Identificación de ataques en redes de cómputo utilizando redes neuronales artificiales

Javier Alberto Carmona Troyo
Departamento de sistemas y Computación
Instituto Tecnológico de La Paz
La Paz B.C.S., México
Email: javier.carmona@itlp.edu.mx

Luis Armando Cárdenas Florido
División de Estudios de Posgrado E Investigación
Instituto Tecnológico de La Paz
La Paz B.C.S., México
Email: armando.cardenas@itlp.edu.mx

José Tadeo Rodríguez Solano
Departamento de sistemas y Computación
Instituto Tecnológico de La Paz
La Paz B.C.S., México
Email: tadeo@itlp.edu.mx

Armando Yuen Coria
Departamento de sistemas y Computación
Instituto Tecnológico de La Paz
La Paz B.C.S., México
Email: yuen@itlp.edu.mx

Alberto Ibarra Berber
Departamento de sistemas y Computación
Instituto Tecnológico de La Paz
La Paz B.C.S., México
Email: berber@itlp.edu.mx

***Resumen*—El presente trabajo busca determinar un método rápido, económico y eficiente para la detección de intrusos en redes de área local. Esta forma de detección que se presenta está basada en técnicas de inteligencia artificial, específicamente en redes neuronales Artificiales.**

La investigación se dividió en cuatro partes. Primero se analizaron diferentes contextos de problemáticas a las que nos enfrentamos en cuanto a seguridad informática. Posteriormente se identificaron los elementos y variables principales que interviene en el tráfico de redes, así como la aplicación de diferentes técnicas como las redes neuronales artificiales.

Posteriormente se presenta una propuesta de solución del problema y el planteamiento de los procedimientos desarrollados. Se concluye con los resultados obtenidos de las observaciones o pruebas realizadas.

***Keywords*—Network Security; Intruder Detection; patterns; artificial intelligence; fuzzy logic; RNA; IDS.**

I. Introducción

Hoy en día, se visualiza la necesidad de desarrollar e implementar nuevos métodos y controles mas eficientes para mantener la información más segura y mas cuando esta es transmitida a través de las tecnologías de información. Esto es, debido al avance de la tecnología en el campo de las comunicaciones y la importancia de mantener la confidencialidad, la integridad y la disponibilidad de los datos, siendo estos los activos más importantes de cualquier organización. Estos controles de seguridad se plantean en redes de área local interna o en las redes perimetrales que dan salida a dicha información. Al aplicar los controles de seguridad necesarios, la empresa, institución o compañía, se vería beneficiada ya que sus activos permanecerían seguros, evitando así, grandes pérdidas económicas o de imagen. Los piratas informáticos, al mismo tiempo en que van incorporando nuevas tecnologías, van adquiriendo más conocimiento acerca de las formas o métodos para llevar a cabo un vector de ataque sobre alguna vulnerabilidad que se tenga en las redes o aplicaciones de las empresas. Estos vectores de ataque que llevan a cabo los cibercriminales pueden ocasionar la pérdida de información, el secuestro digital o atentar contra la privacidad de los usuarios de la información o datos de la empresa o institución misma.

II. Marco Teórico

II-A. Seguridad de la información

La seguridad de la información engloba un conjunto de técnicas y medidas para controlar todos los datos que se manejan dentro de una institución y asegurar que no salgan de ese sistema establecido por la empresa. Principalmente este tipo de sistemas se basan en las nuevas tecnologías, por tanto la seguridad de la información resguardará los datos que están disponibles en dicho sistema y a los que solo tendrán acceso usuarios autorizados. Por otro lado, tampoco se podrán hacer modificaciones en la información a no ser que sea de la mano de los usuarios que tengan los permisos correspondientes.
La seguridad de la información debe responder a tres cualidades principales:

- Crítica

- Valiosa
- Sensible

Por un lado debe ser crítica, ya que es una pieza fundamental para que la empresa pueda llevar a cabo sus operaciones sin asumir demasiados riesgos. También debe ser valiosa, puesto que los datos que se manejan son esenciales para el devenir del negocio y finalmente tiene que ser sensible, ya que al sistema solo podrán acceder las personas que estén debidamente autorizadas. Además, también hay que tener en cuenta que la seguridad de la información debe hacer frente a los riesgos, analizarlos, prevenirlos y encontrar soluciones rápidas para eliminarlos si se diera el caso [1].

II-B. Seguridad informática

La seguridad informática es un conjunto de normas, procedimientos y herramientas, que tienen como objetivo garantizar la disponibilidad, integridad, confidencialidad y buen uso de la información. La seguridad en un ambiente de red es la habilidad de identificar y eliminar vulnerabilidades. Estas normas incluyen horarios de funcionamiento, restricciones a ciertos lugares, autorizaciones, denegaciones, perfiles de usuario, planes de emergencia, protocolos y todo lo necesario que permita un buen nivel de seguridad informática minimizando el impacto en el desempeño de los trabajadores y de la organización en general y como principal contribuyente al uso de programas realizados por programadores [2].

II-C. Sistemas IDS (Sistema identificador de intrusos)

Como herramienta para protección de los intrusos en redes de área local o en infraestructuras de red, se crearon los IDS o sistemas identificadores de intrusos [3]. Estos sistemas, normalmente son colocados en la zona perimetral de un sistema de red. Estos sistemas detectan las posibles amenazas que pueden provocar un ataque a la infraestructura, detectando e identificando intrusos en la red.

Desde que el concepto de Detección de Intrusos fue propuesto por Anderson en 1980 [4]., muchas técnicas para su implementación han sido reportadas y estudiadas. Las técnicas para los sistemas de Detección de Intrusos (o intrusiones) IDS, se clasifican generalmente en dos categorías, la detección de mal uso y la detección de anomalías [5].

Los sistemas de detección de intrusos se pueden clasificar de diversas maneras, según diferentes parámetros que se apliquen; para esta investigación se tomó la clasificación propuesta por Iren Lorenzo Fonseca et al [6].

II-D. Inteligencia artificial

Elaine Rich [7] menciona que la inteligencia artificial, estudia como lograr que las máquinas realicen tareas que, por el momento, son realizadas mejor por los seres humanos.

Así mismo, Margaret Rouse [8] define a la inteligencia artificial como la simulación de procesos de inteligencia humana por parte de máquinas, especialmente sistemas informáticos. Estos procesos incluyen el aprendizaje (la adquisición de información y reglas para el uso de la información), el razonamiento (usando las reglas para llegar a conclusiones aproximadas o definitivas) y la auto corrección [8].

II-E. Técnica de análisis de componentes principales

El análisis de componentes principales o PCA, es una técnica estadística que permite seleccionar información en un conjunto de n variables de interés en m nuevas variables independientes. En relación con las variables seleccionadas, cada una debe explicar una parte especifica de la información y mediante combinación lineal, otorgan la posibilidad de resumir dicha información en pocos componentes que reducen la dimensión del problema [9].

III. JUSTIFICACIÓN

Este tipo de investigación resulta significativa ya que, en el ámbito de las redes de área local, no se tienen los mecanismos de seguridad implementados en su máxima expresión, esto es debido al alto costo que se tiene para la implementación de Sistemas de Identificación de Intrusos o IDS (Intrusion Detection Systems) capaces de detectar en tiempo real una amenaza. Por ello, se propone un mecanismo utilizando técnicas de inteligencia artificial para la detección de intrusos, el cual resulte de bajo costo y pueda escalar hacia un mecanismo mucho más potente y con las mismas herramientas. Estas técnicas de inteligencia artificial están basadas en redes neuronal artificiales las cuales serán capaces de detectar un ataque al protocolo TCP como por ejemplo un mapeo de puertos NMAP.

Existen diversos IDS disponibles actualmente, algunos comerciales y otros de licencia libre. Uno de los más utilizados es SNORT, el cuál trabaja en base a reglas bien definidas. En este caso, la novedad será que, aplicando las técnicas propuestas de inteligencia artificial, permitirá no solo detectar un NMAP, si no también detecten cualquier invasión o ataque persistente a la capa de transporte.

En el caso de muchas instituciones federales de educación superior, no cuentan con mecanismos de seguridad apropiados para la detección y paro de ataques. Es por ello, que nace la idea de generar por medio de inteligencia artificial, un IDS capaz de detectar ataques a la infraestructura de red de área local y a su vez, sea funcional.

Este tipo de trabajos es importante, ya que no hay ningún sistema que pueda ser seguro en cuanto a ataques cibernéticos. Muchas veces tenemos que los sitios de red que manejamos tienen vulnerabilidades las cuales pueden ser aprovechadas por delincuentes cibernéticos.

México es uno de los principales países de Latinoamérica que frecuentemente es atacado en su infraestructura tecnológica, es por ello que debemos de hacer uso de la tecnológica de computación para poder detectar a tiempo a los posibles intrusos de nuestros sistemas para así poder poner un remedio

o solución a esto.

Esto impactará de sobre manera a las organizaciones o empresas ya que la seguridad de sus activos se podrá mantener adecuadamente, evitando grandes pérdidas de los mismos o pérdidas económicas las cuales podrán llevar a las instituciones a un estado de poca credibilidad.

IV. METODOLOGÍA

Para llevar a cabo el análisis de información que es trasmitida a través de una red, se tomaron tramas de datos las cuales fueron capturadas durante una petición simple y otras capturadas durante un ataque NMAP para después obtener las características importantes de ellas, las cuales fueron procesadas por medio de redes neuronales artificiales. Esta técnica ayuda a determinar si se está llevando a cabo un ataque a la infraestructura de red.

El modelo utilizado para llevar acabo el procesamiento de los datos obtenidos de las tramas de red, sigue los siguientes pasos:

- Obtención de datos: Consistió en tomar información captada por un sniffer de red, se utilizaron las herramientas Wireshark y Tcpdump.
- Datos a seleccionar: Se seleccionaron tramas TCP la cuales eran por el tipo de ataques, el punto de partida para siguientes pasos de análisis.
- Extracción de características deseables: Se obtuvieron de las tramas información relevante cuyas características representan a las de un ataque.
- Normalizar y reducir las dimensiones: Se llevó a cabo un preprocesamiento de información sobre el conjunto de datos el cual sirvió para alimentar un conjunto de datos.
- Clasificación: Se llevó a cabo una clasificación sobre un conjunto de datos tomados en el paso anterior.
- Conjunto de datos (DataSet): Para esta etapa, se tomó una parte de la base de datos para entrenar la red y otra parte para validar que esté funcionando correctamente.
- Clasificador: Dependiendo las reglas suministradas, se determina si existe un ataque o no.

V. ADQUISICIÓN DE DATOS

Para este trabajo fue necesario obtener un conjunto de datos sobre el cual se pudiera trabajar, y que permitiera entrenar el algoritmo de aprendizaje. Para esta tarea se utilizaron los programas de distribución gratuita Wireshark para la captura de tramas, VMware para la virtualización, NMAP para escanear puertos, hping3 (bajo plataforma Linux) como generador de ataques de denegación de servicio (DoS). Con estas dos últimas herramientas se simularon ataques en una red compuesta por máquinas virtuales Linux (Kali Linux 2.0), Windows XP SP2 y Windows 7.

Según el modelo propuesto, es necesario generar un conjunto de datos a partir de las señales de ataque capturados con Wireshark. Este conjunto de datos, fue exportado hacia archivos de arreglos en C los cuales posteriormente fueron procesados para ser analizados.

Los datos fueron tomados de una verificación de puertos con NMAP, de un ataque o denegación de servicios con ping y otro ataque con hping2, todos utilizando Kali Linux como máquina atacante. Una muestra de tramas de formato en “C” se muestra como sigue.

/* Frame (58 bytes) */ static const unsigned char pkt1[58] = 0x00, 0x0c, 0x29, 0x20, 0x54, 0xb8, 0x00, 0x0c, /* ..) T... */ 0x29, 0x73, 0xa7, 0x4d, 0x08, 0x00, 0x45, 0x00, /*)s.M..E. */ 0x00, 0x2c, 0x89, 0xf2, 0x00, 0x00, 0x27, 0x06, /* .,....'. */ 0x3b, 0x34, 0xc0, 0xa8, 0x26, 0xab, 0xc0, 0xa8, /* ;4..&... */ 0x26, 0xaa, 0xed, 0x87, 0xd4, 0x38, 0x66, 0x4b, /* &....8fK */

0xf9, 0xbc, 0x00, 0x00, 0x00, 0x00, 0x60, 0x02, /* ......‘. */ 0x04, 0x00, 0xa3, 0xb7, 0x00, 0x00, 0x02, 0x04, /* ..... */ 0x05, 0xb4 /* .. */ ; /* Frame (60 bytes) */ static const unsigned char pkt2[60] = 0x00, 0x0c, 0x29, 0x73, 0xa7, 0x4d, 0x00, 0x0c, /* ..)s.M.. */ 0x29, 0x20, 0x54, 0xb8, 0x08, 0x00, 0x45, 0x00, /*) T...E. */ 0x00, 0x28, 0x1b, 0x93, 0x40, 0x00, 0x80, 0x06, /* .(..@... */ 0x10, 0x97, 0xc0, 0xa8, 0x26, 0xaa, 0xc0, 0xa8, /* ....&... */ 0x26, 0xab, 0xd4, 0x38, 0xed, 0x87, 0x00, 0x00, /* &..8.... */ 0x00, 0x00, 0x66, 0x4b, 0xf9, 0xbd, 0x50, 0x14, /* ..fK..P. */ 0x00, 0x00, 0xbf, 0x60, 0x00, 0x00, 0x00, 0x00, /* ...‘.... */ 0x00, 0x00, 0x00, 0x00 /* .... */ ;

Los anteriores paquetes fueron ataques llevados a cabo desde una máquina con Kali Linux hacia una máquina virtual con sistema operativo Windows 7. Este ataque se llevó a cabo utilizando un escaneo de puertos con NMAP.

Estos datos fueron sujetos a un procesamiento para identificar patrones de ataque por medio del programa llamado “analizador de Tramas”, este programa fue realizado en lenguaje C# y lo que realiza es una selección de tramas TCP tomando todas aquellas cuyo valor del protocolo sea O6H.

Con la ayuda del programa analizador de tramas se convirtieron los valores hexadecimales de la captura en valores decimales para posteriormente pasarlos a un formato delimitado por comas. Cabe señalar que la adquisición y procesamiento de los datos adquiridos no se hace en tiempo real, sino que hay una serie de pasos para poder implementar la información como entrada de datos a la red neuronal artificial.

VI. NORMALIZACIÓN Y REDUCCIÓN DE DIMENSIONES

El proceso de normalización nos permite escalar los valores para que se establezcan en un rango específico y así facilitar

la manipulación de los datos. La normalización en este trabajo se llevó a cabo usando el proceso Z-Score de Matlab®, el cual usa la media y desviación estándar, donde el nuevo valor se halla de la siguiente forma:

$$V' = \frac{V - \bar{V}}{\sigma} \tag{1}$$

Donde:
$V' = Valor\ nuevo$
$V = Valor\ actual$
$\bar{V} = Promedio$
$\sigma = Desviación\ estándar$

Una vez normalizados cada uno de los conjuntos de datos se realizó el proceso de obtención de componentes principales, con el fin de depurar las características y así reducir las variables en cada conjunto de datos (de nuevo este proceso se realizó con el programa Matlab®). Un segmento de los datos normalizados se muestra en la figura 1:

1	2	3	4	5	6	7	8
-0.6320	0.1200	-0.7537	0.7367	2.9103	-2.4014	-0.0485	0.0812
-0.6320	0.1200	1.3511	-1.3687	-2.3700	2.9914	-0.0605	-0.1221
-0.6320	0.1200	1.3511	-1.3687	-2.3700	2.9914	-0.0591	-0.1221
-0.6320	0.1200	-0.7537	0.7367	2.9103	-2.4014	-0.0605	0.0810
-0.6313	0.1200	1.3508	-1.3532	-1.0249	2.1104	-0.0605	0.0812
-0.6320	0.1200	-0.7382	0.7365	2.0476	-1.0276	-0.0605	0.0812
-0.6313	0.1200	1.3509	-1.3532	3.7208	1.1737	-0.0605	0.0812
-0.6320	0.1200	-0.7382	0.7365	1.1305	3.8192	-0.0605	0.0812

Figura 1. Datos Normalizados

Al ser cada conjunto de datos de n x 8 (n, ya que cada conjunto difiere en la cantidad de tramas), el resultado de aplicar el proceso de componentes principales redujo cada uno de estos a una matriz de 8 x 8, donde cada columna representa un componente principal. Estos están ordenados en forma decreciente de importancia, de manera que el primer componente proporciona (o recoge) la mayor variabilidad posible de los datos.
Haciendo uso de la herramienta Neural Network de Matlab® se alimentó la red neuronal (figura 2) para clasificación de la forma que sigue: De 200 datos seleccionados de una trama de

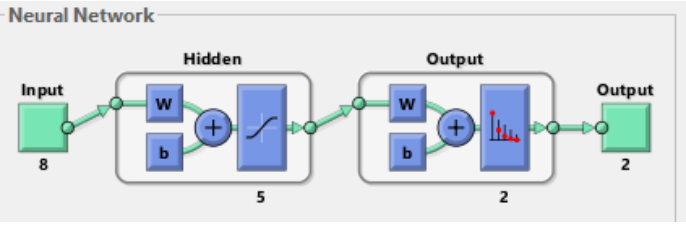

Figura 2. Red neuronal artificial generada con el toolbox de Matlab(R)

1500, se utilizó:

- 120 datos de entrenamiento
- 0 datos de prueba

En este caso, se utilizó para el entrenamiento, 80 datos correspondientes a tráfico normal y 40 datos para tráfico tipo *flooding*. Para la prueba se utilizó, 60 datos de tráfico normal y 20 de tráfico tipo flooding. Cabe señalar que, para este trabajo, se tomaron datos de ataque tipo denegación de servicios (DoS) y diferentes configuraciones de la red neuronal.

VII. ANÁLISIS DE RESULTADOS

Durante las pruebas realizadas con la red neuronal artificial se utilizó una red con aprendizaje supervisado. Se realizaron diferentes pruebas utilizando valores distintos en cada una de ellas con el objetivo de buscar resultados óptimos de la aplicación de la red neuronal.

Se armaron conjuntos de datos con los campos seleccionados a los cuales se les realizó un proceso de normalización y reducción de dimensionalidad por medio del procedimiento de análisis de componentes principales, dando como resultado conjuntos estandarizados en matrices de 8 x 8 (8 filas por 8 columnas) para un total de 64 elementos por conjunto de datos que fueron usados como entradas al clasificador compuesto por una red neuronal con una capa oculta de neuronas variables y una salida de dos neuronas.

VII-A. Caso 1

Para el primer caso de estudio se seleccionaron los siguientes valores (figura2):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 5

Resultado: En este caso, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 63 % de clasificación, así como del 37 % de tráfico de DoS. En general, la clasificación en esta prueba fue del 65 % de clasificación correcta y 35 % de clasificación incorrecta.

VII-B. Caso 2

En el segundo caso de estudio, se seleccionó lo siguiente (figura3):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 6

Resultado: En este caso, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 81 % de clasificación para tráfico normal. Se obtuvo un 77 % de tráfico DoS con respecto al esperado. Se obtuvo en lo general una clasificación correcta del 55 % y 45 % de clasificación incorrecta.

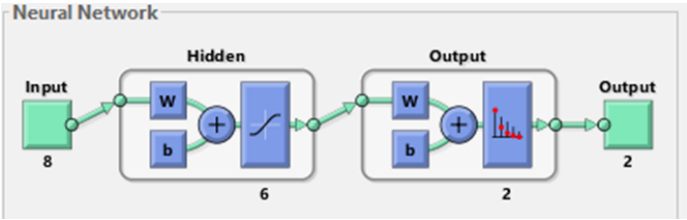

Figura 3. Red neuronal artificial generada con el toolbox de Matlab(R)

VII-C. Caso 3

En el tercer caso, la red neuronal artificial con la que se trabajó fue la siguiente:

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 10

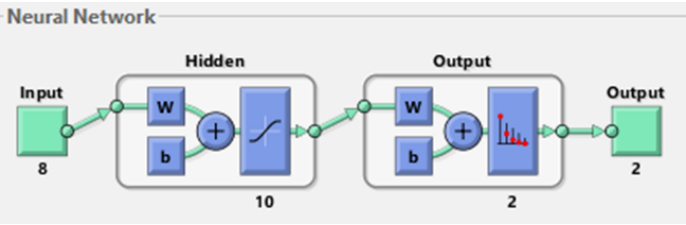

Figura 4. Red neuronal artificial generada con el toolbox de Matlab(R)

Resultado: Después del entrenamiento y dado los datos de prueba, del 80 % de la muestra tomada como tráfico normal, se obtuvo el 81 % de clasificación para datos de tráfico normal. Se obtuvo un 100 % de tráfico DoS con respecto al esperado. Se obtuvo en lo general una clasificación correcta del 95 % y 5 % de clasificación incorrecta.

VII-D. Caso 4

Se realizó una última prueba con los mismos datos de entrada y salida, pero cambiando el número de neuronas en la capa intermedia, quedando la red como sigue (figura5):

- La función utilizada fue: *FeedForward Network*
- Número de entradas: 8
- Número de salidas: 2 (las cuales sirven para clasificar un ataque de un tráfico normal)
- Número de neuronas en la capa intermedia: 7

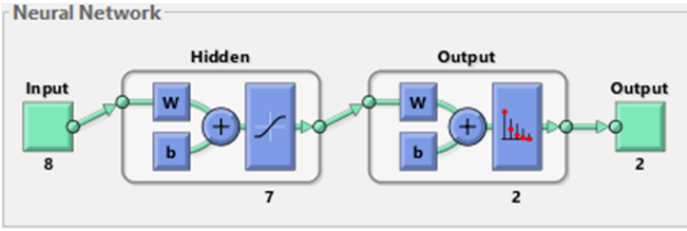

Figura 5. Red neuronal artificial generada con el toolbox de Matlab(R)

Resultado: El resultado obtenido en esta última prueba fue idéntico respecto al análisis con 10 neuronas, no se observó variación en los valores obtenidos.

VIII. CONCLUSIONES

La aplicación herramientas basadas en redes neuronales para la identificación de ataques en redes de cómputo permite detectar peligros que no son considerados en los entornos comunes de operación.

Se analizaron varios casos de prueba con datos recolectados de una red real en operación y se logró obtener un modelo práctico y funcional que puede ser utilizado como apoyo en la identificación de tráfico de red no deseado. La aplicación de redes neuronales para la detección de ataques en las redes de cómputo puede apoyar en gran medida a los sistemas tradicionales dada su capacidad de aprendizaje y adaptación.

El modelo construido basado en redes neuronales artificiales, permite detectar paquetes peligrosos que forman parte de un ataque a la red de cómputo, a partir del conocimiento previamente adquirido con el análisis de anteriores paquetes, lo que no sucede con los sistemas no basados en redes neuronales.

El éxito de un sistema de redes neuronales artificiales para la detección de ataques en una red de cómputo se fundamenta en el adecuado entrenamiento y en la selección de un universo suficiente de paquetes de la red donde se utilice el sistema. Esto para que puedan detectar patrones o comportamiento que para otros sistemas basados en otras técnicas son descartados o no detectados.

REFERENCIAS

[1] B. S. OBS. (2017) Seguridad de la información. [Online]. Available: https://www.obs-edu.com/int/blog-investigacion/sistemas/seguridad-de-la-informacion-un-conocimiento-imprescindible
[2] Icorp. (2018) Seguridad informática. [Online]. Available: http://www.icorp.com.mx/solucionesTI/seguridad-informatica/
[3] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (idps)," National Institute of Standards and Technology, Tech. Rep., 2012.
[4] J. P. Anderson, "Computer security threat monitoring and surveillance," *Technical Report, James P. Anderson Company*, 1980.
[5] C. M. Bonilla, "Sistema neuronal de detección de intrusos," *Tendencias en ingeniería de software e inteligencia artificial*, 2008.
[6] I. Lorenzo-Fonseca, F. Maciá-Pérez, F. J. Mora-Gimeno, R. Lau-Fernández, J. A. Gil-Martínez-Abarca, and D. Marcos-Jorquera, "Intrusion detection method using neural networks based on the reduction of characteristics," in *International Work-Conference on Artificial Neural Networks*. Springer, 2009, pp. 1296–1303.
[7] E. Rich, K. Knight, P. A. González Calero, T. Bodega *et al.*, *Inteligencia artificial*, 1994.
[8] R. Margareth. (2017) Inteligencia artificial o ia. [Online]. Available: https:/https://searchdatacenter.techtarget.com/es/definicion/Inteligencia-artificial-o-AI
[9] D. Peña, *Análisis de datos multivariantes*. McGraw-Hill España, 2013.