International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0442

e ISSN: 2584-2854
Volume: 03
Issue: 09 September 2025
Page No: 2810 - 2813

Intrusion Detection System Using Machine Learning

Krisha J Gowda[1], Mrs. Jaya Kruna B[2], Kiran[3], Gopika[4], M. Yadunandan[5]

[1,3,4,5] UG – Computer Science and Engineering, AMC Engineering College, Bengaluru, Karnataka, India.

[2]Assistant Professor, Department of Computer Science and Engineering, AMC Engineering, College, Bengaluru, Karnataka, India.

***Emails:** 1am22cs097@amceducation.in[1], b.jayakaruna@amceducation.in[2], 1am22cs095@amceducation.in[3], 1am22cs064@amceducation.in[4], 1am22cs104@amceducation.in[5]*

Abstract

In the evolving landscape of cybersecurity, traditional rule-based intrusion detection systems (IDS) struggle to keep pace with the increasing volume, velocity, and sophistication of network attacks. This paper explores the design and implementation of an intrusion detection system that leverages machine learning (ML) techniques to enhance threat detection capabilities. By analyzing network traffic data and identifying patterns indicative of malicious behavior, ML-based IDS solutions offer improved accuracy, adaptability, and automation in identifying both known and unknown threats. The proposed system employs supervised and unsupervised learning algorithms, including Decision Trees, Random Forests, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Neural Networks. Feature selection and data preprocessing are applied to optimize model performance.

***Keywords:** Intrusion Detection System (IDS), machine learning, cyber security, network security, anomaly detection, supervised learning, un supervised learning.*

1. Introduction

With the exponential growth of digital infrastructure and the increasing dependence on net- worked systems, cybersecurity has become a critical concern across all sectors. One of the most persistent and evolving threats to these systems is unauthorized access or intrusion. Intrusion Detection Systems (IDS) are crucial tools designed to monitor network or system activities for malicious actions or policy violations. Traditional IDS methods, which rely on predefined signatures and manual rule sets, often fall short in detecting novel or sophisticated attacks, especially in dynamic and large-scale environments. To address these limitations, researchers and practitioners are turning to Machine Learning (ML) as a promising approach for enhancing IDS capabilities. Machine learning enables systems to learn from historical data, detect com- plex patterns, and adapt to new and emerging threats without requiring explicit programming. By leveraging algorithms such as Support Vec- tor Machines (SVM), Decision Trees, Random Forests, K-Nearest Neighbors (KNN), and Neu- ral Networks, ML-based IDS solutions can offer higher detection accuracy, reduced false positives, and the ability to identify zero-day attacks. In this paper, we propose a machine learning- based intrusion detection framework that utilizes various classification algorithms to distinguish between normal and malicious network behaviour. We focus on key aspects such as data preprocessing, feature selection, model train- ing, and performance evaluation. Benchmark datasets like NSL-KDD and CICIDS2017 are used to assess the effectiveness of different ML models. The results are compared in terms of accuracy, precision, recall, F1-score, and false positive rate. In today’s highly interconnected digital world, the security of information systems is more crit- ical than ever. Organizations, governments, and individuals increasingly rely on networks and internet-based services, making them prime targets for a wide range of cyber threats. Among these threats, unauthorized access, data breaches, denial-of-service attacks, and mal- ware infections pose significant risks to the confidentiality, integrity, and availability of sensi- tive data. As cyberattacks become more fre- quent and complex, defending against them re- quires more than just traditional, rule-based security systems. [1-3]

2. Method

Traditional machine learning algorithms have been extensively used in intrusion detection due to their simplicity and effectiveness on structured datasets. Decision Trees (DT) are widely adopted because of their interpretability, as they classify traffic based on a sequence of rules. Random Forest (RF), an ensemble of decision trees, improves upon this by reducing overfitting and enhancing detection accuracy. Support Vec- tor Machines (SVM) are another powerful approach, capable of distinguishing between nor- mal and malicious traffic using hyperplane separation. Similarly, K-Nearest Neighbours (KNN) classifies new traffic by comparing it with existing labelled data points, though it may struggle with high-dimensional data. Naïve Bayes, based on probabilistic classification, is lightweight and suitable for real-time detection. These classical methods serve as strong baselines for evaluating IDS performance, though their effectiveness may decline in complex, large-scale, or evolving network environments. Deep learning methods have gained significant traction in IDS research due to their ability to model high-dimensional and complex traffic data. Artificial Neural Net- works (ANN) were among the earliest deep learning techniques applied, but more advanced architectures have since emerged. Convolutional Neural Networks (CNN) are particularly effective for identifying spatial patterns in net- work traffic flows, while Recurrent Neural Net- works (RNN) and Long Short-Term Memory (LSTM) networks excel in detecting temporal dependencies, such as patterns in sequential packets. Autoencoders are widely used for anomaly detection by learning compressed representations of normal traffic and flagging deviations as potential intrusions. Generative Adversarial Networks (GANs) are also explored to generate synthetic attack data and improve IDS robustness against adversarial threats. These deep learning models achieve high accuracy but often require large amounts of data Figure 1 shows Flow Chart, Figure 2 shows Intrusion Detection System, Figure 3 IOT shows Systems [4-6]

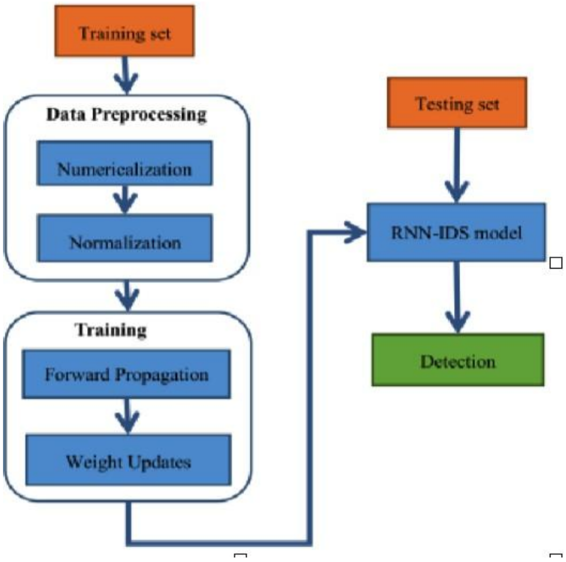

Figure 1 Flow Chart

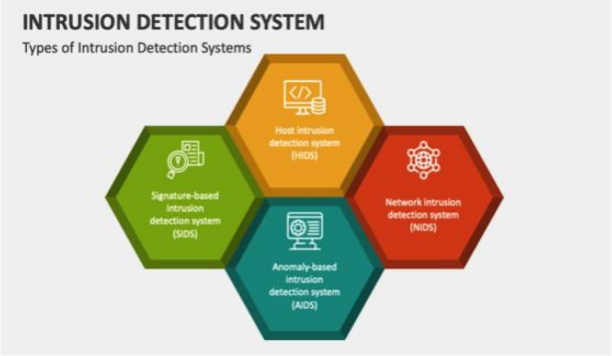
INTRUSION DETECTION SYSTEM
Types of Intrusion Detection Systems
Host intrusion detection system (HIDS)
Signature-based intrusion detection system (SIDS)
Network intrusion detection system (NIDS)
Anomaly-based intrusion detection system (AIDS)

Figure 2 Intrusion Detection System

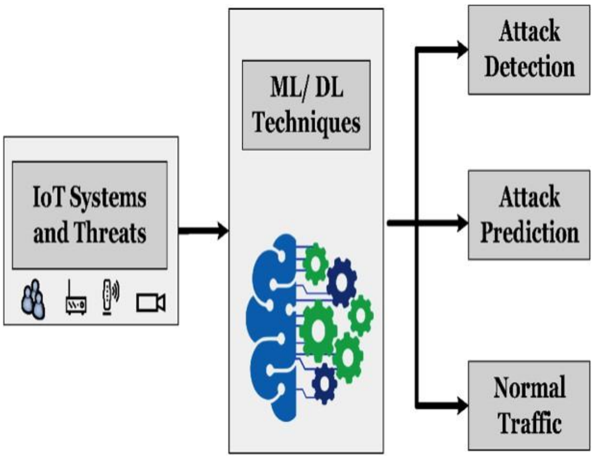

Figure 3 IOT Systems

3. Related Work

Intrusion Detection Systems (IDS) have beenan-active area of research for over two decades, with machine learning emerging as a key enabler for improving detection accuracy and adaptability. Early studies primarily employed traditional classifiers such as Decision Trees, Naïve Bayes, and Support Vector Machines on bench- mark datasets like KDD Cup 1999 and NSL- KDD. These methods demonstrated reasonable accuracy for known attack types but often struggled with scalability and generalization to un- seen attacks. To address these limitations, ensemble learning approaches such as Random Forest, AdaBoost, and Gradient Boosting were later introduced, offering more robust detection by combining multiple weak learners. With the advent of deep learning, researchers began exploring architectures such as Artificial Neural Networks (ANN), Convolutional Neu- ral Networks (CNN), and Recurrent Neural Net- works (RNN) for intrusion detection. CNNs were shown to be effective in capturing spatial features of traffic flows, while RNN and Long Short-Term Memory (LSTM) networks excelled in modelling sequential traffic data, particularly for time-series intrusion patterns. Autoencoders have also been widely applied for unsupervised anomaly detection, where deviations from reconstructed normal traffic patterns are flagged as potential intrusions. More recent works have utilized Generative Adversarial Net- works (GANs) to generate synthetic attack samples, thereby improving IDS robustness against adversarial and zero-day threats. Hybrid models have also gained attention in the literature, combining the strengths of multiple machine learning paradigms. For in- stance, CNN-LSTM hybrids capture both spa- tial and temporal traffic characteristics, while Autoencoder-SVM combinations integrate un- supervised feature learning with supervised classification. Additionally, feature selection and dimensional education techniques such as Principal Component Analysis (PCA) and Information Gain have been applied to enhance efficiency [7-10]

Conclusion

Intrusion Detection Systems (IDS) play a vital role in safeguarding modern networks against increasingly sophisticated cyberattacks, and ma- chine learning has proven to be a powerful tool in enhancing their effectiveness. This paper reviewed various machine learning techniques ranging from traditional classifiers such as Decision Trees, Support Vector Machines, and Random Forests to more advanced deep learning approaches including CNNs, LSTMs, Autoencoders, and hybrid models. The discussion high-lighted that classical ML methods provide simplicity and interpretability, while ensemble and deep learning techniques offer superior detection performance, particularly in handling large- scale and high-dimensional traffic data. Emerging approaches such as Graph Neural Networks, Transformer-based models, and reinforcement learning hold promise for building adaptive and intelligent IDS capable of countering evolving threats. However, challenges remain in terms of reducing false positive rates, ensuring scal- ability in real-time environments, and improving the interpretability of complex models. Future research should focus on integrating explainable AI, federated learning for privacy-preserving detection, and adversarial robustness to strengthen IDS performance in dynamic and distributed network environments. Overall, ma- chine learning continues to pave the way for the development of more accurate, adaptive, and intelligent intrusion detection systems that are essential for securing next-generation cyber infrastructures.

References

[1]. W. Lee and S. Stolfo, "A framework for constructing features and models for intrusion detection systems," ACM Transactions on Information and System Security (TISSEC), vol. 3, no. 4, pp. 227–261, 2000.

[2]. M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in Proc. IEEE Sympo- sium on Computational Intelligence for Se- curity and Defense Applications (CISDA), pp. 1–6, 2009.

[3]. I. Sharafaldin, A. Habibi Lashkari, and A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in Proc. ICISSP, pp.

108–116, 2018. (CICIDS2017 dataset)

[4]. J. Kim, J. Kim, H. Shim, and S. Kim, “Deep learning-based intrusion detection system for real-time network intrusion de- tection,” Electronics, vol. 8, no. 7, p. 826,2019.

[5]. N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems,” in Proc. Military Communications and Information Systems Conference (MilCIS), pp. 1–6, 2015.

[6]. S. Shone, D. Ngoc, V. Phai, and Q. Shi, “A deep learning approach to network in- trusion detection,” IEEE Transactions on Emerging Topics in Computational Intel- ligence, vol. 2, no. 1, pp. 41–50, 2018.

[7]. R. Vinayakumar, M. Alazab, K. Soman, P. Poornachandran, and S. Venkatraman, “Deep learning approach for intelligent in- trusion detection system,” IEEE Access, vol. 7, pp. 41525–41550, 2019.

[8]. T. Diro and N. Chilamkurti, “Distributed attack detection scheme using deep learn- ing approach for Internet of Things,” Fu- ture Generation Computer Systems, vol. 82, pp. 761–768, 2018.

[9]. Y. Zou, C. Leckie, K. Ramamohanarao, and J. Bezdek, “Detection of network at- tacks using neural networks,” in Proc. IEEE Int. Conf. Neural Networks, vol. 3, pp. 1244–1249, 2006.

[10]. X. Zhang, Y. Chen, J. Pei, and C. Tang, “A survey on deep learning for network intrusion detection,” IEEE Communications Surveys & Tutorials, vol. 23, no. 3, pp. 2064–2101, 2021.