Mingzhe Gao

mzgao@njnet.edu.cn · (+86) 151-5053-2663 · My Homepage

Research Interests

My current research interests primarily revolve around Software Security, System Security, and Data Mining. Specifically, I am deeply interested in the following areas: binary analysis, malware detection and family taxonomy (encompassing both binary and script-based malware), static program analysis, and adversarial attacks against learning systems.

Education

Southeast University (SEU), Nanjing, China 2019 – 2022

Master student in Cyberspace Security (CS)

Shandong University of Technology (SDUT), Shandong, China 2015 – 2019

B.S. in Software Engineering (SE)

Research Publication

Rectify the Malware Family Label via Hybrid Analysis 2023

Computer & Security (CCF-B) Corresponding Author

Brief introduction: Rectify the Malware Label bias

- Introduce RecMaL, a malware label correction tool that utilizes hybrid analyses.
- Identify three types of mislabeling issues: error, ontology, and multi-label.
- Rectifying the label results in a 1.9% accuracy improvement using the same features and models.

A Malicious Code Static Detection Framework Based on Multi-Feature Ensemble Learning 2021

Journal of computer research and development Corresponding Author

Brief introduction: Propose a static malware detection framework based on multi-feature ensemble learning.

- Implemented five features: non-PE structure, visible string, assembly code sequences, PE structure, and function call relationship.
- Employed Bagging and Stacking ensemble algorithms to mitigate overfitting.
- Achieved a higher recall rate of 96.99% on packed and obfuscated malware.

Experience

Alibaba Cloud Inc. Hangzhou, China 2022 – Present

Security engineer Xinhuo Sec Lab

Brief introduction: Constructing a Benign Knowledge Base for Script Classification.

- Operating as a Standalone Engine: Achieving a 98% Recall Rate for Benign Samples on Alibaba Cloud, Alleviating 0.08% False Positives in AV Engines.
- Serving as a Feature Preprocessing Method for Malware Detection Models on Web Servers: Enhancing the Recall Rate of SOTA Methods by 5%.

Qi Anxin Technology Research Institute. Nanjing, China 2020 – 2022

Security Research Xingtu Sec Lab

Brief introduction: Malware family classification, Conception shift, Adversarial attack

- Malware family classification via hybrid analysis
- Concept drift detection based on malware classifier
- Malware adversarial sample construction based on static feature

Skills

- Programming Languages: Python > Golang > Java
- Platform: Linux, Mac, Windows
- Tools: Sklearn, IDA Pro, Tensorflow
- Development: Binary, Web

Honors and Awards

4^{th} *Prize*, Award on DataCon Big Data Security Competition	Jan. 2023
1^{st} *Prize*, Award on QiangWang Cup Artificial Intelligence Challenge	Nov. 2021
9^{th} *Prize*, Award on DataCon Big Data Security Competition	Nov. 2021
2^{nd} *Prize*, Award on ZongHeng Cup Network Security Innovation Competition	Nov. 2021
4^{th} *Prize*, Award on Artificial intelligence-based malware family classification Competition	Sep. 2021

Miscellaneous

- Blog: https://mzgao.blog.csdn.net/
- Languages: English - Fluent, Mandarin - Native speaker
- Research Interest: Malware analysis, Static program analysis, System and software security, Software Composition Analysis, Vulnerability Exploitation, etc.