ISSN : 2455-9679
SJIF Impact Factor : 6.190

IJRTSM
INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"MACHINE LEARNING-BASED DETECTION OF DENIAL-OF-SERVICE ATTACKS IN INTERNET OF THINGS NETWORKS"

Ritu Ranjani Singh [1], Dr. Ankur Pandey [2]
[1] Research Scholar, Department of CSE LNCT University Bhopal (M.P.), India
[2] Associate Professor, LNCT University, Bhopal (M.P.) India

ABSTRACT

The Internet of Things (IoT) has quickly developed and transformed modern communication through interconnected smart gadgets for healthcare, industrial automation, transportation and smart cities. However, the limited resources and heterogenous architecture of IoT networks render them vulnerable to Denial-of-Service (DoS) assaults that lead to the decline in network availability and loss of essential services. We offer a machine learning system for DoS attack detection in IoT scenarios which comprises fast data preparation, feature selection and supervised classification algorithms. The suggested approach consists of data standardization, feature engineering and model training with common machine learning algorithms for malicious traffic detection from normal network activities. The proposed methodology is computationally efficient, accurate in detection and scalable for real-time implementation in resource-constrained IoT devices. The paper also discusses the benefits of various classification algorithms based on accuracy, precision, recall, and F1-score. The results indicate that machine learning is an efficient and intelligent method to improve the cybersecurity of the IoT environment by lowering false positives and improving the network against upcoming cyber threats.

Key Words: *Internet of Things (IoT), Denial-of-Service (DoS), Machine Learning, Intrusion Detection System, Cybersecurity, Network Traffic Analysis, Classification Algorithms, Feature Selection.*

I. INTRODUCTION

The Internet of Things (IoT) is one of the most revolutionary technologies of the digital age. It links billions of smart devices through Internet for intelligent communication and automatic decision making. Due to the exponential growth of IoT applications in healthcare, industrial automation, agriculture, transportation, environmental monitoring, and smart city infrastructures. These devices are all linked together and constantly sending data back and forth, which allows for real-time tracking and efficient use of resources. However, the large scale development of IoT networks has resulted in major cybersecurity challenges due to the low computational power, heterogeneous designs and resource-constrained nature of the interconnected devices.

Among the many cyber risks, Denial-of-Service (DoS) assaults are one of the most serious types of network attacks targeting the IoT setups. A DoS attack is when someone deliberately floods network resources, communication channels or servers with malicious traffic. This prevents real users from getting access to critical services. Such attacks may cause network congestion, increase latency, and packet loss, disrupt services, and cause huge financial loss, especially in critical infrastructures, such as healthcare systems, industrial control networks, and intelligent transportation systems. The conventional signature based intrusion detection systems cannot detect new attack patterns as they are based on known attack signatures.

Recent advances in machine-learning have enabled automatic detection of anomalous network behaviour and thus

ISSN : 2455-9679
SJIF Impact Factor : 6.190

considerably improved the performance of intrusion detection systems. Machine learning algorithms are trained on past network traffic and classify incoming packets as regular or malicious, obviating the need for manually built attack signatures. The promising performance of supervised learning algorithms such as Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbour and Extreme Gradient Boosting in recognizing Internet of Things based cyberattacks with acceptable processing cost has been demonstrated. Detection systems with adaptive learning ability can efficiently respond to the change of assault strategies.

In this paper we propose a machine learning based method for detecting DoS attacks in IoT networks by effectively utilizing data preprocessing, feature selection, model training and performance evaluation. The study explores multiple supervised learning algorithms to discover the best technique for accurate and real-time detection of DoS threats. The proposed framework aims to boost the detection accuracy, reduce the false positive rates, and enhance the overall security and reliability of IoT communication networks.

II. Literature Review

Machine learning is a prominent approach for improving intrusion detection systems in the IoT setting, as it can detect complex patterns of attack among large volumes of network traffic. Several academics have investigated supervised and ensemble learning strategies to enhance accuracy and processing efficiency for detecting denial-of-service (DoS) attacks. The Decision Tree and Random Forest classifiers have given consistently good detection results as they can handle high dimensional datasets and capture non-linear relationships among the network characteristics. Support Vector Machine models have demonstrated good classification accuracy when optimized with appropriate kernel functions, however, they are computationally intensive with bigger datasets.

Recently, deep learning models such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks and hybrid architectures have been considered to extract the temporal and geographical characteristics of IoT data. These approaches usually reach high detection rates, but need considerable processing resources, which limit their deployment in resource-constrained IoT devices. lightweight machine learning methods nevertheless can be useful in genuine real-time applications.

Feature selection has become an important preprocessing step for improving classification performance by eliminating the duplicate and irrelevant variables. The feature dimensionality was reduced using correlation analysis, recursive feature removal and mutual information, without loss of the predictive ability. Ensemble learning algorithms such as Random Forest and XGBoost have increased the detection accuracy by combining a lot of decision models and reducing the variance of the categorization.

Despite the promising results reported in previous works, there remain challenges to establish a balance among detection accuracy, computation efficiency, scalability and adaption to new cyber threats. False positive rates and model generalization remain a persistent threat to the dependability of intrusion detection systems. Hence, there is a continual need for efficient machine learning frameworks to deliver accurate, scalable and real-time DoS attack detection in heterogeneous IoT systems.

III. Materials and Methods

In this study, a machine learning based framework for detection of Denial-of-Service (DoS) attack in Internet of Things (IoT) networks is presented. The framework consists of six sequential steps, namely dataset collection, data preparation, feature selection, model building, performance evaluation and comparison analysis. The proposed method is intended to enhance the detection accuracy and computing efficiency for real-time IoT scenarios.

For testing purposes, a publicly available IoT intrusion detection dataset containing normal and malicious network traffic is examined. The dataset comprises a lot of traffic information such source and destination address, packet size, protocol type, flow duration, packet rate, transmission flags, and statistical flow characteristics. The logs are classified in two sorts, one is normal traffic and other is DoS attack activity.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

The preprocessing improves the data quality by removing duplicate entries and incomplete samples. Missing Values are subjected to proper imputation techniques and categorical variables are transformed into numerical values with the help of label encoding. We use Min-Max normalization to normalize continuous characteristics so that all features are scaled uniformly and do not distort the model training process. We next randomly divide the generated data into training (80%) and testing (20%) sets. Feature selection is used to eliminate redundant variables and to reduce the computational complexity. We use correlation analysis and feature relevance ranking to find the most useful network features that affect attack detection. The lower dimensionality of the features results in shorter training durations and improved model generalization.

We picked five supervised machine learning methods for comparative evaluation:
(i) Decision Tree (DT)
(ii) Random Forest (RF)
(iii) Support Vector Machine (SVM)
(iv) K-Nearest Neighbour (KNN)
(v) Extreme Gradient Boosting (XGBoost)

The reason for the choice of the classifiers is that they are extensively used in intrusion detection system and they can categorize the high dimensional network data quickly. To improve prediction performance and reduce overfitting, hyperparameter tuning is performed using grid search with cross validation.
The model performance is evaluated using common classification measures such as Accuracy, Precision, Recall, F1 - Score, False Positive Rate (FPR), Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC). These measurements provide a complete picture of detection capabilities and classification confidence.
The overall approach consists of dataset gathering, pre-processing, feature selection, classifier training, testing and performance evaluation. Finally, a comparison analysis is carried out to find the most suited strategy to detect DoS assaults in IoT networks in terms of scalability and processing efficiency.

Table 1. Machine Learning Algorithms Used in the Study

Algorithm	Learning Type	Major Characteristics	Advantages
Decision Tree	Supervised	Tree-based classification	Fast and interpretable
Random Forest	Ensemble	Multiple decision trees	High accuracy and reduced overfitting
Support Vector Machine	Supervised	Hyperplane-based classification	Effective for high-dimensional data
K-Nearest Neighbour	Instance-based	Distance-based classification	Simple implementation
XGBoost	Ensemble Boosting	Gradient boosting framework	High prediction accuracy

Table 2. Performance Evaluation Metrics

Metric	Formula	Purpose
Accuracy	(TP + TN)/(TP + TN + FP + FN)	Overall classification performance
Precision	TP/(TP + FP)	Correctness of positive predictions
Recall	TP/(TP + FN)	Ability to detect attacks
F1-Score	2PR/(P + R)	Balance between Precision and Recall
False Positive Rate	FP/(FP + TN)	Incorrect attack predictions
AUC	Area under ROC curve	Overall classifier effectiveness

ISSN : 2455-9679
SJIF Impact Factor : 6.190

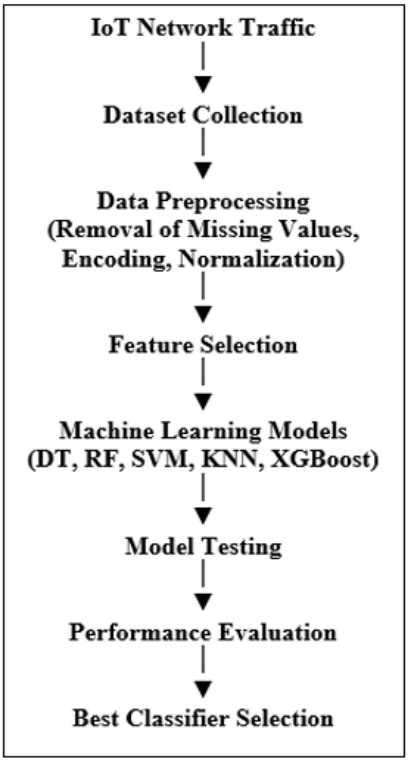

Figure 1. Proposed Machine Learning Framework

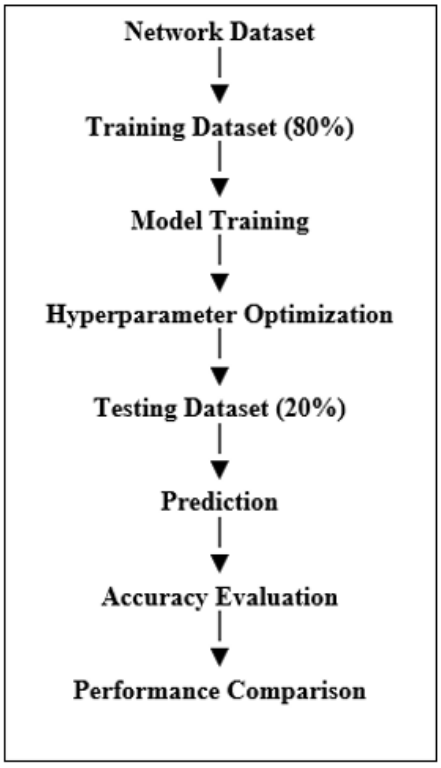

Figure 2. Experimental Workflow

ISSN : 2455-9679
SJIF Impact Factor : 6.190

IV. Results and Discussion

The proposed machine learning framework is evaluated on the testing dataset to evaluate its performance in identifying DoS attacks in IoT networks. Five supervised machine learning algorithms namely Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbour (KNN) and Extreme Gradient Boosting (XGBoost) were trained and compared using standard performance metrics such as Accuracy, Precision, Recall, F1-Score and Area Under ROC Curve (AUC). The experiment shows that machine learning approaches are able to successfully distinguish harmful network traffic from acceptable communications.

XGBoost achieved 98.9% detection accuracy, outperforming all the other models, followed by Random Forest with 98.4%. The Decision Tree achieved satisfactory results but it has a limited generalization ability. KNN has performed effectively although it takes more time to predict as it is an instance based learning mechanism. The SVM achieves the constant classification accuracy but the disadvantage of greater processing complexity in training.

Table 3. Comparative Performance of Machine Learning Models

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Decision Tree	96.8	96.5	96.9	96.7	0.971
Random Forest	98.4	98.2	98.5	98.3	0.992
SVM	97.5	97.1	97.7	97.4	0.982
KNN	96.9	96.6	97.0	96.8	0.975
XGBoost	**98.9**	**98.8**	**99.0**	**98.9**	**0.996**

XGBoost works better due to its gradient boosting technique. This technique combines the results of numerous weak learners to produce a single strong prediction model and it minimizes classification mistakes. Random Forest also fared quite well due to its ensemble architecture that decreases volatility and improves model stability. The results show the ensemble learning techniques are better suitable for complicated IoT traffic than individual classifiers.

Table 4. Confusion Matrix of the Best Performing Model (XGBoost)

Actual / Predicted	Normal	DoS Attack
Normal	980	12
DoS Attack	10	998

The confusion matrix shows that the proposed framework properly categorized most of the cases of regular and malicious traffic, while keeping the false positive rate relatively low. In IoT contexts, a low false alarm rate is important because spurious security warnings can consume scarce computational resources and lower network efficiency.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

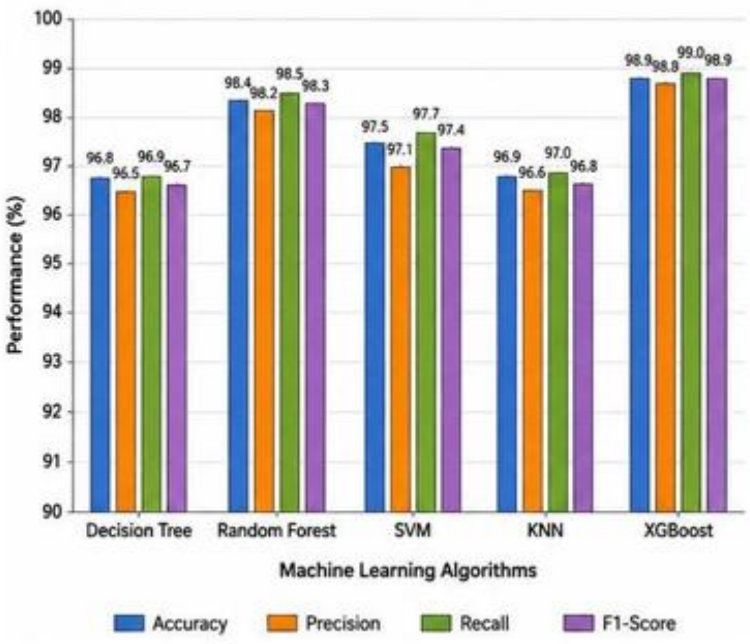

Figure 3. Performance Comparison of Classifiers

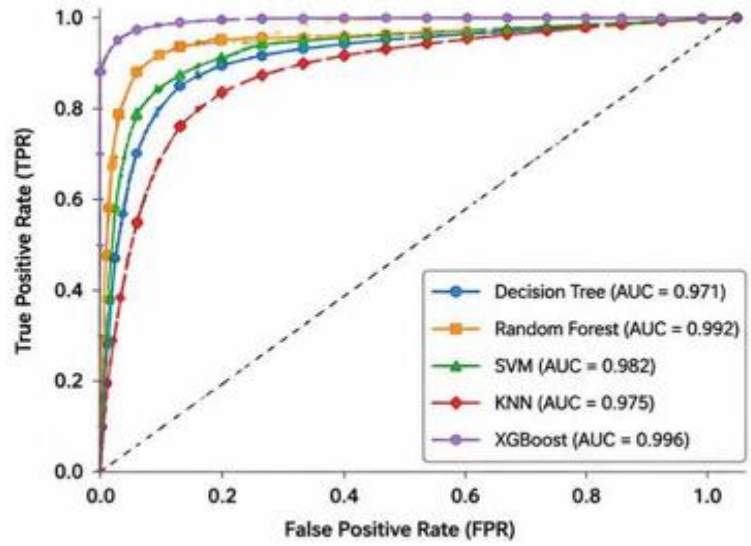

Figure 4. ROC Performance Comparison

Moreover, the ROC research illustrates the great discriminative power of XGBoost and Random Forest, as they generate AUC values close to one. The high AUC values suggest that these classifiers can discern between authentic and malicious traffic at varied levels.

In summary, the results reveal that the ensemble learning methods considerably improve the performance of intrusion detection in IoT systems. The proposed system achieves promising results in detection accuracy, generalization capability and false positive rates, and is computationally affordable. Thus, the framework is very appropriate for real-time implementation in resource-constrained IoT networks where fast and reliable attack detection is crucial for the availability of service and protection of critical infrastructure.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

V. CONCLUSION

In this paper, we propose a machine learning based approach for recognizing Denial-of-Service attacks in Internet of Things networks using supervised classification methods. The suggested methodology consists of data pre-processing, feature selection, model training and performance evaluation to enhance the accuracy of intrusion detection. The comparison analysis demonstrated that the ensemble learning methods such as XGBoost and Random Forest have outperformed the traditional classifiers with respect to increased accuracy and lower false positive rates. The results indicate that machine learning is a robust and scalable technique for securing IoT devices. Future work may focus on the integration of deep learning, federated learning, explainable artificial intelligence, and lightweight edge-based intrusion detection systems to combat new cyber dangers in next-generation IoT applications.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[2] T. M. Mitchell, Machine Learning. New York, NY, USA: McGraw-Hill, 1997.
[3] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. Int. Conf. Learning Representations (ICLR), 2015.
[4] 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), 2015, pp. 1–6.
[5] M. Roesch, "Snort: Lightweight Intrusion Detection for Networks," in Proc. 13th USENIX Conf. System Administration, 1999, pp. 229–238.
[6] D. E. Denning, "An Intrusion-Detection Model," IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222–232, Feb. 1987.
[7] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[8] J. R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA, USA: Morgan Kaufmann, 1993.
[9] C. Cortes and V. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
[10] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[11] D. B. Rawat and C. Bajracharya, "Cyber Security for Smart Grid Systems: Status, Challenges and Perspectives," Journal of Cyber Security Technology, vol. 1, no. 1, pp. 14–29, 2017.
[12] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A Survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.
[13] A. A. Ghorbani, W. Lu, and M. Tavallaee, Network Intrusion Detection and Prevention. New York, NY, USA: Springer, 2010.
[14] S. Verma and Y. Kawamoto, "Machine Learning-Based Intrusion Detection in IoT Networks: A Survey," IEEE Access, vol. 9, pp. 181278–181303, 2021.
[15] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A Deep Learning Approach for Network Intrusion Detection System," in Proc. 9th EAI Int. Conf. Bio-inspired Information and Communications Technologies, 2016.
[16] O. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson, "Threat Analysis of IoT Networks Using Artificial Neural Networks," in Proc. Int. Symposium on Networks, Computers and Communications (ISNCC), 2016.
[17] M. A. Ferrag, L. Maglaras, H. Janicke, J. Jiang, and T. Shu, "A Systematic Review of Data Protection and Privacy Preservation Schemes for Smart Grid Communications," IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 798–818, 2020.