Classification of IoT Devices by Analysing Network Traffic Characteristics Using Deep Learning Techniques

Problem Statement

IoT is an extension of a typical computer based Internet model to a geographically distributed, heterogeneous and constrained model of connected things. The “things” in Internet Of Things can be computers, sensors, actuators and processes. The usage of IoT devices have increased exponentially over the years. This poses a critical problem. It can be difficult for the operators in the smart environments to figure out if the IoT devices are functioning properly and determine which of the IoT devices are connected to the network. Therefore, acquiring ”visibility” into the IoT devices is of utmost importance. In this research, a device classification architecture is proposed that can detect and classify IoT devices based on the characteristics of their network traffic. The deep learning sequential model will use for the classification of IoT devices by using their network characteristics. The results produced by this architecture will helpful for the operator of the smart home environment in order to managing IoT devices by monitoring them and restricts the malicious activities through them.

Background

IoT allows objects to be controlled remotely across existing network infrastructure. It also possesses an autonomous control feature through which any device can be controlled without direct human interaction. Internet of Things has the potential to encompass and instrument an enormous range of connected devices including home appliances and utilities, wearables, buildings, industrial processes, medical devices, law-enforcement devices, military equipment, and other connected applications that today might be barely imaginable. Statistics show that it will exponentially increase to about 50 billion by 2030. The research in this area has gained momentum and is supported by the collaborative efforts from academia, industry, and standardization bodies in several communities such as telecommunication, semantic Web, and informatics. Management of these connected devices has increasingly become challenging, particularly from cyber attacks. It is surprisingly easy to hack a centralised controller connected to an IoT to get the access of complete system.

Methodology

Step 1: Data collection and analysis

The IoT traffic on the network was collected and stored in the form of PCAP files.

Step 2: Data Preprocessing and feature extraction

This step involves separating the IoT traffic from whole network traffic and then extracting the features related to flow, packet and behaviour from PCAP files by using network programming languages or python tools such as scapy.

Step 3: Training and experimentation on datasets

In this step, train the LSTM- a deep learning based sequential classifier for accurately prediction of IoT devices.

Step 4: Deployment and analysis on real life scenario

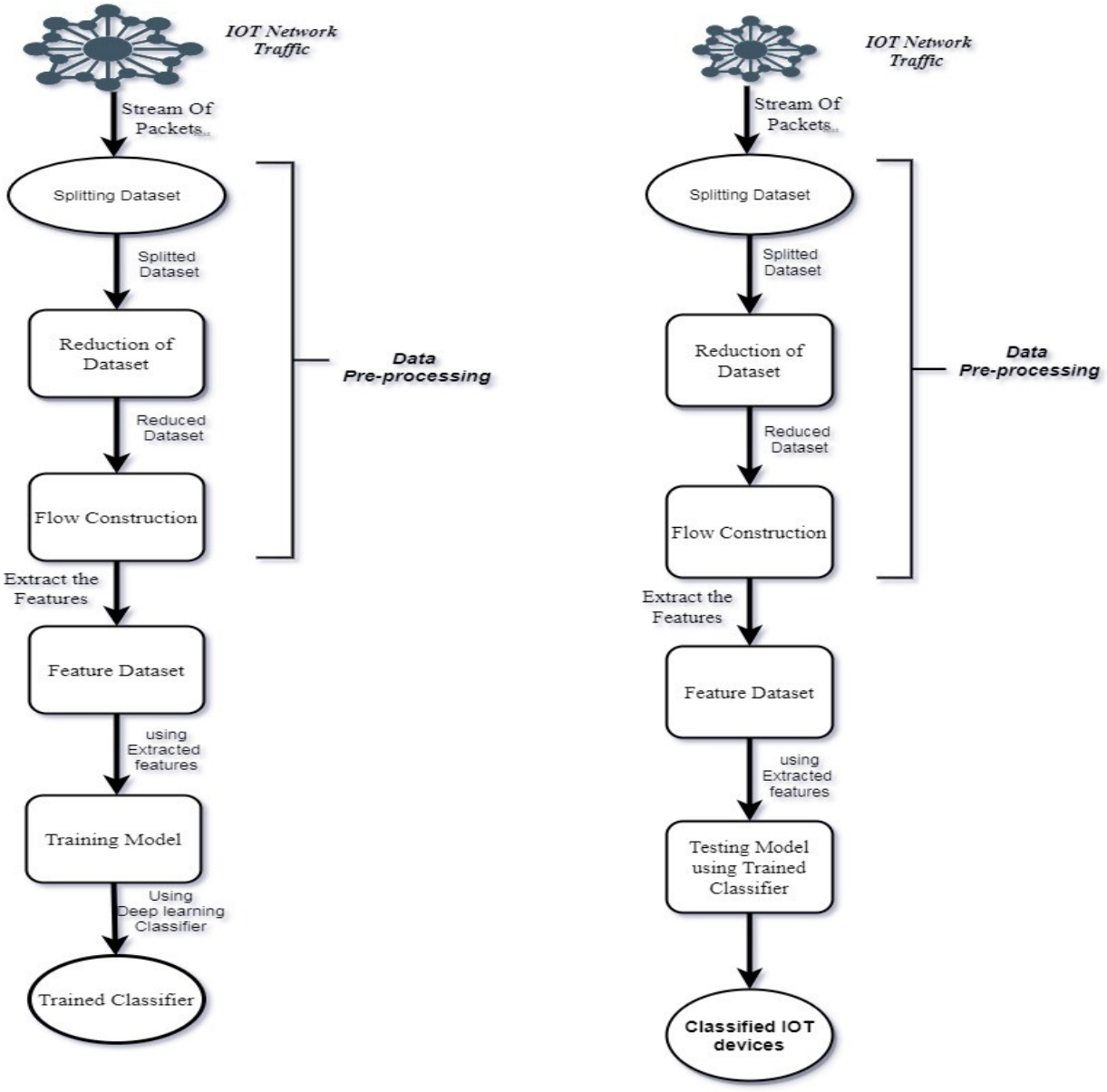

Figure 1: Training and Testing of IoT Traffic

Figure 1 shows the proposed methodology for the classification of IoT devices by analysing the network traffic characteristics.

The training and testing of IoT traffic data over the LSTM- deep learning based classifier and achieved good results for prediction of IoT devices. The real life scenarios leveraged the further improvements in the methodology used.

Experimental Design

Dataset

The real IoT dataset published by University of South Walls. It includes 28-IoT device traces for the period of 20 days. The size of the traffic data ranges from 61MB to 2GB with an average of 365MB in the form of PCAP files.

Evaluation Measures

Evaluation is measured in terms of Accuracy, Precession, Recall, F1 Score, and Errors performed on IoT traffic.

Software Requirements

- Basic knowledge of Python/Java
- Exposure to Linux environment.

Hardware Requirements

- NVIDIA GPU