GW SCHOOL of ENGINEERING & APPLIED SCIENCE

THE GEORGE WASHINGTON UNIVERSITY

Department of Computer Science

Doctoral Dissertation Defense

Friday, July 16, 2021 at 8:00 a.m. - 11:00 a.m.
Zoom: https://gwu-edu.zoom.us/j/5176636495

Fangtian Zhong
George Washington University
School of Engineering and Applied Sciences

Practical Adversarial Malware Example Attacks and Defenses

ABSTRACT

Machine Learning allows computers to learn from experience and to understand the world in terms of a hierarchy of concepts, with each being defined through its relations to simpler concepts. Extensive and beautiful prospects for the global machine learning market spurs the development of various applications in language translation, image recognition, social media, speech recognition, malware detection, etc. It also stimulates the potential of sophisticated adversaries to attack them. Nevertheless, there is no practical adversarial malware example to successfully attack a collection of third-party black-box malware detectors in the real world since they normally combine machine learning techniques with classical techniques. Additionally, existing methods for malware classification, whether static or dynamic, are non-trivial tasks when stubborn knowledge barriers, limited computing resources, and conservative feasibility need to be addressed.

In this thesis, we first propose a convolutional generative adversarial network-based (Conv-GAN) framework titled MalFox, targeting adversarial malware example generation against third-party black-box malware detectors. Motivated by the rival game between malware authors and malware detectors, MalFox adopts a confrontational approach to produce perturbation paths, with each formed by up to three methods (namely Obfusmal, Stealmal, and Hollowmal) to generate practical adversarial malware examples. We then propose a reinforcement learning-based framework called MalInfo, which could generate adversarial malware examples via an adaptive selection of a perturbation path for each malware while MalFox is trained on a collection of malware. To cope with limited computation, MalInfo applies either dynamic programming or temporal difference learning to choose the optimal perturbation path. Finally, we propose a visualization malware classification framework called VisMal, which could provide highly efficient categorization with acceptable accuracy for malware samples that consist of adversarial malware examples and other types of malware samples. VisMal converts malware samples into images and applies a contrast-limited adaptive histogram equalization algorithm to them to enhance the similarity between malware images in the same family. In conclusion, this thesis directs to practical adversarial malware example attacks and defenses.

GW SCHOOL of ENGINEERING & APPLIED SCIENCE

THE GEORGE WASHINGTON UNIVERSITY

Department of Computer Science

Doctoral Dissertation Defense

Friday, July 16, 2021 at 8:00 a.m. - 11:00 a.m.
Zoom: https://gwu-edu.zoom.us/j/5176636495

Fangtian Zhong

B.E. Software Engineering, July 2018, Northeast Normal University, China

DISSERTATION:

Practical Adversarial Malware Example Attacks and Defenses

FIELD OF STUDY: Computer Science

ADMISSION TO DOCTOR OF SCIENCE PROGRAM: Fall 2018

ADVISOR OF THE CANDIDATE'S RESEARCH:

Dr.Xiuzhen Cheng, Professor of Computer Science, GWU

EXAMINING COMMITTEE:

1. Dr. Abdou Youssef, Professor of Computer Science, GWU
2. Dr. Hyeong-Ah Choi, Professor of Computer Science, GWU
3. Dr. Xiuzhen Cheng, Professor of Computer Science, GWU
4. Dr. Qin Hu, Assistant Professor of Computer and Information Science, IUPUI

PRESIDING:

Dr. Abdou Youssef, Professor of Computer Science, GWU