

Topic

IoT Security and Machine Learnin

Student : Abdul Moein Ayman Al-Ali

Course Title: Research Methodology in Computer Science

Dr. Mohammed Anbar

1. Research Title

"A Hybrid Feature Selection Framework for Real-Time Intrusion Detection in Resource-Constrained Internet of Things (IoT) Environments."

2.Problem Statement (Detailed Formulation)

Identification of the Problem Domain 2.1

The domain of this research is Network Security specifically within the Internet of Things (IoT) ecosystem. The IoT domain is characterized by a massive interconnection of heterogeneous devices that communicate via diverse protocols (e.g., MQTT, COAP, Zigbee). These devices are increasingly integrated into critical infrastructures, including smart grids, healthcare systems, and industrial automation (Industry 4.0)

Description of the Problem 2.2

The core problem is the inefficiency of existing Intrusion Detection Systems (IDS) when applied to the unique constraints of IoT networks. Modern cyber-threats, such as the Mirai botnet and Distributed Denial of Service (DDoS) attacks, have become more sophisticated. According to Anbar et al. (2021), IoT devices often lack the computational power (CPU) and memory (RAM) required to run complex deep learning models

Moreover, network traffic data in IoT is often "high-dimensional," meaning it contains dozens of features (e.g., packet size, duration, flags). A large portion of these features is either redundant or irrelevant. Processing this unnecessary data leads to

1.Increased Latency: Delayed detection of attacks which can be fatal in real-time systems

High False Positive Rates (FPR): .2

Incorrectly identifying legitimate traffic as malicious, causing system .disruptions

Significance of the Research 2.3

The significance of addressing this problem is twofold. Technically, it advances the capability of lightweight security mechanisms to operate at the "Edge" of the network rather than relying on the cloud. Socially and Economically, it protects user privacy and prevents financial losses. As stated by Thamilarasu and Chawla (2019), a single breach in a smart healthcare IoT system could lead to life-threatening situations, making the development of an optimized, fast-acting IDS a matter of critical necessity

Existing Solutions and Their 2.4

Limitations

Current solutions primarily involve traditional Machine Learning (ML) algorithms like Support Vector Machines (SVM) or K-Nearest Neighbors (KNN). While these models are effective in controlled environments, they struggle with the dynamic and evolving nature of IoT traffic. Recent research has proposed using Principal Component Analysis (PCA) for dimensionality reduction. However, PCA often loses the semantic meaning of features, which can reduce detection accuracy in specific attack types like "Botnet" or "Injection" (Al-Mhigani et al., 2020). There is a lack of a "hybrid" approach that combines filter-based and

Wrapper-based selection to balance speed .and precision

Justification for Research 2.5

This research is justified by the urgent need for a specialized feature selection framework that identifies the "minimal subset" of traffic features required for high-accuracy detection. By reducing the feature space, we can significantly decrease the computational burden on IoT nodes. This study will fill the gap by proposing a hybrid method that ensures the IDS remains both proactive (detecting threats early) and efficient (consuming minimal battery and processing power)

3.Research Objectives.

Primary Goal 3.1

The primary goal of this research is to propose and evaluate a hybrid feature selection framework that enhances the performance of Machine Learning-based Intrusion Detection Systems in IoT environments

Specific Objectives 3.2

To achieve the primary goal, the following objectives will be pursued

- 1.To analyze and characterize the traffic patterns of common IoT-specific attacks (such as DoS, Scanning, and Probing) using the BoT-IoT and UNSW-NB15 datasets
- 2.To evaluate the effectiveness of various filter-based feature selection techniques (e.g., Information Gain, Chi-Square) in identifying the most influential network parameters
- 3.To design and implement a hybrid feature selection algorithm that combines the speed of filter methods with the accuracy of wrapper methods (e.g., Recursive Feature Elimination)

4.To develop a lightweight Machine Learning model (using Random Forest or LightGBM) integrated with the proposed feature selection .framework

5.To validate and compare the proposed framework against existing state-of-the-art models in terms of Accuracy, Detection Rate, False Positive Rate, and Computational .Time

4.Literature Review (Integrated Summary)

The Literature Review (LR) conducted in Assignment 1 established that the security of IoT is a multi-layered challenge. Anbar et al. (2021) highlighted that the majority of current IDS are designed for traditional networks and fail to account for the "low-power" nature of IoT

Furthermore, Al-Mhigani et al. (2020) provided a taxonomy of IoT attacks, concluding that "Anomaly-based detection" is superior to "Signature-based detection" because it can identify previously unknown (zero-day) attacks. However, they noted that the high dimensionality of data remains the biggest .bottleneck for ML deployment

Research by Thamilarasu and Chawla (2019) proposed a deep learning-based IDS, which achieved 95% accuracy.

However, their model's inference time was too high for real-time applications on small sensors. This research builds upon these findings by focusing on the optimization phase (Feature Selection) to maintain that high accuracy while significantly reducing the processing time, directly addressing .the limitations identified in previous studies

5.References.

Al-Mhiqani, M. N., et al. (2020). "Cyber-security incidents: A review cases in cyber-physical systems." International Journal of Advanced .Computer Science and Applications

Anbar, M., et al. (2021). "A Review of HTTP Flood DoS Detection Approaches Based on Machine .Learning." IEEE Access

Thamilarasu, G., & Chawla, S. (2019).

"Towards Deep-Learning-Driven Intrusion Detection for the Internet of .Things." Sensors