Project Baseline Report

Tyler Heslop, Alma Nkemla, Adrian Rosales

Project Title: Detecting malicious network traffic in IoT devices using machine learning.

Problem Statement: As the number of Internet of Things connected devices grows, so does the exposure to cyber threats. This exposure typically happens in the background of the network and is difficult for the average user to detect, identify and protect against. Many IoT devices are lightweight or barebone linux devices and lack robust security controls. These types of devices are becoming frequent targets for attackers to use these devices for malicious activities like DDos, backdoors, ransomware, and reconnaissance.

Hypothesis: Malicious traffic in IoT networks exhibits statistically detectable patterns that can be classified using machine learning techniques.

Proposed datasets:

- Primary dataset: TON_IoT (https://research.unsw.edu.au/projects/toniot-datasets)
    - Provided by UNSW Canberra Cyber, Australian Centre for Cyber Security
    - Dataset is CSV-formatted network logs, each row is a network flow or session with about 45 features, including:
        - Source and destination IP addresses and ports
        - Packet count, byte count, flow duration
        - Protocol (HTTP, HTTPS, TCP, MQTT)
        - Label of the traffic type (benign, DDoS, DoS, Backdoor, Injection, Ransomware)
    - Why is this data suitable?
        - Labeled for both binary and multi-class classification
        - Diverse attack types and protocols to reflect real-world IoT traffic
        - Scalable from simple ML (random forest) to deep learning (LSTM)
        - Ready to use CSV format
- Secondary dataset: Bot-IoT Dataset (https://research.unsw.edu.au/projects/bot-iot-dataset)
- Secondary dataset: UNSW-NB15 Dataset (https://research.unsw.edu.au/projects/unsw-nb15-dataset)
- Secondary dataset UNB CIC-IDS2017 (https://www.unb.ca/cic/datasets/ids-2017.html)

Potential Challenges:

- Imbalanced data: some attacks are rare, can be mitigated using class weights
- Multi-protocol noise: different devices using different protocols in the same dataset, can be mitigated by focusing on a single protocol

- Overfitting to specific environments: Models may perform well on training data but poorly on new networks due to overfitting to the dataset's environment
- Feature scaling: network features are on different scales, can be mitigated by normalizing or using tree-based models that are robust to scaling
- High data volume and processing overhead: Network captures, especially PCAPs, can be large and slow to process

Summary:
Essentially this project will test the hypothesis that malicious traffic from IoT devices exhibit identifiable patterns that can be detected using machine learning. Using the publicly available TON_IoT dataset provided by UNSW, we will train supervised classifiers to distinguish between benign and malicious network traffic. This project will also explore anomaly detection methods to evaluate their ability to detect unseen threats. Model performance will be evaluated using standard precision, recall, F1-score, as well as live network test to confirm functionality. This work contributes to the cybersecurity field by demonstrating how ML techniques can be applied to network data for proactive network defense and monitoring.