AI-based malware and attack detection for IoT security

Andreas Lyth, Andrii Shalaginov, Guru Bhandari*, Tor-Morten Grønli

Kristiania University College, 0107 Oslo, Norway

Keywords— Cybersecurity, Machine Learning, Malware and Attacks, IoT Security, Artificial Neural Network

In recent years, Internet of Things(IoT) devices have been fostering smarter and greener environments with user-friendly, sustainable, and more efficient interfaces and capabilities. However, in the 2022 Cyber Threat of SonicWall cybersecurity research lab[1] report, a continuously increasing trend of IoT malware threats was reported, with more than 60 million attacks recorded in 2021, the highest ever recorded in a single year. IoT malware attacks, in particular, increased by 6%, with routers being the most targeted devices[1]. These security issues put pressure on enterprises, organizations, and governments to acquire effective threat intelligence, to protect the systems against malware and attacks. Therefore, in this study, we propose an approach with a framework to discover malware attacks on the IoT devices using artificial intelligence (AI) enabled methods. The choice of hardware for setting up the IoT network is representative for typical industrial use, as well as being available off the shelf. All software used is open source and readily available for everyone. The work belongs to the ENViSEC project which aims at developing top-notch solution for bringing the IoT security forward and offering similarity-based detection using AI models. The hardware setup for this project represents a generic application of smart environments deploying the various IoT devices. The IoT network uses MQTT message protocol with Raspberry Pi4 Mod B set up as a gateway and MQTT-broker, relaying messages between the nodes. Arduino UNO WiFi R2 (ATmega4809), Arduino NANO 33 IoT, and ESP32 Huzzah (Tensilica LX6) are used as sensor and actuator nodes, publishing and subscribing messages to topics on the broker. The model deployed on the Raspberry Pi gateway will monitor this traffic.

To check the practical implication of the machine learning approach for malware detection and attacks prediction, we initially applied a basic artificial neural network (ANN) with only four layers on Aposemat IoT-23 dataset[2]. Several other studies [3, 4] have also used the dataset for network traffic analysis, malware, and attack detection applications. Because of the large sample size of the IoT-23 dataset, we only considered a part of the dataset (3394338 samples, or ∼7%). After preprocessing, we ended up with 67652 examples. This data was split into training and testing samples in 80:20 ratio and later trained and tested with 50 epochs. The model gets its fully conversed trained model after a few epochs. The resulting performance of the ANN model looks promising for IoT security. The accuracy, precision, and recall measures for both training and testing are achieved 0.998, where the loss is 0.013.

At present, there are many issues and challenges with implementing AI-enabled methods for IoT security. The publicly available dataset such as IoT-23 and Edge-IIoTset [2] provides insight of and certain behaviour of the real-world malware and cyber attacks artificially generating malware attacks samples. Ground truth data from actual malware attacks would reflect real-world scenarios, but that may suffer from data imbalance between benign samples being really high and malware samples being very low.

References

[1] Bill Conner. *2022 SonicWall Cyber Threat Report*. Tech. rep., p. 66.

[2] Mohamed Amine Ferrag et al. *Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning*. Preprint. Jan. 2022.

[3] Yue Liang and Nikhil Vankayalapati. "Machine Learning and Deep Learning Methods for Better Anomaly Detection in IoT-23 Dataset Cybersecurity" (), p. 6.

[4] Nicolas-Alin Stoian. "Machine Learning for Anomaly Detection in IoT Networks: Malware Analysis on the IoT-23 Data Set" (), p. 10.

[1] https://www.sonicwall.com

[2] https://www.stratosphereips.org/datasets-iot23

Please mark your interests of contribution during the conference(choose multiple options that you are interested):

✓	I like to give an oral presentation
✓	I like to submit original unpublished scientific work in Nordic Machine Intelligence Journal
	I like to present a poster / demo