



EDUCATION

Xiamen University, School of Informatics

Bachelor of Engineering

Xiamen, China

Sep 2015 - Jun 2019

Relevant Coursework: Algorithms, Principles of Operating Systems, Principles of Compilers

WORK

QI-ANXIN Technology Research Institute

Research and Development Engineer, Supervisor: *Lingyun Ying*

Beijing, China

Aug 2019 - Now

- **MacOS Sandbox for Malware Analysis:** Designed and implemented the **first** macOS sandbox system in China. For macOS versions below 10.15, I read the source code of **XNU kernel**, used the Mandatory Access Control Framework(MACF) to monitor **process behavior** and **file operations**, wrote Network Kernel Extension(NKE) to capture **network traffic and behavior** respectively, and use threaded to send the monitoring data from kernel mode to user mode, and wrote a client at user mode to receive the behavior data from kernel mode. For macOS versions 10.15 and above, the previous mechanism was deprecated, so I developed a new system using the EndPoint Security Framework(ESF) and Network Extension(NE) for monitoring and capturing behavioral data. **Apple Script** was used to write a dynamic traversal tool for macOS apps to simulate user interaction and trigger malicious behavior.
- **Static and Dynamic Analysis Tool for Android Apps:** Developed a static analysis tool based on **AndroGuard**, with the same analysis capability as **VirusTotal**. Combined the depth-first search(DFS) algorithm and breadth-first search(BFS) algorithm to traverse the UI of Android App.
- **Continuous Fuzzing Platform:** A **fuzzing infrastructure** is used to schedule and allocate resources, perform **continuous** and **large-scale** fuzzing of target software, and obtain various monitoring data and performance metrics. Various frameworks and middleware are used, such as MongoDB, InfluxDB, Fluentd, Redis, Amazon S3. Integrated multiple **fuzzing engines**, and supports fuzzing multiple targets on multiple operating system platforms.
- **TianWen: Dependency Analysis Platform for Software Supply Chain:** Developed a crawler to continuously crawl software binaries. Built and deployed **graph database cluster**. Optimized database query performance to support querying **extremely complex** graph information in seconds, and wrote tools to visualize large amounts of graph data.
- **CVE Information Extraction Tool:** Based on Named Entity Recognition(NER) theory, we use **BiLSTM-CRF** model to extract the **vulnerability function**, **vulnerability version** and **vulnerability source path** from the unstructured official description information of CVE vulnerabilities, with an accuracy rate of about **88%**.
- **pySnoopSnitch: Android Firmware Patch Existence Detection Tool:** Rewrote the **SnoopSnitch** project code in Python. Supports checking the presence of patches in Android firmware using the **full core performance** of the server and generating a heat map of patch misses.
- **String filter:** A tool for filtering strings dumped from memory which contain a large number of strings that are neither recognizable nor readable by humans, and also contain many duplicate strings that we don't care about that are contained in dependent libraries. I combine **Shannon entropy** and **Markov chains** to filter strings, and use **bloom filters** to filter library strings with excellent time and space efficiency.

Institute of Information Engineering, Chinese Academy of Sciences

Research Intern, Supervisor: *Feng Li*

Beijing, China

Jul 2018 - Sep 2018

- **IO2BO Vulnerability Automatic Detection:** Performed the **infra-procedural** analysis and **inter-procedural** analysis on the source code based on **Klee**. Modified the Klee to read target files as input for **concolic execution**. When executing, add constraints to the state that may have vulnerabilities and pass the symbolic expression to the **constraint solver**, and determine whether there is an Integer-Overflow-to-Buffer-Overflow(IO2BO) vulnerability based on the result.

PUBLICATIONS

- Detecting Logical Bugs of DBMS with Coverage-based Guidance, Y. Liang, **S. Liu**, H. Hu, **USENIX 2022**
- Large-scale Security Measurements on the Android Firmware Ecosystem, Q. Hou, W. Diao, Y. Wang, X. Liu, **S. Liu**, L. Ying, S. Guo, Y. Li, M. Nie, H. Duan, **ICSE 2022**

PROJECTS

- **CTF Wiki:** The CTF Wiki is an open source knowledge base about CTF competitions. I contributed most of the content of the **reverse engineering** chapter of the Wiki ([github](#), 5.6k stars)
- **Awesome-Binary-Similarity:** An awesome list of binary code similarity papers ([github](#), 178 stars)
- **Awesome-Binary-Rewriting:** An awesome list of binary rewriting papers ([github](#), 99 stars)
- **PyPi-Typosquatting-Graph:** Analyzing and visualizing **typosquatting** for Python packages hosted on PyPi.org. I obtained all python package names and created a **trie tree**, then calculated the **edit distance** between the names and used a **force-directed algorithm** to draw the graph ([github](#))
- **Fugitive Consortium:** A blockchain platform based on **Hyperledger Fabric**. It is written in Java and starts a network topology consisting of docker containers. Chaincode is written to add, delete, query and change information in the ledger ([github](#))