IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

An Intelligent Intrusion Detection System Using Machine Learning for Network Traffic Analysis

Mrs. Pradeepa G[1], Ashika P[2], Keerthika S[3], Kirthiga M[4]
Associate Professor, Vivekanandha College of Engineering for Women, College (Autonomous),Tiruchengode [1]
UG Scholars, Vivekanandha College of Engineering for Women, College (Autonomous),Tiruchengode[2-4]

Abstract: *The rapid advancement of communication technologies and the widespread use of internet-based services have significantly increased the vulnerability of network systems to cyber threats. Organizations today face a wide range of attacks, including malware, phishing, denial-of-service, and unauthorized access attempts. Traditional intrusion detection systems (IDS), which rely primarily on signature-based detection techniques, are often unable to identify newly emerging threats and zero-day attacks.*

This paper proposes an intelligent intrusion detection system that leverages machine learning algorithms to analyze network traffic patterns and detect anomalies effectively. The system incorporates various stages such as data preprocessing, feature extraction, model training, and classification. Algorithms like Random Forest and Support Vector Machines are used to improve detection accuracy and system reliability. The proposed approach not only detects known attack patterns but also identifies unknown threats by learning behavioral patterns from historical data. Experimental results demonstrate improved accuracy, reduced false positives, and enhanced adaptability, making the system suitable for real-time cybersecurity applications..

Keywords: Intrusion Detection System, Machine Learning, Cybersecurity, Network Traffic Analysis, Random Forest, Anomaly Detection

I. INTRODUCTION

In recent years, the increasing dependence on digital infrastructure has made network security a critical concern for individuals, organizations, and governments. With the growth of cloud computing, IoT devices, and online services, networks have become more complex and vulnerable to cyberattacks. Attackers continuously develop new techniques to exploit system vulnerabilities, making it challenging for traditional security mechanisms to keep up.

Intrusion Detection Systems (IDS) play a vital role in identifying malicious activities within a network. These systems monitor network traffic and generate alerts when suspicious behavior is detected. However, conventional IDS approaches are primarily rule-based and depend on predefined attack signatures. As a result, they are ineffective in detecting previously unseen attacks and require frequent manual updates.

Machine learning provides a powerful alternative by enabling systems to automatically learn patterns from data and make intelligent decisions. By analyzing large volumes of network traffic, machine learning models can identify subtle anomalies that may indicate potential intrusions. This paper focuses on designing a robust machine learning-based IDS that enhances detection capabilities while minimizing false alarms.

II. LITERATURE REVIEW

The development of intrusion detection systems has evolved significantly over the years. Early IDS models were based on signature detection techniques, where known attack patterns were stored in a database and matched against incoming network traffic. While effective for known threats, these systems failed to detect new or modified attacks.

To address this limitation, anomaly-based detection systems were introduced. These systems establish a baseline of normal network behavior and identify deviations from this baseline as potential threats. Although anomaly-based

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

systems can detect unknown attacks, they often suffer from high false positive rates due to variations in normal behavior.

Recent research has focused on integrating machine learning techniques into IDS. Algorithms such as Decision Trees, Support Vector Machines, and Random Forest have been widely used for classification tasks. Ensemble learning methods, in particular, have shown improved performance due to their ability to combine multiple models and reduce overfitting.

Deep learning approaches, including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), have also been explored for intrusion detection. These models can capture complex patterns in data but require large datasets and high computational power. Despite these advancements, challenges such as data imbalance, scalability, and real-time processing remain areas of ongoing research.

III. PROBLEM STATEMENT

Despite the availability of various intrusion detection techniques, existing systems still face several critical challenges. Traditional IDS solutions are heavily dependent on predefined rules and signatures, which limits their ability to detect new and unknown threats. As cyberattacks continue to evolve, this limitation becomes increasingly significant.

Another major issue is the high rate of false positives, where normal network activities are incorrectly classified as malicious. This leads to unnecessary alerts and increases the workload for network administrators. Additionally, many systems lack adaptability and require frequent manual updates to remain effective.

Handling large volumes of network data is also a challenge, as modern networks generate massive amounts of traffic. Conventional systems often struggle with scalability and performance when processing such data. Therefore, there is a need for an intelligent, automated, and scalable intrusion detection system that can accurately detect threats in real time while adapting to changing network environments.

IV. PROPOSED SYSTEM

The proposed system is designed to provide an efficient and intelligent solution for intrusion detection using machine learning techniques. It focuses on analyzing network traffic data and classifying it into normal or malicious categories based on learned patterns.

The architecture of the system consists of multiple modules that work together to ensure accurate detection. The data collection module gathers network traffic data from various sources, including datasets and real-time monitoring systems. The preprocessing module cleans and prepares the data for analysis by handling missing values and transforming categorical features.

The feature selection module plays a crucial role in identifying the most relevant attributes, which helps in reducing computational complexity and improving model performance. The machine learning module is responsible for training and testing models using selected features. Finally, the prediction module analyzes incoming data and generates alerts if any suspicious activity is detected.

The system is designed to operate in real time, allowing continuous monitoring of network traffic and immediate response to potential threats.

V. METHODOLOGY

The methodology adopted in this system involves several important steps to ensure accurate and efficient intrusion detection.

5.1 Data Preprocessing

Data preprocessing is a crucial step in the machine learning pipeline. Raw network data often contains missing values, noise, and inconsistencies that can affect model performance. In this stage, missing values are handled, and irrelevant data is removed. Categorical variables are converted into numerical format using encoding techniques, and normalization is applied to scale the data for better model performance.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

5.2 Feature Selection

Feature selection helps in identifying the most important attributes that contribute to accurate classification. By removing redundant and irrelevant features, the system reduces computational complexity and improves efficiency. This step also helps in avoiding overfitting and enhances the generalization capability of the model.

5.3 Machine Learning Models

The system uses multiple machine learning algorithms to improve detection accuracy. Random Forest is used due to its robustness and ability to handle large datasets. Support Vector Machine is effective for classification tasks with high-dimensional data. Decision Trees provide interpretability and ease of implementation.

5.4 Model Evaluation

To evaluate the performance of the system, several metrics are used. Accuracy measures the overall correctness of the model, while precision and recall provide insights into classification performance. The F1-score balances precision and recall, offering a comprehensive evaluation of the model.

VI. IMPLEMENTATION

The implementation of the proposed system is carried out using modern tools and technologies. The backend is developed using Python, which provides powerful libraries for data analysis and machine learning. Flask is used to create APIs that enable communication between the frontend and backend.

The frontend is developed using Flutter and Dart, providing a user-friendly interface for interaction. Users can input network data and view results in an intuitive format. The system processes the data on the backend, applies the trained model, and returns predictions to the frontend.

The modular design of the system ensures flexibility and scalability, allowing it to be extended or modified based on future requirements.

VII. RESULTS AND DISCUSSION

The experimental results demonstrate that the proposed system performs effectively in detecting network intrusions. The use of machine learning algorithms significantly improves detection accuracy compared to traditional methods. The system is able to identify both known and unknown attack patterns, which is a major advantage in dynamic network environments.

The reduction in false positives enhances the reliability of the system and reduces unnecessary alerts. Among the algorithms used, Random Forest shows superior performance due to its ensemble nature and ability to handle complex data patterns.

The results indicate that the system is suitable for real-time deployment and can be used to strengthen network security in various applications.

VIII. ADVANTAGES

- Detects both known and unknown attacks by learning patterns instead of relying only on signatures
- Significantly reduces false positives, improving system reliability and trust
- Continuously improves performance through retraining with new data
- Eliminates the need for manual rule creation and frequent updates
- Efficiently processes large volumes of network traffic data
- Enables real-time monitoring and faster detection of threats
- Provides high accuracy due to the use of advanced machine learning algorithms
- Scalable architecture suitable for enterprise and cloud environments

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

IX. CONCLUSION

This paper presents a machine learning-based intrusion detection system designed to enhance network security. The proposed system overcomes the limitations of traditional IDS by providing intelligent, adaptive, and accurate detection of cyber threats. By leveraging machine learning techniques, the system can identify both known and unknown attacks while minimizing false positives.

The results demonstrate that the system is effective, scalable, and suitable for real-time applications. It provides a strong foundation for developing advanced cybersecurity solutions in modern network environments.

X. FUTURE WORK

Future work can focus on enhancing the system by integrating deep learning techniques for improved accuracy. Real-time data streaming using big data technologies can be implemented to handle large-scale networks. Cloud-based deployment can increase accessibility and scalability.

Additionally, automated response mechanisms can be developed to take immediate action against detected threats. Advanced visualization tools can also be integrated to provide better insights into network activities and attack patterns.

REFERENCES

[1] K. Kendall, "A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems," *Massachusetts Institute of Technology*, 1999.

[2] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," *Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009.

[3] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed., Morgan Kaufmann, 2011.

[4] T. M. Mitchell, *Machine Learning*, McGraw-Hill Education, 1997.

[5] L. Breiman, "Random Forests," *Machine Learning Journal*, vol. 45, no. 1, pp. 5–32, 2001.

[6] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning Journal*, vol. 20, pp. 273–297, 1995.

[7] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[8] W. Lee and S. J. Stolfo, "A Framework for Constructing Features and Models for Intrusion Detection Systems," *ACM Transactions on Information and System Security*, vol. 3, no. 4, pp. 227–261, 2000.

[9] D. E. Denning, "An Intrusion-Detection Model," *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222–232, 1987.

[10] N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," *Military Communications and Information Systems Conference*, 2015. Intrusion Detection," *International Conference on Platform Technology and Service*, 2016.