IOT Botnet Detection using Machine Learning

Bhagat Inamdar
Asst. Professor, Dept. of CSE KLS VDIT,
Haliyal, Karnataka, India

Deepa Kenchanagoudar
Dept. of CSE, KLS VDIT,
Haliyal

Muriel Pinto
Dept. of CSE, KLS VDIT,
Haliyal

Sahana Malin
Dept. of CSE, KLS VDIT,
Haliyal

Sai Samiksha Budihal
Dept. of CSE, KLS VDIT,
Haliyal

Abstract—**The rapid expansion of interconnected smart devices has enabled widespread automation and connectivity across smart homes, healthcare, manufacturing, and industrial control systems. However, this large-scale connectivity has significantly increased vulnerabilities, leading to a sharp rise in IoT botnet attacks. Compromised devices are remotely controlled and used to perform cyber-crimes such as Distributed Denial of Service (DDoS), unauthorized data access, and aggressive network flooding. Traditional signature-based detection methods fail to identify evolving botnet families due to rapid mutation of attack patterns. This paper introduces a machine learning-based intelligent detection framework capable of identifying IoT botnet behavior through traffic flow analysis. The proposed system preprocesses collected network traffic, extracts relevant features, classifies flows as benign (0) or malicious (1), and triggers an automated alert notification when malicious activity is detected. Experimental evaluation demonstrates highly reliable performance and strong prediction accuracy, proving the suitability of machine learning methods for real-time IoT botnet detection and prevention. Additional discussions include feature selection techniques, model optimization, and deployment considerations for scalable and robust IoT security solutions.**

***Index Terms*—IoT Security, Botnet Detection, Machine Learning, Network Traffic Analysis, Intrusion Detection System, Real-Time Alerting, Flow-Based Detection.**

I. Introduction

The ecosystem of interconnected IoT devices has grown tremendously, with billions of smart devices communicating autonomously through networked environments. These devices span smart homes, industrial control systems, healthcare monitoring, and smart city infrastructures. Despite offering convenience and automation, most IoT devices are developed with minimal built-in security and rely on weak authentication mechanisms, making them extremely vulnerable to cyberattacks. Cybercriminals exploit these security loopholes and convert devices such as cameras, sensors, routers, smart lights, and medical equipment into remotely controlled bots. These botnets are coordinated using command-and-control (C2) servers and are used for launching massive cyber-attacks capable of disrupting essential services, financial systems, and global internet infrastructure. The Mirai botnet attack, which compromised millions of devices worldwide, highlighted the destructive potential of IoT botnets and the urgent need for intelligent security solutions.

Detecting IoT botnet activity is challenging because malicious traffic is often disguised as normal behavior, allowing attackers to bypass traditional intrusion detection systems. Signature-based methods fail to detect zero-day attacks or evolving threats, as they depend on predefined attack patterns. Machine learning provides an effective solution by enabling behavior-based anomaly detection that learns from network traffic characteristics and identifies unusual patterns through data-driven analysis. By analyzing traffic flows, packet statistics, and protocol behavior, ML models can distinguish subtle differences between benign and malicious traffic that are often invisible to conventional systems.

This study aims to address these challenges by proposing a machine learning-based detection framework that improves IoT security through real-time threat identification, binary classification, and automatic alert notification to users. The system emphasizes modularity, allowing future integration of more advanced deep learning techniques, encrypted traffic analysis, and edge-based deployment for faster detection and lower latency.

II. IoT Botnets Overview

A botnet refers to a network of infected IoT devices remotely controlled after malware compromise. Once infected, these devices operate as part of a coordinated system and perform malicious tasks ranging from attacking remote servers to silently exfiltrating sensitive information.

The most common real-world botnet attacks include:

- **Distributed Denial of Service (DDoS):** Overloading servers with massive traffic until services fail. IoT botnets can generate enormous traffic from thousands of devices simultaneously, making mitigation extremely difficult.
- **Port Scanning:** Scanning connected devices to find open ports and exploitable vulnerabilities. This often serves as the initial phase for further compromise and malware propagation.
- **Brute-force Attacks:** Repeatedly trying passwords to gain unauthorized access, exploiting weak authentication on IoT devices.
- **Data Exfiltration:** Stealing sensitive information from compromised devices or networks, including personal user data, device credentials, and corporate information.

Modern IoT botnets use advanced features such as encrypted communication, fast mutation, and decentralized control strategies, making detection more difficult. They can remain dormant for long periods, only activating under specific conditions to evade security measures. Additionally, botnets often exhibit traffic patterns that mimic legitimate IoT communication, requiring more sophisticated detection strategies.

To counter these threats, automated machine learning-based detection frameworks are necessary. By leveraging statistical, temporal, and behavioral traffic features, ML models can detect subtle deviations from normal network activity and classify suspicious behavior even in the presence of obfuscation techniques. This ensures both scalability and robustness for protecting critical IoT infrastructures.

III. Related Work

Several studies have explored techniques to detect and mitigate IoT botnet attacks. Early approaches relied on firewall rules, deep packet inspection, or heuristic anomaly detection. However, these methods struggle with the high volume of IoT traffic and cannot adapt quickly to new and evolving threats. Signature-based detection approaches are particularly limited, as they cannot identify previously unknown attack patterns.

Soe et al. [1] employed a sequential architecture using machine learning to detect IoT botnet activity, demonstrating that behavior-based analysis improves detection over signature-based systems. Su et al. [8] implemented a Random Forest classifier with flow-based traffic features, achieving high accuracy in detecting malicious IoT traffic.

While both studies contribute valuable insights into botnet detection, they have limitations: Soe et al.'s model can be computationally intensive for real-time implementation, and Su et al.'s work does not incorporate automated alert notifications or comprehensive feature selection techniques. Our proposed system addresses these gaps by combining efficient flow-based traffic analysis with multiple supervised ML classifiers and automated alerting.

Additionally, preprocessing techniques such as normalization, categorical encoding, and outlier removal are incorporated to improve model generalization and robustness. Compared to previous work, the proposed system supports real-time detection and is designed for practical deployment in heterogeneous IoT networks.

IV. Proposed System

The proposed IoT botnet detection system classifies network traffic behavior using machine learning algorithms after analyzing raw flow data. The framework consists of multiple stages: data collection, preprocessing, feature extraction, model training, and real-time traffic classification. Automatic alerts are generated when malicious activity is detected, allowing rapid incident mitigation.

A. Description of Dataset

The dataset comprises real IoT network traffic captured from diverse device types under both normal and attack conditions. Each record represents statistical flow attributes derived from bidirectional communication sessions between source and destination nodes. Key features include:

- Packet size distribution and communication flow duration
- Total transmitted and received bytes
- Header lengths and packet counts
- Inbound and outbound traffic statistics
- Protocol identifiers and connection flag indicators
- Temporal features such as inter-arrival times and session durations

The dataset includes diverse attack scenarios, including DDoS, scanning, brute-force attempts, and data exfiltration. Preprocessing involves missing data handling, normalization, feature selection, and train-test splitting. This ensures models learn from clean, representative data, reducing noise and improving generalization.

B. Machine Learning Models Used

Multiple supervised classifiers are implemented, each offering unique advantages for IoT botnet detection:

- **Logistic Regression (LR):** Logistic Regression is a linear classification algorithm that predicts the probability of malicious behavior based on weighted feature contributions. It is easy to interpret, making it useful for understanding which features influence attack detection. Although it forms a strong baseline, its performance reduces when handling complex or nonlinear IoT traffic patterns.
- **Random Forest (RF):** Random Forest is an ensemble method that constructs multiple decision trees and aggregates their outputs for improved accuracy. It effectively captures nonlinear relationships between features and handles noisy IoT traffic with robustness. Its ability to learn complex attack behaviors makes it the highest-performing model in this study, with excellent generalization and reduced overfitting.
- **Support Vector Machine (SVM):** SVM identifies an optimal hyperplane to separate benign and malicious traffic. Using kernel functions, it can model complex boundaries in high-dimensional IoT datasets. SVM excels in detecting subtle anomalies but can be computationally demanding for large datasets, making real-time deployment more challenging compared to lighter models.
- **Decision Tree (DT):** Decision Trees classify traffic based on hierarchical if-else conditions using feature thresholds. They are highly interpretable, allowing clear visualization of decision paths. While effective for quick classification, they tend to overfit on noisy IoT traffic. This limits their standalone performance, especially when detecting sophisticated or evolving botnet attacks.
- **K-Nearest Neighbor (KNN):** KNN is an instance-based model that classifies traffic by examining the closest neighbors in the feature space. It works well when IoT traffic forms natural clusters but becomes computationally expensive on large datasets. Its sensitivity to feature scaling and noise makes it less suitable for real-time botnet detection despite its simplicity.
- **Naive Bayes (NB):** Naive Bayes is a probabilistic classifier that assumes independence among features and

predicts attacks using conditional probability. It trains extremely fast and performs effectively on high-dimensional IoT traffic. However, its simplifying assumptions may reduce classification accuracy when actual feature correlations are strong, limiting its performance in complex attack scenarios.

Random Forest emerged as the most effective model in this study due to its ensemble-based architecture, which combines multiple decision trees to capture complex relationships among features in IoT network traffic. This multi-tree mechanism allows the model to handle highly variable and evolving attack patterns, providing robustness against noisy, incomplete, or imbalanced datasets. By aggregating the predictions of numerous individual trees, Random Forest achieves greater stability and reduces the risk of overfitting, ensuring consistent performance across diverse attack scenarios. Furthermore, careful hyperparameter tuning—such as optimizing the number of trees, maximum depth, and minimum samples per leaf—combined with cross-validation techniques, enhanced the model's ability to generalize and improved its detection accuracy. Compared to other supervised learning approaches used in this study, Random Forest consistently demonstrated superior reliability, precision, and effectiveness in identifying IoT botnet attacks, making it a highly suitable choice for real-time intrusion detection systems.

C. System Architecture and Flow Representation

1) **Data Collection:** Capturing real-time traffic flows from IoT devices connected within the network. Traffic is collected in the form of bidirectional flow records, containing packet statistics, timestamps, protocol usage, and communication patterns that serve as the primary input for analysis.
2) **Preprocessing:** Cleaning, normalization, and encoding of features to ensure consistent data quality. This stage removes duplicates, handles missing values, scales numerical attributes, and converts categorical protocol labels into machine-readable formats for improved learning performance.
3) **Feature Extraction:** Selecting relevant flow-based patterns by analyzing statistical attributes such as packet length, inter-arrival times, byte counts, and connection flags. This step enhances model accuracy by eliminating redundant features and emphasizing discriminative characteristics of malicious behavior.
4) **Model Training:** Training and validating multiple classifiers using labeled traffic data. The models learn underlying behavioral differences between benign and malicious flows, applying techniques such as cross-validation, hyperparameter tuning, and performance optimization to achieve reliable detection.
5) **Real-Time Classification:** Predicting benign or malicious traffic using the trained model under live network conditions. Incoming flows are processed instantly, and classification results are generated with minimal latency, enabling continuous monitoring and proactive detection of potential IoT attacks.
6) **Alert Notification:** Automatic alerts and logging of detected attacks, providing immediate warnings to administrators via system logs or email notifications. This ensures rapid response to malicious activities, allowing timely mitigation and detailed forensic analysis of compromised devices.

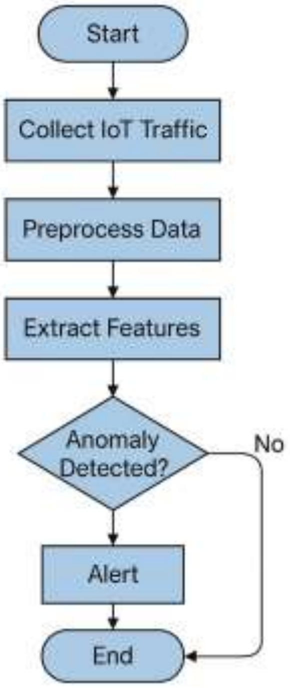

Fig. 1. Overall System Flowchart for IoT Botnet Detection

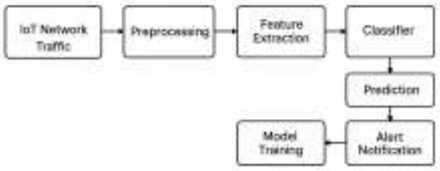

Fig. 2. Block Diagram of System Architecture

D. Performance Evaluation and Classification Results

Models were evaluated using accuracy, precision, recall, F1-score, and confidence scores. Random Forest achieved:

- Accuracy: 98.76%
- Precision: 98.12%
- Recall: 97.45%
- F1-score: 97.70%
- Confidence Score: 99.02%

Comparisons with other classifiers show that ensemble models outperform simpler approaches due to robust feature learning. Cross-validation and confusion matrices demonstrate minimal false positives and high detection reliability.

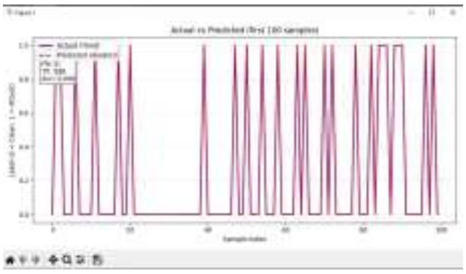

Fig. 3. Accuracy Comparison - Actual vs Predicted

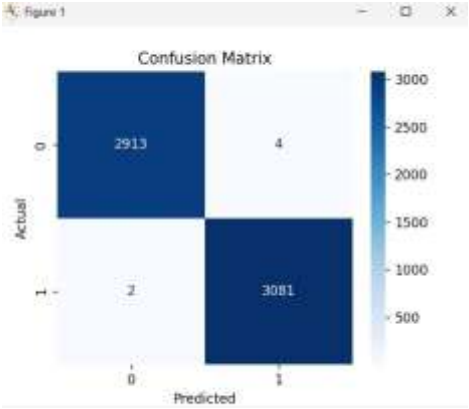

Fig. 4. Confusion Matrix for Model Prediction

V. Conclusion

The proposed IoT botnet detection framework successfully demonstrates the capability of machine learning-based models to identify malicious traffic patterns in IoT environments with high accuracy and reliability. By utilizing flow-based statistical features and multiple supervised classifiers, the system exhibits strong performance in detecting a wide range of botnet behaviors such as DDoS, port scanning, brute-force attempts, and abnormal communication bursts.

A key strength of our model is its integration of an automated alert notification mechanism. When malicious activity is detected, the system triggers an immediate email alert to the registered user or network administrator. This real-time notification capability significantly reduces incident response time, enabling quicker mitigation before the attack spreads or causes large-scale damage. The alert system also logs the attack details, helping in forensics and further analysis.

Additionally, the system is modular, lightweight, and designed for practical deployment in real-world IoT networks. It can operate on gateways, local servers, and even cloud-based platforms without requiring high computational resources. Through effective preprocessing, feature selection, and model optimization, the framework achieves efficient runtime performance while maintaining high detection accuracy.

Overall, this work highlights the effectiveness of combining flow-based traffic analysis with machine learning to create a proactive IoT security solution. The automated email alerting, binary classification design, and scalable architecture make this system suitable for homes, enterprises, and industrial IoT networks seeking stronger cyber defense. This research forms a foundation for future integration of advanced deep learning models, encrypted traffic analysis, and federated learning techniques to further enhance IoT botnet detection and overall cybersecurity.

Fig. 5. Real-Time Classification Output Showing Detected Botnet Attacks and Benign Traffic

This figure displays the real-time classification results generated by the deep learning model. Each network traffic instance is labeled as either benign or a specific botnet attack, demonstrating how the system processes live IoT traffic and assigns accurate predictions instantly.

VI. Future Enhancements

Although the proposed system performs effectively in detecting IoT botnet activity and generating real-time email alerts, several enhancements can be incorporated to increase scalability, robustness, and adaptability in future implementations. The following improvements are suggested:

- **Integration of Advanced Deep Learning Models:** Future work may involve implementing deep learning architectures such as CNN, LSTM, GRU, or hybrid CNN-LSTM models to capture complex spatial and temporal dependencies in network traffic flows. These models can improve detection of sophisticated botnet families that use stealthy or multi-stage attack patterns.
- **Edge-Level Deployment for Low-Latency Detection:** Deploying the detection engine on edge devices such as routers, gateways, or IoT hubs can significantly reduce detection latency. Local processing ensures faster response and eliminates the need to send large volumes of traffic data to central cloud servers.
- **Federated Learning for Privacy-Preserving Detection:** Instead of collecting data in a centralized location, federated learning can be used to train models across multiple distributed IoT nodes without exposing raw data. This ensures better privacy, compliance with data protection regulations, and collaborative learning across environments.
- **Adaptation to Encrypted Traffic Analysis:** Modern botnets increasingly use traffic encryption to evade signature-

based tools. Future systems may incorporate techniques such as TLS fingerprinting, metadata-based analysis, and encrypted flow behavior modeling to detect threats with- out decrypting packets.

- **Reinforcement Learning for Autonomous Defense:** Incorporating reinforcement learning may enable the sys- tem to automatically learn optimal mitigation strategies such as throttling malicious traffic, isolating compromised devices, or applying dynamic firewall rules.
- **Enhanced Visual Dashboards and Monitoring Tools:** A real-time dashboard can be implemented to visualize traffic flow statistics, model predictions, detected attacks, and alert logs. Interactive visual analytics will help network administrators understand threat patterns more intuitively.

These enhancements would significantly strengthen the system's accuracy, adaptability, and suitability for real-world deployment across diverse IoT environments, including smart homes, enterprises, smart cities, and industrial automation networks.

References

[1] Y. N. Soe et al., “Machine Learning-Based IoT-Botnet Attack Detection with Sequential Architecture,” *Sensors*, vol. 20, no. 16, 2020.

[2] S. M. K. Quadri et al., “Review of Botnet Attack Detection in SDN-enabled IoT Using Machine Learning,” *PeerJ Computer Science*, 2021.

[3] Y. Meidan et al., “Flow-based Identification of IoT Botnet Attacks Using Deep Learning,” *arXiv:1805.03409*, 2018.

[4] A. Naeem et al., “Efficient IoT Intrusion Detection with CNN-BiLSTM Architecture,” *arXiv:2503.19339*, 2025.

[5] R. Ullah et al., “Attack Categorization in IoT Networks Using Machine Learning,” *arXiv:2101.12270*, 2021.

[6] M. Ferrag et al., “Deep Learning-Based Intrusion Detection: A Comparative Study,” *Journal of Information Security and Applications*, 2020.

[7] S. Althubiti et al., “Detecting IoT Botnet Attacks Using Machine Learning,” *Security and Communication Networks*, 2020.

[8] J. Su et al., “Random Forest-Based IoT Botnet Detection Using Traffic Features,” *IEEE Access*, 2020.

[9] P. Mishra et al., “Machine Learning for IoT Security: A Survey,” *IEEE Internet of Things Journal*, 2022.

[10] H. HaddadPajouh et al., “RNN-Based IoT Malware Threat Hunting,” *Future Generation Computer Systems*, 2021.