Master’s Thesis Report

Federated learning approach for IoT network intrusion detection

NGUYEN VAN TUAN

Supervisor BEURAN, Razvan Florin

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
(Information Science)

September, 2024

Abstract

The rise of the Internet of Things (IoT) has revolutionized various human life aspects by interconnecting numerous information devices, but this has also increased the cyber attack surface to more sophisticated intrusions. An intrusion is defined as any activity that attempts to compromise the CIA characteristics (confidentiality, integrity, and availability) or bypass the security of a computer or network. The intrusion detection system is one of the most popular and effective solutions to secure modern information systems. Host-based and network-based approaches are the main categories of intrusion detection systems. In which, a host-based method observes the information on individual computers in the system. Meanwhile, a network-based method captures and analyzes network traffic to protect multiple hosts in network segments. Leveraging the outbreak of machine learning, there are so many studies trying to improve the performance in IoT intrusion detection, which can be categorized as supervised, semi-supervised, and unsupervised approaches depending on the availability of training data. To tackle the existing challenges such as the evolution of attacking techniques, and the lack of labeled anomalous data, the semi-supervised and unsupervised approaches are adopted more popularly.

Besides developing an effective learning data model, addressing resource constraints and ensuring privacy preservation are critical concerns in modern IoT networks. Traditional machine learning methods operate within a centralized paradigm, where a single model is trained on a server using all the data collected from connected devices. This faces challenges related to computational resources, data latency, data transfer cost, and particularly privacy preservation when collecting all data to a single server.

This thesis addresses these problems of IoT intrusion detection by proposing a novel network-based approach based on federated learning that leverages the computational power of distributed IoT devices while training the local model itself and protecting the local data from being accessed directly. I propose a semi-supervised federated learning approach using the combination of the Shrink Autoencoder model and Centroid one-class classifier (SAE-CEN) to enhance IoT intrusion detection. The Shrink Autoencoder tries to represent the normal network data in a new optimal data space around the origin in the latent layer, then, the Centroid algorithm can detect the unseen data point based on its distance to the origin. This makes the detection task more effective and efficient. I develop a novel mean square error-based aggregation algorithm (MSEAvg) to improve global model performance when prioritizing the more accurate local model in aggregating the global model. Our approach aims to address issues such as data heterogeneity, unbalanced and noisy data, and the scarcity of labeled abnormal data, which are prevalent in IoT environments.

One of the most important aspects is applying the experimental study to real-world scenarios. Some existing research using federated learning creates experimental environments that do not accurately reflect real-world conditions. An IoT network is divided into multiple sub-networks, each having some IoT devices of different types and innovating over time. During the lifetime, some new devices are added or removed, changing seriously the network topology, and leading to a change in the data distribution. In this thesis, I construct the experimental scenarios to be more practical by including both IID (Independent and Identically Distributed) and non-IID settings using the N-BaIoT dataset and the Dirichlet distribution.

The experimental outcomes in this setup demonstrate that our SAE-CEN model, combined with the MSEAvg aggregation algorithm, significantly improves detection accuracy and robustness in heterogeneous IoT networks. I also conduct some investigations in different federated learning settings to examine the robustness of my approach. The results also expose that my approach not only can boost the performance but also reduce the learning costs of federated intrusion detection and adapt strongly in large-scale IoT networks.

This work contributes to the field by presenting a practical federated learning framework for IoT intrusion detection, highlighting the capability of tailored aggregation methods and the potential of semi-supervised learning techniques in addressing real-world cybersecurity challenges.

Keywords. IoT, Intrusion Detection, Machine Learning, Federated Learning.