



Next-Generation Intrusion Detection Systems: An Exhaustive Review Of Techniques And Evaluation Benchmarks

Deepti Kalawat¹, Prof. Ruchi Dronawat²

M. Tech. Scholar, Department of CSE, SIRT, Bhopal, India¹, Assistant Professor, Department of CSE, SIRT, Bhopal, India²

Abstract — In order to address the deficiencies of the existing intrusion detection models, a number of researchers have proposed the implementation of an intrusion detection system that is founded on data mining technology. It tackles the hazards of misreporting or omission, as well as the systems' inability to self-adapt to changing circumstances. In spite of this, an intrusion detection system that is based on data mining technology will need to consume an increasing number of resources when dealing with big volumes of data, which will slow down the detection speed. When we take into consideration the technique that was discussed before, we run into another problem: when we construct clusters of the same item for the purpose of future pattern development, we do not always receive a consequence that is suitable. In light of the fact that the support is different every time, the pruning value shifts in accordance with them. As a result, we must process the rule generation dynamic, which solves the aforementioned challenge and allows us to obtain a variety of results under different conditions.

Keywords: IDS, Dos, Probe, R2L, U2R, NSL-KDD.

I. Introduction

Intrusion detection is the ongoing monitoring and analysis of events in a computer system or network to seek for signs of an intrusion. There are two types of intrusion detection systems: misuse detection and anomaly detection. There are several threats/attacks today that have a variety of effects on network security. An intrusion detection system is used to protect against such attacks. This study included Intrusion Detection Systems and methodologies such as ALAD, PHAD, and SNORT. The NSL-KDD dataset, a more sophisticated variant of the KDD99 dataset, was another topic we discussed. The proposed method uses an SVM algorithm based on HIDS to process accurate findings.

1. Misuse Detection: In order to detect intrusions, misuse intrusion detection employs well defined attack patterns that take advantage of flaws in system and application software. These patterns are pre-coded and compared to user activity in order to identify infiltration.

2. Anomaly Detection: Anomaly intrusion detection uses the typical use patterns to find the intrusion. The typical usage patterns are created using statistical evaluations involving system variables, including CPU and I/O usage by an individual or program. Assault is defined as any deviation from the consumer's constructed normal behavior. There are two ways to adequately safeguard the system: either we can steer

clear of dangers and risks brought on by flaws in the OS and application developers, or we can find them and take action to stop them in future releases while also fixing the damage. In reality, it is difficult to create a completely safe system, even if it were feasible, it would be very costly and complicated. The method of collecting and examining information from several key points on a network or system of computers is known as intrusion detection, or Ids. It could find proof of an attack or infractions of safety regulations on an internet connection or computer. A software and hardware surveillance system is called an Intrusion Detection System (IDS) [2]. Modern intrusion detection systems use anomaly and abuse detection techniques to evaluate intrusion detection. Throughout as well as following testing, it can be carried out.

The features of IDS include complicated, fluid, active, and real-time. How could a vast quantity of data be used to determine a user's their own interests? Numerous academics put out a number of innovative ideas and algorithms. The application of data mining techniques to detection systems for intrusions was initially proposed by Wenke Lee [3] of Columbia University in the United States. Identification based on data mining techniques has been a prominent area of research with several theoretical conclusions [4][5][6]. primarily by using data categorization, correlation analysis, and algorithmic pattern mining in order to manage massive amounts of security audit data, extract security-related activity, create intrusion detection rules, and create atypical activity identification model. The issue of inefficiency and sluggish reaction times may be resolved with the use of data mining techniques in IDS. It can adjust to the rapid development of network intrusion detection technologies, lowering the requirement for human intervention while improving the effectiveness of detection.

Generally speaking, an intrusion detection system has three practical components. A data origin, also known as an event power source, is the initial component of an intrusion detection system. Host-based, network-based, application-based, and target-based monitors are the four types of data sources. The evaluation engine is the second component of an intrusion detection system. This component retrieves data from the data source and analyses it to search for signs of assaults or other policy infractions. Any or all of the following analytical approaches may be used by the analysis engine:

The detection according to Misuse/Signature: This type of surveillance system identifies intrusions that adhere to established pattern of attack (or signatures) that take use of known software flaws. This method's main drawback is that it just looks for known vulnerabilities; it might not be interested in spotting any future intrusions.

Statistical/Anomaly Detection: An abnormality recognition machine will search for unusual or uncommon items. We examine system event flows and search for anomalous activity patterns using statistical methods. This system's primary drawbacks are its high cost and the possibility that, in the absence of data, it could misinterpret intrusive activity for normal behavior. The third part of an intrusion detection system is the reaction manager. To put it plainly, the response administrator will only step in if errors (possible intrusion attacks) are found on the system, informing everyone or something through a reply.

II. Literature Survey

In this paper [1] application of artificial intelligence to increase the security breach detection process's precision is the main topic of this essay. In actuality, machine learning makes it simple to recognize patterns and trends, which can then be used to network environments to identify intrusions. It can be quite helpful. To enhance decision-making and intrusion detection, we employ support vector machines (SVM) and Radial Basis Function (RBF) neural networks. Key data characteristics are retrieved by using RBF neural networks. Please verify that the author information and affiliations are displayed accurately. Please make any required changes. Improve the method's general effectiveness and the SVM technique's efficiency. Lastly, the Python programming language with the KDD99 data set were used to assess the suggested approach. The results of the assessment show that the suggested method outperforms the other options in terms of precision as well as accuracy. Thus, the precision is over 99% and the accuracy is 97%.

This paper develops an adaptive neuro-fuzzy inference system for intrusion detection based on the Crow Search Optimizations approach [2]. Any anomalies in a computer network or systems are found using the IDS. The crow search optimization approach is used to enhance the efficiency of the ANFIS approach, which is a mix of an artificial neural network and a fuzzy interference system. The NSLKDD data set was employed to verify the proposed model's intrusion detection capabilities. The NSL-KDD dataset's detection of breaches results was superior to those of previous techniques.

The recommended work in this paper [3] employs an extraction and barrier-based approach that selects characteristics via the firefly technique within the wrapper, which has an impact on the analysis's performance. The characteristics gathered are applied to a C4.5 and Bayesian Networks (BN) dependent classification using the KDD CUP 99 dataset. The trials' findings imply that 10 traits are sufficient to identify the intrusion with higher precision. The proposed work reveals possible improvements over previous efforts.

Liang et al. [9] created an integrated repositioning method for an intrusion detection system that depends on a system with multiple agents. The setup comprises the reaction module, an evaluation module, an information management segment, and a data collection module. The investigation's evaluation component is constructed utilizing a deep neural network security detection technique. According to the results, deep learning techniques are good at spotting transport-layer dangers.

A modified K-mean approach was proposed by Laftah et al. [10] to provide an excellent training collection that greatly improves classifier performance. An intrusion detection system can train classifiers in a quarter of the time by using altered K-mean to create fresh small datasets for training which replicate every bit of initially generated training data.

Ali et al. [11] proposed a hybrid machine learning approach for identifying network breaches that combines Repetitive Minimal optimization (SMO) classifying with K-medium clusters.

To identify network intrusions, Feng et al. [12] employed an artificial intelligence information categorization technique. To minimize failures in classification within a network protocol, the primary objective is to categorize network activity as either regular or anomalous connection data. While several classification models, including vector machine techniques, have been developed for a network attack detector, every offers benefits and drawbacks of their own.

Ma T et al. [13] presented a unique method named KDSVM to detect intrusion networks. It established training capability using a deep neural network (DNN) model and SVM, classifier, as well as k-means. The KDSVM is divided into two stages. In the initial stage, the data collection is divided into k subsets based on the distance of each sample from the cluster centers of the k-means approach. In the following phase, the test information set splits from the precise same cluster center, and in the final stage, it is fed into the model created by the DNN using SVM.

A deep learning-driven DDoS attack detection technique (Deep Defence) was proposed by Yuan et al. [14]. The use of deep learning produces powerful representation and conclusions by proactively separating higher-level features from low-level functions. A recurring neural network learning project is provided to watch network attack behavior and learn trends from network traffic patterns. On the larger information set, the algorithm performed better than normal machine learning designs, with a rate of error of 2.103 percent as opposed to 7.517 % for typical machine learning methods.

Chuanlong Yin et al. [15] examined the algorithm's effectiveness in binary categorisation and the classification of several classes. The outcome of this evaluation performs better in both binary and multiple-class classification than conventional classification methods for neural networks, and it is ideally suited to building a system for classification with high precision.

An attack detection method based on probability neural networks (PNN) and deep belief networks (DBN) was proposed by Zhao et al. [16]. The raw data is first converted into small data while retaining the key characteristics of the original data by using DBN's linear retraining capability. Second, the amount of node with concealed levels per level is optimised using a swarm of particles algorithms for optimization to increase learning rate. Low-dimensional data is then categorized using PNN.

A UNSW-NB15 dataset was developed by Moustafa et al. [17] for detecting intrusions. This dataset includes new standard traffic patterns and nine types of current attack mechanisms. It uses 49 characteristics, such as host-flow and networking packet management, to distinguish between typical and unusual data. In this paper, we demonstrate the intricate nature of the UNSW-NB15 dataset in three ways. The mathematical examination of outcomes and attributes is then explained. It then does a standard association analysis. Fifth, the complexity is measured in terms of accuracy and false alarm rate (FAR) using five classifiers that are currently in use. UNSW-NB15 is more complex than KDD99, according to experiments, and an additional dataset is being suggested for NIDS assessment.

III. Problem Domain

Certain intrusion behaviours are similar to those observed in various kinds of incursions and normal settings. Moreover, a number of methods, such as K-Means, are unable to distinguish between invasion and normal instances. An intrusion detection system built on the FP-Growth tree was developed [7]. In the laboratory, their suggested method produces favourable outcomes. But at the data analysis stage, the rule is set, and if the rule generation is set, so is the grouping of data that will be created for recognizing patterns. Nonetheless, we observe a notable change in user or customer behavior every day in the current environment. Consequently, analysis of data relies on unpredictable behavior or selection at random. For evaluation, a pattern's approach of identifying shift is more accurate. This makes it possible for us to spot odd or surprising patterns and respond appropriately. A record is categorized as a known assault if it matches one in the rule database. The notification alarm unit of the Admin Processing Module (APM) is going to be triggered [18]. The Rule Repository is automatically updated with a new intrusion detection criterion. In order to save a period of time we also use comparable patterns to form clusters. Additionally, this reduces the possibility of false positive data being submitted into the connection research module by helping to eliminate noisy and redundant data. A stronger bond with the homogeneous element is made possible as a result. Our Dynamic approach allows you to change the cluster when an item set is often. Consequently, the search has become more precise [19][20].

1. A lot of systems have a high false alarm rate and low precision.
2. To obtain a precise outcome, use fewer parameters for the value comparison.
3. Only compare a small number of kinds of attacks where additional detection techniques for abuse and anomaly detection can be applied.
4. A few parameters for evaluating the outcome. For instance, we may utilise additional metrics like precision, recall, TP (True Positive), TN (True Negative), and accuracy and false alarm rate.
5. The identification rate of some algorithms is poor.

IV. Propose Methodology

In our approach, we aggregated or clustered the same collection of items by combining data mining with a clustering algorithm. Any given point is assigned to one among multiple clusters whenever a dataset is clustered; each group may be identified by a single reference point, which is often the mean of the cluster's points. The technique of grouping every point in a dataset into clusters is known as segmentation. among the most renowned applications of clustering is the classification of plants or animals into distinct groups or species. Reducing the size and level of detail in the dataset is the main objective of grouping Landsite data. Reducing the size and level of detail in the information set is the main objective of grouping Landsite information. To decrease data, the geographic coordinates of the cluster's center point are substituted for the parameters of every single point in the cluster. Compared to raw data, organized information can be

manipulated considerably more quickly and requires a lot less storage space. How accurately the reference points describe the data as well as how fast the method runs affect how beneficial a grouping strategy is. Implementing the clustering technique, we establish the results based upon the data set; certain data sets are detailed in the exact same setting, whereas other sets are placed in different locations. The data was then categorized according to the behavior of the IDS. We consistently classify data using dynamic rule creation, which produces more accurate and dependable results [8].

V. Conclusion

The combination of wide research subjects, such as data mining, and the more recent and constantly developing security detection system, on the one hand. The use of data mining techniques to detect systems of intrusion is a key avenue for future study. Building an IDS model and applying appropriate data mining techniques are helpful tactics to improve systems detection efficiency.

VI. References

- [1] Zhao, W., Zhao, Z. providing a hybrid approach to increase the accuracy of intrusion detection systems in computer networks. J. Eng. Appl. Sci. 71, 123 (2024). <https://doi.org/10.1186/s44147-024-00404-y>.
- [2] S Manimurugan , Al-qdah Majdi , Mustaffa Mohmmmed, C Narmatha , R Varatharajan "Intrusion Detection in Networks using Crow Search Optimization algorithm with Adaptive Neuro-Fuzzy Inference System", "Microprocessors and Microsystems" Elsevier 6 September 2020. <https://doi.org/10.1016/j.micpro.2020.103261>.
- [3] Selvakumar B, Muneeswaran K, "Firefly algorithm based feature selection for network intrusion detection", Computers & Security, Volume 81, 2019, Pages 148-155, <https://doi.org/10.1016/j.cose.2018.11.005>.
- [4] Dubey, S., Gupta, C. (2024). An Effective Model for Binary and Multi-classification Based on RFE and XGBoost Methods in Intrusion Detection System. Cyber Security and Digital Forensics. Lecture Notes in Networks and Systems, vol. 896. Springer, Singapore. https://doi.org/10.1007/978-981-99-9811-1_3.
- [5] Gupta, C., Kumar, A. & Jain, N.K. (2024). An Enhanced Hybrid Intrusion Detection Based on Crow Search Analysis Optimizations and Artificial Neural Network. Wireless Personal Communication 134, 43–68. <https://doi.org/10.1007/s11277-024-10880-3>.
- [6] Gupta, C., Kumar, A., Jain, N.K. (2023). A Detailed Analysis on Intrusion Detection Systems, Datasets, and Challenges. Advances in Data Science and Computing Technologies. ADSC 2022. Lecture Notes in Electrical Engineering, vol 1056. Springer, Singapore. https://doi.org/10.1007/978-981-99-3656-4_26.
- [7] G., C., Kumar, A. & Jain, N.K. Intrusion defense: Leveraging ant colony optimization for enhanced multi-optimization in network security. Peer-to-Peer Netw. Appl. 18, 98 (2025). <https://doi.org/10.1007/s12083-025-01911-2>.
- [8] Jain, T., Gupta, C. (2022). Multi-Agent Intrusion Detection System Using Sparse PSO K-Mean Clustering and Deep Learning. In: Mathur, G., Bundele, M., Lalwani, M., Paprzycki, M. (eds) Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-6332-1_10.
- [9] Namdev, P., Gupta, C., Dubey, S. (2023). An Improved Intrusion Detection System Using Data Clustering and Support Vector Machine. In: Buyya, R., Misra, S., Leung, YW., Mondal, A. (eds) Proceedings of International Conference on Advanced Communications and Machine Intelligence. MICA 2022. Studies in Autonomic, Data-driven and Industrial Computing. Springer, Singapore. https://doi.org/10.1007/978-981-99-2768-5_37.
- [9] Chao Liang, Bharanidharan Shanmugam, Sami Azam, Mirjam Jonkman, Friso De Boer, Ganthan Narayansamy, "Intrusion Detection System for Internet of Things based on a Machine Learning approach", International Conference on Vision Towards Emerging Trends in Communication and Networking, IEEE, 2019.
- [10] Wathiq Laftah Al-Yaseen , Zulaiha Ali Othman ,Mohd Zakree Ahmad Nazri, "Multi- Level Hybrid Support Vector Machine and Extreme Learning Machine Based on Modified K-means for Intrusion Detection System", International Journal in Expert Systems With Applications, Elsevier, 2017.

- [11] Saad Mohamed Ali Mohamed Gadal and Rania A. Mokhtar, “Anomaly Detection Approach using Hybrid Algorithm of Data Mining Technique”, International Conference on Communication, Control, Computing and Electronics Engineering, IEEE, 2017.
- [12] Feng, W., “Mining Network data for Intrusion Detection through Combining SVMs with Ant Colony Networks”, Future Gener. Comput. Syst., 2014, 37, 127–140.
- [13] Ma T et al, “A Hybrid Methodologies for Intrusion Detection Based Deep Neural Network with Support Vector Machine and Clustering Technique”, International conference on frontier computing. Springer, 2016.
- [14] Yuan X, Li C, Li X, “Deep Defense: Identifying D-DoS Attack via Deep Learning”, IEEE international conference on smart computing (SMARTCOMP), 2017.
- [15] Yin C et al, “A Deep Learning Approach for Intrusion Detection using Recurrent Neural Networks”, IEEE Access 5:21954–2196, 2017.
- [16] Zhao, G.; Zhang, C.; Zheng, L. “Intrusion Detection Using Deep Belief Network and Probabilistic Neural Network”, In Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, China, 21– 24 July 2017; Volume 1, pp. 639–642.
- [17] Nour Moustafa & Jill Slay, “The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 Data Set and the comparison with the KDD99 Data Set”, Information Security Journal: A Global Perspective, 2015.
- [18] Solanki, S., Gupta, C., & Rai, K. (2020). A survey on machine learning based Intrusion Detection System on NSL-KDD dataset. Int. J. Comput. Appl, 176, 36-39.
- [19] G., C., Sinhal, A., Kamble, R. (2015). An “Enhanced Associative Ant Colony Optimization Technique-based Intrusion Detection System”. Artificial Intelligence and Evolutionary Algorithms in Engineering Systems. Advances in Intelligent Systems and Computing, vol 325. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2135-7_58.
- [20] G., C., Sinhal, A Sinhal, R Kamble, “Intrusion detection based on k-means clustering and ant colony optimization: A survey”, International Journal of Computer Applications, 20 Volume 79 – No 6, October 2013.

