Detection of Unauthorized IoT Devices Using Machine Learning Techniques

Shafiq Hussain[1] and Hudaa Gull[1]

[1]Affiliation not available

August 2, 2022

Detection of Unauthorized IoT Devices Using Machine Learning Techniques

Hudaa Gull, Shafiq Hussain

ABSTRACT

Security experts have demonstrated numerous risks imposed by Internet of Things (IoT) devices on organizations. Due to the widespread adoption of such devices, their diversity, standardization obstacles, and inherent mobility, organizations require an intelligent mechanism capable of automatically detecting suspicious IoT devices connected to their networks.

INTRODUCTION

The Internet of Things (IoT) is globally expanding, providing diverse benefits in nearly every aspect of our lives. Unfortunately, the IoT is also accompanied by a large number of information security vulnerabilities and exploits[1-3].

SYSTEM AND ATTACK MODEL: In this research, the system we assume is a typical large enterprise, facing an ever growing range of IoT-related cyber threats. *Untargeted:* The connected IoT device has been previously infected by a malware of indiscriminate nature, virally spreading among as many devices as possible[4, 5]. Cross-contamination provides a mechanism for this kind of attack. *Specifically targeted:* The malware was intentionally implanted on the IoT device by an attacker, based on the assumption that the device would likely be connected to a specific organizational network in the future.

White Listing For IOT Devices:

White list of authorized device types marked as safe is much smaller than the ever growing list of presumably insecure types, unauthorized by default. As a result, a shorter list contributes to the increased efficiency of the machine learning (ML) processes underlying the proposed white listing method, including model training, validation, testing, and deployment.

PROPOSED METHOD:

Given a set of authorized device types (i.e., the white list) and a structured set of traffic data, we treat the task of IoT device type identification as a multi-class classification problem. That is, we wish to map each IP stream to the type of IoT device that is most likely to have produced it.

Classifier Training:

The Random Forest supervised ML algorithm is selected for model training. According to a recent survey on ML methods in cyber security, this algorithm which combines decision tree induction with ensemble learning has several advantages relevant to our study, including:

• There is no need for prior feature selection.

• It requires just a few input parameters.

• The algorithm is resistant to over fitting.

• When the number of trees increases, the variance is decreased without resulting in bias.

Parameter Tuning

Application for Device Type Identification

[1] M. Heydari and K. K. Lai, "A study on risk and expense evaluation of agility supply management of machinery," *Discrete Dynamics in Nature and Society,* vol. 2020, 2020.

[2] M. Heydari, K. K. Lai, and Z. Xiaohu, *Risk Management in Public-Private Partnerships*. Routledge, 2020.

[3] M. Heydari, K. K. Lai, and Z. Xiaohu, "How to Manage Red Alert in Emergency and Disaster Unit in the Hospital? Evidence From London," *Frontiers in Public Health,* vol. 9, 2021.

[4] S. A. Shah and N. Mazher, "A review on security on internet of things," in *November 2018 Conference: 1st International Multi-Disciplinary Research Conference (IMDRC 2017)*.

[5] N. Mazher and I. Ashraf, "A Systematic Mapping Study on Cloud Computing Security," *International Journal of Computer Applications,* vol. 89, no. 16, pp. 6-9, 2014.