UNIVERSITY of GUELPH

COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

PhD Seminar 2

Tuesday December 6, 2022 at 2pm via Zoom

Hamed Haddadpajouh

An Adversarially Robust Multi-view Multi-kernel Framework for IoT Malware Threat Hunting

Advisor: Dr. Ali Dehghantanha
Co-Advisor: Dr. Hadis Karimipour (SoE)
Advisory: Dr. Xiaodong Lin

Abstract:

Today, cyber threats are becoming more complicated than anytime before. These threats cause a massive loss in social, political, and financial resources. Malware attacks have a significant share of cyber threats in different platforms ranging from computer devices to critical infrastructure. The emerging trends of using the Internet of Things (IoT) in our daily life bring up both promising prospects and security challenges. With diversified and numerous applications, IoT systems are now facing more security challenges than ever before. Malware is among the primary tool of cybercriminals to infect and exploit IoT devices.

Detecting malware threats (also known as threat hunting) allows security analysts to design a robust security posture against attackers' tactics, techniques, and procedures (TTPs). However, timely threat hunting is a complicated task as not only do the security mechanisms encounter new malicious payloads that do not include single behavior, also threat actors use evading techniques like generating adversarial examples to bypass Artificial Intelligence (AI) powered defensive mechanisms. In this research, we propose an adversarially robust multi-view multi-kernel malware threat hunting framework for IoT environments. This framework consists of three elements 1) a multi-kernel IoT malware threat hunting module; a light-weight threat hunting module that works with different static properties of executable files 2) a malware example generative module based on code-cave vulnerability for evaluating machine learning malware threat hunting module against adversarial attacks in order to make an adversarial robust threat hunter model, and 3) an adversarial malware example prevention module based on statistical bytecode analysis for evaluating originality of incoming malware samples in ML-based malware threat hunting mechanism.

The multi-kernel approach uses an aggregation function that grabs all kernel information to detect malicious payloads from a different view (bytecodes and opcodes). The generative model tries to bypass deep neural network models like the Malconv that use adversarial techniques such as Code-caves. Finally, the prevention mechanism tries to detect adversarial examples based on their feature spaces and an ensemble structure for deciding whether incoming samples are adversarially generated or not.

We evaluate the elements of our framework by Precision, Recall, and Confidence Interval metrics to assess the accuracy of the framework in hunting malware samples. Also, we adopt an evasion rate metric for assessing the robustness of the ML-based threat-hunting model against adversarial examples. We test our framework with an IoT cloud-edge malware dataset.