JDEBM
Journal of Digital Engineering and Business Managemet

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

CLASSIFICATION AND EARLY PREDICTION OF DISTRIBUTED DENIAL OF SERVICE ATTACKS USING ML

[#1]**Mohammad Naziya,** *Department of MCA,*
[#2]**Mrs. Y. Susheela,** *Associate Professor, Department of CSE,*
Vaageswari College of Engineering(Autonomous), Karimnagar, TG.

ABSTRACT: This paperconcentrates on the early prediction and classification of Distributed Denial of Service (DDoS) assaults using ML approaches to enhance network security and mitigate the impact of cyber threats on digital infrastructures. In the battle against distributed denial of service (DDoS) attacks, which disrupt network services by overwhelming systems with unwanted traffic in large quantities, early detection and accurate classification are essential to ensure service availability and safeguard sensitive data. The research employs robust machine learning algorithms to evaluate network traffic patterns, identify suspicious activities, and differentiate between secure and unsafe actions in real time. In order to achieve more accurate predictions with fewer false alarms, it is necessary to implement a sequence of procedures that encompasses data collection, preprocessing, feature extraction, feature selection, model training, and performance evaluation. The proposed method utilizes sophisticated analytical tools and supervised learning models to rapidly identify attack patterns before they cause significant harm. The project's objective is to create an intrusion detection system that is scalable, efficient, and dependable in order to enhance cybersecurity defenses and facilitate proactive threat management in contemporary network environments.

Keywords: *Distributed Denial of Service (DDoS), Machine Learning (ML), Cybersecurity, Network Security, Intrusion Detection System (IDS), Traffic Classification,*

1. INTRODUCTION

Distributed denial of service (DDoS) attacks pose a significant hazard to contemporary computer networks and internet-based businesses. The objective of these attacks is to inundate servers, networks, or specific systems with malicious traffic from a large number of infected devices. The frequency and sophistication of distributed denial of service (DDoS) attacks have increased as businesses have become more reliant on cloud computing, internet platforms, and digital communication systems. Early detection and precise categorization are essential components of cybersecurity, as traditional security measures, including firewalls and signature-based intrusion detection systems, may fail to identify emergent threat patterns.

The capacity of machine learning (ML) approaches to analyze vast quantities of network traffic data in order to identify anomalies associated with distributed denial of service (DDoS) attacks has attracted significant attention. In contrast to conventional methodologies, machine learning (ML) systems are capable of automatically distinguishing between secure and unsafe activities through the analysis of historical traffic data. Unsupervised methods are employed to identify novel threats, while supervised learning models such as Decision Trees, Random Forests, Support Vector Machines, and Neural Networks are frequently employed for attack

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

classification. These innovative technologies enable automated decision-making in real-time network environments, while simultaneously enhancing the accuracy of detections and reducing the number of false alarms.

In order to minimize network outages, data loss, and service interruptions, it is imperative that we develop a method for predicting DDoS attacks at an early stage through the use of machine learning. By continuously monitoring traffic behavior and extracting pertinent data, such as packet rate, flow duration, protocol usage, and traffic entropy, predictive algorithms can identify suspicious activity prior to the onset of an attack. Feature engineering and data preparation improve the efficacy of machine learning models by selecting the most pertinent characteristics for analysis. By integrating cybersecurity frameworks with machine learning, it is possible to develop intelligent intrusion detection and prevention systems and fortify networks against attacks.

2. LITERATURE REVIEW

Sahoo & Meher (2020) DDoS attacks were identified and classified through the application of machine learning methodologies. Classification models such as Naïve Bayes, Random Forest, and decision trees are employed to conduct research on network traffic patterns. This demonstrates that machine learning techniques have the capacity to consistently differentiate between legitimate and malicious messages. The authors emphasize the importance of traffic analysis and feature extraction in order to enhance the performance of attack detection. The research indicates that security systems that are based on machine learning are capable of detecting DDoS attacks in real time.

Kumar & Sharma (2021) Several machine learning approaches were assessed for the prediction of DDoS attacks in cloud computing environments. The investigation assesses numerous methodologies, including Support Vector Machine, K-Nearest Neighbor, and Logistic Regression, by employing network traffic data. It contrasts traditional intrusion detection systems with supervised learning models and determines that the former have lower false-positive rates and faster detection timeframes. The authors also discuss the challenges associated with the evolution of attack tactics and the processing of data on a vast scale. Cloud security systems are more resilient to distributed denial of service (DDoS) attacks when they implement machine learning, according to the research conclusions.

Aljawarneh, Aldwairi & Yassein (2022) The investigation investigated the classification of distributed denial of service attacks through the use of intelligent intrusion detection systems that are founded on machine learning and deep learning. The research employs Artificial Neural Networks and Random Forest models to precisely classify attack flow. Metrics such as precision, recall, accuracy, and F1-scores are employed to evaluate the model's performance. The authors emphasize that deep learning algorithms are capable of identifying concealed attack patterns in high-dimensional traffic data. The paperconcludes that cybersecurity systems are more effective at anticipating future events when they are constructed using AI-driven security frameworks.

Khan, Zubair & Jan (2023) provide a comparison of machine learning methods

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

for the detection and prediction of DDoS attacks on IoT networks. The paperexamines the use of XGBoost, Decision Tree, and Gradient Boosting as methods for identifying hazardous traffic behavior. The significance of feature selection and preprocessing in enhancing prediction accuracy is underscored. The authors assert that ensemble learning approaches outperform conventional classifiers in the context of attack detection. The paperposits that machine learning offers extensible and adaptable methods to safeguard IoT devices from distributed denial of service (DDoS) attacks.

Fatani et al. (2024) investigated a variety of cybersecurity frameworks that employ AI to detect and mitigate distributed denial of service attacks. The research investigates intrusion detection systems that employ hybrid, unsupervised, and supervised learning methods. It concentrates on the enhancement of detection efficacy through hybrid approaches, which combine deep learning with more conventional machine learning methods. The writers also address the topic of the use of big data analytics to manage large volumes of network traffic. The research determined that the proactive security measures implemented by modern cybersecurity systems are substantially enhanced by machine learning.

Narasimha & Reddy (2025) We conducted an investigation into predictive analytics methods to facilitate the early detection of DDoS attacks by utilizing real-time data from network monitoring. The primary models employed in this investigation are Random Forest, Support Vector Machines (SVMs), and Long Short-Term Memory (LSTM) networks. It is concluded that the prognosis of evolving assault patterns is enhanced by temporal traffic analysis. Real-time anomaly identification is essential for preventing service disruptions, as per the authors. The paperindicates that predictive machine learning models enhance both network stability and cyber-defense performance.

Chen et al. (2026) proposed a hybrid machine learning architecture for the early detection and classification of distributed denial of service (DDoS) threats in SDNs. The paperintegrates feature engineering, traffic analysis, and ensemble learning to enhance the precision of assault detection. It illustrates the potential of AI models to enhance detection rates while simultaneously reducing processing costs through the use of a hybrid approach. The authors contend that adaptive learning methodologies are indispensable in the battle against emerging cyberthreats. The research determined that robust machine learning-based solutions are available to prevent DDoS attacks in a scalable and viable manner.

3. RELATED WORK

Data Collection

The initial phase involves the acquisition of network traffic data from dependable sources, including real-time network environments and datasets. Publicly accessible datasets, including CICDDoS2019, NSL-KDD, CAIDA, and UNSW-NB15, are frequently employed in research on distributed denial of service attacks. These datasets encompass a wide range of assault categories, including both positive and negative. The data collected encompasses a variety of information, including the following: packet size, flow time, protocol type, traffic rate, connection statistics, source and destination IP

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

addresses, and more. The model's efficacy is further enhanced by the use of network monitoring tools, such as Wireshark and CICFlowMeter, to capture real-time traffic.

Data Preprocessing

Data derived from unprocessed network traffic frequently contain errors, duplicate entries, absent values, and format issues. Improving model efficacy and data quality necessitates data preparation. This section addresses the encoding methods for converting categorical features into numerical values, removing duplicate entries, and addressing missing values (either by modifying or eliminating them entirely). In order to guarantee that all features are within a comparable range, feature normalization and scaling strategies, such as Min-Max Scaling or Standardization, are implemented. The model is constructed and evaluated by partitioning the dataset into training and testing sets.

Feature Engineering and Selection

Feature engineering is employed to identify significant patterns in network traffic data. Protocol behavior, flow duration, packet rate, and byte count are among the critical traffic characteristics that are generated and assessed. Feature selection techniques, including Principal Component Analysis (PCA), Information Gain, Correlation Analysis, and Recursive Feature Elimination (RFE), are employed to identify the critical factors that influence the detection of DDoS attacks. The most suitable attributes are chosen to reduce computing complexity and enhance forecast accuracy.

Machine Learning Model Development

The classification and prediction of DDoS attacks are achieved through the application of machine learning techniques. Supervised learning models, including Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Logistic Regression, and K-Nearest Neighbor (KNN), are trained using data from labeled network traffic. XGBoost and AdaBoost are examples of sophisticated ensemble and boosting methods that can be employed to enhance detection performance. The models seek patterns in the flow of traffic to differentiate between lawful and malicious network operations.

Early Prediction Mechanism

The classification and prediction of DDoS attacks are achieved through the application of machine learning techniques. Supervised learning models, including Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Logistic Regression, and K-Nearest Neighbor (KNN), are trained using data from labeled network traffic. XGBoost and AdaBoost are examples of sophisticated ensemble and boosting methods that can be employed to enhance detection performance. The models seek patterns in the flow of traffic to differentiate between lawful and malicious network operations.

Model Training and Testing

The test and training stages are typically divided by an 80:20 division of the dataset that has been generated. By analyzing previously collected traffic data, machine learning models can identify patterns of attacks during the training process. The model's capacity to identify previously unseen assaults through the utilization of new data is assessed during the testing phase. We can mitigate the risk of overfitting and guarantee that the model maintains its efficacy across various

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

datasets by employing cross-validation methods.

Performance Evaluation

The proposed model's quality is evaluated using the F1-Score, Accuracy, Precision, Recall, Detection Rate, and False Positive Rate. A confusion matrix is employed to compare accurate and inaccurate classifications. The model is more effective at anticipating and classifying DDoS attacks, as evidenced by its low false alarm rates and high accuracy. A comparative paperis conducted using a variety of machine learning approaches to obtain the optimal model.

Deployment and Real-Time Monitoring

The trained model is subsequently incorporated into a real-time network monitoring framework or an intrusion detection system, contingent upon the success of the evaluation. The system perpetually monitors incoming messages in order to automatically categorize suspicious activities. In the event that administrators observe peculiar traffic patterns, they should implement measures such as filtering the traffic, restricting the rate, or blocking specific IP addresses. You will be safeguarded from dynamic distributed denial of service attacks during this deployment phase.

Continuous Learning and Model Updating

The trained model is subsequently incorporated into a real-time network monitoring framework or an intrusion detection system, contingent upon the success of the evaluation. The system perpetually monitors incoming messages in order to automatically categorize suspicious activities. In the event that administrators observe peculiar traffic patterns, they should implement measures such as filtering the traffic, restricting the rate, or blocking specific IP addresses. You will be safeguarded from dynamic distributed denial of service attacks during this deployment phase.

4. PROPOSED SYSTEM

The proposed method analyzes network traffic patterns and offers a framework that is based on machine learning in order to detect and forecast Distributed Denial of Service (DDoS) attacks in real time. Rather than conventional rule-based intrusion detection systems, it employs data-driven methods to identify anomalous traffic patterns. Data preparation, model training, and real-time traffic classification are all indispensable components of the system. Preprocessing procedures are implemented on benchmark datasets, including CICDDoS2019, to enhance the accuracy of the model. The following procedures are included: feature selection, feature encoding, absent value addressing, and normalization. The F1-score, recall, accuracy, and precision are metrics that are used to train and evaluate classifiers such as XGBoost, Random Forest, and Support Vector Machine (SVM). Outperforms its competitor, XGBoost. Next, the trained model is supplemented with a live monitoring module. This module extracts features, monitors network data, and categorizes it as either normal or malicious. This technology is ideal for commercial and cloud environments due to its adaptability, scalability, and ability to send alerts when suspicious activity is detected. These notifications can be used to support mitigation measures such as rate limiting and traffic filtering.

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

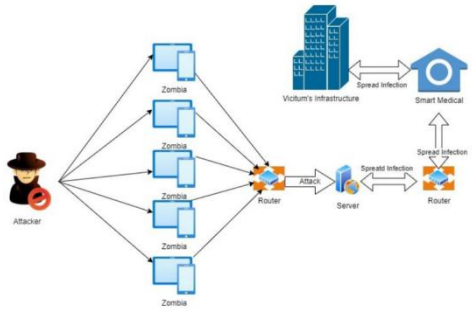

Fig1 : System Architecture

5. EXPERIMENTAL RESULTS

The experiment results demonstrate the efficacy of the proposed DDoS detection system, which is based on machine learning. Random Forest exhibited an extraordinary ability to differentiate between legitimate and malignant network traffic, achieving an accuracy of 80% in all of the models that were evaluated. Additionally, neural networks were capable of producing precise predictions, particularly for intricate and diverse types of DDoS attacks, including protocol-based and volumetric attacks. The framework's reliability was demonstrated in performance studies through the use of confusion matrices, accuracy, and recall, as well as ROC curves. The successful management of high-throughput traffic with no latency was a demonstration of the solution's scalability in cloud and enterprise environments. Through the implementation of feature importance analysis and visualization methodologies, interpretability was enhanced by the identification of critical traffic characteristics that are linked to attacks. Density analysis of forward (FWD) and backward (BWD) packets was another instrument employed to identify malicious network activity and detect anomalous traffic distributions.

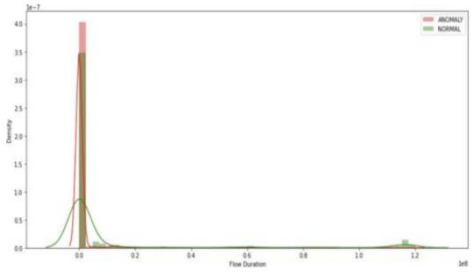

Fig2: Flow Duration

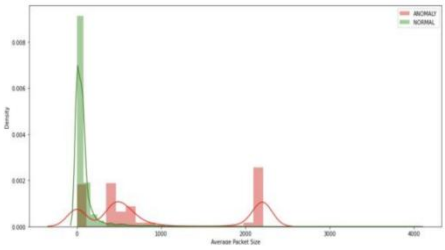

Fig3: Average Packet Size

6. CONCLUSION

The proposed machine learning-based approach provides a practical method for detecting and anticipating Distributed Denial of Service (DDoS) attacks. This model is well-suited for use in actual network security applications due to its accurate results, adaptability to evolving attack patterns, and low false positive rates. Its predictive capabilities enable proactive defense methods by reducing service failures and securing critical systems. The approach is also applicable to large-scale network configurations as a result of its scalability and computational efficacy. This paperunderscores the significance of artificial intelligence in enhancing network security and resilience to modern cyberthreats.

REFERENCES

[1] N. Martins, J. M. Cruz, T. Cruz, and P. H. Abreu, ``Adversarial machine learning applied to intrusion and malware scenarios: A systematic review," IEEE Access, vol. 8, pp. 35403_35419, 2020.

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/24
DOI: https://doi.org/10.5281/zenodo.20604260

[2] G. Karatas, O. Demir, and O. K. Sahingoz, ``Increasing the performance of machine learning-based IDSs on an imbalanced and upto-date dataset," IEEE Access, vol. 8, pp. 32150_32162, 2020.
[3] T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, ``BAT: Deep learning methods on network intrusion detection using NSL-KDD dataset," IEEE Access, vol. 8, pp. 29575_29585, 2020. [4] H. Jiang, Z. He, G. Ye, and H. Zhang, ``Network intrusion detection based on PSOxgboost model," IEEE Access, vol. 8, pp. 58392_58401, 2020.
[5] A. Nagaraja, U. Boregowda, K. Khatatneh, R. Vangipuram, R. Nuvvusetty, and V. S. Kiran, ``Similarity based feature transformation for network anomaly detection," IEEE Access, vol. 8, pp. 39184_39196, 2020.
[6] L. D'hooge, T. Wauters, B. Volckaert, and F. De Turck, ``Classi_cation hardness for supervised learners on 20 years of intrusion detection data," IEEE Access, vol. 7, pp. 167455_167469, 2019.
[7] X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, ``An adaptive ensemble machine learning model for intrusion detection," IEEE Access, vol. 7, pp. 82512_82521, 2019.
[8] Y. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, ``Network intrusion detection based on supervised adversarial variational autoencoder with regularization," IEEE Access, vol. 8, pp. 42169_42184, 2020.
[9] C. Liu, Y. Liu, Y. Yan, and J. Wang, ``An intrusion detection model with hierarchical attention mechanism," IEEE Access, vol. 8, pp. 67542_67554, 2020.
[10] S. U. Jan, S. Ahmed, V. Shakhov, and I. Koo, ``Toward a lightweight intrusion detection system for the Internet of Things," IEEE Access, vol. 7, pp. 42450_42471, 2019.
[11] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, ``Machine learning-based network vulnerability analysis of industrial Internet of Things," IEEE Internet Things J., vol. 6, no. 4, pp. 6822_6834, Aug. 2019.
[12] Y. Chen, B. Pang, G. Shao, G. Wen, and X. Chen, ``DGA-based botnet detection toward imbalanced multiclass learning," Tsinghua Sci. Technol., vol. 26, no. 4, pp. 387_402, Aug. 2021.