

# INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE OCCIDENTE

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación el 29 de noviembre de 1976.

Departamento de Electrónica, Sistemas e Informática  
MAESTRÍA EN SISTEMAS COMPUTACIONALES



HAGA CLIC AQUÍ PARA ESCRIBIR TEXTO.

Uso de Algoritmos de Aprendizaje Automático  
para Seguridad informática.

Reporte de Avance de Trabajo de Obtención de Grado  
IDI II

Presenta: Nelson Victor Cruz Hernández

Asesor: Mtro. Alvaro Iván Parres Peredo

Tlaquepaque, Jalisco. Julio de 2016.

## RESUMEN

Las diferentes técnicas de seguridad en redes de computadoras descuidan los ataques que provienen del interior de la red, por lo que los usuarios pueden ser agentes de ataques de forma consciente o indirecta.

Este trabajo tiene como **objetivo prevenir problemas de seguridad interna en redes CAN** a través del perfilado de usuarios de la red, mediante la implementación de un algoritmo de clustering que permita la definición de patrones de comportamiento normal y anómalo dentro de la red para establecer el perfil de un usuario normal y un atacante.

# TABLA DE CONTENIDO

<b>1. INTRODUCCIÓN .....</b>	<b>7</b>
1.1. ANTECEDENTES.....	8
1.2. JUSTIFICACIÓN .....	8
1.3. PROBLEMA .....	9
1.4. HIPÓTESIS .....	9
1.5. OBJETIVOS .....	9
1.5.1.      Objetivo General:.....	9
1.5.2.      Objetivos Específicos:.....	9
1.6. NOVEDAD CIENTÍFICA, TECNOLÓGICA O APORTACIÓN .....	10
<b>2. ESTADO DEL ARTE.....</b>	<b>11</b>
2.1. ALGORITMOS DE MACHINE LEARNING Y SUS APLICACIONES .....	12
2.2. SELECCIÓN DE VARIABLES PARA LA CORRECTA DETECCIÓN DE ATAQUES .....	13
2.3. DATASETS AUTOGENERADOS .....	14

## LISTA DE FIGURAS

No se encuentran elementos de tabla de ilustraciones.

## LISTA DE TABLAS

No se encuentran elementos de tabla de ilustraciones.

## LISTA DE ACRÓNIMOS Y ABREVIATURAS


# 1. INTRODUCCIÓN

*Rapidos* Las técnicas actuales de seguridad en redes de computadoras trabajan principalmente en la frontera de la red, descuidando en mayor forma las amenazas que se originan en el interior *E93* *x Ref.*

*an Antecedentes* Diversos trabajos se han desarrollado buscando prevenir y contener los ataques por parte de usuarios internos a través de técnicas como ~~sólo~~; Machine Learning y Sistemas de Detección de Intrusos (IDS) [2] [3], algoritmos de clustering [2, 4, 5], generación de infraestructuras de cloud computing y minería de datos [4].

Dichas propuestas se enfrentan a altos porcentajes de falsos positivos, una tasa de detección de ataques variante, y una alta dependencia a las bases de firmas, las cuales son difíciles y costosas de mantener actualizadas para ambientes de producción.

La ejecución de los algoritmos de machine learning de clustering permite la generación de patrones de comportamiento de un usuario dentro de la red, permitiendo identificar usuarios con comportamientos normales conocidos dentro de la red y usuarios con comportamientos no definidos o anómalos. *Reducción*

*x Ref.* Un usuario tiene un comportamiento similar dentro de la red, mismo que es consistente en el tiempo [6], haciendo viable la generación de perfiles de comportamiento a través de algoritmos de clustering en un periodo corto de tiempo, permitiendo identificar usuarios con comportamientos normales conocidos dentro de la red y usuarios con comportamientos no definidos o anómalos cuyas características pueden ser: ataques conocidos, nuevas formas de ataques o un nuevo patrón de usuario seguro dentro de la red.

## 1.1. Antecedentes

En base al reporte de seguridad de Cisco publicado en 2015 [1] los usuarios y los equipos de Tecnologías de información se han vuelto clave en la seguridad informática dado que estos comprometen la seguridad de forma inconsciente a través del uso indiscriminado de los navegadores y la participación en campañas de publicidad [1], principales medios para la distribución de *malware*.

Las técnicas actuales de seguridad en redes de computadoras trabajan principalmente en la frontera de la red, buscando prevenir los ataques que son ejecutados desde el exterior descuidando en mayor forma las amenazas que se originan en el interior.

Se han desarrollado diversos trabajos buscando prevenir y contener los ataques por parte de usuarios internos con diferentes propuestas como son; técnicas de *Machine Learning* e IDS [2] [3], algoritmos de *clustering* [2, 4, 5], generación de infraestructuras de *cloud computing* y minería de datos [4].

Dentro de las propuestas de nuevas técnicas de IDS existe la problemática de los altos índices de falsos positivos, una tasa de detección que varía en base al tipo de ataque, y una dependencia completa a la creación de firmas que muchas veces pueden no ser óptimas en el ambiente de producción o sus tiempo de creación es demasiado largo, siendo las principales razones de esto la extensa cantidad de información que debe ser analizada, el poco tiempo que se tiene para su análisis, el uso incorrecto de las técnicas de procesamiento aplicadas o el tiempo que se debe invertir para la creación de firmas para el ambiente vulnerable.

Tian [3] propone que es posible hacer uso de selección de variables, para reducir el conjunto de información a analizar lo que impacta directamente en el tiempo que se toma el análisis y es posible analizar de forma más detallada la información. Por su parte Kamarularin [2] explora la eficiencia de los procesos. (indicar cuáles procesos).

Por su lado Yamada [5] y Kamarularin [2] ejecutan diferentes algoritmos de *machine learning* cuyos resultados permitan conocer y comparar resultados en la detección de ataques y la retroalimentación para la generación de firmas de forma automatizada.

Es posible observar que existe un campo de estudio enfocado a los IDS, utilizando técnicas de machine learning con un alto nivel de fiabilidad, por lo que su aplicación permitiría combatir de forma efectiva la brecha de seguridad en redes causada por los mismos usuarios.

## 1.2. Justificación

Los algoritmos de clustering de machine learning son una opción para la generación de patrones de comportamiento de un usuario dentro de la red, basado en el flujo de datos que este transmite, de tal forma que es posible realizar un análisis de comportamiento en un periodo corto de tiempo, permitiendo identificar usuarios con comportamientos normales conocidos dentro de la red y usuarios con comportamientos no definidos o anómalos cuyas características pueden ser: ataques conocidos, nuevas formas de ataques o un nuevo patrón de usuario seguro dentro de la red.

### 1.3. Problema

Las diferentes técnicas de seguridad en redes de computadoras descuidan los ataques que provienen del interior de la red, ya que estas funcionan bajo el supuesto de que un usuario debe tener un comportamiento seguro dentro de la misma, lo que da paso a dos tipos de ataques: el primero consciente por parte del usuario y el segundo es aquel donde el usuario es víctima de un atacante, pero no lo sabe y vulnera de forma inconsciente aquellas redes donde tiene actividad.

+ Ref.

En el caso donde el usuario es consciente del tipo de actividades que está realizando podemos encontrar ataques básicos que van muy apegados al comportamiento de un scriptkiddie hasta una persona más experimentada que es capaz de escribir su propio código y hacer uso de herramientas a nivel más avanzado.

?

De igual manera un usuario puede verse afectado por algún tipo de malware [1] y su capacidad de distribución a través de navegadores y campañas publicitarias, ya que su falta de conocimiento técnico le impide ver que su dispositivo puede estar comprometido y por lo tanto comprometer las redes en las que se encuentra, permitiendo la distribución del mismo hacia otros hosts.

Ambos comportamientos tienen un impacto directo para la seguridad de la organización, ya que son vulnerables a cualquier tipo de ataque por parte de sus usuarios.

### 1.4. Hipótesis

Un usuario tiene un comportamiento similar dentro de la red, mismo que es consistente en el tiempo [6], haciendo viable la generación de perfiles de comportamiento. Estos perfiles permiten generar grupos de usuarios, clúster, con comportamientos similares. Una acción no identificada dentro del perfil se considera como una actividad sospechosa que requiere ser revisada con el fin de determinar si es un ataque interno o no independientemente si es una acción consciente o inconsciente.

La redacción no corresponde a una hipótesis.

### 1.5. Objetivos

#### 1.5.1. Objetivo General:

Prevenir problemas de seguridad interna en *campus area networks* (CAN) a través del perfilado de usuarios de la red, mediante técnicas de *clustering* y *machine learning*, haciendo posible la detección de ataques de usuarios internos en un periodo corto de tiempo, y la reducción de falsos positivos.

Distinguir entre lo normal y lo anómalo.

#### 1.5.2. Objetivos Específicos:

Para poder lograr el objetivo general es necesario desarrollar las siguientes fases:

- 1) Implementar un algoritmo de *clustering* cuyo resultado nos permita identificar grupos de usuarios con comportamientos similares.
- 2) Hacer uso de los clusters obtenidos para definir patrones de comportamiento normal dentro de la red, con el fin de establecer el perfil de un usuario normal y recurrente de la red.
- 3) Identificar un usuario anómalo, basado en la comparación del clúster y el perfil de un usuario normal en la red.
- 4) Clasificar el nivel de conocimiento de un usuario malicioso en: 1) básico, aquel usuario cuya actividad en la red sea para investigación referente a seguridad en redes de computadoras,

¿En qué se basa el uso de Machine Learning?

- descarga de scripts y ejecución de estos, o manejo de sistemas operativos enfocados a *pen testing* 2) intermedio o 3) avanzado
- 5) Retroalimentar el *set* de clusters existentes con el comportamiento de algún usuario anómalo, con el objetivo de actualizar los clusters existentes con la nueva información o crear un nuevo clúster con un comportamiento propio definido.

## 1.6. Novedad científica, tecnológica o aportación

Los trabajos relacionados con seguridad informática en especial en el área de detección de intrusos con algoritmos de machine learning como son redes neuronales, Support Vector Machine (SVM), árboles de decisión o lógica difusa, se centran en analizar el desempeño que estos tienen, o la forma en que estos podrían trabajar de manera más eficiente y con menos falsos positivos generalmente a través de la selección de variables. Sin embargo las implementaciones y las mejoras a estos IDS son probadas y evaluadas utilizando datasets de pruebas que muchas veces son ineficientes en un ambiente de producción [5]. Este trabajo propone una nueva implementación que haga uso de un dataset capturado y optimizado en el ambiente de producción, lo que permite que el algoritmo de machine learning implementado funcione de una forma más óptima al ser entrenado con información real del ambiente, que estará monitoreando.

Hay que tener cuidado al usar estos términos.  
Posteriormente tendrás que garantizar, definir, y medir: eficiencia, optimización

✓ Implementación de qué?

Revisión de ideas y fundamentación (argumentos sólidos).

Falta un párrafo para hacer la transición a la siguiente sección del estado del arte.

---

## 2. ESTADO DEL ARTE

---

Los IDS basados en algoritmos de machine learning presentan problemas principalmente en el volumen de información que necesitan manejar para el análisis, las altas tasas de falsos positivos y la difícil implementación de los mismos en ambientes de producción. Múltiples autores [2] [5] [4] [7] abordan estas problemáticas a través de diferentes puntos de vistas como: optimización de algoritmos de machine learning, selección y reducción de variables, y generación de datasets de forma automática.

~~For the reason para los autores~~  
~~modificar redacción.~~

Chaudhari y Prasad [4] establecen los principales problemas que se presentan actualmente en los IDS que son: 1) El origen de los datos con los que un IDS trabaja y el recorrido que estos deben hacer previo a su análisis, lo que causa que en repetidas ocasiones se haga una mala interpretación de los datos debido a la falta de contexto de la información; 2) el exceso de uso de recursos extras para su funcionamiento continuo; y, 3) la dependencia de los IDS hacia otros módulos de que pueden ser susceptibles a ataques y por lo tanto dejar temporalmente o de forma definitiva inhabilitado el sistema.

Redacción (~~original~~) es la idea principal del párrafo?

Autores como Kamarularin [2], Chaudhari [4], Xu [6] y Muda[7], abordan diferentes técnicas de machine learning, mientras Pilabutr [8], Tian [3] y Al-Jarrah [9] se enfocan en la selección de variables, adicionalmente Yamada [5] nos presenta metodologías para la generación de datasets de entrenamiento y pruebas.

## 2.1. Algoritmos de machine learning y sus aplicaciones

Kamarularin [2] propone evaluar el desempeño del algoritmo de árboles de decisión y compararlo con los algoritmos de redes neuronales y SVM con el objetivo de evaluar que algoritmo de estos es más efectivo para la detección de ataques. Utilizando para esto cuatro criterios a evaluar: 1) la precisión en la detección; 2) la tasa de detección; 3) los falsos positivos que presento cada uno de los algoritmos; 4) la precisión para clasificar los ataques en su correspondiente categoría: *Probe*, *Denial of Service* (DoS), *User to root* (U2R), *Remote to local* (R2L). Basado en esto el algoritmo que mejor se desempeña es Arboles de Decisión, mostrando una tase de detección del 98.55%.

Xu, Wang y Gu [6] proponen una nueva forma de hacer un perfilamiento del tráfico de red identificando y analizando clusters de hosts y aplicaciones que tienen un comportamiento similar, reduciendo significativamente el costo del análisis del tráfico. Dicho análisis se logra en cuatro fases:

- 1) Creación de grafos bipartitos para modelar el tráfico de red y representar patrones de comunicación entre hosts origen y hosts destinos.
- 2) Generación de *one-mode projections* a partir de los grafos bipartitos, con el fin de descubrir de forma eficiente relaciones ocultas entre nodos con los mismos vértices y descubrir comportamientos similares.
- 3) Construcción de matrices de similitud basados en las proyecciones previamente obtenidas, donde la caracterización de estas, está basada en el número de host destinos, y hosts orígenes comparten.
- 4) Aplicación de un algoritmo de clustering, para agrupar comportamientos similares basados en la caracterización que poseen las matrices, donde cada clúster consiste en host orígenes que se comunican con hosts destinos con los mismos prefijos de red, servidores, o inclusive clientes.

Con el análisis anterior es posible encontrar un número finito de clusters con un comportamiento definido, que es más fácil de analizar comparado a un análisis de tráfico de hosts individuales,

además de que permite encontrar el comportamiento de un grupo de hosts en el mismo prefijo de red. Xu et. al [6] Logran mostrar que mediante su técnica de perfilamiento es posible crear clusters con comportamientos e intercambio de información diferentes que son consistentes en el tiempo.

Debido a la alta tasa de falsos positivos que los algoritmos de machine learning arrojan, Muda [8] propone crear un método híbrido de detección de intrusos basado en K-Means y la técnica de clasificación OneR, dicho algoritmo funciona a través de la agrupación de datos, basado en el comportamiento que presentan (comportamiento normal y comportamiento malicioso) a través del algoritmo de clustering K-means, una vez agrupados, el algoritmo de clasificación OneR permite identificar las diferentes clases de ataques que están presentes. El desempeño del algoritmo es comparado con otros algoritmos: k-NN, Hierarchical Clustering + SVM y ESC-IDS, mostrando una reducción en la tasa de falsos positivos del 4.56% respecto a los demás algoritmos.

Chaudhari [4] y Tian W. [3] presentan la propuesta de hacer uso del algoritmo de *Particle Swarm Optimization* (PSO), una técnica heurística de optimización cuya funcionalidad está basada en el comportamiento social que tienen las aves cuando vuelan o los peces cuando necesitan protección de depredadores más grandes. En este algoritmo cada partícula o elemento intenta buscar la mejor posición basado en la posición de sus compañeros alrededor. Chaudhari y Tian W. aprovechan las características del algoritmo de auto aprendizaje y rápida convergencia para aprender patrones típicos de un ataque a la red.

## 2.2. Selección de variables para la correcta detección de ataques

Los sistemas de detección de intrusos anómalos (*Anomaly Intrusion detection systems*, AIDS) basan su funcionamiento en los comportamientos normales de los usuarios o las aplicaciones y verifican si el sistema está siendo utilizado de forma correcta, generalmente a través de algoritmos de *machine learning*, desafortunadamente un AIDS debe procesar una cantidad extensa de datos, incluso para una pequeña red [8].

El manejo y análisis de datos incurre en el problema de una gran cantidad de variables, que muchas veces son irrelevantes y/o redundantes teniendo impacto en el desempeño, entendido como el tiempo que toma en ejecutarse, y la eficiencia, la calidad de la detección realizada, de los algoritmos.

Pilabutr, Sonwang y Srinoy, [8] proponen refinar los datos del set de pruebas KDCUP99 usando el análisis de componentes independientes (*Independent component analysis*, ICA), para obtener una mejor clasificación por parte del algoritmo de SVM. ICA selecciona variables que tienen independencia estadística, separando aquellas variables que no tienen impacto en la clasificación final. Pilabutr et.al [8] muestran que esta propuesta tiene un índice de detección de amenazas del 97.51% y un índice de falsos positivos del 2.49%.

Por su parte Tian y Liu [3] analizan el algoritmo de redes neuronales y plantean dos principales problemas en su implementación: 1) la correcta selección de los parámetros para el funcionamiento óptimo; 2) en un ambiente real, existen muchas variables irrelevantes en los datos de prueba, por lo que esta redundancia de información genera cálculos innecesarios, y reduce la capacidad del algoritmo para detectar ataques en un periodo corto de tiempo real. Con el fin de buscar una solución para ambos problema Tian y Liu [3] implementan una red neuronal cuyos argumentos de entrada solo evalúan ocho características de los paquetes y los pesos en las conexiones entre partículas son

definidos a través del algoritmo de PSO. Este método mejora la tasa de detección en un 10% mostrando que es posible mejorar el desempeño y la eficiencia cuando los parámetros son elegidos de forma correcta.

Al-Jarrah, Siddiqui, Elsalamouny [9] proponen un sistema de detección basado en Random Forest (RF), el trabajo consiste en 4 fases principales: 1) selección de *dataset*; 2) la selección de variables; 3) selección del modelo; 4) evaluación. Los autores concluyen que la selección de variables es importante en especial si se toma en cuenta el tamaño de la red donde el volumen de información es mucho y la velocidad en que los datos se generan es demasiado alta.

### 2.3. Datasets autogenerados

Yamada [5] menciona dos problemas principales en el uso de los IDS: 1) el esfuerzo requerido para mantener actualizadas las bases de datos de firmas usadas por los IDS; 2) la implementación de los IDS basados en algoritmos de *Machine Learning* debido a su alta dependencia a *datasets* de entrenamiento y prueba que en la mayoría de los casos no son útiles, ya que el dataset no fue creado con el tráfico de la red en la que será utilizado. Yamada [5] busca una solución para ambos problemas creando una nueva arquitectura de detección de ataques el cual no necesita datos de entrenamiento, utiliza la salida de un IDS basado en firmas para la creación de un *dataset* que posteriormente es utilizado en el entrenamiento de un algoritmo de *machine learning*. Para la evaluación de esta nueva propuesta de arquitectura se utilizó SNORT como el IDS y el árboles de ejecución decisión como el algoritmo de *machine learning*. Esta nueva arquitectura fue evaluada utilizando peticiones HTTP con los datos de 1999 DARPA y los datos generados por el escáner de seguridad Nessus, logrando detectar 3 tipos de ataques que no estaban incluidos en DARPA 1999 y 2687 ataques que el IDS SNORT no fue capaz de detectar.

Es posible encontrar una gran variedad de trabajos combinando técnicas de *machine learning* y detección de intrusos, sin embargo es posible explorar aún más el área de clustering, ya que por la naturaleza de la misma es posible generar clusters cuyo contenido sea dinámico y no esté restringido a los comportamientos anómalos y legítimos.

## BIBLIOGRAFÍA

- [2] K. A. Jalil and N. M. Masrek, "Comparison of Machine Learning Algorithms Performance in Detecting Network," *2010 International Conference on Networking and Information Technology*, p. 6, 2010.
- [3] Al-Jarrah, Siddiqui, Elsalamouny, Yoo, Muhaidat and Kim, "Machine-Learning-Based Feature Selection Techniques for LargeScale," *2014 IEEE 34th International Conference on Distributed Computing Systems Workshops*, p. 5, 2014.
- [4] B. Chaudhari and R. Prasad, "Particle Swarm Optimization Based Intrusion Detection for Mobile," *9th International Conference on Computer Engineering and Applications*, p. 6, 2015.
- [5] CISCO, "Annual Security Report," 2015.
- [6] L. Zhe, S. Weiqing and W. Lingfeng, "A neural network based distributed intrusion detection system on cloud platform," *IEEE 2nd International Conference on Cloud Computing and Intelligence Systems*, vol. 1, pp. 75-79, 2012.
- [7] O. Y. S. A. E. M. Y. P. D. M. S. & K. K. Al-Jarrah, "Machine-learning-based feature selection techniques for large-scale network intrusion detection," *IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pp. 177-181, 2014.
- [8] S. S. P. & S. S. Pilabutr, "Integrated soft computing for Intrusion Detection on computer network security," *Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference*, pp. 559-563, 2011.
- [9] W. & L. J. Tian, "A new network intrusion detection identification model research.," *Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference*, vol. 2, pp. 9-12, 2010.
- [10] A. M. Y. T. K. & T. T. Yamada, "Intrusion detection system to detect variant attacks using learning algorithms with automatic generation of training data," *International Conference on Information Technology: Coding and Computing (ITCC'05)-Volume II*, vol. 1, pp. 650-655, 2005.
- [11] B. S. & P. R. S. Chaudhari, "Particle Swarm Optimization Based Intrusion Detection for Mobile Ad-hoc Networks".

