INGENIERÍA INVESTIGA

Vol. 5, e840, año 2023
ISSN – Online: 2708-3039
DOI: https://doi.org/10.47796/ing.v5i0.840

Artículo de revisión

Análisis integral de los sistemas de detección de intrusos y sus algoritmos asociados en la seguridad de la información

Comprehensive Analysis of Intrusion Detection Systems and their Associated Algorithms in Information Security

JOEL RENATO ENCISO SUÁREZ[1]
https://orcid.org/0009-0008-5535-0717

JACK EDINSON PORTILLA RODRIGUEZ[2]
https://orcid.org/0009-0003-4479-2177

ALBERTO CARLOS MENDOZA DE LOS SANTOS[3]
https://orcid.org/0000-0002-0469-915X

Recibido: 17/09/2023
Aceptado: 19/10/2023
Publicado: 8/11/2023

[1,2,3]Escuela de Ingeniería de sistemas, Universidad Nacional de Trujillo, La Libertad, Perú

E-mail: [1]to53300320@unitru.edu.pe, [2]to33300320@unitru.edu.pe, [3]amendozad@unitru.edu.pe

Facultad de Ingeniería
Publicación Oficial

Resumen

El estudio se enfoca principalmente en el análisis y la comparación de las técnicas de detección de intrusiones en entornos de red, con el objetivo de evaluar el impacto de los sistemas de detección de intrusiones (IDS) en la protección de datos. También se busca comprender cómo estas técnicas se han adaptado a las amenazas emergentes y evaluar su eficacia en la detección de actividades maliciosas. Para lograrlo, se realizó una revisión sistemática de documentos almacenados en las bases de datos de IEEE Xplore, Redalyc y ScienceDirect entre los años 2019 y 2023. El análisis revela que las técnicas de detección de intrusiones han evolucionado de manera significativa para enfrentar las amenazas cibernéticas en constante cambio. En particular, las técnicas basadas en el aprendizaje automático y el análisis de comportamiento han demostrado ser eficaces en la reducción de falsos positivos. Sin embargo, para mantenerse al día con las amenazas, se requiere una gestión constante y la actualización de estas técnicas. Además, se ha determinado que la detección de intrusiones es crucial para la seguridad cibernética. Esto se debe a que, en un entorno cibernético en constante evolución, donde las amenazas avanzan rápidamente, los IDS ofrecen una defensa crítica al proporcionar visibilidad y protección continua contra posibles intrusiones o actividades maliciosas en tiempo real.

Palabras clave: seguridad de información; protección digital; IDS; algoritmos; amenazas cibernéticas.

Abstract

The study primarily focuses on the analysis and comparison of intrusion detection techniques in network environments, with the objective of evaluating the impact of Intrusion Detection Systems (IDSs) on data protection. It also aims to understand how these techniques have adapted to emerging threats and evaluate their effectiveness in detecting malicious activities. To achieve this, a systematic review of documents stored in the IEEE Xplore, Redalyc, and ScienceDirect databases between 2019 and 2023 was conducted. The analysis reveals that intrusion detection techniques have evolved significantly to address constantly changing cyber threats. Specifically, machine learning-based techniques and behavior analysis have proven to be effective in reducing false positives. However, to keep up to date with the threats, constant management and updating of these techniques is required. Furthermore, it has been determined that intrusion detection is crucial for cybersecurity. This is because in a rapidly evolving cyber environment where threats swiftly advance, IDSs provide critical defense by offering continuous visibility and protection against potential intrusions or malicious activities in real-time.

Keywords: information security; digital protection; IDS; algorithms; cyber threats.

1. Introducción

En la actual era digital, la creciente dependencia de la tecnología y la interconexión de sistemas han convertido a la Seguridad de la Información (SI) en un campo que requiere atención constante y evolución continua para contrarrestar los diversos tipos de ataques de seguridad. Según Dhillon et al. (2021), este hecho se divide en un sistema social y un sistema técnico. El sistema social se subdivide en estructura, que comprende la política y el cumplimiento de la gestión, y personas. Por otro lado, el sistema técnico se subdivide en tecnología y tareas, que incluyen la gestión de vulnerabilidades y el diseño de sistemas.

Dentro de este marco de referencia, los Sistemas de Detección de Intrusiones (IDS; Por sus siglas en inglés) adquieren una relevancia significativa. Según Alhowaide et al. (2021), es una herramienta de seguridad informática diseñada para monitorear y analizar el tráfico en una red o sistema en busca de actividades maliciosas o no autorizadas. Su objetivo principal es detectar y alertar sobre posibles intentos de intrusión, ataques cibernéticos u otras actividades sospechosas que podrían comprometer la seguridad de los sistemas, lo que proporciona una línea adicional de defensa al dificultar que el atacante acceda a la red sin ser detectado. Por lo tanto, son una opción muy utilizada dentro del sistema técnico para ayudar en la SI. Su uso es crucial para mantener la integridad y confidencialidad de los datos, así como para garantizar el correcto funcionamiento de los sistemas y redes.

No obstante, existen desafíos asociados con su implementación y uso. Hay diversos tipos de IDS, cada uno con sus propias fortalezas y debilidades, además de distintos enfoques para la detección de intrusiones, lo que puede dificultar que las organizaciones determinen cuál es el más adecuado para sus necesidades específicas.

Además, según Saranya et al. (2020), los IDS tienen una amplia variedad de áreas de aplicación, que incluyen la red, los dispositivos, las aplicaciones, el Internet de las cosas, el big data y la nube. No obstante, cada una de estas áreas presenta sus propios desafíos y vulnerabilidades únicas que los IDS deben ser capaces de abordar.

Por lo tanto, la revisión sistemática tuvo como objetivo principal evaluar el impacto de los Sistemas de Detección de Intrusiones (IDS) en la protección de datos, al mismo tiempo que se buscaba recopilar información significativa sobre su aplicabilidad en diferentes campos y los desafíos que surgen al aplicarlos en diversas áreas. Este enfoque proporcionará una visión integral de la importancia de los IDS en la seguridad de la SI.

2. Metodología

2.1. Fundamentación metodológica

Se llevó a cabo un análisis exhaustivo de la literatura científica respaldado por la metodología PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*). Según Page et al. (2021), esta metodología se presenta como un marco que facilita la realización de un estudio sistemático basado en trabajos académicos previos. Su enfoque metodológico se establece con el propósito de abordar la pregunta planteada de manera rigurosa y completa.

Además, los autores indican que las revisiones sistemáticas cumplen diversas funciones críticas. Pueden proporcionar una síntesis del estado del conocimiento en un campo, a partir

de la cual se pueden identificar futuras prioridades de investigación. También permiten identificar deficiencias en la investigación primaria que deben abordarse en investigaciones futuras y generan la posibilidad de evaluar teorías sobre cómo o por qué ocurren fenómenos particulares.

2.2. Criterios de inclusión, exclusión y calidad

Los criterios de inclusión se basaron en la disponibilidad de información en inglés, portugués y español, relacionada con algoritmos para sistemas de detección de intrusos publicada en los últimos cuatro años (2019-2023), con el propósito de optimizar la seguridad de los datos.

Por otro lado, los criterios de exclusión se aplicaron para descartar información que no tratara sobre la seguridad de los datos o que no incluyera la palabra clave "seguridad". También se excluyó la información relacionada con ciencias sociales.

En cuanto a los criterios de calidad, se priorizó la actualidad de los recursos, aceptando aquellos publicados en un período de hasta cinco años antes de la revisión, ya que se consideró que abordaban un campo tecnológico en constante evolución. Se evaluó la coherencia de los enfoques, se valoró la originalidad de los trabajos y se tuvo en cuenta el rigor metodológico empleado para reducir posibles sesgos, errores e interpretaciones incorrectas. Estos aspectos en conjunto contribuyeron a fortalecer la robustez de los resultados obtenidos.

2.3. Proceso de recolección de información

La búsqueda y extracción de información se realizaron de manera individual, y se resolvieron las discrepancias entre los colaboradores a través de un proceso de consenso, con el objetivo de llevar a cabo una revisión sistemática, como se muestra en la Figura 1 y Tabla 1.

Figura 1

Esquema del proceso de extracción de datos

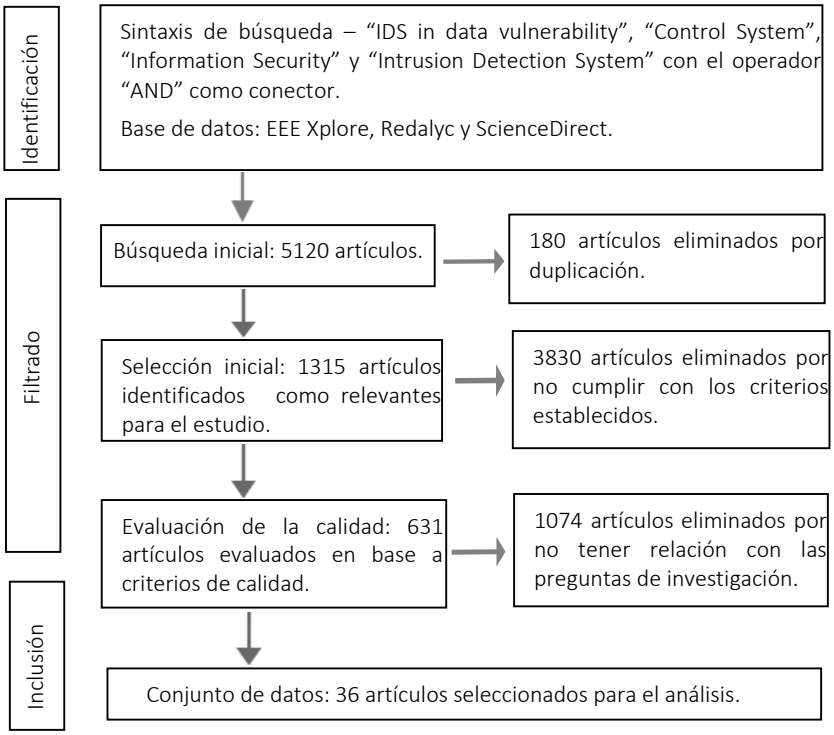


Tabla 1

Motores de búsqueda académicos y sus respectivos términos

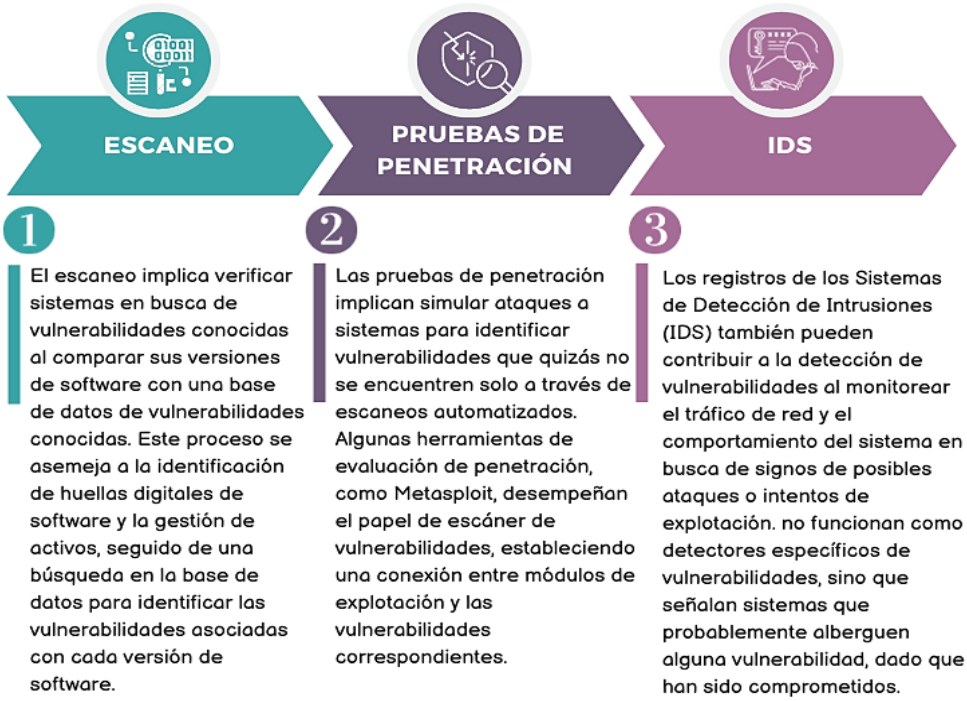
Base de datos	Términos de Búsqueda	Resultados	Seleccionados
EEE Xplore	("All Metadata": IDS in data vulnerability" AND "All Metadata": Control System)	62	2
Redalyc	("Intrusion Detection System" AND "Information Security")	2 299	4
ScienceDirect	("Intrusion Detection System" AND "Information Security")	2 759	30

3. Resultados

Actualmente, Dhillon et al. (2021) señalan la presencia de diversos problemas de seguridad, como el malware, el phishing, la inyección SQL, la pérdida de disponibilidad de datos, la gestión insuficiente del control de acceso, la falta de monitoreo, la falta de actualizaciones y parches, entre otros. Estos problemas pueden categorizarse en ataques a la seguridad de la información, vulnerabilidades de la infraestructura del sistema y la gestión de la seguridad de los sistemas de información, además de cuestiones reglamentarias. En este contexto, se destaca la importancia de la detección de vulnerabilidades y su gestión. Spring (2023) sostiene que la detección de vulnerabilidades puede llevarse a cabo de al menos tres maneras: mediante el escaneo, pruebas de penetración y registros de sistemas de detección de intrusiones (IDS), como se ilustra en la Figura 2.

Figura 2

Formas de detección de vulnerabilidades



Nota. Adaptado de Spring (2023).

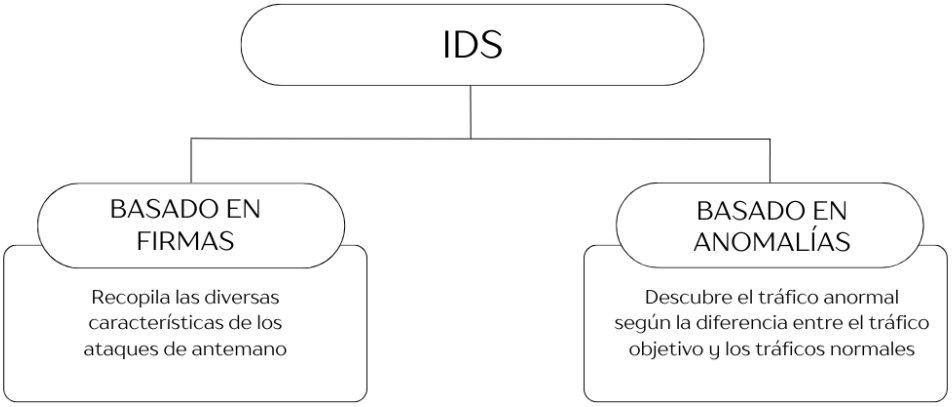
Las medidas para reducir el riesgo asociado a las vulnerabilidades se dividen en dos categorías principales: mitigación y remediación. La mitigación tiene como objetivo reducir el impacto causado por una vulnerabilidad sin eliminarla por completo. Ejemplos de esto incluyen restringir el acceso a la red para componentes susceptibles o limitar el acceso a un componente vulnerable mediante modificaciones en la configuración del software.

Por otro lado, la remediación implica acciones dirigidas a eliminar o erradicar la vulnerabilidad. Ejemplos de ello son la aplicación de parches o la desinstalación del sistema que presenta la vulnerabilidad.

Un IDS es una técnica de defensa ampliamente utilizada contra ciberataques. Se refiere específicamente a un dispositivo o software estratégicamente ubicado en un punto específico de una red para monitorear todo el tráfico (Alhowaide et al., 2021). En el estudio realizado por Lee et al. (2022), los dividen en dos grandes grupos según su modo de detección, como se detalla en la Figura 3.

Figura 3

Clasificación de IDS



Nota. Adaptado de Masdari & Khezri (2020).

Los sistemas basados en host se encargan de supervisar la actividad del sistema, como los cambios en los archivos o el uso de la memoria. Este monitoreo interno depende en gran medida de las pistas de auditoría y los registros del sistema para determinar si un sistema ha sido comprometido. Este enfoque se lleva a cabo en dispositivos individuales que monitorizan su propio funcionamiento y detectan cualquier uso inapropiado de los recursos disponibles. Algunos ejemplos destacados de *Host-based Intrusion Detection Systems* (HIDS) son OSSEC y Tripwire, que se encargan de analizar registros y verificar la integridad de los archivos, respectivamente.

Por otro lado, los sistemas basados en red se concentran en supervisar la actividad de la red, las comunicaciones y auditar la información de los paquetes para proteger un sistema de amenazas basadas en la red mediante la detección de comportamientos sospechosos en los paquetes entrantes. Este enfoque suele operar en modo promiscuo, interceptando y analizando paquetes sin exponerlos a posibles amenazas. Dos de los *Network-based Intrusion Detection Systems* (NIDS) más populares en el campo de la ciberseguridad son Snort y Suricata.

3.1. IDS basado en firmas

La detección de intrusos basada en firmas es un mecanismo de seguridad utilizado para identificar actividades maliciosas en una red. Este enfoque implica la comparación de patrones de tráfico de red con una base de datos que contiene firmas conocidas de ataques previos. Cuando se encuentra una coincidencia, se genera una alerta que indica la detección de un posible ataque. Las firmas pueden ser definidas por el usuario o proporcionadas por el proveedor de software de detección de intrusos (Li et al., 2019). Sin embargo, la investigación llevada a cabo por Masdari M & Khezri H (2020), como se detalla en su artículo, sugiere que, si bien esta técnica es eficaz para detectar ataques conocidos, su eficacia puede disminuir en la identificación de ataques nuevos o de naturaleza desconocida.

3.1.1. Snort

Snort es una de las soluciones IDPS basadas en firmas más utilizadas que admite tanto el modo IDS como el IPS. Es una valiosa herramienta NIDPS que es relativamente fácil de configurar. Puede monitorear el tráfico en la red, comparar los paquetes recibidos y detectar ataques de fuerza bruta. En el modo IDS, solo genera alertas en función de la detección, mientras que bloquea los paquetes maliciosos en el modo IPS (Waleed et al., 2022). En un estudio de investigación realizado por Adiwal et al. (2023), se utilizó Snort para crear firmas innovadoras destinadas a detectar la amplificación DNS, la tunelización y los ataques de DoS. Se capturaron patrones de tráfico de aplicaciones y herramientas de explotación de DNS disponibles públicamente para luego agregar estas firmas novedosas al archivo de reglas de DNS existente de Snort. Esto permitió la detección de varios tipos de ataques de DNS, y la solución resultante se denomina Detección de Intrusiones (DID).

3.1.2. Suricata

Suricata es una herramienta NIDPS de código abierto que admite tanto el modo IDS como el IPS. Es conocida por su alto rendimiento y utiliza un motor de reglas basado en firmas y un motor de detección basado en comportamiento (Idrissi et al., 2023). Suricata es capaz de procesar cargas de red más altas que Snort y Zeek, aunque esto se traduce en un uso más intensivo de los recursos de hardware. Este algoritmo introduce la detección basada en scripts y estructuras de datos bien diseñadas para analizar y registrar información de flujo para investigaciones posteriores (Adiwal et al., 2023). Un estudio realizado por Hu et al. (2020) destaca que Suricata incorpora soporte extendido para el filtro de paquetes BSD (eBPF) y XDP en su última versión 4.1.4, lo que le permite procesar la captura de paquetes a un ritmo de más de 33,000 flujos por segundo. Además, el artículo compara el rendimiento de Suricata con Snort, otro IDS, y discute cómo Suricata utiliza la detección basada en patrones y scripts para detectar ataques.

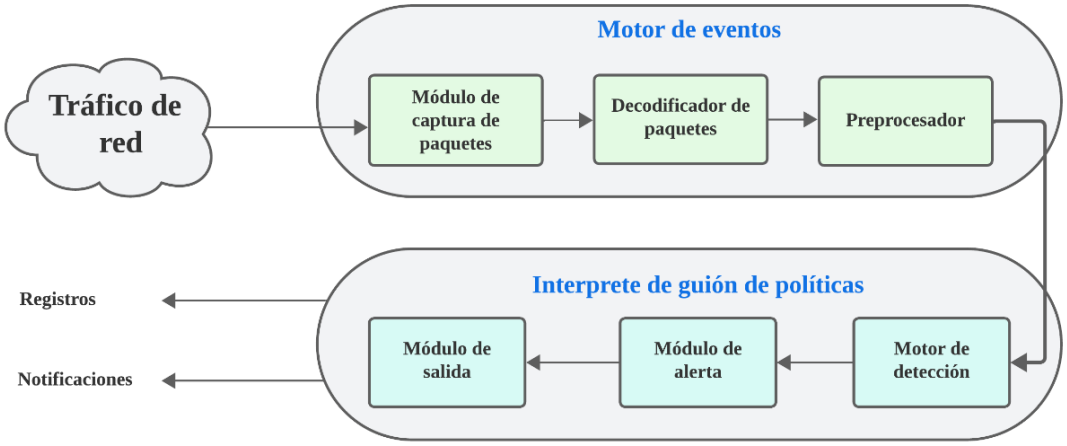
3.1.3. Zeek

Zeek, por otro lado, es otra herramienta de NIDPS de código abierto que solo admite el modo IDS. Utiliza una arquitectura altamente escalable en la que es posible mejorar el rendimiento dedicando más recursos de hardware a los trabajadores y al administrador (Tiwari et al., 2023). Un estudio realizado por Adiwal et al. (2023) resalta que Zeek resulta especialmente útil en entornos con ataques de día cero, ya que admite la detección basada en anomalías, una

característica que falta tanto en Snort como en Suricata, que se limitan a la detección de uso indebido. Zeek emplea trabajadores que se implementan en los dispositivos de red y estos trabajadores envían sus registros al administrador. A continuación, en la Figura 4 se presenta un diagrama de nivel de bloque de Zeek que ilustra la estructura de la herramienta.

Figura 4

Diagrama de nivel de bloque Zeek



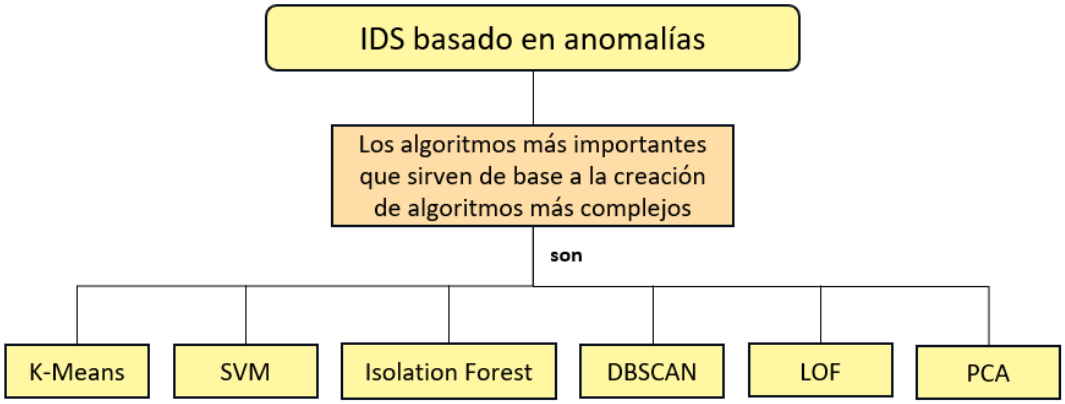
Nota. Adaptado de Waleed et al. (2022).

3.2. IDS basado en anomalías

Los algoritmos basados en anomalías son técnicas de aprendizaje automático utilizadas para la detección de intrusiones. Estos algoritmos se centran en la identificación de patrones anómalos en los datos de red, lo cual puede indicar la presencia de una intrusión (Mohammadi et al., 2021). De acuerdo con la investigación de Aksu & Aydin (2022), se ha destacado que los sistemas de detección de intrusos (IDS) basados en anomalías han demostrado ser de vital importancia en el campo de la ciberseguridad. En su estudio, los autores resaltan que entre los enfoques más ampliamente empleados y efectivos se encuentran los que se presentan en la ilustración que se hace referencia en la Figura 5.

Figura 5

Diagrama IDS basado en anomalías



Nota. Adaptado de Gu & Lu (2021).

3.2.1. K-Means

Según Hu et al. (2023), el algoritmo K-Means no es comúnmente empleado para la detección de intrusos en sistemas de seguridad, ya que su aplicación principal se centra en la agrupación de datos en clústeres en contextos de aprendizaje no supervisado. No obstante, es factible utilizar variantes o enfoques adaptados de K-Means para la detección de anomalías o intrusiones en conjuntos de datos. A continuación, en la Tabla 2 se presentan algunos algoritmos basados en K-Means que han demostrado ser altamente eficaces en la protección de datos.

Tabla 2

IDS basados en K-means

Autor	Nombre	Características	Resultado
Hu et al (2023)	Clustering LK-means	El objetivo del clustering K-means es dividir n datos en k clases, donde cada dato pertenece al centro de cluster más cercano, como el centro del cluster.	Los resultados experimentales muestran que el algoritmo LK-means tiene mayor precisión y mejores centroides de agrupamiento que los otros algoritmos de referencia.
Wang et al. (2022)	R-reference points-based k-means (r-RPKM)	Este algoritmo utiliza puntos de referencia para acelerar el proceso de cálculo de distancias entre los datos y los centros de los clusters. El r-RPKM reduce el tiempo de cálculo al solo calcular las distancias entre cada dato y los centros de los clusters más cercanos.	Los resultados experimentales sugieren que el algoritmo r-RPKM es una alternativa prometedora al algoritmo k-means tradicional para el clustering de grandes conjuntos de datos.
Ay et al. (2023)	K-means++	El algoritmo se basa en el concepto de minimizar la suma de las distancias al cuadrado entre los puntos de datos y sus centros de conglomerados asignados.	El algoritmo tiene las ventajas de brevedad, eficiencia y velocidad.

3.2.2. Support vector machine (SVM)

Según Baldomero, Martínez & Rodríguez (2021), el algoritmo SVM es un tipo de algoritmo de aprendizaje supervisado que se puede aplicar con éxito en la detección de intrusos. Estos algoritmos son eficaces para separar datos en dos clases, incluso cuando las clases no están perfectamente equilibradas o cuando los datos son no lineales. A continuación, en la Tabla 3 se presentan los algoritmos basados en SVM que se consideran más eficaces.

Tabla 3

IDS basados en SVM

Autor	Nombre	Características	Resultado
Santhi & Srinivasa (2023)	Duo Autoencoder-SVM	Es una combinación de un autoencoder disperso y SVM para distinguir eficazmente entre ciberataques y fallas en los sistemas de monitoreo inteligente	Logró una alta eficiencia de detección, con una tasa de detección de amenazas del 99,89 % y una tasa de detección de FDIAs del 98,79 %.
Gu & Lu (2021)	Incrustación con Nave Bayes	Propone una nueva técnica de mejora de características llamada incrustación de características de Naïve Bayes, que se utiliza para transformar los datos del espacio de características original a nuevas características de alta calidad	Se muestra que el algoritmo propuesto supera a otros métodos de detección de intrusiones, incluyendo los métodos de detección no basados en aprendizaje automático, en términos de las tres medidas de rendimiento.
Baldomero, Martínez, & Rodríguez (2021)	Detección de valores atípicos con SVM (RL -FS -M)	Los autores presentan un modelo basado en SVM que intenta eliminar los efectos adversos de los valores atípicos utilizando un número reducido de características relevantes	Se puede ver la mejora de (RL -FS -M) con respecto al resto de los modelos en un 5 % de ruido de etiqueta y un 5 % de valores atípicos de SVM. Además, se proporciona la mejora promedio de la precisión y el AUC.

3.2.3. Isolation forest

Para Anantha Krishnan & Senthil Kumar (2022), Isolation Forest es un algoritmo de detección de anomalías que aísla instancias anómalas mediante la construcción de árboles de decisión. Cabe destacar que es especialmente efectivo en la detección de anomalías en conjuntos de datos grandes y multidimensionales. Además, cuando se integra con otros algoritmos, sus propiedades mejoran. A continuación, en la Tabla 4 se presentan algoritmos basados en el Isolation Forest.

Tabla 4

IDS basados en Isolation Forest

Autor	Nombre	Características	Resultado
Lifandali, Abghour, & Chiba (2023)	Optimización de Colonias de Hormigas con Algoritmos de Bosques Aleatorios.	En la primera fase, se utilizan los algoritmos ACO y Random Forest para seleccionar las características relevantes del conjunto de datos. En la segunda fase, se utiliza el algoritmo de Isolation Forest para detectar intrusiones en el conjunto de datos.	El sistema tiene una efectividad del 98,2 %. Sin embargo, un defecto fundamental es que no logra maximizar la tasa de detección y la precisión al tiempo que minimiza las falsas alarmas y la incapacidad para detectar correctamente tipos específicos de ataques.
Anantha Krishnan & Senthil Kumar (2022)	Garra rufa Fish optimization (GRFO) - Isolation Forest (iForest)	El algoritmo funciona mediante la selección aleatoria de parámetros del conjunto de datos y la elección de un punto de división y atributo que se encuentra entre el valor máximo y mínimo. Luego, se calcula la aptitud y se separa el nodo.	Los resultados son superiores a los métodos convencionales de ajuste de controladores PI y técnicas de optimización como Ant Lion Optimization (ALO) y Modified Ant Lion optimization based Artificial Neural Network (MALANN) respecto a estabilidad y rendimiento del sistema de microred.

3.2.4. Density-based spatial clustering of applications with noise (DBSCAN)

Para Bai et al. (2023), el DBSCAN es un algoritmo de agrupación o clustering utilizado en minería de datos y análisis de datos espaciales. Los autores sostienen que, a diferencia de los métodos de clustering tradicionales, como K-Means, DBSCAN no presupone que los clústeres sean necesariamente esféricos o de forma similar; en su lugar, identifica clústeres en función de la densidad de los datos en el espacio. A continuación, en la Tabla 5 se presentan algoritmos basados en el DBSCAN.

3.2.5. Local outlier factor (LOF)

Según Zhang, You & Jia (2020), LOF se basa en el concepto de "factor local de atipicidad", que mide cuán inusual es un punto de datos en relación con sus vecinos más cercanos. A continuación, en la Tabla 6 se presentan algoritmos que se han desarrollado basados en LOF.

Tabla 5

IDS basados en Isolation Forest

Autor	Nombre	Características	Resultado
Huang et al. (2023)	GriT-DBSCAN	Es un algoritmo de agrupamiento espacial basado en cuadrícula que utiliza un árbol de cuadrícula para organizar las cuadrículas no vacías y mejorar la eficiencia de las consultas de cuadrículas vecinas no vacías.	Los resultados de los experimentos indican una mayor eficiencia en los nuevos algoritmos frente a los existentes basados en DBSCAN, aunque se identifican limitaciones que requieren investigación adicional para su mejora.
Bai et al. (2023)	K-DBSCAN	K-DBSCAN optimiza DBSCAN al enfocarse en puntos centrales, mejorando la eficiencia y adaptabilidad de parámetros.	El algoritmo K-DBSCAN se ejecuta con una diferencia de tiempo de un solo sello de 3,21 s comparado con el algoritmo DBSCAN, y los resultados de agrupamiento son precisos en un 99 %.
Hanafi & Saadatfar (2022)	AnyDBC	AnyDBC comprime los datos en subconjuntos de densidad más pequeños llamados clusters primitivos y etiqueta los objetos en otro cluster según el procedimiento mencionado.	El algoritmo propuesto tiene un tiempo de ejecución más corto en comparación con otros métodos competitivos, especialmente en el caso de grandes conjuntos de datos.

Tabla 6

IDS basados en LOF

Autor	Nombre	Características	Resultado
Zhang, You & Jia (2020)	LOF-KNN	El algoritmo LOF-KNN detecta valores atípicos en datos de obleas mediante la combinación de KNN y LOF, enfocándose en la detección local de anomalías espaciales	Se realizaron pruebas con datos de obleas y se verificó la eficacia del método, que en comparación con el LOF normal mejor en un 67,3 %.
Asniar & Surendro (2022)	SMOTE-LOF	La combinación de SMOTE y LOF detecta y mejora la precisión al manejar datos desequilibrados mediante la identificación del ruido en datos sintéticos generados por SMOTE	SMOTE-LOF ofrece mejor rendimiento predictivo que SMOTE en la mayoría de los casos. Sin embargo, en conjuntos con menos ejemplos, SMOTE destaca con una AUC superior en el manejo de datos desequilibrados

3.2.6. Principal Component Analysis (PCA)

Según Priya et al. (2020), el algoritmo PCA implica la reducción de la dimensionalidad de los datos de tráfico de red o de registros de eventos de seguridad. Esto puede facilitar la detección de anomalías y la identificación de patrones inusuales. Los autores presentan un algoritmo eficaz que combina las propiedades de LOF y PCA, lo que resulta en un algoritmo de alta eficacia. A continuación, en la Tabla 7 se presentan algoritmos que se han desarrollado basados en PCA.

Tabla 7

IDS basados en PCA

Autor	Nombre	Características	Resultado
Priya et al. (2020)	PCA GWO híbrido	El algoritmo híbrido propuesto en este artículo extrae características de alto impacto y excluye las características de impacto negativo, lo que aumenta la precisión del modelo de clasificación.	La metodología propuesta mejora la detección en entornos IoMT con dispositivos médicos comunicándose a través de direcciones IP únicas, aumentando la eficacia en un 45,2 %.

Como se puede observar, estos algoritmos han servido como base para la creación de enfoques más avanzados en la detección de intrusiones. La evolución de estos algoritmos ha impulsado el desarrollo de algoritmos de detección de intrusos más complejos. Estas nuevas iteraciones han aprovechado los fundamentos establecidos por los enfoques previamente mencionados, lo que ha permitido una detección más precisa y adaptable de las amenazas cibernéticas en constante evolución. La sinergia entre los algoritmos originales y las innovaciones resultantes ha contribuido significativamente a la mejora continua de la seguridad cibernética.

3.3. Áreas de aplicación de IDS

3.3.1. Entornos en Internet de las cosas (IoT)

Conforme a Ghasempour (2019), el concepto de Internet de las Cosas (IoT) hace referencia a una configuración compuesta por entidades físicas o dispositivos con identificadores únicos, capaces de capturar, retener y transmitir datos en la red global sin la necesidad de intervención humana o interconexión directa entre humanos y máquinas. Estos artefactos del IoT operan con un consumo energético moderado y utilizan protocolos de comunicación altamente eficientes.

Martins et al. (2022) enfatizan la necesidad de enfoques computacionales eficaces para abordar las limitaciones en tiempo real de la seguridad en IoT. Señalan que las soluciones actuales a menudo descuidan el retraso entre el entrenamiento, la predicción y la respuesta. Destacan que un sistema de detección de anomalías eficaz debe ser en línea, continuo y adaptable. En este contexto, Jasim & Kurnaz (2023) llevaron a cabo un estudio reciente que se alinea con estas necesidades y desarrollaron nuevas técnicas de aprendizaje profundo para la

detección de IDS en una red de sensores inalámbricos. Aplicaron algoritmos de optimización basados en autocodificadores y clasificadores profundos para asegurar la privacidad y los datos del usuario. Sin embargo, destacaron que las redes neuronales artificiales, aunque son eficaces para detectar características complejas, tienen un alto requerimiento de cálculo que limita su uso en dispositivos IoT.

3.3.2. Sistemas industriales (ICS y SCADA)

La Supervisión, Control y Adquisición de Datos Automatizada (SCADA) se utiliza ampliamente en sistemas industriales para supervisar y controlar procesos y dispositivos en entornos industriales y de infraestructura crítica. Según un estudio realizado por Zolanvari et al. (2019), se han identificado varios tipos de ataques principales que pueden afectar a los sistemas SCADA, como desbordamiento de buffer, inyección de código, validación incorrecta de entrada, denegación de servicio, reconocimiento y acceso no autenticado. Para contrarrestar estos ataques, se emplean diversas técnicas de seguridad, como el cifrado y la detección de actividades anómalas de usuarios verificados. Estas medidas son esenciales para garantizar la integridad y la confidencialidad de los sistemas SCADA y para proteger los procesos y dispositivos en entornos industriales y de infraestructura crítica.

En este contexto, Friha et al. (2023) proponen un IDS (2DF-IDS) basado en aprendizaje federado (FL) seguro, descentralizado y diferencialmente privado (DP) para proteger instalaciones industriales inteligentes de diversas vulnerabilidades, como DDoS ICMP, DDoS UDP, DDoS HTTP, inyección SQL, DDoS TCP, secuestro de datos, XSS y MITM.

Además, Shokry et al. (2022) señalan la importancia del uso de IDS en los sistemas de la Infraestructura de Medición Avanzada (AMI) debido a que son susceptibles a ataques en varias capas. Por ejemplo, en la capa de hardware, pueden ocurrir ataques de Inyección SQL, DoS o DDoS. En la capa de datos, se pueden realizar ataques de inyección de datos falsos, y en la capa de comunicación, se pueden llevar a cabo ataques de secuestro de sesión y "Man in the Middle" (MITM). Según Corallo et al. (2020), los impactos negativos de estos ataques pueden ser graves, como el robo de secretos industriales y propiedad intelectual, lo que podría resultar en la reducción de la ventaja competitiva de la empresa, daños a la imagen y reputación de la empresa y violación de acuerdos comerciales con socios industriales sobre la confidencialidad de datos.

Por otro lado, para mejorar la seguridad de la red en los Sistemas de Control Industrial (ICS), Al-Abassi et al. (2020) sugieren desarrollar un modelo para identificar diferentes tipos de ataques y sus ubicaciones. Este enfoque evitaría fallos críticos del sistema y mejoraría la seguridad de la red de los sistemas de control industrial (ICS) contra ciberataques, ya que reduciría el tiempo de inactividad del procesamiento y mejoraría la eficiencia informática una vez que se detecta un ataque.

3.3.3. Redes (SDN)

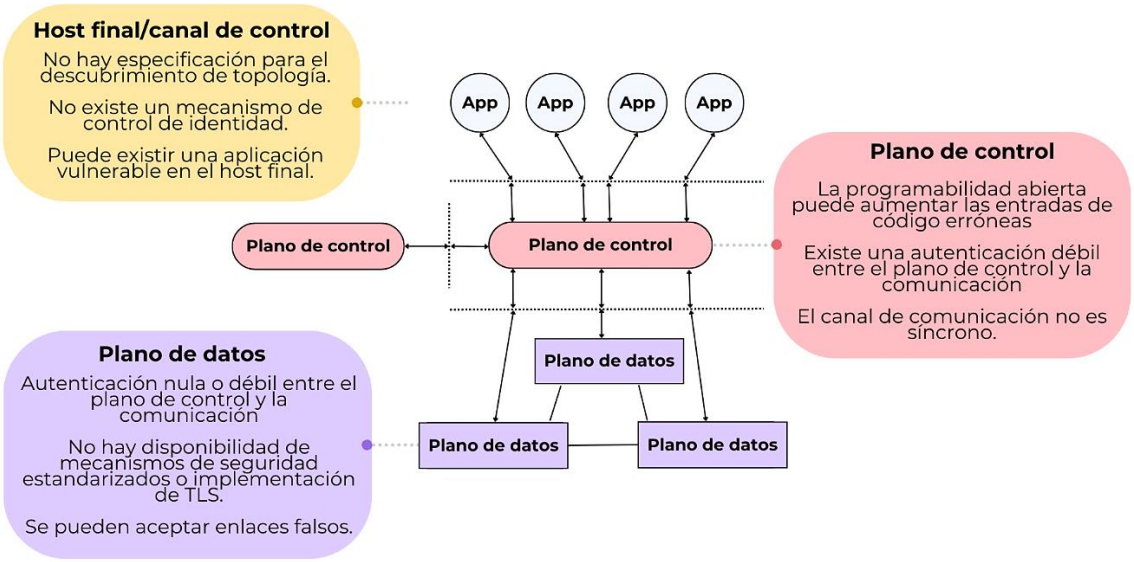
Kumar et al. (2023) mencionan que las Redes Definidas por Software (SDN) representan una tecnología informática avanzada que ofrece una amplia gama de beneficios, como supervisión centralizada, reprogramación de la red y adopción de estándares abiertos, entre otros. Las SDN desempeñan un papel fundamental en la mejora del rendimiento en comparación con las redes tradicionales y han demostrado ser especialmente valiosas al implementar con éxito redes

inalámbricas, capitalizando las ventajas de un rendimiento de red mejorado de manera significativa.

Además, Deb & Roy (2022) explican que las SDN se componen de un plano de control, un plano de datos y un host final/canal de control, como se muestra en la Figura 6. Según su estudio, es posible aplicar un Sistema de Detección de Intrusiones (IDS) en el host final/canal de control. Esta aplicación sería particularmente útil cuando un host final interno malicioso intenta acceder a servicios de red no autorizados. Por lo tanto, los IDS en SDN proporcionan una capa adicional de seguridad al monitorear el tráfico y detectar comportamientos anómalos, lo que alerta a los administradores sobre posibles amenazas.

Figura 6

Arquitectura conceptual de SDN



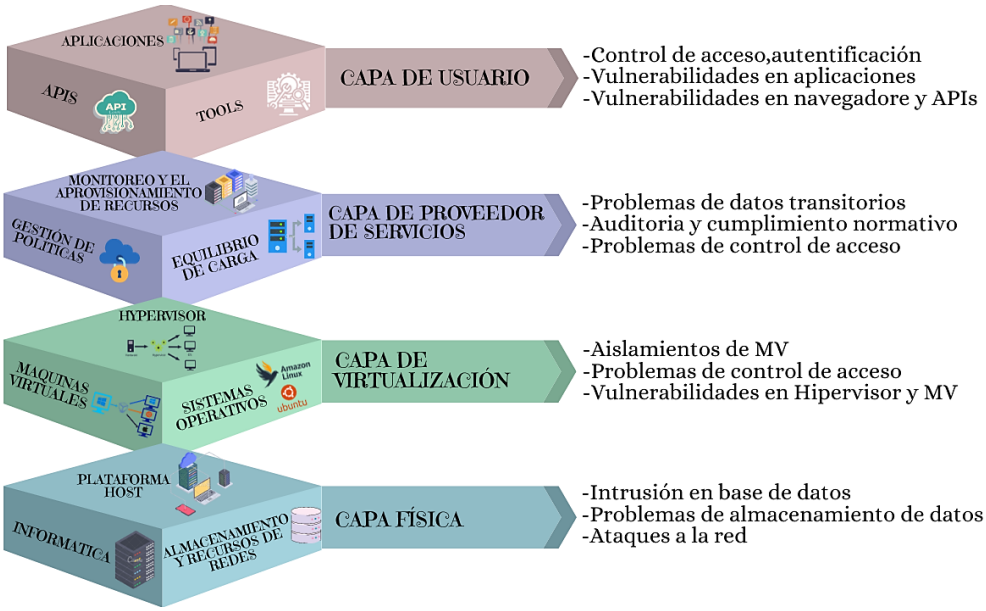
Nota. Adaptado de Deb & Roy (2022).

3.3.4. Cloud y flog computing

La computación en la nube es un modelo que proporciona servicios tecnológicos a través de Internet. Los usuarios pueden acceder a recursos como almacenamiento y procesamiento a través de proveedores en línea en lugar de poseer su propia infraestructura física. Esto ofrece flexibilidad y eficiencia en el uso de recursos. Sin embargo, junto con estas ventajas, también existen desafíos de seguridad. Belal & Sundaram (2022) indican que en cada capa de la arquitectura de seguridad en la nube existen vulnerabilidades, como se muestra en la Figura 7. En particular, mencionan que, en la nube móvil que implica una red de clientes heterogénea, un Sistema de Detección de Intrusiones (IDS) basado en Aprendizaje Automático podría ser útil para la fusión de datos, por ejemplo, en bases de datos. Este IDS podría detectar ataques del tipo "Man in the Middle" (MITM) y ataques de denegación de servicio (DDOS) en la nube móvil. Por lo tanto, aunque la computación en la nube ofrece numerosos beneficios, también es crucial tener en cuenta las consideraciones de seguridad.

Figura 7

Arquitectura conceptual de Cloud



Nota. Adaptado de Belal & Sundaram (2022).

En este contexto, Almiani et al. (2020) sugieren que los ataques cibernéticos en entornos nebulosos podrían prevenirse mediante la implementación de un sistema de detección de intrusiones plenamente autónomo. Proponen el uso de redes neuronales multicapas recursivas como una alternativa viable para identificar ataques del tipo Probe, DoS, R2L y U2R.

4. Discusión

De acuerdo a los datos obtenidos, se puede observar respecto a los algoritmos de Sistemas de Detección de Intrusos (IDS) que, en primer lugar, realizando un análisis comparativo entre los estudios de Hu et al. (2020) y Adiwal et al. (2023), que se enfocaron en evaluar el desempeño de tres herramientas de Prevención y Detección de Intrusos de Código Abierto (NIDPS), arrojan resultados notables. En particular, estas investigaciones ilustran que Suricata, mediante una explotación altamente eficiente de la arquitectura de subprocesos múltiples de la infraestructura subyacente, supera a Snort y Zeek en los modos de Detección y Prevención de Intrusos (IDS e IPS); este estudio concluye que Suricata es el mejor algoritmo de código abierto que existe actualmente debido a que es capaz de procesar cargas de red más altas que todos los algoritmos de su tipo; además, este algoritmo no es muy complejo y utiliza una menor cantidad de recursos de hardware.

En segundo lugar, en lo que respecta a los IDS basados en anomalías, autores como Huang et al. (2023), Hadem et al. (2021), junto con Gu & Lu (2021), señalan que los algoritmos basados en anomalías alcanzan su mayor potencial y efectividad (más del 98 %) cuando se integran entre sí. Estos autores concluyen que los algoritmos basados en anomalías reducen su capacidad de protección si se utilizan de manera individual.

También se destaca el trabajo de Zhang et al. (2020), Santhi & Srinivasan (2023), Wang et al. (2022), Bai et al. (2023) y Priya et al. (2020), quienes han contribuido significativamente a

la evolución de las técnicas de detección de intrusiones. Han mejorado la precisión, la eficiencia y la adaptabilidad de estos algoritmos para hacer frente a las amenazas cibernéticas, desarrollando algoritmos complejos basados en enfoques simples, teniendo en cuenta el uso y el campo específico en el que se aplicarán.

Aunque los IDS han evolucionado significativamente, todavía existen áreas que requieren mejoras. Martins et al. (2022) enfatizan la necesidad de mejorar los IDS en el contexto del Internet de las Cosas (IoT), para que sean en línea, continuos y adaptables. Esta idea es respaldada por Jasim & Kurnaz (2023), quienes, a pesar de reconocer la eficacia de las redes neuronales, advierten sobre su limitación debido al alto requerimiento de cálculos.

En el contexto de los sistemas de AMI, Shokry et al. (2022) destacan la importancia de los IDS debido a la gran cantidad de ataques en sus distintas capas. Corroborando esta idea, Corallo et al. (2020) sugieren que la implementación de IDS podría disminuir el robo de secretos industriales y propiedad intelectual.

Además, Deb & Roy (2022) subrayan la importancia de los IDS en el host final/canal de control, proporcionando así una capa adicional de seguridad. Belal & Sundaram (2022) enfatizan la importancia de los IDS basados en los datos de las organizaciones y su eficacia contra ataques del tipo Man-in-the-Middle (MITM). Finalmente, Almiani et al. (2020) proponen una solución factible para los ataques en entornos nebulosos mediante el uso de IDS basados en redes neuronales multicapa.

Las investigaciones futuras deberían centrarse en explorar la interacción de los IDS con diversas tecnologías, como su integración con la Gestión de Identidad y Acceso (IAM), con el objetivo de mejorar aún más la seguridad de la información.

Este estudio presenta algunas limitaciones importantes que deben destacarse. En primer lugar, la revisión podría no haber cubierto de manera equitativa todos los enfoques de IDS debido a la amplia cantidad de información disponible en este campo. Por otro lado, las fortalezas de la revisión se derivan de su impacto significativo en la era digital actual, donde la protección de datos es fundamental.

5. Conclusiones

Los Sistemas de Detección de Intrusos (IDS) desempeñan un papel fundamental en la seguridad de la información debido a su diversidad y capacidad de adaptación. Esta diversidad hace referencia a la existencia de diversos tipos de algoritmos basados en dos principios fundamentales: firmas y anomalías. Los basados en firmas demuestran una alta eficacia en la detección de ataques de naturaleza conocida, respaldados por las bases de datos del algoritmo, pero presentan una eficacia reducida cuando se enfrentan a ataques de naturaleza desconocida. Además, se observó que son particularmente efectivos en campos como la informática en la nube y sistemas industriales, donde la cantidad de información y la necesidad de gestionar incidentes relacionados con ataques son más prominentes.

En cuanto a los algoritmos basados en anomalías, tienden a mostrar una eficacia limitada cuando se aplican individualmente. Sin embargo, cuando se integran con otros algoritmos según los requisitos específicos de cada campo y su uso previsto, pueden superar

en eficacia a los basados en firmas. Los algoritmos basados en anomalías se pueden aplicar en una amplia gama de áreas, siempre y cuando se realice un análisis para determinar los requisitos específicos y garantizar su correcta integración a través de modificaciones o en combinación con otros algoritmos.

Debido a lo mencionado anteriormente, se recomienda el uso de IDS basados en anomalías, ya que, cuando se combinan con la aplicación de técnicas de correlación de eventos y análisis de tráfico de red, se obtienen algoritmos altamente eficaces y confiables. No obstante, es necesario tener en cuenta que, incluso con la aplicación de diversas técnicas, aún existen áreas de mejora, como los altos requisitos de cálculo.

En última instancia, se enfatiza la importancia de implementar sistemas de seguridad en las organizaciones. Por lo tanto, se recomienda a las organizaciones que consideren la implementación de estas medidas, teniendo en cuenta el entorno al que están sujetas y los posibles ataques a los que podrían estar expuestas.

6. Referencias Bibliográficas

Abubakar, I. R., Maniruzzaman, K. M., Dano, U. L., AlShihri, F. S., AlShammari, M. S., Ahmed, S. M. S., Al-Gehlani, W. A. G., & Alrawaf, T. I. (2022). Environmental sustainability impacts of solid waste management practices in the Global South. *International Journal of Environmental Research and Public Health, 19*(19), 12717. https://doi.org/10.3390/ijerph191912717

Adiwal, S., Rajendran, B., Shetty, P., & Sudarsan, S. D. (2023). DNS Intrusion Detection (DID) — A SNORT-based solution to detect DNS Amplification and DNS Tunneling attacks. *Franklin Open, 2*(100010), 100010. https://doi.org/10.1016/j.fraope.2023.100010

Aksu, D., & Aydin, M. A. (2022). MGA-IDS: Optimal feature subset selection for anomaly detection framework on in-vehicle networks-CAN bus based on genetic algorithm and intrusion detection approach. *Computers & Security, 118*(102717), 102717. https://doi.org/10.1016/j.cose.2022.102717

Al-Abassi, A., Karimipour, H., Dehghantanha, A., & Parizi, R. M. (2020). An ensemble deep learning-based cyber-attack detection in industrial control system. *IEEE access: practical innovations, open solutions, 8,* 83965–83973. https://doi.org/10.1109/access.2020.2992249

Alhowaide, A., Alsmadi, I., & Tang, J. (2021). Ensemble detection model for IoT IDS. *Internet of Things, 16*(100435), 100435. https://doi.org/10.1016/j.iot.2021.100435

Almiani, M., AbuGhazleh, A., Al-Rahayfeh, A., Atiewi, S., & Razaque, A. (2020). Deep recurrent neural network for IoT intrusion detection system. *Simulation Modelling Practice and Theory, 101*(102031), 102031. https://doi.org/10.1016/j.simpat.2019.102031

Anantha Krishnan, V., & Senthil Kumar, N. (2022). Robust soft computing control algorithm for sustainable enhancement of renewable energy sources based microgrid: A hybrid Garra rufa fish optimization – Isolation forest approach. *Sustainable Computing Informatics and Systems, 35*(100764), 100764. https://doi.org/10.1016/j.suscom.2022.100764

Aslam, B., Maqsoom, A., Tahir, M. D., Ullah, F., Rehman, M. S. U., & Albattah, M. (2022). Identifying and ranking landfill sites for municipal solid waste management: An

integrated remote sensing and GIS approach. *Buildings, 12*(5), 605. https://doi.org/10.3390/buildings12050605

Asniar, Maulidevi, N. U., & Surendro, K. (2022). SMOTE-LOF for noise identification in imbalanced data classification. *Journal of King Saud University - Computer and Information Sciences, 34*(6), 3413–3423. https://doi.org/10.1016/j.jksuci.2021.01.014

Ay, M., Özbakır, L., Kulluk, S., Gülmez, B., Öztürk, G., & Özer, S. (2023). FC-kmeans: Fixed-centered K-means algorithm. *Expert Systems with Applications, 211*(118656), 118656. https://doi.org/10.1016/j.eswa.2022.118656

Bai, X., Xie, Z., Xu, X., & Xiao, Y. (2023). An adaptive threshold fast DBSCAN algorithm with preserved trajectory feature points for vessel trajectory clustering. *Ocean Engineering, 280*(114930), 114930. https://doi.org/10.1016/j.oceaneng.2023.114930

Baldomero M., Martínez I., & Rodríguez- M. (2021). A robust SVM-based approach with feature selection and outliers detection for classification problems. *Expert Systems with Applications, 178*(115017), 115017. https://doi.org/10.1016/j.eswa.2021.115017

Belal, M. M., & Sundaram, D. M. (2022). Comprehensive review on intelligent security defences in cloud: Taxonomy, security issues, ML/DL techniques, challenges and future trends. *Journal of King Saud University - Computer and Information Sciences, 34*(10), 9102–9131. https://doi.org/10.1016/j.jksuci.2022.08.035

Corallo, A., Lazoi, M., & Lezzi, M. (2020). Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts. *Computers in Industry, 114*(103165), 103165. https://doi.org/10.1016/j.compind.2019.103165

Deb, R., & Roy, S. (2022). A comprehensive survey of vulnerability and information security in SDN. *Computer Networks, 206*(108802), 108802. https://doi.org/10.1016/j.comnet.2022.108802

Dhillon, G., Smith, K., & Dissanayaka, I. (2021). Information systems security research agenda: Exploring the gap between research and practice. Journal of Strategic Information Systems, 30(4), 101693. https://doi.org/10.1016/j.jsis.2021.101693

Friha, O., Ferrag, M. A., Benbouzid, M., Berghout, T., Kantarci, B., & Choo, K.-K. R. (2023). 2DF-IDS: Decentralized and differentially private federated learning-based intrusion detection system for industrial IoT. *Computers & Security, 127*(103097), 103097. https://doi.org/10.1016/j.cose.2023.103097

Ghasempour, A. (2019). Internet of things in Smart Grid: Architecture, applications, services, key technologies, and challenges. *Inventions, 4*(1), 22. https://doi.org/10.3390/inventions4010022

Gu, J., & Lu, S. (2021). An effective intrusion detection approach using SVM with naïve Bayes feature embedding. *Computers & Security, 103*(102158), 102158. https://doi.org/10.1016/j.cose.2020.102158

Hadem, P., Saikia, D. K., & Moulik, S. (2021). An SDN-based intrusion detection system using SVM with selective logging for IP traceback. *Computer Networks, 191*(108015), 108015. https://doi.org/10.1016/j.comnet.2021.108015

Hanafi, N., & Saadatfar, H. (2022). A fast DBSCAN algorithm for big data based on efficient

density calculation. *Expert Systems with Applications, 203*(117501), 117501. https://doi.org/10.1016/j.eswa.2022.117501

Hu, H., Liu, J., Zhang, X., & Fang, M. (2023). An effective and adaptable K-means algorithm for big data cluster analysis. *Pattern Recognition, 139*(109404), 109404. https://doi.org/10.1016/j.patcog.2023.109404

Hu, Q., Yu, S.-Y., & Asghar, M. R. (2020). Analysing performance issues of open-source intrusion detection systems in high-speed networks. J*ournal of Information Security and Applications, 51*(102426), 102426. https://doi.org/10.1016/j.jisa.2019.102426

Huang, J.-C., Zeng, G.-Q., Geng, G.-G., Weng, J., Lu, K.-D., & Zhang, Y. (2023). Differential evolution-based convolutional neural networks: An automatic architecture design method for intrusion detection in industrial control systems. *Computers & Security, 132*(103310), 103310. https://doi.org/10.1016/j.cose.2023.103310

Huang, X., Ma, T., Liu, C., & Liu, S. (2023). GriT-DBSCAN: A spatial clustering algorithm for very large databases. *Pattern Recognition, 142*(109658), 109658. https://doi.org/10.1016/j.patcog.2023.109658

Idrissi, M. J., Alami, H., El Mahdaouy, A., El Mekki, A., Oualil, S., Yartaoui, Z., & Berrada, I. (2023). Fed-ANIDS: Federated learning for anomaly-based network intrusion detection systems. *Expert Systems with Applications, 234*(121000), 121000. https://doi.org/10.1016/j.eswa.2023.121000

Jasim, A. F. J., & Kurnaz, S. (2023). New automatic (IDS) in IoTs with artificial intelligence technique. *Optik, 273*(170417), 170417. https://doi.org/10.1016/j.ijleo.2022.170417

Krishna, M. V. B. M., Ananth, C. A., & Krishnaraj, N. (2023). Detection of intrusions in clustered vehicle networks using invasive weed optimization using a deep wavelet neural networks. *Measurement. Sensors, 28*(100807), 100807. https://doi.org/10.1016/j.measen.2023.100807

Kumar, R., Venkanna, & Tiwari, V. (2023). Optimized traffic engineering in Software Defined Wireless Network based IoT (SDWN-IoT): State-of-the-art, research opportunities and challenges. *Computer Science Review, 49*(100572), 100572. https://doi.org/10.1016/j.cosrev.2023.100572

Lee, J.-S., Chen, Y.-C., Chew, C.-J., Chen, C.-L., Huynh, T.-N., & Kuo, C.-W. (2022). CoNN-IDS: Intrusion detection system based on collaborative neural networks and agile training. *Computers & Security, 122*(102908), 102908. https://doi.org/10.1016/j.cose.2022.102908

Li, W., Tug, S., Meng, W., & Wang, Y. (2019). Designing collaborative blockchained signature-based intrusion detection in IoT environments. *Future Generations Computer Systems: FGCS, 96*, 481–489. https://doi.org/10.1016/j.future.2019.02.064

Lifandali, O., Abghour, N., & Chiba, Z. (2023). Feature selection using a combination of ant colony optimization and random forest algorithms applied to isolation forest based intrusion detection system. *Procedia Computer Science, 220*, 796–805. https://doi.org/10.1016/j.procs.2023.03.106

Martins, I., Resende, J. S., Sousa, P. R., Silva, S., Antunes, L., & Gama, J. (2022). Host-based IDS:

A review and open issues of an anomaly detection system in IoT. *Future Generations Computer Systems: FGCS, 133*, 95–113. https://doi.org/10.1016/j.future.2022.03.001

Masdari, M., & Khezri, H. (2020). A survey and taxonomy of the fuzzy signature-based Intrusion Detection Systems. *Applied Soft Computing, 92*(106301), 106301. https://doi.org/10.1016/j.asoc.2020.106301

Mohammadi, M., Rashid, T. A., Karim, S. H. T., Aldalwie, A. H. M., Tho, Q. T., Bidaki, M., Rahmani, A. M., & Hosseinzadeh, M. (2021). A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. *Journal of Network and Computer Applications, 178*(102983), 102983. https://doi.org/10.1016/j.jnca.2021.102983

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology, 134*, 178–189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Priya, S., Maddikunta, P. K. R., Parimala, Koppu, S., Gadekallu, T. R., Chowdhary, C. L., & Alazab, M. (2020). An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. *Computer Communications, 160*, 139–149. https://doi.org/10.1016/j.comcom.2020.05.048

Santhi, T. M., & Srinivasan, K. (2023). A duo autoencoder-SVM based approach for secure performance monitoring of industrial conveyor belt system. *Computers & Chemical Engineering, 177*(108359), 108359. https://doi.org/10.1016/j.compchemeng.2023.108359

Saranya, T., Sridevi, S., Deisy, C., Chung, T. D., & Khan, M. K. A. A. (2020). Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science, 171*, 1251–1260. https://doi.org/10.1016/j.procs.2020.04.133

Shokry, M., Awad, A. I., Abd-Ellah, M. K., & Khalaf, A. A. M. (2022). Systematic survey of advanced metering infrastructure security: Vulnerabilities, attacks, countermeasures, and future vision. *Future Generations Computer Systems: FGCS, 136*, 358–377. https://doi.org/10.1016/j.future.2022.06.013

Spring, J. M. (2023). An analysis of how many undiscovered vulnerabilities remain in information systems. *Computers & Security, 131*(103191), 103191. https://doi.org/10.1016/j.cose.2023.103191

Tiwari, A., Saraswat, S., Dixit, U., & Pandey, S. (2022). Refinements in Zeek intrusion detection system. *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS).* https://doi.org/10.1109/ICACCS54159.2022.9785047

Waleed, A., Jamali, A. F., & Masood, A. (2022). Which open-source IDS? Snort, suricata or Zeek. *Computer Networks, 213*(109116), 109116. https://doi.org/10.1016/j.comnet.2022.109116

Wang, C.-L., Chan, Y.-K., Chu, S.-W., & Yu, S.-S. (2022). r-Reference points based k-means algorithm. *Information Sciences, 610*, 204–214.

https://doi.org/10.1016/j.ins.2022.07.166

Zhang, J., You, H., & Jia, R. (2020). Reliability hazard characterization of wafer-level spatial metrology parameters based on LOF-KNN method. *Microelectronics and Reliability, 107*(113599), 113599. https://doi.org/10.1016/j.microrel.2020.113599

Zolanvari, M., Teixeira, M. A., Gupta, L., Khan, K. M., & Jain, R. (2019). Machine learning-based network vulnerability analysis of industrial internet of things. *IEEE internet of things journal, 6*(4), 6822–6834. https://doi.org/10.1109/jiot.2019.2912022