Avances en Sistemas de Detección de Intrusos con Auto-aprendizaje - Revisión Sistemática de la Literatura

Javier Maldonado
Laboratorio de Computación de Alto Rendimiento (LCAR)
Universidad Nacional Experimental del Táchira (UNET)
San Cristóbal, Venezuela
jmaldo@unet.edu.ve

***Resumen*—La presente investigación está orientada en hacer una revisión sistemática de la literatura (SLR), de los avances realizados en el área de los Sistemas de Detección de Intrusos (IDS), con técnicas de aprendizaje automático. Determinar tendencias y características relevantes de los modelos, experimentos e implementaciones realizadas en el área. Se logró evidenciar la existencia de múltiples estrategias y aproximaciones, siendo las más notorias las técnicas de machine learning, luego las metaheurísticas como apoyo a los procesos de aprendizaje y clasificación. Ésta revisión sirve como punto de partida para investigar acerca de técnicas de machine learning y/o metaheurísicas que permitan resolver problemas de detección/clasificación de intrusos en el campo de ciberseguridad.**

***Palabras Clave*—Sistemas de Detección de Intrusos (IDS), Auto-aprendizaje, Machine Learning, Metaheurísticas, Systematic Literature Review.**

I. Introducción

Las redes de datos se han convertido en un soporte para muchas actividades en diferentes áreas de la evolución humana y social. Así como es un apoyo para compartir gran cantidad de información, también se ha convertido en un punto de interés para usuarios maliciosos, ya sea en la obtención de un beneficio propio a partir de información o afectando el funcionamiento de servicios objeto de ataque, con algún fin propio, en detrimento de las funcionalidades y disponibilidad para usuarios legítimos. Se hace necesario crear mecanismos de defensa que puedan mitigar éste riesgo latente por el simple hecho de estar conectado, por ejemplo, a Internet o a una intranet corporativa, donde pueden existir sistemas de autenticación y seguridad previos, sin embargo no la hace exenta de los riesgos mencionados.

Existen numerosos mecanismos y dispositivos de seguridad implementados para tratar de asegurar, en lo posible, las actividades de los usuarios y prestadores de servicios. Uno de estos son los *Sistemas de Detección de Intrusos* o también conocidos por su denominación en inglés *Intrusion Detection Systems (IDS)*, el cual se define como el proceso de identificar y responder a actividades maliciosas dirigidas a computadores y recursos de red [1]. Existen distintas implementaciones y clasificaciones dependiendo de su funcionamiento, elementos a vigilar, modo de vigilancia, sin embargo todos ellos dependen de la actualización de su base de conocimiento o firmas para actualizar sus habilidades de detección.

Los ataques a los sistemas informáticos están en constante evolución, por lo cual los IDS también deben actualizar sus reglas o firmas para mantener su capacidad de detección, deseablemente al menos tan rápido como evolucionan las amenazas. Sin embargo esto puede ser un problema, las actualizaciones generalmente dependen de administradores, fabricantes o licencias. Si éstas demoran en ser incorporadas a los IDS en funcionamiento, pudiera ocurrir que ataques nuevos no podrían ser detectados oportunamente, debido a que no tienen la información suficiente para hacerlo. Es por esto que surgen nuevas tendencias de implementar componentes de aprendizaje automático (ej: machine learning, computación evolutiva, metaheurísticas o combinaciones de éstas), para mitigar ésta debilidad, haciéndolos entonces capaces de actualizar la información interna de los ataques conocidos, capturando y analizando el tráfico de red durante su funcionamiento para luego incorporar información de ataques detectados por ellos mismos.

El presente trabajo se centrará en una revisión sistemática de la literatura referente a IDS con componentes de aprendizaje automático. Se hará una exploración de trabajos relacionados, distintas aproximaciones y técnicas empleadas en la clasificación de los ataques, así como también indicadores de rendimiento obtenido para identificar algunos mecanismos adecuados en el procesamiento de los datos de red. La intención de realizar esta revisión es fijar un punto de partida para la búsqueda de una estrategia que pueda servir de solución para algún problema específico del campo de la detección/clasificación de ataques o intrusiones.

El documento está organizado en cinco partes: Sección II enfoque de la investigación, donde se plantea la motivación y algunos conceptos básicos que sirven de punto de partida, así como también las preguntas de investigación, descripción de como se hará la recolección de información y los criterios bajo los cuales serán tomadas en cuenta las publicaciones. Sección III resultados obtenidos, se presenta un sumario de las publicaciones efectivamente incluidas en la revisión, las áreas de interés identificadas y un resumen del resultado

de la investigación. Sección IV exploración, se responden las preguntas planteadas soportado en los datos recolectados durante la revisión, es posible identificar posibles temas de desarrollo en los trabajos revisados. Sección V conclusiones, se hace el cierre de las ideas colectadas en la investigación identificando avances y aportes en el área de estudio.

II. Enfoque de la Investigación

Un IDS es un proceso o dispositivo activo que analiza la actividad del sistema (HIDS, Host Intrusion Detection System), de red (NIDS, Network Intrusion Detection System) o de un conjunto de puntos de interés en una red (DNIDS, Distributed Network Intrusion Detection System), en busca de actividad no autorizada o maliciosa, con el fin de alertar a los administradores y evitar daños mayores en los sistemas. Sin embargo, los IDS no tienen la capacidad de evitar el ataque, sólo se limitan a alertar del problema detectado al momento de su ocurrencia [2].

Con el fin de profundizar en el área de los IDS con aprendizaje automático, se realiza una revisión sistemática de la literatura. En este sentido, se plantea las siguientes preguntas de investigación las cuales servirán para guiar la revisión:

- P1: ¿Cuáles son las técnicas más relevantes utilizadas en la adquisición de conocimiento de los IDS con autoaprendizaje?
- P2: ¿Cuáles son los ataques más comunes clasificados por los IDS, tanto en su fase de entrenamiento como en operación?
- P3: ¿Qué técnicas son las utilizadas en los comportamientos registrados?

Para responder P1: se revisaran las técnicas de inteligencia artificial más utilizadas en las implementaciones halladas durante la revisión bibliográfica. Con respecto a P2: se identificarán los ataques más comunes detectados por los experimentos vistos en la literatura explorada. Y en cuanto a P3: se obtiene la información agrupando los ataques o comportamientos detectados contra las técnicas evidenciadas en la revisión.

II-A. Términos de Búsqueda Utilizados

Para conseguir los artículos y trabajos relacionados con el tema objeto de este estudio, se realizó una búsqueda sistemática en distintas fuentes de información científica utilizando algunos términos relevantes al tema, con el fin de conseguir documentos referentes al tema en estudio. Se utilizaron términos de búsqueda en inglés, dado que es el idioma donde más fuentes y objetos de información se consiguen:

- Intrusion detection
- Self - Learning
- Security

Se utilizaron los términos con conectores lógicos AND y OR para delimitar los resultados a los IDS que tuviesen mecanismos de autoaprendizaje incorporados. Se alternaron en función de delimitar las aplicaciones halladas en los campos de HIDS (Host IDS), NIDS (Network IDS) y DIDS (Distributed IDS).

II-B. Criterio de Inclusión de los Artículos

Con el fin de delimitar los resultados obtenidos en la búsqueda de información, se establecen criterios para incluir los artículos considerados como relevantes en la revisión, esto con base a los aportes en los temas que pudieran refinar la investigación y orientarla a trabajos que permitan identificar aquellos con mayores aportes en el área:

- Publicaciones después de Enero de 2013.
- Publicaciones de IDS con técnicas de autoaprendizaje orientadas a la detección de ataques a nivel de aplicación o red.
- Publicaciones donde se detalle los ataques detectados y la efectividad del método empleado.

II-C. Data Recolectada

La información recolectada de cada uno de los trabajos y que se mostrará en un resumen es la siguiente:

- Referencia
- Resumen
- Técnicas de IA utilizadas
- Tipos de comportamientos detectados/reportados (TCD)

III. Resultados Obtenidos

En ésta sección se condensa la información obtenida a partir de los trabajos de investigación hallados según los criterios definidos para ejecutar la revisión sistemática. Se hicieron consultas en 3 fuentes de información científica:

- IEEEXplore
- ScienceDirect
- Elseiver

Utilizando las palabras de búsqueda y el rango de fecha definido, se obtuvo un conjunto de documentos de las fuentes de información consultadas. En la Tabla I se pueden ver los primeros resultados obtenidos.

Tabla I: Resultados de la Consulta

Fuente	Resultados
IEEExplorer	27
Elsevier	134
ScienceDirect	755

A continuación se hizo una primera revisión para depurar los resultados hallados, donde se descartaron los artículos no relacionados con el tema, ya sea porque pertenecen a otra área o porque fueron incluidos por coincidencia parcial de la búsqueda pero sin relevancia al tema en estudio. Luego de ésta primera depuración quedaron 58 documentos, los cuales fueron seleccionados para una segunda revisión.

Para determinar la pertinencia del documento, se procedió a revisar el abstract y los resultados mostrados de las técnicas utilizadas. Cumpliendo con los criterios de inclusión establecidos, se consiguieron un total de 29 artículos de investigación, los cuales están agrupados en 11 conferencias y 18 revistas. Los detalles se resumen la Tabla II.

TABLA II: RESUMEN DE LOS ARTÍCULOS ENCONTRADOS

Ref	Resumen	Tec. IA	TCD
[3]	Utilizan un método híbrido para la detección y selección de características en el tráfico de red, tomando éstas últimas como las unidades más pequeñas de clasificación. Su combinación definen un comportamiento positivo o negativo.	CON, CFS, IG y GR, Naïve Bayes, C4.5.	Dos, U2R, R2L, Prob
[4]	Emplea una aproximación basada en agentes colaborativos utilizando el paradigma de sistema inmune artificial, como mecanismo de implementación de un sistema IDS distribuido. El paradigma de AIS es muy efectivo en la detección de anomalías.	Sistema inmune artificial, agentes colaborativos	Detección de anomalías
[5]	Usa una técnica de aprendizaje semi-supervisado. Utiliza conjuntos de datos no etiquetados asistido por un algoritmo de aprendizaje supervisado. Una red neuronal de una capa proporciona el vector de clasificación a ser procesado por el controlador difuso.	Red neural de una sola capa y un controlador difuso.	Dos, U2R, R2L, Prob
[6]	Un systematic review acerca de los IDS en sistemas de cloud computing.		
[7]	Se hace una revisión completa de un conjunto de papers mediante un systematic review.		
[8]	Systematic review del área de IDS.		
[9]	Utiliza un método bio-inspirado de sistema inmune humano apoyado en lógica difusa. Está construido en módulos especializados en distintas tareas. Se hacen comparativas con otras técnicas disponibles en términos de precisión en la detección.	Co-FAIS	DDoS
[10]	Hace una revisión de distintas técnicas de detección y clasificación de anomalías en redes de datos.	SVM, BN, NN, Rule-based.	
[11]	Está basado en una técnica de sistema inmune artificial para crear un DNIDS	Aprendizaje de máquina no supervisado, Algoritmo de selección negativa, Naïve bayes, SVM, J48, NSA-GA	Detección de anomalías.
[12]	Utiliza un método de votación mayoritaria para aceptar un comportamiento, basado en una serie de características que modelan cada uno de los comportamientos detectados.	Análisis de correlación, votación, comportamientos normales de las características.	Detección de anomalías
[13]	Se enfoca en la creación de un IDS no supervisado, capaz de generar sus reglas a partir de la clusterización de tráfico. Se propone el uso de heurísticas para asignar etiquetas a los datos observados tomando en cuenta los datos de origen y contenido de la petición HTTP.	Heurísticas	Ataques HTTP
[14]	Utilizan una combinación de técnicas de machine learning para mitigar el volumen de falsos positivos.	K-means, J48, self-organized map	DoS, U2R, R2L y Probing.
[15]	Plantea un esquema híbrido, donde combina maquina de soporte vectorial, algoritmo genético y mapa de características auto-organizado para incrementar la precisión y reducir la cantidad de falsos positivos en grandes volúmenes de datos.	SVM, SOFM, GA.	Dos, U2R, R2L, Prob.
[16]	Utiliza una máquina de soporte vectorial auto-entrenada para clasificar el tráfico observado y detectar ataques, sin intervención humana alguna.	SVM	Dos, U2R, R2L, Prob.
[17]	Emplean HSOM como técnica de aprendizaje en el reconocimiento de patrones.	HSOM	MiM, DoS, Node Replication y Jamming.
[18]	Presentan una nueva técnica de clasificación de patrones llamada EGHSOM, la cual utiliza a GHSOM con un entrenamiento inicial orientado a la clasificación de patrones detectados en redes.	GHSOM, EGHSOM	Falsos/Verdaderos Positivos/Negativos
[19]	Utilizan la caracterización del comportamiento normal de los protocolos presentes en cada capa del modelo OSI, como base para detectar anomalías.	Granular IDS.	Anomalías.
[20]	Utilizan una técnica GHSOM para procesar y generar clúster de datos que permitan detectar comportamientos regulares/irregulares en el tráfico de la nube en estudio.	GHSOM	Se establecieron 39 categorías de ataque.
[21]	Se enfoca en la clasificación del tráfico mediante el uso de redes neuronales artificiales.	MLP, SOM	Dos, U2R, R2L, Prob
[22]	Utiliza una técnica de detección de ataques en función de su distribución temporal, mediante el uso de una HQSOM modificada.	HQSOM	IP Spoofing, DoS, Teardrop
[23]	Usa detección de características como método para identificar y clasificar ataques. Hace uso de un algoritmo genético como búsqueda de una estrategia de detección y regresión logística para como algoritmo de aprendizaje.	Algortimo genético, LR	Normal, Dos, Probe, R2L, U2R, Backdoor, Analysis, Fuzzers, Shellcode, Reconnaissance, Exploits, Worms.
[24]	Realiza una categorización de los IDS/IDS según técnicas de inteligencia artificial tradicionales con soporte multiagente.	FS, NN, GA, AIS, CI, MCI	

Tabla II – *Continuación de la página anterior*

Ref	Resumen	Tec. IA	TCD
[25]	Tiene como objetivo detectar y clasificar ataques orientados a sistemas SCADA. Se hace la detección basado en características del tráfico detectado.	IWP-CSO, HNA-NN	Datos erróneos de sensores
[26]	Emplea una técnica de clasificación de características, con el fin de hacer clusters que permitan determinar la pertenencia de un tráfico a una clase determinada.	CANN	DoS, R2L, U2R
[27]	Plantea una estrategia basada en teorías de juegos junto con lógica difusa. Los actores de un ataque son modelados como un juego cooperativo.	Fuzzy Q-Learning,	DDoS
[28]	Se basa en una estrategia de votación de varios IDS distribuidos en una red, evaluados desde la tasa de aciertos individuales de cada uno, lo que hace que el sistema tenga más confianza en un IDS dependiendo de su calificación interna.	CVW, majority voting	Traf. malicioso.
[29]	Utiliza un método híbrido para clusterización del tráfico, llamado D-FICCA orientado a sensores de redes inalámbricas.	D-FICCA, ICA, Fuzzy logic	DDoS
[30]	Muestra el uso de un algoritmo Modified Open Path Forest (OPF) y un k-means para generar grupos de entrenamiento adicionales en grandes cantidades de datos no etiquetados.	Open Path Forest (OPF), Modified Open Path Forest (MOPF), K-Means	DoS, U2R, R2L y Probing.
[31]	Se propone un modelo de intrusión híbrido multinivel para clasificar el tráfico en normal y anormal. Asi mismo se apoya en SVM y extreme learning machine para detectar ataques conocidos y desconocidos. Se utiliza k-means modificado para generar conjuntos de entrenamiento de alta calidad, lo que reduce el tiempo de entrenamiento para los clasificadores.	Support vector machine (SVM), Extreme learning machine	Dos, U2R, R2L, Prob

En cuanto a la fecha de publicación de los artículos mostrados, se agrupan como se muestra en la Tabla III.

TABLA III: DISTRIBUCIÓN DE LOS ARTÍCULOS POR AÑO

Año	Artículos	Ref.
2013	9	[6], [8], [13], [15], [16], [18], [22], [24], [28]
2014	7	[3], [4], [9], [12], [20], [27], [29]
2015	3	[19], [21], [26]
2016	4	[7], [10], [11], [14]
2017	6	[5] , [17], [23], [25], [30], [31]

Se puede observar en ésta selección de artículos que, aparentemente, hubo un menor avance en este campo entre los años 2014 y 2016, donde se comienza a ver un aumento de la actividad, incluso en lo que va de año 2017 se ve una producción importante respecto a años anteriores.

Éstos resultados son arrojados por las consultas y los criterios de selección utilizados, no es el total de trabajos existentes en el área. Utilizando otros criterios de búsqueda o realizando el SLR en otro momento, pudieran variar los resultados.

IV. EXPLORACIÓN

Se responden las preguntas planteadas en el enfoque de la investigación, utilizando como argumentos la bibliografía revisada, según los criterios de inclusión definidos.

IV-A. ¿Cuáles son las técnicas más relevantes utilizadas en la adquisición de conocimiento de los IDS con autoaprendizaje?

Según éste systematic review se pueden detectar que, desde el 2013, se han realizando un buen número de investigaciones en el área, las cuales siguen avanzando. En la Tabla IV se muestra un resumen de las técnicas detectadas en la literatura, así como también la proporción en la que se encuentran dentro de la revisión.

TABLA IV: RESUMEN DE LAS TÉCNICAS DE IA EMPLEADAS

Técnica	Apariciones
Machine Learning	20
Metaheurísticas	5
Fuzzy Logic	4

Como se puede observar en la Tabla IV, existe un número importante de trabajos relacionados con técnicas de machine learning (cerca del 69 % de las publicaciones analizadas), esto se debe al poder de clasificación que tienen éstas técnicas sobre un gran volumen de datos, tal como ocurre en el tráfico de las redes de computadores. Por otro lado, se pudo evidenciar el uso de otras técnicas tales como metaheurísticas y fuzzy logic como complemento de las técnicas de machine learning empleadas.

IV-B. ¿Cuáles son los ataques más comunes clasificados por los IDS, tanto en su fase de entrenamiento como en operación?

Revisados los artículos finalmente seleccionados, se identifican los ataques presentes en los ellos. Muchos de los trabajos revisados utilizan el set de datos KDD 99 [32], como conjunto de entrenamiento y pruebas. Esto se debe a que éste set tiene sus datos etiquetados según las categorías de ataques (con cuatro sub categorías: DoS, U2L, R2L y Probing), y tráfico normal. Adicional a esto, se revisaron trabajos que utilizan set de datos sintéticos, así como también extraídos de algún ambiente controlado.

Los resultados se muestran en la Tabla V.

TABLA V: RESUMEN DE COMPORTAMIENTOS DETECTADOS

Ataque	Apariciones
DoS	17
Anomalías	11
U2R	10
R2L	10
Probbing	9

TABLA VI: RESUMEN DE COMPORTAMIENTOS DETECTADOS Y TÉCNICAS IA EMPLEADAS

Ataque	Machine Learning	Metaheurísticas	Fuzzy Logic
DoS	[3], [5], [14], [15], [16], [17], [22], [20], [21], [26], [30], [31]	[9], [23]	[27], [29]
Anomalías	[11], [12], [17], [18], [19], [22], [25]	[4], [13], [23], [28]	
U2R	[5], [14], [15], [16], [20], [21], [26], [30], [31]	[23]	
R2L	[5], [14], [15], [16], [20], [21], [26], [30], [31]	[23]	
Probbing	[5], [14], [15], [16], [20], [21], [30], [31]	[23]	

Tal como se puede ver en la Tabla V, la mayoría de los trabajos revisados se enfocan a la clasificación y detección de ataques DoS y anomalías.

IV-C. ¿Qué técnicas son las utilizadas en los comportamientos registrados?

Para responder la pregunta, se enfoca la extracción de datos en función de los comportamientos detectados y las técnicas utilizadas para detectarlos. Se hace un cruce de información entre las Tablas IV y V, con el fin de identificar las técnicas empleadas para cada tipo de comportamiento. Así mismo se obtienen los artículos a los que hace referencia. En la Tabla VI se muestra dicha categorización.

Como se puede observar en la Tabla VI, se ha desarrollado un mayor trabajo en el área de machine learning, debido a que en ésta área se encuentran gran variedad de algoritmos especializados en clasificar datos tanto etiquetados (aprendizaje supervisado), como no etiquetados (aprendizaje no supervisado). Se emplean algoritmos y combinaciones de estos para llevar a cabo la tarea de clasificación de gran cantidad de datos como lo es el tráfico de red. Así mismo, se observa poca actividad en el área de metaheurísticas, sin embargo estas sirven de apoyo para optimizar y mejorar el rendimiento de las implementaciones experimentales revisadas en éste trabajo.

V. CONCLUSIONES

Luego de obtenidos los resultados de la cadena de búsqueda en cada fuente de información, se hizo una primera revisión, donde se descartaron los artículos no relacionados con el tema, ya sea porque pertenecen a otra área o porque fueron incluidos por coincidencia parcial de la búsqueda pero sin relevancia al tema en estudio. Aplicando los criterios de búsqueda, se pudo reducir sustancialmente la cantidad de artículos a revisar, enfocando la bibliografía al objeto de investigación planteado.

Se pudo evidenciar una preferencia por las técnicas basadas en aprendizaje de máquina (machine learning) para agregar el componente de aprendizaje automático a los IDS convencionales. En algunos casos se utilizó ensamblaje entre varias técnicas, tanto de machine learning como de metaheurísticas, con la finalidad de resolver situaciones particulares del problema general.

Existe un trabajo emergente muy activo en el área de IDS con aprendizaje automático, donde se destacan las aplicaciones de técnicas de machine learning, combinadas entre ellas o apoyadas en otros algoritmos metaheurísticos, los cuales pueden servir para optimizar el funcionamiento de las técnicas de machine learning o para verificar las respuestas ante un conjunto de datos y retroalimentar información al sistema, reforzando entonces de forma automática el aprendizaje del IDS y contribuyendo finalmente a su funcionamiento general.

La presente revisión presenta un comienzo para identificar algún problema de detección/clasificación de ataques recientes, con mayor complejidad en su estructura y comportamiento. Las áreas de machine learning y metaheurísticas muestran una herramienta muy poderosa por su capacidad de clasificación y generalización como por su adaptación y resilencia, respectivamente. Sirve entonces como un buen comienzo para realizar la revisión de un problema abierto específico en el campo de la detección de intrusos que pueda ser resuelto con estas técnicas.

REFERENCIAS

[1] C. Kruegel, F. Valeur, and G. Vigna, *Intrusion Detection and Correlation - Challenges and Solutions*, ser. Advances in Information Security. Springer, 2005, vol. 14. [Online]. Available: http://dx.doi.org/10.1007/b101493

[2] N. Alrajeh, S. Khan, and B. Shams, "Intrusion Detection Systems in Wireless Sensor Networks: A Review," *International Journal of Distributed Sensor Networks,*, vol. 2013, no. 6, p. 7, 2013.

[3] P. Ahmed *et al.*, "A Hybrid-Based Feature Selection Approach for IDS," in *Networks and Communications (NetCom2013)*. Springer, 2014, pp. 195–211.

[4] N. A. Seresht and R. Azmi, "MAIS-IDS: A Distributed Intrusion Detection System using Multi-Agent AIS Approach," *Engineering Applications of Artificial Intelligence*, vol. 35, pp. 286–298, 2014.

[5] R. A. R. Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness Based Semi-Supervised Learning Approach for Intrusion Detection Dystem," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[6] A. Patel, M. Taghavi, K. Bakhtiyari, and J. C. Junior, "An Intrusion Detection and Prevention System in Cloud Computing: A Systematic Review," *Journal of network and computer applications*, vol. 36, no. 1, pp. 25–41, 2013.

[7] K. Singh, P. Singh, and K. Kumar, "A Systematic Review of IP Traceback Schemes for Denial of Service Attacks," *Computers & Security*, vol. 56, pp. 111–139, 2016.

[8] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion Detection System: A Comprehensive Review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 16–24, 2013.

[9] S. Shamshirband, N. B. Anuar, M. L. M. Kiah, V. A. Rohani, D. Petković, S. Misra, and A. N. Khan, "Co-FAIS: Cooperative Fuzzy Artificial Immune System for Detecting Intrusion in Wireless Sensor networks," *Journal of Network and Computer Applications*, vol. 42, pp. 102–117, 2014.

[10] M. Ahmed, A. N. Mahmood, and J. Hu, "A Survey of Network Anomaly Detection Techniques," *Journal of Network and Computer Applications*, vol. 60, pp. 19–31, 2016.

[11] O. Igbe, I. Darwish, and T. Saadawi, "Distributed Network Intrusion Detection Systems: An Artificial Immune System Approach," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 2016, pp. 101–106.

[12] D. Hock and M. Kappes, "A Self-Learning Network Anomaly Detection System using Majority Voting." in *INC*, 2014, pp. 59–69.

[13] H. Pereira and E. Jamhour, "A Clustering-based Method for Intrusion Detection in Web Servers," in *2013 20th International Conference on Telecommunications (ICT)*. IEEE, 2013, pp. 1–5.

[14] A. D. Landress, "A Hybrid Approach to Reducing The False Positive Rate in Unsupervised Machine Learning Intrusion Detection," in *SoutheastCon, 2016*. IEEE, 2016, pp. 1–6.

[15] S. Anil and R. Remya, "A Hybrid Method Based on Genetic Algorithm, Self-organised Feature Map, and Support Vector Machine for better Network Anomaly Detection," in *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. IEEE, 2013, pp. 1–5.

[16] H. Sallay, A. Ammar, M. B. Saad, and S. Bourouis, "A Real Time Adaptive Intrusion Detection Alert Cassifier for High Speed Networks," in *2013 12th IEEE International Symposium on Network Computing and Applications (NCA)*. IEEE, 2013, pp. 73–80.

[17] K. Huang, Q. Zhang, C. Zhou, N. Xiong, and Y. Qin, "An Efficient Intrusion Detection Approach for Visual Sensor Networks Based on Traffic Pattern Learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017.

[18] M. Salem and U. Buehler, "An Enhanced gHSOM for IDS," in *2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2013, pp. 1138–1143.

[19] P. Satam, "Cross Layer Anomaly Based Intrusion Detection System," in *2015 IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops (SASOW)*. IEEE, 2015, pp. 157–161.

[20] S.-Y. Huang, Y. Huang, and N. Suri, "Event Pattern Discovery on IDS Traces of Cloud Services," in *Proceedings of the 2014 IEEE Fourth International Conference on Big Data and Cloud Computing*, ser. BDCLOUD 2014. Washington, DC, USA: IEEE Computer Society, 2014, pp. 25–32.

[21] G. A. Haidar and C. Boustany, "High Perception Intrusion Detection System using Neural Networks," in *2015 Ninth International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS),*. IEEE, 2015, pp. 497–501.

[22] J. Cannady, "The Detection of Temporally Distributed Network Attacks using an Adaptive Hierarchical Neural Network," in *2013 World Congress on Nature and Biologically Inspired Computing (NaBIC)*. IEEE, 2013, pp. 5–9.

[23] C. Khammassi and S. Krichen, "A GA-LR Wrapper Approach for Feature Selection in Network Intrusion Detection," *Computers & Security*, 2017.

[24] S. Shamshirband, N. B. Anuar, M. L. M. Kiah, and A. Patel, "An Appraisal and Design of a Multi-Agent System Based Cooperative Wireless Intrusion Detection Computational Intelligence Technique," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 9, pp. 2105–2127, 2013.

[25] S. Shitharth *et al.*, "An Enhanced Optimization Based Algorithm for Intrusion Detection in SCADA Network," *Computers & Security*, vol. 70, pp. 16–26, 2017.

[26] W.-C. Lin, S.-W. Ke, and C.-F. Tsai, "CANN: An Intrusion Detection System Based on Combining Cluster Centers and Nearest Neighbors," *Knowledge-based systems*, vol. 78, pp. 13–21, 2015.

[27] S. Shamshirband, A. Patel, N. B. Anuar, M. L. M. Kiah, and A. Abraham, "Cooperative Game Theoretic Approach Using Fuzzy Q-learning for Detecting and Preventing Intrusions in Wireless Sensor Networks," *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 228–241, 2014.

[28] Y.-D. Lin, Y.-C. Lai, C.-Y. Ho, and W.-H. Tai, "Creditability-based Weighted Voting for Reducing False Positives and Negatives in Intrusion Detection," *Computers & security*, vol. 39, pp. 460–474, 2013.

[29] S. Shamshirband, A. Amini, N. B. Anuar, M. L. M. Kiah, Y. W. Teh, and S. Furnell, "D-FICCA: A Density-based Fuzzy Imperialist Competitive Clustering Algorithm for Intrusion Detection in Wireless Sensor Networks," *Measurement*, vol. 55, pp. 212–226, 2014.

[30] H. Bostani and M. Sheikhan, "Modification of Supervised OPF-based Intrusion Detection Systems using Unsupervised Learning and Social Network Concept," *Pattern Recognition*, vol. 62, pp. 56–72, 2017.

[31] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, "Multi-level Hybrid Support Vector Machine and Extreme Learning Machine Based on Modified K-means for Intrusion Detection System," *Expert Systems with Applications*, vol. 67, pp. 296–303, 2017.

[32] M. Lichman, "UCI Machine Learning Repository," 2013. [Online]. Available: http://archive.ics.uci.edu/ml