International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 12 | Dec 2022 www.irjet.net p-ISSN: 2395-0072

Automatic Intrusion Detection based on Artificial Intelligence Techniques: A Review

Raunit Singh[1], Dinesh J[2], Mohammed Rehaan[3], Madhushree P[4] , Dr. John T Mesia Dhas[5]

[1234]*Students, Computer Science and Engineering, T. John Institute of Technology, Bengaluru, Karnataka, India*
[5]*Assoc. Professor, Computer Science and Engineering, T. John Institute of Technology, Bengaluru, Karnataka, India*

Abstract – *An Intrusion Detection System (IDS) is a system that is used to detect and prevent unauthorized access to or use of a computer system or network. IDSs are commonly used in businesses and organizations to protect their computer systems and networks from unauthorized access or use. This study presents the conduction and results of a review.*

Investigating: Automatic Intrusion Detection based on Artificial Intelligence Techniques. This study found that single, hybrid, and ensemble classification algorithms, along with some unsupervised learning techniques, has been used in the development of intrusion detection systems. In addition, soft computing techniques are getting considerable attention, as many have applied them in the context of IDS. The most popular datasets used are the variants of NSL-KDD. This study also found that the most common metrics for evaluating the performance of IDS are Accuracy, Precision, Recall, AUC and F1 Score. However, the focus on the classification of known intrusion attacks may pose a problem in detecting anomalous intrusions, which may include new or modified intrusion attacks. To overcome this, it was recommended that future research should focus on developing novel techniques and algorithms that are capable of detecting such intrusions.

Key Words: Machine learning, Classification, Clustering Ensemble, Intrusion Detection System, Soft Computing,

1. INTRODUCTION

Intrusion Detection Systems (IDSs) can detect attacks that firewalls cannot detect, and it can be used as a preventive measure against attacks [1] . IDSs are mainly divided into two categories: signature-based IDS and anomaly-based IDS. Signature-based IDS is used to detect malicious activities by comparing data with a database of known attack signatures. It has high accuracy for detecting known attacks but it cannot detect unknown attacks. Anomaly-based IDS is used to detect abnormal network activities which may indicate malicious activity or an attack. It can detect unknown attacks that signature-based IDSs can't detect. However, it has the disadvantage of generating a lot of false positives [2].

To increase the effectiveness of IDSs, various techniques have been proposed such as combining different types of IDSs together. A hybrid IDS [3] combines two or more IDSs to create a system that is more effective than either one alone. The hybrid IDS can reduce the number of false positives and improve the detection accuracy of unknown attacks.

Intrusion Detection Systems (IDSs) are used to detect both known and unknown attacks on networks from internal and external sources. However, due to the ever-evolving nature of cyber threats, current network security systems are proving insufficient for protecting computer systems. As a result, it is necessary to develop new methods and improve existing IDS technologies. This paper provides a literature review to investigate and analyse the state of IDSs.

2. BACKGROUND

Intrusion Detection (IDS) is a process of monitoring for suspicious or malicious behaviour and misuse in networks. This practice was developed in the 1980s with the rise of the internet and the need to monitor potential risks, leading to an increased demand for security infrastructures. In the 1980s and 1990s, James P. Anderson worked on Intrusion Detection Expert System (IDES) [4]. During this period, the U.S. government-funded much of the research. Multiple projects such as Haystack, Discovery, Network Audit Director, Intrusion Reporter (NADIR) [5], and others were used to identify intrusions. This led to improved detection techniques that verified data and resulted in better operating systems. IDS and HIDS were first introduced, and around the 1990s, they started to generate revenue and the intrusion detection market began to grow. Genuine Secure, created by ISS, is an example of an intrusion detection network. Additionally, Cisco recognized the need for network intrusion detection and purchased Wheel Group to improve their security arrangements [6].

2.1 Intrusion Detection System

An Intrusion Detection System (IDS) is a security service that monitors and assesses system activities and access challenges in order to identify system resources and unauthorized activities. An intrusion in a network is defined as a series of actions that aim to compromise the confidentiality, integrity, and availability of resources.

IDSs are typically used in conjunction with other protective security techniques, such as authentication and access control. Numerous research works have concluded that the IDS are an essential part of a comprehensive defence system. Historically, many conventional systems and applications were developed without security features. IDS system is a combination of hardware and software that can detect any unauthorized activities from both external and internal users. According to the NIST [7], an IDS is a process of controlling activities within a network or system .Log files can be used to identify intrusions and in various environments, applications and systems are developed to handle different task.

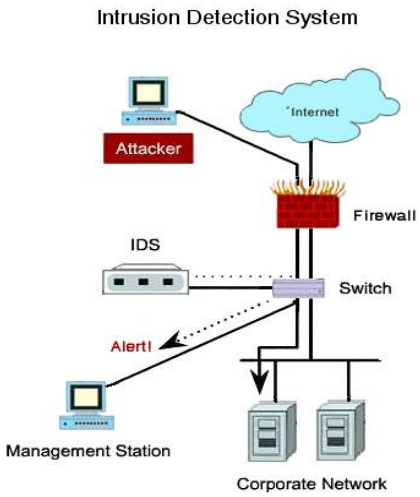

Fig-1: Intrusion detection system[8]

There are two main types of IDS, namely:

- Host Based IDS
- Network Based IDS

Network-based Intrusion Detection System (NIDS) detects network attacks by analysing the payload. Host-based Intrusion Detection System (HIDS) are employed to sense local attacks before they reach the network. Various techniques are used to detect intrusion in networks. To ensure accurate detection of network disruption, multiple detection methods or hybrid methods are used by IDS. The IDS detection models can be classified into [9]:

- Misuse based IDS
- Anomaly based IDS

Misuse identification relies upon the known characteristics in the database. Anomaly detection, on the other hand, is used to discover both computer and network loop-holes. It surveys and sorts the system activities that are out of the ordinary and can distinguish between the typical intrusions and the anomalous ones. Unfortunately, misuse IDS have a high detection rate for known intrusions, but are unable to detect unknown attacks. In the following section, all the models and types of IDS are described:

2.1.1 Host Based Intrusion Detection system

The purpose of Host-Based Intrusion Detection Systems (HIDS) [10] is to observe and monitor the condition and functioning of a computer system. HIDS inspect all data packets in the network to see what resources are being used and by which programs. When alterations or changes are detected on the network, system administrators will be sent warnings. As cyber security continues to become more essential, these systems are increasingly being added to host computer networks to detect any suspicious behaviour, application outages, and threats from external sources while also protecting critical data systems from attackers.

A HIDS is responsible for overseeing what goes on inside a computer system, and identifying any abnormal activities. It is vitally important to have a secure environment for any system; thankfully these systems are capable of providing that safety. Common occurrences with security violations occur due to malicious code and unauthorized events passing through the barriers of a system. These potentially destructive activities and misused code can have a detrimental effect on the system overall. Therefore, the HIDS preventing unapproved access provides greater protection of users' information.

2.1.2 Network Based Intrusion Detection System

Network Intrusion Detection Systems (NIDS) [11] is an attribute function of the target system, which allows one to observe its functioning modules in a network. NIDS can be analysed with either manual or automated techniques and its usage is essential in reinforcing the security infrastructure of any system. To ensure that incoming and outgoing traffic is managed properly, anti-virus software should be installed on servers. NIDS are important for a variety of sectors such as government services, businesses, industries and educational institutions. Classification occurs through signature detection, where patterns normal or abnormal behaviour will be compared to pre-existing log files or signatures. Anomaly-based technique will also detect potential misuse or computer abnormalities by looking at heuristics of signatures. With NIDS it is possible to control incoming and outgoing packets depending on their host/network based classification.

2.1.3 Misuse Attack Detection or Signature Based Detection of IDS

This system of detection is based upon recognizing anomalous behaviour through the components of known security breaches and recognized framework

vulnerabilities [12]. Detection by misuse compares already-known attack signatures to observed actions as a way to distinguish potential intrusions. The downside of this method of detection is that unknown intrusions cannot be noticed due to lack of signature identification. Signature-based intrusion detection systems do have a means of updating the database with new attacks, yet other protection products such as anti-virus software have their own limitations since their signature files are only updated on either a daily or weekly basis. This creates a major problem for computer users, who then become exposed to multiple, unknown intrusions in between these intervals – an issue which has been further posed with the number of new attacks which can circulate online within minutes.

Advantages of Misuse Attack Detection

• Quickly monitors incoming intrusion attempts and identifies them with ease.

• System administrators are also able to enjoy a reduced false positive rate as they wouldn't have to spend much time attempting to handle it.

Disadvantages of Misuse Attack Detection

• While it allows for detection of known attacks, it may not identify unknown attacks due to the fact that its operation relies heavily on static pattern-matching.

• False alarms are also frequent due to this static based approach adopted by this system.

2.1.4 ANOMALY BASED DETECTION OF IDS

Anomaly-based IDS [13] includes three categories, namely semi-supervised, supervised and unsupervised discovery. Statistical-based methods are frequently utilized to recognize network activity from two datasets. The first database is based on the current network profiles obtained over time; the latter being a profile that has been trained statistically in advance. As soon as activities have been observed for the present profile, an anomaly score is computed by comparing the two activities - usually representing the degree of abnormality for that particular activity. Knowledge based strategies are also highly used in this system; wherein using a collection of rules helps label the input data and does so in two stages. First, it identifies dissimilar classes and actions for input training data. Then, a grouping of classification events is extracted from normal operations [32, 33].

Supervised anomaly detection systems attempt to build models separately assigned to identify abnormal records while having normal ones as well – semi-supervised systems do likewise but make use of abnormal data too, though this type requires a labelled database which increases false positive rate [34]. Unsupervised approaches don't require any dataset yet still detect new anomalies [34].

Advantages of Anomaly Based Intrusion Detection System

• The system does not require regular updates to identify any new attacks.

• After the program is installed, some maintenance can be done regularly.

• The application will also analyse the behaviours of the network and create profiles for every activity taking place in it.

• This method supports friendly identification of potential threats in a massive system more efficiently.

Disadvantages of Anomaly-Based IDS

• It cannot detect normal traffic which may appear abnormal, so the admin will not receive any warning signals from the system.

• Due to its setup, there are higher chances of false positives being triggered by the IDS.

2.2 Overview of Machine Learning (ML) Algorithms

Machine Learning (ML) is an approach wherein models are trained to learn and improve their performance parameters automatically, as opposed to having to be programmed with previous experience or example. Based on attributes, the ML model focuses on training datasets to identify different class labels [14]. Generally, ML can be divided into three categories: supervised learning, unsupervised learning, and reinforcement learning [15]

2.2.1 Supervised Learning

In this type of learning, the dataset used for training contains examples of input vectors, each with its corresponding desired output vector. Algorithms used in this category include Naive Bayes, KNN, ANN, Decision Tree, SVM, ensemble methods (bagging, voting classifier, Adaboost, Gradient Boosting), and logistic regression [16]. Machine Learning Classifiers [17] - ML Classifiers are classified as single classifiers when they contain only one classification algorithm. Several intrusion detection systems have adopted single machine classification models. SVM, ANN, DT, KNN, and NB are all made up of one ML algorithm.

Support Vector Machine (SVM) is a machine learning algorithm used for both classification and regression tasks. The hyperplane that separates the various classes to be predicted depends on the dimensionality of the dataset. For example, if it is one-dimensional, then a point on the

line is the hyperplane. If two dimensional, then a separating line and for three dimensional, a plane. SVM has been used in many intrusion detection systems due to its ability to make accurate predictions [18]–[22] .

Artificial Neural Networks (ANNs) are a type of ML algorithms based on the working of the human brain's biological nervous system. ANNs consist of an input layer, one or more hidden layers, and an output layer .The hidden layers weigh and process the inputs so that the output can be determined. A gradient-descent back propagation of error is used as the learning rule to adaptively adjust the weights and biases of the neurons in the hidden and output layers to obtain the desired output [23]–[27].

Decision Trees (DT) are also a form of machine learning and are used for categorical and numeric classifications. They are composed of a root node at the top, intermediate nodes (nodes), and leaves at the bottom. The flow of learning is from top to bottom. The leaf nodes are the endpoints of the decision tree and data samples are split into homogeneous sets (subsamples) based on the most significant splitter. Decision trees require less data cleaning and are popularly used for exploring data in classification problems, which is why they are widely used in intrusion detection systems and other applications [28]–[32] .

K-Nearest Neighbour (KNN) is a distance-based machine learning model used to solve classification problems. KNN classifies each unlabelled example by the majority label among its K-nearest neighbours in the training set. The nearest neighbours are identified using distance metrics such as Euclidean distance. KNN is time efficient and easy to interpret, making it a popular choice for solving classification problems in intrusion detection systems and other applications [33]–[37].

2.2.3 Unsupervised Learning

In unsupervised Learning, the learning algorithm is not given labels, so it has to find structure in its own input. This is also known as learning without a teacher. Self-Organizing Map (SOM) [38] , Apriori algorithm, Éclat algorithm, outlier detection [39], hierarchical clustering, and cluster analysis (K-Means clustering, Fuzzy clustering) [40] are some of the algorithms used in unsupervised learning.

2.2.4 Reinforcement Learning

In reinforcement learning, the model is taught to make a series of decisions. The goal is achieved in an uncertain and potentially intricate way. The model uses trial and error to find a solution to the problem. Deep Q Network (DQN), Q-Learning, State Action-Reward-State-Action (SARSA), Deep Deterministic Policy Gradient (DDPG) are some of the reinforcement learning algorithms [41] .

3. Review of Related Works

Liu et al. [42] in their paper examined a Weighted-Multi-Random-Decision-Tree (WMDRT) method for intrusion detection. This new approach was designed to tackle the challenge of handling large, complex data sets that traditional decision trees cannot process on their own. To further enhance this model, a multi-agent IDS model was also proposed. Experiments demonstrated excellent classification accuracy, flexibility and minimal system resource cost for this WMDRT algorithm in comparison with other conventional models.

Duque et al. [43] conducted a study to propose a model for Intrusion Detection System (IDS) with higher efficiency and fewer false positives/negatives. The NSL-KD data set, which consisted of 25,192 entries and 22 different types of data, was analysed using k-means unsupervised machine learning. Results indicated that the highest efficiency rate (81.61%) occurred when 11 clusters were used; however false positive rates increased progressively from 0.74% to 31.91%, while the false negative rate decreased correspondingly from 99.82% to 95.70%. Interestingly, the best performance occurred when the number of clusters matched the number of data types in the set. In light of these findings it was recommended other data mining techniques be explored, such as a combination of k-means algorithms and signature-based approaches - this may help to decrease false negatives further still; also an automated system for cluster identification should be developed.

Bohara et al. [44] had explored an approach for detecting anomalous behaviour in unlabelled system and network log data using unsupervised machine learning in their paper. The analysis was based on the monitors available in the VAST 2011 Mini Challenge 2 dataset, where they first developed a threat model and extracted meaningful features from their log data that could help detect attacks. To combine this information, their method utilized specialized features to perform anomaly detection over both network-level and host-level log datasets. They then used k-means and DBSCAN clustering algorithms to assign each entry into different usage profiles, comparing feature distributions among them to identify any abnormal activities. Their cluster difference metric further classified these clusters according to their malicious tendencies. Manual analysis was then done to correlate them with known attacks as well as evaluate their approach's performance which discovered all threats present except those requiring additional information for detection.

Erbacher et al. [45] performed a research aimed to design a multi-layered architecture for the detection of various

existing and new botnets. It was not intended to rely on just one technique, but rather support multiple techniques in order to be able to detect a broader array of bots and botnets than used with only one technique. The open structure and API enabled integration of any techniques created by other researchers. The goal was twofold: Utilizing signature type techniques to spot well known bots and botnets, as well as data mining tactics to identify new varieties or anomalies that can be regarded as misuse detection category tactics.

Jabez et al. [46] Presented a new approach to intrusion detection in computer networks called the Outlier Detection Approach. This training model used big datasets with distributed environments to improve performance on Intrusion Detection Systems (IDS). The proposed method was tested using KDD datasets from realistic sources and yielded good results compared to machine learning approaches which take longer execution time and storage. The IDS system proposed here was able to detect anomalies more efficiently, leading to improved results than existing methods. It was concluded that their future research may focus on further incorporating this into distance computation functions between trained models and testing data. In total, their findings can be leveraged to increase efficiency of the IDS system.

Chen et al. [47] proposed a novel NIDS system based on CNN.

Kumar et al. [48] in their study mentioned Converging K-Means, Fuzzy Rule System and Neural Network had helped in achieving a robust architecture. The deficiencies perceived in these individual techniques towards obtaining an effective intrusion detection algorithm were rectified by blending them appropriately. Analysis of the characteristics of both abnormal as well as normal packets aided recognition of their patterns and discrimination effectively.

3.1 Datasets used in the previous works

The most commonly used dataset in the work being studied is KDD-NSL. Generally speaking, the datasets used to evaluate various algorithms applied in the studied works include KDDCup, NSL-KDD, Kyoto2006+, UGR2006, CICIDS'17, and UNSW-NB'15.

KDDCup is the dataset applied in the 3rd International Knowledge Discovery and Data Mining Tools Competition and consists of 41 columns of attributes.

The NSL-KDD data set is an upgraded version of the KDD'99 dataset which has removed redundant records in order to avoid any bias effects of classification. This dataset consists of 38 numeric features and 3 nominal features.

Kyoto2006+ was developed using real traffic data gathered for three years at Kyoto University using 348 honeypots. It contains 24 features, 14 of which are similar to the KDD Cup'99 dataset with 10 additional columns containing six characteristics relevant to knowledge.

AWID is composed of real benign and attack traffic collected from real network environments and published in 2015.

CIC-IDS2017 is another dataset consisting of standard and recent typical attacks founded in 2017 by the Canadian Institute for Cyber Security. It has 3,119,345 rows and 84 labelled features.

3. Discussions and Future Work

It was seen in the review that ensemble and hybrid classifiers have a higher predictive accuracy and detection rate than single classifiers. Consequently, future research should focus on the classification methods of hybrid and ensemble ML in order to enhance the efficiency of IDSs. Experiments should also be conducted to establish models which are successful across multiple datasets. Furthermore, attribute extraction should be incorporated into the classification phase and all irrelevant, unnecessary and redundant features should be removed to boost the reliability and detection rate of intrusion detection systems. Additionally, more recent datasets must be used to evaluate the algorithms utilised in order to stay abreast with modern malicious intrusions and threats.

4. Conclusion

The Intrusion Detection System is a tool that helps protect computer systems from hackers and attackers. Researchers are working on making the IDS better so it can help protect us better. The integration of machine learning into intrusion detection systems has led to the emergence of new techniques and approaches, with researchers and academics developing a range of classifiers to build IDS models. This paper surveyed the research papers concerning the use of ML classifiers for intrusion detection and listed its findings in respective sections.

REFERENCES

[1] H. Cavusoglu, S. Raghunathan, and H. Cavusoglu, “Configuration of and Interaction Between Information Security Technologies: The Case of Firewalls and Intrusion Detection Systems,” *Inf. Syst. Res.*, vol. 20, no. 2, pp. 198–217, Jun. 2009, doi: 10.1287/isre.1080.0180.

[2] Y. Otoum and A. Nayak, “AS-IDS: Anomaly and Signature Based IDS for the Internet of Things,” *J.*

Netw. Syst. Manag., vol. 29, no. 3, p. 23, Mar. 2021, doi: 10.1007/s10922-021-09589-6.

[3] S. Pan, T. Morris, and U. Adhikari, "Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov. 2015, doi: 10.1109/TSG.2015.2409775.

[4] J. R. Yost, "The March of IDES: Early History of Intrusion-Detection Expert Systems," *IEEE Ann. Hist. Comput.*, vol. 38, no. 4, pp. 42–54, Oct. 2016, doi: 10.1109/MAHC.2015.41.

[5] J. S. Sherif and T. G. Dearmond, "Intrusion detection: systems and models," in *Proceedings. Eleventh IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, Jun. 2002, pp. 115–133. doi: 10.1109/ENABL.2002.1029998.

[6] A. S. Ashoor and S. Gore, "Importance_of_Intrusion_Detection_System," vol. 2, no. 1, p. 4, 2010.

[7] W. Burr, H. Ferraiolo, and D. Waltermire, "NIST and Computer Security," *IT Prof.*, vol. 16, no. 2, pp. 31–37, Mar. 2014, doi: 10.1109/MITP.2013.88.

[8] @NexWebSites, "Intrusion Detection & Prevention Systems : The Ultimate Guide." https://datasilk.com/intrusion-detection-prevention/ (accessed Dec. 08, 2022).

[9] T. Sarkar and N. Das, "Survey on Host and Network Based Intrusion Detection System," *Int. J. Adv. Netw. Appl.*, vol. 6, pp. 2266–2269, Sep. 2014.

[10] S. A. Maske and Thaksen. J. Parvat, "Advanced anomaly intrusion detection technique for host based system using system call patterns," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, Aug. 2016, vol. 2, pp. 1–4. doi: 10.1109/INVENTIVE.2016.7824846.

[11] A. Shah, S. Clachar, M. Minimair, and D. Cook, "Building Multiclass Classification Baselines for Anomaly-based Network Intrusion Detection Systems," in *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, Oct. 2020, pp. 759–760. doi: 10.1109/DSAA49011.2020.00102.

[12] S. M. Hussein, "Performance Evaluation of Intrusion Detection System Using Anomaly and Signature Based Algorithms to Reduction False Alarm Rate and Detect Unknown Attacks," in *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*, Dec. 2016, pp. 1064–1069. doi: 10.1109/CSCI.2016.0203.

[13] A. Fadhlillah, N. Karna, and A. Irawan, "IDS Performance Analysis using Anomaly-based Detection Method for DOS Attack," in *2020 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*, Jan. 2021, pp. 18–22. doi: 10.1109/IoTaIS50849.2021.9359719.

[14] G. Shang-fu and Z. Chun-lan, "Intrusion detection system based on classification," in *2012 IEEE International Conference on Intelligent Control, Automatic Detection and High-End Equipment*, Jul. 2012, pp. 78–83. doi: 10.1109/ICADE.2012.6330103.

[15] S. Thirimanne, L. Jayawardana, P. Liyanaarachchi, and L. Yasakethu, "Comparative Algorithm Analysis for Machine Learning Based Intrusion Detection System," in *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*, Aug. 2021, pp. 191–196. doi: 10.1109/ICIAfS52090.2021.9605814.

[16] A. Singh, N. Thakur, and A. Sharma, "A review of supervised machine learning algorithms," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2016, pp. 1310–1315.

[17] S. Prakash and R. B, "Detection and Classification of Thoracic Diseases in Medical Images Using Artificial Intelligence Techniques: A Systematic Review," *ECS Trans.*, vol. 107, no. 1, p. 307, Apr. 2022, doi: 10.1149/10701.0307ecst.

[18] L. Khan, M. Awad, and B. Thuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering," *VLDB J.*, vol. 16, no. 4, pp. 507–521, Oct. 2007, doi: 10.1007/s00778-006-0002-5.

[19] S.-J. Horng *et al.*, "A novel intrusion detection system based on hierarchical clustering and support vector machines," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 306–313, Jan. 2011, doi: 10.1016/j.eswa.2010.06.066.

[20] 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 Syst. Appl.*, vol. 67, pp. 296–303, Jan. 2017, doi: 10.1016/j.eswa.2016.09.041.

[21] D. S. Kim and J. S. Park, "Network-Based Intrusion Detection with Support Vector Machines," in *Information Networking*, Berlin, Heidelberg, 2003, pp. 747–756. doi: 10.1007/978-3-540-45235-5_73.

[22] E. A. Shams and A. Rizaner, "A novel support vector machine based intrusion detection system for mobile

ad hoc networks," *Wirel. Netw.*, vol. 24, no. 5, pp. 1821–1829, Jul. 2018, doi: 10.1007/s11276-016-1439-0.

[23] E. Hodo *et al.*, "Threat analysis of IoT networks using artificial neural network intrusion detection system," in *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, May 2016, pp. 1–6. doi: 10.1109/ISNCC.2016.7746067.

[24] A. Shenfield, D. Day, and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks," *ICT Express*, vol. 4, no. 2, pp. 95–99, Jun. 2018, doi: 10.1016/j.icte.2018.04.003.

[25] A. Drewek-Ossowicka, M. Pietrołaj, and J. Rumiński, "A survey of neural networks usage for intrusion detection systems," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 497–514, Jan. 2021, doi: 10.1007/s12652-020-02014-x.

[26] M. R. Norouzian and S. Merati, "Classifying attacks in a network intrusion detection system based on artificial neural networks," in *13th International Conference on Advanced Communication Technology (ICACT2011)*, Feb. 2011, pp. 868–873.

[27] M. Al-Zewairi, S. Almajali, and A. Awajan, "Experimental Evaluation of a Multi-layer Feed-Forward Artificial Neural Network Classifier for Network Intrusion Detection System," in *2017 International Conference on New Trends in Computing Sciences (ICTCS)*, Oct. 2017, pp. 167–172. doi: 10.1109/ICTCS.2017.29.

[28] N. B. Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs decision trees in intrusion detection systems," in *Proceedings of the 2004 ACM symposium on Applied computing*, New York, NY, USA, Mar. 2004, pp. 420–424. doi: 10.1145/967900.967989.

[29] S. Sahu and B. M. Mehtre, "Network intrusion detection system using J48 Decision Tree," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Aug. 2015, pp. 2023–2026. doi: 10.1109/ICACCI.2015.7275914.

[30] M. Kumar, M. Hanumanthappa, and T. V. S. Kumar, "Intrusion Detection System using decision tree algorithm," in *2012 IEEE 14th International Conference on Communication Technology*, Nov. 2012, pp. 629–634. doi: 10.1109/ICCT.2012.6511281.

[31] B. Ingre, A. Yadav, and A. K. Soni, "Decision Tree Based Intrusion Detection System for NSL-KDD Dataset," in *Information and Communication Technology for Intelligent Systems (ICTIS 2017) - Volume 2*, Cham, 2018, pp. 207–218. doi: 10.1007/978-3-319-63645-0_23.

[32] C. Kruegel and T. Toth, "Using Decision Trees to Improve Signature-Based Intrusion Detection," in *Recent Advances in Intrusion Detection*, Berlin, Heidelberg, 2003, pp. 173–191. doi: 10.1007/978-3-540-45248-5_10.

[33] Y. Liao and V. R. Vemuri, "Use of K-Nearest Neighbor classifier for intrusion detection11An earlier version of this paper is to appear in the Proceedings of the 11th USENIX Security Symposium, San Francisco, CA, August 2002," *Comput. Secur.*, vol. 21, no. 5, pp. 439–448, Oct. 2002, doi: 10.1016/S0167-4048(02)00514-X.

[34] K. Atefi, H. Hashim, and M. Kassim, "Anomaly Analysis for the Classification Purpose of Intrusion Detection System with K-Nearest Neighbors and Deep Neural Network," in *2019 IEEE 7th Conference on Systems, Process and Control (ICSPC)*, Dec. 2019, pp. 269–274. doi: 10.1109/ICSPC47137.2019.9068081.

[35] S. Malhotra, V. Bali, and K. K. Paliwal, "Genetic programming and K-nearest neighbour classifier based intrusion detection model," in *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Jan. 2017, pp. 42–46. doi: 10.1109/CONFLUENCE.2017.7943121.

[36] Y. Y. Aung and M. Myat Min, "Hybrid Intrusion Detection System using K-means and K-Nearest Neighbors Algorithms," in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, Jun. 2018, pp. 34–38. doi: 10.1109/ICIS.2018.8466537.

[37] M. Govindarajan and RM. Chandrasekaran, "Intrusion detection using k-Nearest Neighbor," in *2009 First International Conference on Advanced Computing*, Dec. 2009, pp. 13–20. doi: 10.1109/ICADVC.2009.5377998.

[38] M. Li and W. Dongliang, "Anormaly Intrusion Detection Based on SOM," in *2009 WASE International Conference on Information Engineering*, Jul. 2009, vol. 1, pp. 40–43. doi: 10.1109/ICIE.2009.240.

[39] M. Kumar and R. Mathur, "Unsupervised outlier detection technique for intrusion detection in cloud computing," in *International Conference for Convergence for Technology-2014*, Apr. 2014, pp. 1–4. doi: 10.1109/I2CT.2014.7092027.

[40] "Clustering Approach Based on Mini Batch Kmeans for Intrusion Detection System Over Big Data | IEEE Journals & Magazine | IEEE Xplore."

https://ieeexplore.ieee.org/document/8304564 (accessed Dec. 08, 2022).

[41] Q.-V. Dang and T.-H. Vo, "Studying the Reinforcement Learning techniques for the problem of intrusion detection," in *2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, May 2021, pp. 87–91. doi: 10.1109/ICAIBD51990.2021.9459006.

[42] Y. Liu, N. Li, L. Shi, and F. Li, "An intrusion detection method based on decision tree," in *2010 International Conference on E-Health Networking Digital Ecosystems and Technologies (EDT)*, Apr. 2010, vol. 1, pp. 232–235. doi: 10.1109/EDT.2010.5496597.

[43] S. Duque and Mohd. N. bin Omar, "Using Data Mining Algorithms for Developing a Model for Intrusion Detection System (IDS)," *Procedia Comput. Sci.*, vol. 61, pp. 46–51, Jan. 2015, doi: 10.1016/j.procs.2015.09.145.

[44] A. Bohara, U. Thakore, and W. H. Sanders, "Intrusion detection in enterprise systems by combining and clustering diverse monitor data," in *Proceedings of the Symposium and Bootcamp on the Science of Security*, New York, NY, USA, Apr. 2016, pp. 7–16. doi: 10.1145/2898375.2898400.

[45] R. Erbacher, A. Cutler, P. Banerjee, and J. Marshall, "A Multi-Layered Approach to Botnet Detection.," Jan. 2008, pp. 301–308.

[46] J. Jabez and B. Muthukumar, "Intrusion Detection System (IDS): Anomaly Detection Using Outlier Detection Approach," *Procedia Comput. Sci.*, vol. 48, pp. 338–346, Jan. 2015, doi: 10.1016/j.procs.2015.04.191.

[47] L. Chen, X. Kuang, A. Xu, S. Suo, and Y. Yang, "A Novel Network Intrusion Detection System Based on CNN," in *2020 Eighth International Conference on Advanced Cloud and Big Data (CBD)*, Dec. 2020, pp. 243–247. doi: 10.1109/CBD51900.2020.00051.

[48] K. S. Anil Kumar and V. N. Mohan, "Adaptive Fuzzy Neural Network Model for intrusion detection," in *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, Nov. 2014, pp. 987–991. doi: 10.1109/IC3I.2014.7019811.