Sistema de Detección de Intrusos mediante Técnicas de Aprendizaje Automático: Una revisión*

Angie P. Solórzano[1,*,†]

[1]Universidad Laica Eloy Alfaro de Manabí - Extensión en El Carmen (ULEAM)), Eloy Alfaro Delgado St, El Carmen, 130401, Federación Ecuatoriana

Abstract

El rápido crecimiento en el uso de redes informáticas plantea desafíos en cuanto a la disponibilidad, integridad y confidencialidad de las mismas. Como respuesta, los administradores de redes adoptan sistemas de detección de intrusiones (IDS) para monitorear y alertar sobre actividades maliciosas. Estos IDS pueden detectar amenazas mediante la comparación con patrones. Este artículo presenta una revisión de investigaciones sobre IDS eficientes utilizando clasificadores de aprendizaje automático, evaluados en múltiples conjuntos de datos. Los resultados y comparaciones proporcionan una guía para futuras investigaciones en este campo.

Keywords

Sistema de Detección de Intrusos, aprendizaje Automático, seguridad, detección de anomalías, clasificadores: Conjunto, Híbrido, Débil, ULEAM.

1. Introducción

Los Sistemas de Detección de Intrusiones (IDS) detectan amenazas en redes informáticas mediante firmas o la detección de anomalías. Pueden ser basados en host (HIDS) o en red (NIDS).

Hasta hace poco, los sistemas de detección de intrusiones (IDS) se basaban en firmas, dependiendo de colecciones predefinidas de ataques conocidos. No obstante, este enfoque tenía un inconveniente importante: las firmas debían actualizarse continuamente debido a las nuevas tácticas de los atacantes.

Pero con la llegada del aprendizaje automático, se abrió una nueva puerta: la detección de anomalías. Esta técnica permite detectar ataques desconocidos al comparar los parámetros de actividad de usuarios legítimos con eventos que se desvían de esas actividades benignas. A lo largo de los años, se han adoptado varias técnicas de aprendizaje automático con el propósito de mejorar la tasa de detección, reducir falsos positivos y aumentar la precisión predictiva de los IDS.

El artículo consta de cuatro partes que incluyen una introducción, una revisión de investigaciones, comparaciones y discusiones sobre técnicas de aprendizaje automático en IDS.

P. Angie (Eds.), Sistema de Detección de Intrusos mediante Técnicas de Aprendizaje Automático: Una revisión (IEEE 2023), ULEAM, El Carmen, Manabí, Ecuador, Octubre 05, 2023.

*You can use this document as the template for preparing your publication.

*Corresponding author.

angieponce364@gmail.com (A. P. Solórzano)

0009-0007-4554-5281 (A. P. Solórzano)

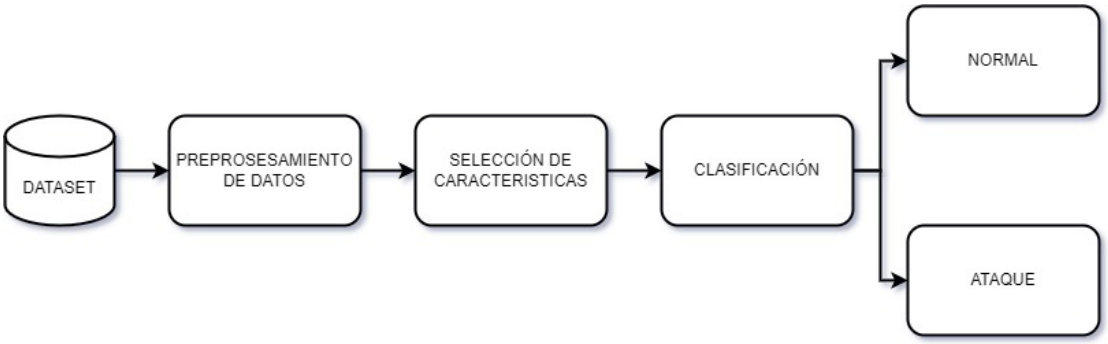

Figure 1: Diagrama de bloques del sistema de detección de intrusos

2. Resumen del Trabajo de Investigación

2.1. Aprendizaje Automático (ML)

Como rama de la Inteligencia Artificial (IA), el ML puede definirse como una técnica en la que los ordenadores son entrenados para tener la capacidad de aprender y mejorar u optimizar automáticamente criterios de rendimiento sin ser programados expresamente, utilizando experiencias pasadas o datos.

Por ejemplo, Imagina que tienes un robot al que no le dices exactamente lo que debe hacer, sino que le enseñas cosas y él aprende a hacerlas por sí solo.

2.2. Clasificadores Individual

Un clasificador de aprendizaje automático único es simplemente un tipo de programa que ayuda a identificar cosas. Algunos sistemas de seguridad utilizan estos programas para detectar intrusiones en redes.

Algunos ejemplos de estos programas son SVM, redes neuronales, árboles de decisión, KNN y Naïve Bayes. Se utilizan en diferentes sistemas de seguridad que hemos revisado.

2.3. Clasificadores Híbridos

Un clasificador híbrido es un equipo de algoritmos que trabajan juntos para protegernos de los ataques cibernéticos de manera más eficiente. Combina diferentes algoritmos de aprendizaje automático para mejorar la detección de intrusos.

2.4. Clasificador Conjunto

Un clasificador ensemble es un conjunto de múltiples clasificadores de aprendizaje automático, a veces llamados aprendices débiles. Estos clasificadores toman decisiones individuales que se combinan de alguna manera para producir un rendimiento predictivo más efectivo a través de

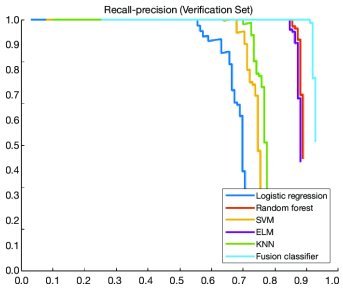

Figure 2: Clasificador Simple

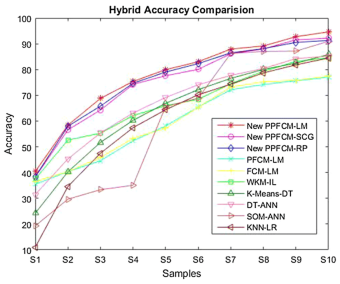

Figure 3: Clasificador Híbrido

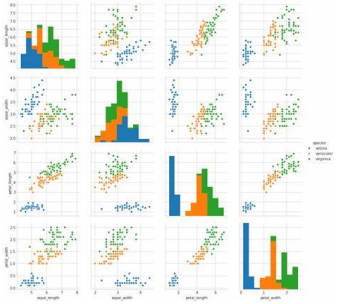

Figure 4: Clasificador Conjunto

un consenso. En otras palabras, un clasificador ensemble mejora su rendimiento al agregar los resultados de varios clasificadores débiles.

3. Revisión y Comparación de Trabajos Relacionados

Cada autor analiza varios enfoques y técnicas en el campo de los Sistemas de Detección de Intrusiones (IDS) en el contexto de mejorar la precisión y eficiencia en la identificación de intrusos en redes.

Alkasassbeh y Almseidin: Usaron tres técnicas de clasificación, que incluyen árboles J48, Perceptrón Multicapa y algoritmos de redes Bayes, para mejorar la precisión de los IDS al tratar con ataques de baja frecuencia. La mejor precisión se logró con los árboles J48, pero enfrentaron desafíos con la selección de características.

Bhavani et al: Construyeron un IDS utilizando técnicas de bosque aleatorio y árboles de decisión en el conjunto de datos KDD-NSL, logrando una alta precisión del 95,323%. Sin embargo, no pudieron abordar completamente las tasas de detección baja y falsos positivos.

Ponthapalli et al: Usaron algoritmos de aprendizaje automático únicos como árboles de decisión, regresión logística, bosque aleatorio y máquina de vectores de soporte para detectar intrusiones en redes en el conjunto de datos KDD-NSL. El bosque aleatorio funcionó mejor, pero el sistema funcionaba eficientemente solo con un conjunto de datos

Marzia Z. y Chung-Horng L.: Implementaron un IDS basado en conjuntos que utilizaba múltiples algoritmos de aprendizaje automático, agregando resultados con un clasificador de votación. Si bien esto mejoró la precisión, tuvo problemas con una alta tasa de falsos negativos.

Dutt I. et al: Propusieron un enfoque de detección de intrusiones híbrido en tiempo real que utilizaba tanto el enfoque de uso incorrecto como el de anomalía para detectar ataques

conocidos y novedosos. Se logró una alta tasa de detección, pero la detección lenta en datos grandes seguía siendo un desafío.

Verma et al Aplicaron Extreme Gradient Boosting y Adaptive Boosting en el conjunto de datos NSL-KDD, obteniendo una precisión del 84.253%. Sugirieron la necesidad de enfoques híbridos o de conjunto para mejorar aún más el rendimiento.

Kazi Abu Taher et al: Evaluaron diferentes modelos de aprendizaje automático con selección de características en el conjunto de datos NSL-KDD, logrando una precisión mejorada. Sin embargo, el modelo tuvo dificultades para la detección de ataques desconocidos y se centró solo en ataques basados en firmas.

Zhou et al: Presentaron un IDS que combinaba clasificadores de conjunto con selección de características. Logrando una alta precisión en tres conjuntos de datos, incluidos NSL-KDD, CIC-IDS2017 y AWID.

Ahmad Iqbal y Shabib Aftab: Utilizaron redes neuronales de avance y reconocimiento de patrones con técnicas de entrenamiento de regularización bayesiana y gradiente conjugado escalado. Las redes neuronales de avance proporcionaron una mejor precisión del 98.0742%. Pero se necesitaba probar el sistema en diferentes conjuntos de datos.

Vinoth Y. K y Kamatchi K.: Emplearon un enfoque de conjunto que combinaba varios clasificadores para manejar datos desequilibrados. El modelo funcionó bien en NSL-KDD, pero se requería realizar más pruebas en conjuntos de datos actualizados.

Maniriho et al: Utilizaron técnicas de aprendizaje automático único y en conjunto en los conjuntos de datos NSL-KDD y UNSW NB-15, centrándose en la selección de características. El enfoque en conjunto superó al aprendizaje automático único, pero se debían abordar el tamaño de los datos y la dimensionalidad.

Rajagopal et al Propusieron un enfoque de conjunto apilado utilizando varios conjuntos de datos, incluidos UNSW NB-15 y UGR '16. El modelo logró una alta precisión, especialmente con UGR '16, pero se necesitaban más experimentos con categorías de ataques recientes.

Perez D. et al: Presentaron un IDS híbrido basado en redes que utilizaba múltiples técnicas híbridas de aprendizaje automático en el conjunto de datos NSL-KDD. La combinación de SVM y agrupamiento K-means con selección de características proporcionó la mejor precisión, y se sugirió la construcción de más modelos híbridos para mejorar las tasas de falsos positivos.

A. Comparación de trabajos relacionados

En esta revisión de investigación analiza varios artículos publicados entre 2015 y 2020 que se centran en sistemas de detección de intrusiones. Estos trabajos emplearon principalmente

clasificadores individuales, híbridos y de conjunto. La tabla 1 resume la comparación, principalmente en términos de precisión, entre los distintos algoritmos utilizados en los artículos de investigación estudiados. Como se observa en la figura 5, el clasificador por conjuntos ofrece la máxima precisión cuando se emplea a lo largo de los años.

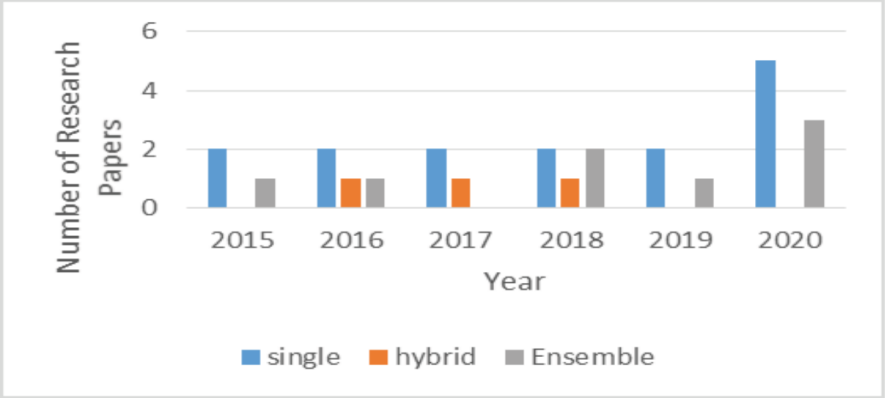

Figure 5: Agrupación de trabajos de investigación en función del tipo de clasificador utilizado

B. Conjuntos de Datos Utilizados en los Trabajos de Investigación

Un conjunto de datos es una colección de instancias, donde una instancia se refiere a una sola fila de datos. Cada instancia está compuesta por múltiples características, que a menudo se llaman atributos de una instancia de datos. En los trabajos de investigación estudiados, se utilizó principalmente el conjunto de datos KDD-NSL, aunque se mencionaron siete conjuntos de datos diferentes en total, que incluyen KDD Cup '99, KDD-NSL, Kyoto2006+, AWID, CIC-IDS2017, UNSW NB-15 y UGR'16.

El conjunto de datos KDD Cup '99, por ejemplo, se usó en la competencia KDD Cup '99 y consta de 41 características que representan conexiones TCP, tanto cualitativas como cuantitativas. El conjunto de datos modificado NSL-KDD se creó como una versión mejorada del KDD Cup '99 para resolver algunos de sus problemas, y contiene 41 características, 38 de las cuales son atributos numéricos y tres son atributos nominales. También contiene 24 tipos de ataques en el conjunto de entrenamiento, con 14 tipos adicionales en el conjunto de datos. El conjunto de entrenamiento consta de 125,973 puntos de datos, donde el 53.4% son conexiones normales y el 47.6% son ataques. Además, se menciona el conjunto de datos Kyoto2006+, que se basó en datos reales de tráfico recopilados durante tres años utilizando 348 honeypots en la Universidad de Kyoto. Este conjunto de datos contiene 24 características, de las cuales 14 son las mismas que las del conjunto de datos original KDD Cup '99, y las 10 características adicionales aportan información adicional sobre los desafíos que a menudo se enfrentan cuando se utiliza el conjunto de datos KDD Cup '99.

Título	Algoritmo	Conjunto de Datos	Hallazgo	Fallo
IDS mediante bagging con clasificador basado en árbol de decisión parcial	1) Algoritmo genético (GA) basado en selección de características. 2) Clasificador en bolsas con árbol de decisión	NLS-KDD99	Reducción de las falsas alarmas	Se necesitó mucho tiempo para construir el modelo
IDS basado en la combinación de centros de conglomerados y vecinos más próximos	1) k-Nearest Neighbor (k-NN) 2) Cluster Center and Nearest Neighbor (CANN) 3) Support Vector Machine	KDD-Cup99	La representación de características se aplicó a las conexiones normales y a los ataques	Los ataques U2L y R2L no fueron detectados detectados por CANN
Comparación de técnicas de clasificación aplicadas a la detección y clasificación de intrusiones en la red	1. Árbol de amplitud forestal (BFTree) y NBTree). 2. J48 y Naïve Baye 3. Árbol forestal aleatorio (RFT). 4. Perceptrón multicapa (MLP).	NSL-KDD	Reducción lograda de falsos positivos	Es necesario evaluar el modelo en los conjuntos de datos actualizados.
Modelado Random Forest para IDS de red	Clasificador conjunto basado en Random Forest (RF)	NSL-KDD	El modelo es eficaz, ya que produce pocas falsas alarmas y un alto índice de detección.	Aplicar un método de selección de características como la computación evolutiva para mejorar la precisión
Detección de anomalías basada en la firma de perfiles en la red mediante técnicas de ML	1. Algoritmo genético (AG). 2. Máquina de vectores de apoyo (SVM). 3. Modelo híbrido.	KDDCup '99	Baja tasa de falsos positivos	Es necesario realizar ensayos con diferentes conjuntos de datos
Clasificadores KNN rápidos para sistemas de detección de intrusiones en redes	K-Nearest Neighbor (KNN)	NSL-KDD	Alta precisión alcanzada.	Tiempo de cálculo elevado debido a la imposibilidad de aplicar la selección de características.

Table 1
Distribución del uso de los conjuntos de datos a lo largo de los años

AÑO	KDD Cup 99	NSL-KDD	CIC-IDS 2017	UNSW NB-15	UGR‘16
2015	1	1	0	0	0
2016	1	2	0	0	0
2017	0	3	0	0	0
2018	2	1	1	0	0
2019	0	3	0	1	0
2020	0	4	0	0	2

Table 2
Distribución del uso de los conjuntos de datos a lo largo de los años

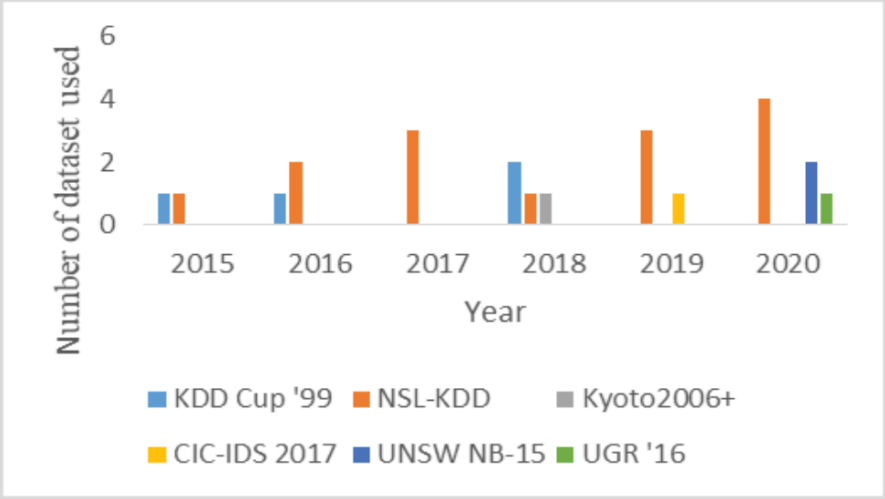

Figure 6: Distribución del uso de los conjuntos de datos a lo largo de los años

Se mencionan varios conjuntos de datos utilizados en la investigación, como AWID, CIC-IDS2017 y UNSW NB-15. La tabla 2 y la figura 6 muestran cómo estos conjuntos de datos se han utilizado a lo largo del tiempo en varios artículos. El conjunto de datos NSL-KDD se utilizó con mayor frecuencia (14 veces), seguido por el conjunto de datos original KDD Cup 99 (4 veces), mientras que otros conjuntos se utilizaron una o dos veces.

4. Debate y Trabajo Futuro

De acuerdo a la Figura 3, se observa que los clasificadores combinados y híbridos muestran una mejor capacidad de predicción y detección en comparación con los clasificadores individuales. Para futuras investigaciones, se han identificado ciertos aspectos clave que deben ser considerados para mejorar el desempeño de los sistemas de detección de intrusiones. Esto incluye la necesidad de emplear más frecuentemente técnicas híbridas y de conjunto para combinar clasificadores de aprendizaje automático, ya que algunos funcionan mejor en ciertos conjuntos de datos. Además, es esencial aplicar la selección de características antes de la clasificación con el objetivo de eliminar características irrelevantes y redundantes, lo que mejora la eficiencia y

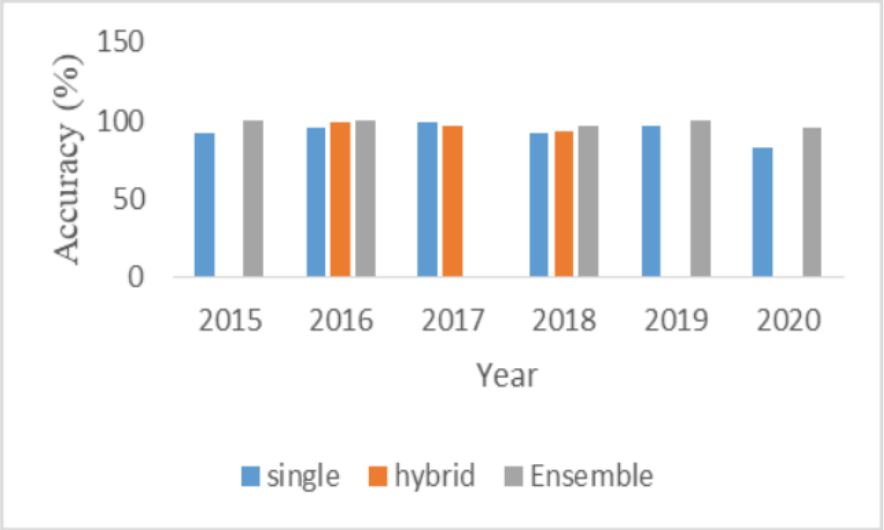

Figure 7: Comparación de clasificadores en términos de precisión

la tasa de detección de los sistemas de detección de intrusiones.

También se señala que una parte significativa de los trabajos de investigación se basó en el conjunto de datos KDD-NSL, por lo que se recomienda el uso de conjuntos de datos más actualizados en futuras investigaciones para abordar amenazas y ataques más recientes.

5. Conclusión

El período comprendido entre 2015 y 2020 presenció avances significativos en el campo de la detección de intrusiones gracias a la aplicación del aprendizaje automático.

La investigación examinada en este artículo revela que los clasificadores de conjunto e híbridos han surgido como protagonistas destacados, superando a los clasificadores individuales en términos de precisión predictiva y tasas de detección.

Estos resultados subrayan la importancia de considerar enfoques de clasificación más avanzados en la construcción de sistemas de detección de intrusiones.

Además, se enfatiza la necesidad de trabajar con conjuntos de datos más actuales para mantenerse al tanto de las amenazas y ataques en constante evolución. El futuro de la detección de intrusiones parece prometedor, con un enfoque cada vez más centrado en el aprendizaje automático y en la mejora de la eficiencia y precisión de los sistemas.

References

[1] D. Gaikwad, R. C. Thool, Intrusion detection system using bagging with partial decision tree base classifier, Procedia Computer Science 49 (2015) 92–98.

[2] W. in, K. Shih-Wen, T. Chih-Fong, Intrusion detection system based on combining cluster centers and nearest neighbors, Knowledge-Based Systems 78 (2015) 13–21.

[3] A. Aziz, Comparison of classification techniques applied for network intrusion detection and classification, Journal of Applied Logic 24 (2016) 109–118.

[4] N. Farnaaz, M. Jabbar, Random forest modeling for network intrusion detection system, in: International Multi-conference on information processing (IMCIP), volume 12, Elsevier, 2016, pp. 213–217.

[5] K. A. Saadiah, R. Amirali, S. Hazyanti, Anomaly detection based on profile signature in network using machine learning techniques, in: IEEE TENSYMP, 2016, pp. 71–76.

[6] B. Brao, K. Swathi, Fast knn classifiers for network intrusion detection system, Indian Journal of Science and Technology 10 (2017) 1–10.

[7] L. Chie-Hong, S. Yann-Yean, Y.-C. Lin, S.-J. L., Machine learning based network intrusion detection, in: 2nd IEEE International Conference on Computational Intelligence and Applications, 2017, pp. 79–83.

[8] P. Deyban, A. Miguel A., A. David P., S. Eugenio, Intrusion detection in computer networks using hybrid machine learning techniques, in: XLIII Latin American Computer Conference (CLEI), 2017, pp. 1–10.

[9] M. Alkasassbeh, M. Almseidin, Machine learning methods for network intrusions, in: International Confrernce on Computing, Communication (ICCCNT), 2018. arXiv:1803.04741.

[10] M. Z., C.-H. L., Evaluation of machine learning techniques for network intrusion detection, in: IEEE, 2018, pp. 1–5.

[11] e. a. Dutt I., Real time hybrid intrusion detection system, in: International Conference on Communication, Devices and Networking (ICCDN), Springer, 2018, pp. 885–894.

[12] S. A. S. B. Verma P., Shadab K., Network intrusion detection using clustering and gradient boosting, in: International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1–7.

[13] M. R. Kazi A., Billal M., Network intrusion detection using supervised machine learning technique with feature selection, in: International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 2019, pp. 643–646.

[14] S. J. M. D. Yuyang Z., Guang C., Building an efficient intrusion detection system based on feature selection and ensemble classifier, Computer Networks (2020). doi:https://doi.org/10.1016/j.comnet.2020.107247.

[15] M. Iqbal, M. Aftab, A feed-forward ann and pattern recognition ann model for network intrusion detection, International Journal of Computer Network and Information Security 4 (2019) 19–25.

[16] V. Y., K. K., Anomaly based network intrusion detection using ensemble machine learning technique, International Journal of Research in Engineering, Science and Management IJRESM (2020) 290–296.

[17] K. M. R. Bhavani T. T, M. A. R, Network intrusion detection system using random forest and

decision tree machine learning techniques, in: International Conference on Sustainable Technologies for Computational Intelligence (ICSTCI), Springer, 2020, pp. 637–643.

[18] e. a. Maniriho, Detecting intrusions in computer network traffic with machine learning approaches, International Journal of Intelligent Engineering and Systems INASS (2020) 433–445.

[19] e. a. Ponthapalli R., Implementation of machine learning algorithms for detection of network intrusion, International Journal of Computer Science Trends and Technology (IJCST) (2020) 163–169.

[20] K. S. H. Rajagopal S., Poornima P. K., A stacking ensemble for network intrusion detection using heterogeneous datasets, Journal of Security and Communication Networks Hindawi (2020) 1–9.

[21] C. Bolon, Feature selection and classification in multiple class datasets-an application to kdd cup 99 dataset, https://doi.org/10.1016/j.eswa.2010.11.028 (2012).

[22] e. a. Farah N. H, Application of machine learning approaches in intrusions detection systems, International Journal of Advanced Research in Artificial Intelligence IJARAI (2015) 9–18.

[23] e. a. Haq, N. F., An ensemble framework for anomaly detection using hybridized feature selection approach (hfsa), in: Intelligent System Conference, 2015, pp. 989–995.

[24] S. Thapa, A. Mailewa, The role of intrusion detection/prevension systems in modern computer networks: A review, in: Midwest Instruction and Computing Symposium (MICS), volume 53, 2020, pp. 1–14.