Real Time Anomaly Detection in Network Traffic: A Comparative Analysis of Machine Learning Algorithms

Aswathy M C[1], *Rajkumar T*[2]

[1,2]*Department of Computer Science and Engineering, College of Engineering, Kallooppara, Kerala, India*

Emails: *mc.aswathy@gmail.com*[1], *rajkumar@cek.ac.in*[2]

Abstract

In the constantly changing field of cybersecurity, real-time intrusion detection using machine learning algorithms has become crucial for protecting network infrastructures. This paper presents a comprehensive literature survey focusing on the comparative study of diverse machine learning algorithms employed for anomaly detection in network traffic. The objective is to critically evaluate the effectiveness of various algorithms in identifying and mitigating threats in real-time scenarios. The study delves into the nuances of prominent machine learning models, including Decision Trees, Random Forests, Support Vector Machines, Neural Networks, and ensemble methods, as they apply to the domain of anomaly detection. Each algorithm is scrutinized based on its ability to adapt to dynamic network behaviors, handle imbalanced datasets, and provide accurate real-time threat assessments. Throughout the survey, key research contributions are analyzed, encompassing methodologies, datasets, and performance metrics. Comparative insights are provided to emphasize the strengths and weaknesses of each algorithm, elucidating their appropriateness for real-time intrusion detection in network traffic. Notably, the examination extends beyond traditional approaches, exploring recent advancements such as deep learning and ensemble techniques. The findings from this comparative study aim to provide practitioners and researchers with valuable insights into selecting the most suitable machine learning algorithm for real-time anomaly detection in the context of network security. By understanding the comparative performance of these algorithms, organizations can make informed decisions to enhance their cybersecurity posture and fortify their defenses against emerging threats.

Keywords: *Anomaly Detection; Cybersecurity; Decision Tree; Deep Learning; Dynamic Behaviors; Ensemble Methods; Intrusion Detection Imbalanced Datasets; Neural Networks; Performance Metrics; Random Forests; Real-Time Network Traffic; Machine Learning Algorithms; Support Vector Mach.*

1. Introduction

In today's ever-evolving landscape of cybersecurity, the protection of network infrastructures against malicious intrusions remains paramount. (Satish Kumar et.al;2021) explains two styles of detection techniques are used by the network intrusion detection systems(NIDS). One is signature based intrusion detection system(SIDS) and another one is anomaly based intrusion detection system(AIDS). SIDS uses previously known signatures to identify it threats in network traffic. According to (Tushar Rakshe et.al;2017) AIDS compares the normal traffic threats in network traffic. According to (Tushar Rakshe et.al;2017) AIDS compares the normal traffic pattern with incoming traffic if there is any deviations from the normal traffic happened the system recognize it as anomaly. (Mohammed Hussaein Thwaini et.al;2022) explains the importance of real-time anomaly detection, empowered by machine learning algorithms, has emerged as a pivotal defense mechanism to combat sophisticated threats. This literature survey embarks on a comprehensive exploration, focusing on the comparative study of diverse machine learning algorithms employed for anomaly detection in network traffic. By discerning patterns and deviations from normal behavior, these algorithms enable proactive threat mitigation and bolster the resilience of network defenses. The aim of this literature survey is to critically assess the effectiveness of different machine learning algorithms for real-time anomaly detection. It explores the intricacies of notable models such as Decision Trees, Random Forests, Support Vector

Machines, Neural Networks, and ensemble methods, evaluating their adaptability to dynamic network behaviors and their capability to manage imbalanced datasets. Moreover, recent advancements such as deep learning and ensemble techniques are explored to provide a comprehensive understanding of the evolving landscape. Contributions includes: Key research contributions are examined, including methodologies, datasets, and performance metrics, to provide comparative insights into each algorithm's strengths and limitations. Practical implications derived from the findings aim to empower practitioners and researchers in selecting the most suitable machine learning algorithm for real-time intrusion detection in network traffic. By equipping organizations with the knowledge to make informed decisions, this survey seeks to fortify cybersecurity postures and mitigate emerging threats in an increasingly interconnected world.

1.1 Related Work

The field of cybersecurity is constantly advancing, and as cyber threats grow more sophisticated, the need for robust Network Intrusion Detection Systems (NIDS) has become ever more critical. According to (Oluwadamilare Harazeem Abdulganiyu et al;2023) NIDS are essential for protecting network infrastructures by continuously monitoring and detecting anomalous activities that could indicate potential security breaches. Machine learning algorithms are crucial for anomaly detection in cybersecurity because of their adaptability, scalability, and real-time processing capabilities. These algorithms can analyze vast amounts of data and detect complex patterns, making them ideal for identifying new and emerging threats. With the capability to process data in real-time, machine learning algorithms enable timely detection of anomalies, thereby minimizing the impact of security breaches on network infrastructures Machine learning algorithms enhance cybersecurity by learning from new data and continuously updating their models, enabling them to identify unknown or zero-day threats and diminishing the dependence on signature-based systems. They also handle imbalanced datasets effectively, minimize false positives, and automate the threat detection process. This automation frees security personnel to concentrate on strategic tasks and speeds up incident response times. Notable among these is the work by (Roesch;1999), who introduced Snort, a widely adopted open-source NIDS. Snort employs a rule-based detection engine but has since evolved to incorporate machine learning techniques for improved accuracy. A comprehensive study by (Muda et al;2017) explored the application of machine learning in NIDS, emphasizing the significance of feature selection and classification algorithms. The authors highlighted the effectiveness of ensemble methods, such as Random Forests and AdaBoost, in improving the overall performance of intrusion detection systems. In the pursuit of anomaly detection, the work of (Patcha and Park et al;2007) delved into the challenges and opportunities presented by machine learning techniques in detecting unknown and emerging threats. The authors emphasized the effectiveness of ensemble methods, including Random Forests and AdaBoost, in enhancing the overall performance of intrusion detection systems. Optimizing hyperparameters has been a focal point in enhancing the performance of machine learning models. (Bergstra et al;2011) proposed the use of Bayesian optimization for hyperparameter tuning, demonstrating its effectiveness in optimizing complex models. This methodology has influenced subsequent research, including its application in the current study. (Ali Bou Nassif et.al;2021) surveyed research papers on a period of 2000 to 2020 and discussed different machine learning techniques used for anomaly detection and mentioned the advantages and disadvantages of each technique. He also compared the precision of each techniques used. They included journals and conference papers only and excluded other sources. (Bhuyan et.al;2013) focussed on network anomalies, they tried to identify different network attacks and detected by the intrusion detection systems and comparing the effectiveness of different anomaly detection methods. (Zhen Yang et al;2022) conducted a survey of 119 highly cited papers on anomaly-based intrusion detection, focusing on application domains, attack detection techniques, evaluation metrics, and datasets. The survey employed the Systematic Literature Review (SLR) method. They noted that traditional machine learning techniques, such as Decision Trees, Random Forests, and SVM, as well as deep learning techniques like Convolutional Neural Networks (CNN), Deep Learning, Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM), have become increasingly popular in modern intrusion detection. (V. Jyothsna et.al;2011) reviews the different categories of intrusion detection systems, anomaly detection techniques, discussed about different models, current state of art, mainly elaborates the foundations of the main anomaly-based network intrusion detection technologies and their operational architectures and also done a classification based on the processing related to each model. The study observed that majority of the surveyed works did not meet their requirements. This literature survey offers an in-depth review of machine learning algorithms used for anomaly detection in network traffic, delivering essential insights, methodological advice, and practical applications for researchers, practitioners, and policymakers in cybersecurity. From this survey, several research questions arise that can direct future research and exploration in the domain of machine learning for network traffic anomaly detection. How do different machine learning algorithms, such as Decision Trees, Random Forests, Support Vector Machines, Neural Networks, and ensemble methods, compare in terms of their effectiveness for real-time anomaly detection in network traffic? What are the key factors influencing the performance of machine learning algorithms in anomaly detection, including adaptability to dynamic network behaviors, handling of imbalanced datasets, and provision of accurate real-time threat assessments? What are the existing research gaps and unresolved issues in the area of anomaly detection using machine learning algorithms, and what future research directions can tackle these challenges to enhance the state-of-the-art in cybersecurity? These research questions provide a framework for further investigation and exploration into the effectiveness, methodologies, and practical implications of machine learning algorithms for anomaly detection in network traffic. For this survey we used a systematic literature review methodology because it gives a comprehensive coverage, transparency and reproducibility, structured data extraction, quality assessment and also provide clear reporting. This approach offers valuable insights and helps to guide future research directions in the field of cybersecurity.

2. Anomaly Detection Methods

2.1 General Classification

Anomaly detection methods used in network intrusion detection encompass a range of approaches designed to identify abnormal activities or behaviors within network traffic. Here are some of the different methods commonly employed in this domain:

Statistical Methods: Statistical anomaly detection techniques model normal behavior based on statistical measures such as mean, median, variance, and standard deviation of network traffic features. Deviations from these statistical norms are flagged as anomalies. Examples include Z-score, Gaussian distribution models, and time-series analysis.

Machine Learning: Machine learning algorithms are widely employed for anomaly detection in network traffic. Supervised learning algorithms categorize instances as normal or anomalous using labeled training data. Unsupervised learning algorithms identify anomalies without prior data labeling by detecting patterns or outliers. Semi-supervised learning merges elements of both supervised and unsupervised learning to enhance anomaly detection.

Rule-Based Methods: Rule-based anomaly detection involves defining a set of rules or thresholds that characterize normal behavior. Any deviations from these established rules are identified as anomalies. Rules can be based on protocol specifications, known attack patterns, or predefined thresholds for network metrics such as packet size, frequency, or protocol violations.

Signature-Based Detection: Signature-based methods identify known attack patterns or signatures within network traffic. These signatures are typically predefined based on known attack behaviors or malicious patterns. Matching network traffic against these signatures can help detect known attacks, but it may be less effective against novel or unknown threats.

Protocol Analysis: Protocol analysis involves monitoring network protocols for deviations from standard specifications or expected behavior.

Anomalies such as protocol misuse, unauthorized protocol usage, or abnormal protocol sequences can indicate potential security threats.

Graph-Based Methods: Graph-based anomaly detection techniques model network traffic as a graph, with nodes representing network entities (e.g., hosts, IP addresses) and edges representing connections or interactions between them. Anomalies are detected based on deviations from expected graph structures or patterns, such as unexpected communication patterns or changes in network topology.

Behavioral Analysis: Behavioral anomaly detection methods focus on analyzing patterns of behavior within network traffic to identify deviations from normal behavior. This may involve profiling typical user or system behavior and flagging deviations as anomalies. Behavioral analysis can include user behavior analysis, application behavior analysis, or system behavior analysis.

Heuristic Methods: Heuristic anomaly detection techniques use domain-specific knowledge or heuristics to identify anomalies within network traffic. These methods may involve expert-defined rules, thresholds, or patterns specific to the network environment or application domain.

2.2 Machine Learning Approach

Various machine learning methods commonly used for anomaly detection in network traffic:

Supervised Learning: Supervised learning algorithms learn from labeled training data, where each instance is annotated as either normal or anomalous. These algorithms build a model that maps input features (e.g., network traffic features) to their corresponding labels. Supervised learning algorithms such as Decision Trees, Random Forests, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Naive Bayes classifiers can be used for anomaly detection in network traffic. Supervised learning is suitable when labeled training data is available, making it possible to train models to distinguish between normal and anomalous network traffic patterns.

Unsupervised Learning: Unsupervised learning algorithms detect anomalies in network traffic without requiring labeled training data. These algorithms identify patterns or clusters within the data and flag instances that deviate significantly from these patterns as anomalies. Common unsupervised learning algorithms used for anomaly detection in network traffic include k-means clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), Isolation Forest, and One-Class SVM. Unsupervised learning is particularly useful when labeled training data is limited or unavailable. These algorithms can detect anomalies based solely on the inherent characteristics of the data, without prior knowledge of what constitutes normal or anomalous instances.

Semi-Supervised Learning: Semi-supervised learning integrates aspects of both supervised and unsupervised learning. It utilizes a small portion of labeled training data combined with a larger set of unlabeled data to train models for anomaly detection. Semi-supervised learning algorithms such as Self-Training, Co-Training, and Generative Adversarial Networks (GANs) can be applied to anomaly detection in network traffic. Semi-supervised learning is useful when limited labeled data is available, as it allows leveraging both labelled and unlabeled data to improve model performance.

Deep Learning: Deep learning methods use neural networks with multiple layers to automatically learn hierarchical representations of data. These models can capture complex patterns and relationships within network traffic data, making them suitable for anomaly detection. Deep learning architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Autoencoders, are used for anomaly detection in network traffic. Deep learning techniques excel at identifying complex patterns and anomalies in high-dimensional network traffic data. They are especially valuable when working with large and intricate datasets.

Ensemble Learning: Ensemble learning enhances prediction accuracy and robustness by combining multiple base models. These methods aggregate the predictions from individual models to make collective decisions. Ensemble techniques such as Random Forests, Gradient Boosting Machines (GBMs), and Stacked Generalization (Stacking) are commonly used for anomaly detection in network traffic. Ensemble learning can improve the performance of anomaly detection models by leveraging the diversity of multiple base models, thereby reducing overfitting and biases present in individual models. Each of these machine learning methods provides unique advantages and can be applied to various scenarios depending on the availability of labeled data, the complexity of network traffic patterns, and the specific requirements of the anomaly detection task. This study explores the intricacies of prominent machine learning models in the context of anomaly detection.

2.3 Current State of Art

(Mohammad Al-Omari et.al;2021) introduced an Intelligent tree-based intrusion detection model for cyber security. This model is designed using Decision trees which can predict and detect attacks in cyber space. Real dataset is applied in this system. The validation is based on predefined performance evaluation matrices such as accuracy, recall, precision and F1-score. The benefit of this model is to predict effectively and reduce the computational processes compared to other machine learning algorithms. (Azidine Guezzar et.al;2021) propose an NIDS model using decision tree with enhanced data quality. This model attains an accuracy of 99.42% with the NSL-KDD dataset and 98.80% accuracy with the CICIDS2017 dataset. The accuracy detection rate and false alarm rate is also high in this approach. He suggests to integrate with other efficient machine learning approach such as deep learning will give better results. (Nabila Farnaaz et.al ;2016) built a model for intrusion detection using Random forest classifier. Random forest is an ensemble classifier that performs efficiently compared to other classifiers. This model used NSL-KDD dataset for experiments. The experiments showed that this model attains a high detection rate and low false alarm rate. For DoS attacks, it achieved an accuracy of 99.67%. DoS, probe, R2L, and U2R attack types are detected using this model. (Prashil Negandhi et.al;2019) used a Random forest classifier for modelling and an NSL-KDD dataset for experimental studies. To increase the classification accuracy this model employ datamining techniques of feature selection using Gini importance. This model is fast and have high accuracy. (Tao Wa et.al;2022) proposed an IDS based on enhanced Random forest and Synthetic Minority Oversampling technique(SMOTE) algorithm. First this method combines K- K-means clustering with SMOTE sampling to increase the number of minor samples to achieve a balanced dataset. The enhanced random forest model, trained and tested using the NSL-KDD dataset, achieved a training accuracy of 99.72% and a testing accuracy of 78.47%. (S. Krishnaveni et.al;2020) proposed an anomaly-based intrusion detection model using SVM for cloud computing. The model uses NSL-KDD dataset for experiment and got an accuracy of 96.24%. (ND Patel et.al;2022) put forward two models, SVM and DNN(Deep Neural Network, an artificial neural network model). The accuracy, precision and recall were calculated and compared using NSL-KDD training and validation data in which DNN shows slightly high accuracy than SVM. DNN proves to be more effective than SVM. (Dong Seong Kim et.al;2003) also put forward a method of applying SVM to network-based intrusion detection systems. Experiments have been done using the 1999KDD intrusion detection dataset and found that SVM is effective in applying intrusion detection (Quamar Niyaz et.al;2015) studied about deep learning-based approach for developing an efficient and flexible NIDS. The author used Self Taught Learning (STL) on the NSL-KDD benchmark dataset for network intrusions. They compared their approaches performance with previous work, considering performance evaluation matrices such as accuracy, precision, recall and F – measure. (Kantagha Edmond et.al;2022) conducted a survey summerises various deep learning techniques that applied in intrusion detection systems. He addressed various problems encountered in network security. The author had given a deep summary of benchmark datasets used in deep learning techniques. He did a comparison of performances of different techniques. (R. Vinayakumar et.al;2017) proposed a DNN model for IDS and classify unseen and unpredicatable cyber-attacks. The experimental study evaluates DNN with other models using benchmark datasets. The DNN performed well on KDD-CUP99 than NSL-KDD, UNSW-NB15, Kyoto, WSN-DS and CICIDS2017 and it performs well than other Machine learning models.

Table 1 Comparison of Previous Studies

AUTHOR & YEAR	METHODOLOGY	DATASET
S. Snehal Mulay et.al; 2010	Decision tree, SVM	KDD CUP 99
A. Khraisat.et al; 2014	Decision Tree	NSL KDD
Muhammad Farhan et al; 2022	Random Forests	UNSW-NB15
Tao Wu et al; 2022	Random Forests, SMOTE	NSL KDD
Md.Al Mehedi Hasan et al; 2014	SVM,RandomForests	KDD 99
Mokhtar Mohammadi et al;2021	SVM	KDD 99 NSL KDD UNSW NB-15
Lirim Ashiku et al;2021	Deep Learning (DNN)	UNSW NB-15
Sunanda Gamage et.al; 2020	Learning	NSL KDD, KDD99, CICIDS2017
Salam Allawi Hussain et.al; 2021	Ensemble Learning	NSL-KDD, UNSW NB15
Sabrine Ennaji et al;2021	Ensemble Learning	NSL KDD

(Ali H Mirza et.al;2018) combined different machine ldearning algorithms such as Neural network, Decision tree, and Logistic Regression. Then boost the overall performance using ensemble learning by employing majotity voting scheme based on individual classifier performance using KDD-CUP datasets. (Nitesh Singh Bhati et.al2015) built an ensemble model to provide better accuracy for intrusion detection. Three based learning technique, Adaboost, Random-forest, Logistic regression have integrated to voting ensemble model. This model acquires an accurat

Table 2 Advantages and Disadvantages of Existing Studies

AUTHOR(S) & YEAR	ADVANTAGES	DISADVANTAGES
C. Snehal Mulay. et al; 2010	Utilizes both Decision Trees and SVM, offering diverse analysis.	- Relies solely on the KDD Cup 1999 dataset.
	Evaluates performance using widely accepted metrics.	- Does not explore ensemble methods or deep learning techniques.
A. Khraisat. et al; 2014	- Provides comprehensive survey of IDS techniques, datasets and challenges.	- No proper methods for zero day attacks.A.
	- Evaluates performance using NSL-KDD dataset.	- May overlook effectiveness of other machine learning algorithms.
Muhammad Farhan et al; 2022	- Explores effectiveness of Random Forests for intrusion detection.	- May lack comparison with other machine learning algorithms.
	- Utilizes the UNSW-NB15 dataset, enhancing realism of evaluation.	
Tao Wu et al; 2022	- Conducts comparative analysis between Random Forests and SMOTE Algorithm	- Focuses solely on comparing two specific algorithms.
	- Utilizes NSL-KDD dataset for evaluation.	- May overlook effectiveness of other machine learning algorithms.
Md.Al Mehedi	- Investigates effectiveness of SVM and	- May not explore other machine learning

Hasan et al; 2014	Random Forests.	algorithms.
	- Utilizes KDD99 dataset, enhancing realism of evaluation.	
Mokhtar Mohammadi et. Al; 2021	- Provides comprehensive survey of SVM-based systems.	- Focuses solely on SVM-based methods.
	- Utilizes KDD Cup 1999 dataset for evaluation.	- May overlook effectiveness of other machine learning algorithms.
Lirim Ashiku et.al; 2021	- Explores application of deep learning techniques for intrusion detection.	- May not provide comparisons with traditional machine learning algorithms.
	- Utilizes UNSW-NB15 dataset, enhancing realism of evaluation.	
Sunanda Gamage et al; 2020	- Provides comprehensive survey of deep learning techniques.	- May lack detailed comparisons with traditional machine learning algorithms.
	- Utilizes NSL KDD, CICIDS2017, dataset for evaluation.	
Salam Allawi Hussein et.al	- Explores effectiveness of ensemble learning approaches.	- May not provide comparisons with individual machine learning algorithms.
	- Utilizes NSL-KDD and UNSW NB-15 dataset for evaluation.	

3. Results and Discussions

3.1 Results

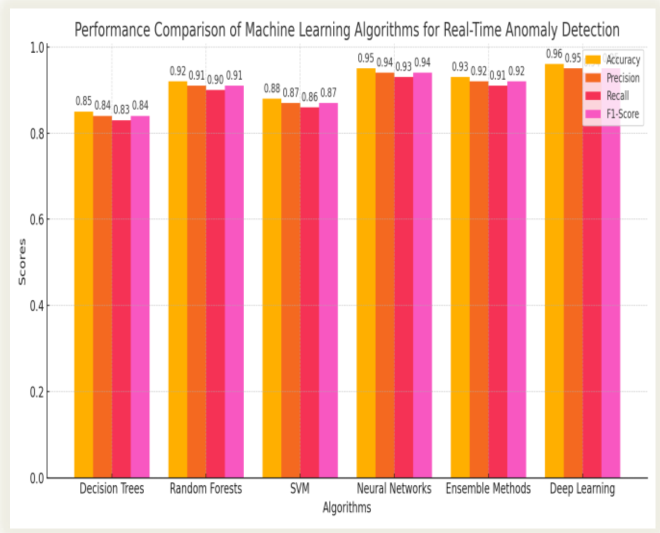

Figure 1 Performance Comparison of Different Machine Learning Models in Anomaly Detection

The performance comparison of machine learning algorithms for real-time anomaly detection highlights significant differences across several key metrics, including accuracy, precision, recall, and F1-score. Decision Trees, known for their simplicity and interpretability, tend to overfit on training data, achieving moderate performance. Random Forests, as an ensemble method, provide a robust improvement in accuracy and are particularly effective due to their ability to combine multiple decision trees, enhancing both precision and recall.

3.2 Discussions

Support Vector Machines (SVM) perform well in high-dimensional spaces and offer versatility in classification tasks, though they can struggle with very large or imbalanced datasets. Neural Networks and Deep Learning techniques exhibit superior accuracy and recall, capable of capturing complex patterns within the data. However, they require significant computational resources and may not be as efficient for real-time processing. Ensemble Methods strike a balance by leveraging the strengths of multiple models, thus achieving high overall performance. This comparative analysis underscores that while advanced techniques like Deep Learning offer high accuracy, their computational demands must be carefully considered in real-time applications.

The mentioned intrusion detection system (IDS) studies face several limitations. Many rely on specific datasets such as NSL-KDD or KDD Cup 1999, limiting generalization to diverse real-world network traffic and varying attack scenarios. Algorithmic bias is prevalent due to a narrow focus on specific machine learning techniques like Decision Trees, Random Forests, or Support Vector Machines, often overlooking the effectiveness of alternative methods. Data quality issues, including missing values, noise, and imbalanced class distributions, also impact IDS model performance. Scalability is another concern, particularly in large-scale networks, due to the computational complexity of certain algorithms and limitations in processing power and memory resources. Additionally, deep learning and ensemble models' increased complexity can reduce model interpretability, making it difficult to understand decision-making processes and identify the root causes of false positives or negatives. Real-time constraints pose further challenges, as some models may not meet the latency requirements needed for timely threat response, especially with large traffic volumes or complex algorithms. Moreover, IDS models are vulnerable to adversarial attacks, where sophisticated manipulation of network traffic can evade detection, highlighting potential cybersecurity weaknesses. Lastly, significant computational resources, expertise, and infrastructure support are required for deploying and maintaining IDS models, posing difficulties for small organizations or resource-constrained environments.

Conclusion

This study provides a comprehensive analysis of various machine learning algorithms for real-time anomaly detection in network traffic. Through a detailed comparative evaluation, identified the strengths and weaknesses of algorithms such as Decision Trees, Random Forests, Support Vector Machines, Neural Networks, and Ensemble Methods. Our findings highlight the trade-offs between model accuracy, computational efficiency, scalability, and interpretability. While advanced techniques like Deep Learning exhibit high accuracy and robustness, they also demand substantial computational resources and pose challenges in real-time implementation due to their complexity and latency issues. Furthermore, the study underscores the critical importance of using diverse and high-quality datasets to ensure the generalizability and effectiveness of IDS models across different network environments and attack scenarios. It also emphasizes the need for continuous updates and adaptability of IDS models to address evolving threats and adversarial attacks. By addressing the identified limitations and leveraging the strengths of various machine learning approaches, we can move closer to achieving robust, scalable, and interpretable IDS solutions that effectively safeguard network integrity in dynamic and complex environments.

Acknowledgements

I would like to express my heartfelt gratitude to Mrs. Nisha Kuruvila, our esteemed principal, for her unwavering support and encouragement throughout this study. Special thanks are also due to Mr. Rajkumar T, my dedicated guide, whose expertise and guidance were invaluable in completing this research. I am deeply thankful to the College of Engineering, Kallooppara, for providing the necessary resources and conducive environment for conducting this study.

References

[1]. Roesch, M. (1999). “Snort - Lightweight Intrusion Detection for Networks.”

[2]. Muda, Z et.al(2017). “Machine Learning in Intrusion Detection: A Comprehensive Survey”. Journal of Network and Computer Applications, 78, 42-60.

[3]. Patcha, A et.al (2007),” An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends”. Computer Networks, 51(12), 3448-3470.

[4]. Bergstra et.al(2012). “Random Search for Hyper-Parameter Optimization”. Journal of Machine Learning Research, 13, 281-305.

[5]. Oluwadamilare Harazeem Abdulganiyu et.al(2023), “ A systematic literature review for network intrusion detection system”, Rsearchgate,International Journal of Information security, Article Vol. 22, pages 1125-1162(2023)

[6]. Tushar Rakshe et.al(2017), "Anomaly based network intrusion detection using machine learning techniques", International journal of engineering research and technology(IJERT), ISSN: 2278-0181, vol.6, Issue 05, May2017.

[7]. Satish Kumar et.al(2021), "Research trends in network based intrusion detection systems: A review", IEEE Access , Vol.9,DOI: 10.1109/ACCESS.2021.3129775.

[8]. Mohammed Hussein Thwaini et.al(2022), "Anomaly detection in network traffic using machine learning for early threat detection", Research gate, December 2022, Data&Metadata, DOI: 10.56294/dm202272.

[9]. Ali Bou Nassif et,al(2021), "Machine learning for anomaly detection: A Systematic Review", IEEE ACCESS, Vol.9(2021).

[10]. M.H Bhuyan et.al (2013), "Network Anomaly Detection: Methods, Systems and Tools", IEEE Common. Serveys, Tuts, Vol.16, No.1, pp. 303-336, 1st quart, 2013, http://ieeeexplore.ieee.org/document/6524462/

[11]. Zhen Yang et.al (2022), "A systematic literature review of methods and dataset for anomaly-based network intrusion detection", Elsevier, Computers and security, Vol.116, May 2022, 102675

[12]. V. Jyothsna et.al (2011), " A Review of Anomaly based intrusion detection system", International Journal of Computer applications, 2011, Vol.28, no:7, September 2011

[13]. Mohammed Al-Omari et.al(2021), "An Intelligent Tree based Intrusion detection model for cyber security", Springer, Journal of networks and systems management, 2021, Vol.29, Article No:20(2021)

[14]. Azidine Guezzar et.al(2021), "A Reliable network intrusion detection approach using Decision tree with enhanced data quality", Hindawi, Security and Communication networks, vol.2021/ Article ID: 1230593/ https://doi.org/10.1155/2021/1230593

[15]. Nabila Farnaaz et.al(2016), " Random forest modelling for network intrusion detection system", Elsevier, Proceda, Computer science, Vol.89, 2016, pages 213-217.

[16]. Prashil Negandhi et.al() "Intrusion detection system using random forest on the NSL-KDD dataset", Springer, Emerging research in Computing, Information, communication and Applications, pp.519-531.

[17]. Tao Wu et.al(2022), "Intrusion detection system combined enhanced random forest with SMOTE Algorithm.", Springer, EURASIP Journal on advances in signal processing, 2022, Article No: 39(2022)

[18]. S. Krishnaveni et.al(2020), "Anomaly based intrusion detection system using Support Vector Machine", Springer, Artificial Intelligence and Evolutionary Computations in Engineering systems, pp.723-731

[19]. ND Patel et.al(2022), "Detection of Intrusions using Support Vector Machines and Deep Neural network", IEEE, 2022, 10th International conference in Reliability, Infocom Technologies and Optimization, DOI: 10.1109/ICRJT056286, 2022,9964756.

[20]. Dong Seong Kim et.al(2003), "Network based intrusion detection with Support vector machines", Springer, International conference on information networking, ICOIN2003: Information Networking pp. 747-756.

[21]. Quamar Niyaz et.al(2015), " A deep learning approach for network intrusion detection system", BICT2015, December 03-05, New York City, United States, DOI:10.4108/eai.3-12-2015.2262516.

[22]. Kantagha Edmond et.al (2022), "Intrusion Detection System using deep learning approaches: A Survey" , Springer, International Conference on Innovative computing and Communication pp.777-790.

[23]. R Vinayakumar et.al(2019), "Deep Learning Approach for Intelligent intrusion detection System", IEEE Access, DOI: 10.1109/ACCESS.2019.2895334.

[24]. Ali. H. Mirza et.al(2018), " Computer network intrusion detection using various classifiers and ensemble learning", IEEE, 26thb signal processing and communications applications conference, DOI: 10.1109/SIU.2018.8404704.

[25]. Nitesh Singh Bhati et.al, "An ensemble model for network intrusion detection using

Adaboost, random forest and Logoistic regression”, Springer, Applications of artificial intelligence and machine leraning, pp. 777-789.

[26]. Snehal Mulay et.al, “Intrusion Detection System using Support vector machine and Decision Tree”, International Journal of computer educations, June 2010, DOI:10.5120/758-993

[27]. A.Khraisat et.al ,” Survey of Intrusion Detction systems: Techniques, datasets and challenges” , Springer, Cybersecurity, Article Number 20,(2019).

[28]. Sabrine Ennaji et.al, “A powerful Ensemble learning apPrecproach for Improving network intrusion detection system(NIDS)” , In Prec: 2021 fifth international conference on intelligent computing in data sciences, DOI:10.1109/ICDS53782.2021.9626727

[29]. Muhammad Farhan et.al , “ Network Intrusion detection by using the Random forest with Extra tree classifier”, International Jouranl of Applied engineering Reasearch, Vol 7, No: 1, June 2022, ISSN: 2666-2795.

[30]. Md Al Mehedi Hasan et.al ,” Support vector machine and Random forest modeling for Intrusion Detection System(IDS)”, Journal of Intelligent Learning Systems and Applications, 2014,6, 45-52. DOI:10.4236/jilsa.2014.61005

[31]. Sunanda Gamage et.al, “Deeep learning methods in network intrusion detection: A survey and an objective comparison”, Elsevier, Journal of network and computer applications, Vol.169, November2020, 102767.

[32]. Lirim Ashiku et.al, “Network Intrusion Detection Sustem using Deep Learning” , Elsevier, Procedia Computer Science, Vol. 185, 2021, Pages 239-247.