after laser treatment.

The highest effects were registered 48 hr after laser treatment. Particularly the highest incorporation of  $[^3\text{H}]$ -thymidine was obtained for an intensity  $I=3.75\text{mW}/\text{cm}^2$  and an exposition time of 30s. The incorporation of  $[^3\text{H}]$ - thymidine was 210% higher than their respective control (Figure 5).



**Figure 5. Fibroblasts proliferation for an intensity of 3.75mW/cm<sup>2</sup> and four different exposition times: 15s, 30s, 60s and 120s.**

## 5. Conclusions

From the results above obtained it is possible to conclude that the wavelength, the intensity and the exposition time to laser radiation are important parameters to produce stimulation effects. In the case of wheat seeds the highest increase, 21.7% with respect to control, was obtained for a wavelength  $\lambda=850$  nm, an intensity of 12.5mW/cm<sup>2</sup> and an exposition time of 120s. In the case of cucumber seeds an important increase, 43.7% over the control level, was obtained for a wavelength  $\lambda=650$  nm, an intensity of 10mW/cm<sup>2</sup> and an exposition time of 60s. For the human fibroblasts, an

important increase with a value of 210% over the control level was registered for a wavelength  $\lambda=650$  nm, an intensity of 3.75mW/cm<sup>2</sup> and an exposition time of 30s.

## 6. References

- [1] A. Hernández, C.A. Carballo, A. Artola, A. Michtchenko, "Laser irradiation effects on maize seed file performance", *Journal Seed Science and Technology*, 34, 1, 193-197, 2006.
- [2] R.W. Hunt, A. Zavalin, A. Bhatnagar, S. Chinnasamy, K.C. Das, "Electromagnetic Biostimulation of Living Cultures for Biotechnology, Biofuel and Bioenergy Applications", *Int.J.Mol. Sci*, 10, 4515-4558, 2009.
- [3] Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, "A light switchable gene promoter system", *Nature Biotechnology*, 20, 1041-1044, 2002.
- [4] A. Levskaia, O.D. Weiner, W. A. Lim, C. A. Voigt, "Spatiotemporal control of cell signaling using a light-switchable protein interaction", *Nature Biotechnology*, 461, 15, 997-1001, 2009.
- [5] T. Karu, "The science of low power therapy", London: Harwood, 1998.
- [6] Yi-Ping Chen, Yong-Jun liu, Xun-Ling Wang, Zhao-Yu Ren, Ming Yue, "Effect of microwave and He-Ne laser on enzyme activity and biophoton emission of *Isatis indigotica* Fort", *Acta Botanica Sinica*, 47, 7, 849-855, 2005.
- [7] S. Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, "A light switchable gene promoter system", *Nature Biotechnology*, 20, 1041-1044, 2002
- [8] Markolf H. Niemz, "Laser-Tissue interactions", Springer, 2007.

## ***Instrucciones para los autores***

Los artículos que se someten a **RISCE** deben contener resultados inéditos y originales, no haber sido publicados con anterioridad ni haber sido sometidos simultáneamente a otra revista científica. Si el artículo ha sido presentado, sometido o publicado en alguna otra parte, deberá informarse al coordinador editorial. Los artículos deben ajustarse a las siguientes especificaciones:

- Idioma Inglés (anexar un resumen y palabras clave en español)
- Idioma Español (anexar un resumen y palabras clave en Inglés)
- Procesador de texto admitido: MS-Word.
- Tamaño de página: carta, utilizar un solo lado de la hoja. Máximo 10 páginas.
- Márgenes: izquierdo 2.5 cm y derecho 2 cm., superior 2.5 cm e inferior 2.5 cm.
- Autores: primer nombre seguido de los dos apellidos (sin abreviaturas), abajo: afiliación y e-mail.
- Tipo de letra del texto regular: Times o Times New Roman de 10 pt (título original 22 pt; secciones 11.5 pt, subsecciones 11.5 pt, en negritas).
- Texto: a una columna y con espaciado sencillo (renglón seguido).
- Resumen/Abstract: entre 70 y 150 palabras, colocado al principio del texto, seguido del de Español o inglés según sea el caso.
- Palabras clave/Keywords: colocadas después del resumen en negritas, y no más de 10.
- Imágenes y fotografías: deben ser de alta calidad, con colores bien definidos y contrastantes, en mapa de bits (no sectorizadas) en formato JPG e incrustadas en el texto de forma que se puedan manipular independiente.
- Fórmulas: Deberán de presentarse en formato de tabla sin bordes, centradas y la numeración de c/u justificada a la derecha con negritas en mapa de bits, no vectorizadas.
- Pies de figura. Deben mencionarse dentro del texto y numerarse de manera consecutiva con un tipo de letra Times New Roman 9 puntos
- Cabecera de tabla. Deberá presentarse en la parte superior de la tabla un numeración consecutiva y descripción con tipo de letra Times New Roman 9
- Referencias:

En cualquier caso el nombre del autor del artículo o publicación web deberá mostrarse al principio. Deberán ordenarse conforme aparezcan dentro del texto encerradas entre paréntesis cuadrado —[ ]—. A continuación algunos ejemplos:

- [1]. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. *Int. J. Digit. Libr.* 1 (1997) 108–121
- [2+]. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): *Theoretical Aspects of Computer Software. Lecture Notes in Computer Science*, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415–438
- [3]. van Leeuwen, J. (ed.): *Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science*, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)
- [4]. Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)

### **Instrucciones:**

Enviar el archivo en extenso a la siguiente dirección electrónica: [ebustosf@gmail.com](mailto:ebustosf@gmail.com)

Los revisores técnicos le harán llegar sus observaciones y modificaciones, las cuales deberá realizar y reenviar el archivo corregido al correo arriba mencionado.

El comité editorial se comunicara mediante correo electrónico indicándole la aceptación o rechazo del artículo.

Se le solicitará autorización para publicación; en caso de aceptar se le indica la cuenta donde debe hacer el depósito por cobro de publicación y el costo, el cual no debe exceder de \$1000.00 pesos mexicanos.

**Reserva de Derechos 04-2008-062613190500-203**