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: concept drift of malware detection, static analysis, code similarity, malware family taxonomy (encompassing both binary and script-based malware), 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

RecMaL: 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 Group. Hangzhou, China 2022 – Present

Security engineer Xinhuo Sec Lab

① Brief introduction: Mitigate False Positives of Web Shell Detection

- Established an expansive collection of code fragments as a knowledge base, enabling the identification of benign samples through similarity computations.
- Evaluation: Achieved an impressive 98% recall rate for benign samples on the Alibaba public cloud, effectively mitigating false positive incidents in AV engines by 0.81%.
- Submitted Work for Review at ICSE 2025

② Brief introduction: Anti-Bot Detection and Captcha Defense

- Engineered a sophisticated anti-bot detection system leveraging anomaly detection algorithms, behavioral analytics, and intelligent rule generation to combat automated attacks.
- Enhanced bot detection rates by 30% in high-stakes scenarios such as flash sales, login/registration processes, and data scraping activities by utilizing browser fingerprinting and user behavior analytics to strengthen defense against aggressive automated attacks.

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

- Traffic Analysis, Program Analysis
- Programming Languages: Python > Golang > C
- Platform: Linux, Mac, Windows
- Tools: Sklearn, IDA Pro, Tensorflow, GDB

Honors and Awards

4th Prize, Award on DataCon Big Data Security Competition Jan. 2023
1st Prize, Award on QiangWang Cup Artificial Intelligence Challenge Nov. 2021
9th Prize, Award on DataCon Big Data Security Competition Nov. 2021
2nd Prize, Award on ZongHeng Cup Network Security Innovation Competition Nov. 2021
4th 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 analysis, System and software security, Software Composition Analysis, Vulnerability Exploitation, etc.