Yiwei Hou

No.30 Shuangqing Road, Haidian District, Beijing, 100084, P.R.China
https://m1-llie.github.io · houyw22@mails.tsinghua.edu.cn · yiweihou233@gmail.com
M.S., Tsinghua University

Education

M.S. in Software Engineering, Tsinghua University, Beijing, China — Sep. 2022 – Jun. 2025
Software System Security Assurance Group, advised by Prof.Yu Jiang.
Software Security, System Security, Program Analysis, Privacy

B.S. in Cybersecurity, Sichuan University, Chengdu, China — Sep. 2018 – Jun. 2022
GPA: 3.95 / 4.0, Ranking: 1 / 172. Advised by Prof. Haizhou Wang.
Thesis: Design and Implementation of Binary Testing Tool based on Fuzzing

Summer School, University of California, Berkeley, CA, USA — Jul. 2019
Undergraduate Summer School in EECS

Publications

Preventing Disruption of System Backup Against Ransomware Attacks
Yiwei Hou, Lihua Guo, Chijin Zhou, Quan Zhang, Yu Jiang
In Submission, The 34th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA'25)

Ransomware Detection Through Temporal Correlation Between Encryption and I/O Behavior
Lihua Guo, **Yiwei Hou**, Chijin Zhou, Quan Zhang, Yu Jiang
In Submission, The 33rd ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE'25)

An Empirical Study of Data Disruption by Ransomware Attacks
Yiwei Hou, Lihua Guo, Chijin Zhou, Yiwen Xu, Zijing Yin, Shanshan Li, Chengnian Sun, Yu Jiang
2024, The 46th IEEE/ACM International Conference on Software Engineering (ICSE'24)

Limits of I/O Based Ransomware Detection: An Imitation Based Attack
Chijin Zhou, Lihua Guo, **Yiwei Hou**, Zhenya Ma, Quan Zhang, Mingzhe Wang, Zhe Liu, Yu Jiang
2023, The 44th IEEE Symposium on Security and Privacy (S&P Oakland'23)

MIDAS: Safeguarding IoT Devices Against Malware via Real-Time Behavior Auditing
Yiwen Xu, Zijing Yin, **Yiwei Hou**, Jianzhong Liu, Yu Jiang
2022, The 22nd ACM SIGBED International Conference on Embedded Software (EMSOFT'22)

Identification of Chinese Dark Jargons in Telegram Underground Markets Using Context-oriented and Linguistic Features
Yiwei Hou, Hailin Wang, Haizhou Wang
2022, Information Processing and Management (IP&M 59(5))

Research Experience

Software System Security Assurance Group, Tsinghua University — Sep. 2022 – Present
Graduate Student Researcher, advised by Prof. Yu Jiang
Topic Keywords: Software Security, Program Analysis, System Security, Ransomware Attacks, Data Protection, Dynamic Analysis, Runtime Defense, Mitigation Strategies, Empirical Study.

- Presented an imitation-based ransomware attack that mimics the behavior of benign programs to disguise its encryption tasks, and theorized and experimentally explored the limitations of ransomware detection techniques based on I/O behavior.
- Implemented an experimental pipeline to analyze large-scale ransomware samples, built a real-world dataset of 7,796 active ransomware samples, analyzed their behavior in disrupting data within victim systems, collected perspectives from six categories of dynamic behavior, and presented six critical findings spanning three phases of data disruption in ransomware attacks.

- Proposed two ransomware mitigation strategies, including preventing the disruption of system backups and the temporal correlation between encryption and I/O behavior when ransomware executes, and validated their performance on existing and zero-day ransomware samples.

Grant Ho Security Lab, University of Chicago Jul. 2024 – Present

Visiting Student Researcher, advised by Prof. Grant Ho
Topic Keywords: AI for Security, LLM for Security, System Audit Logs, APT Attack Analysis, Prompt Engineering, LLM Evaluation Criteria, Measurement, ATT&CK Matrix, Benchmark Construction.

- Proposed a framework for systematically evaluating the utility of publicly available state-of-the-art LLMs for log auditing tasks, and investigated the impact of different prompts.
- Constructed a well-rounded dataset consisting of malicious system events in the ATT&CK Matrix, and performed evaluation on the self-generated log dataset along with existing APT log datasets such as DARPA TC3 and ATLAS.
- Proposed various criteria to evaluate the utility of LLMs for system log investigation, presented the difficulties encountered by LLMs, and suggested promising potential in leveraging LLMs to support security log auditing.

Software System Security Assurance Group, Tsinghua University Nov. 2021 – Sep. 2022

Research Assistant, advised by Prof. Yu Jiang
Topic Keywords: IoT Security, Real-time Threat Detection, Dynamic Malware Defense, Behavior Auditing, Embedded System, Firmware, Program Analysis, Hardware Debugging.

- Designed the anti-malware protection framework with a real-time behavior auditing mechanism for Linux-based IoT devices, developed the backend system for OpenWrt, and compiled the corresponding firmware.
- Evaluated system viability on routers such as EZVIZ W3, by disassembling the router, establishing the TFTP serial port communication pipeline using a USB-to-TTL adapter and DuPont cables, and physically updating the firmware for the routers.
- Deployed large-scale virtual IoT devices worldwide to test the performance of our defense system against malware, operated compromised devices, and analyzed results.

Wang Research Group, Sichuan University Jun. 2020 – Nov. 2021

Undergraduate Student Researcher, advised by Prof. Haizhou Wang
Topic Keywords: Online Crime Modeling, Information Forensics, Jargon Identification, Natural Language Processing, Feature Engineering, Word Embedding, Transfer Learning, Contextual Word Vectors, Semantic Projections.

- Developed a legitimate crawler to collect chat history from Telegram groups related to the underground market and constructed the corpus TUMCC, the first Chinese corpus in jargon identification.
- Proposed an approach to generate high-quality word vectors based on a relatively small amount of chat history, and applied a transfer learning method to further improve the quality of the vectors, effectively capturing and utilizing the context of the jargon to improve the identification performance.
- Extracted seven new features in three categories to identify jargon from commonly used words, performed statistical outlier detection to determine whether a word qualifies as jargon, and analyzed the results.

Honors & Awards

- **Deng Feng Fund for International Conference Travel**, Tsinghua University (2024)
- **Execellent Comprehensive Scholarship**, Tsinghua University (2022, 2023, 2024)
- **Outstanding Undergraduate Award**, Sichuan Province (2022)
- **Top 10 Undergraduate of 2022**, Sichuan University (Top 0.03% in SCU) (2022)
- **National Scholarship**, Ministry of Education of China (Top 0.2% in China) (2019, 2020)
- **Cybersecurity Merit Student Scholarship**, Sichuan University (2019)

Community Involvement

- **Teaching Assistant:** Model-driven Software Development@Tsinghua, Fall 2024, Course No.44100662
- **Artifact Evaluation Committee (AEC):** CCS'23, ISSTA'24
- **Journal Reviewer/External Reviewer:** IP&M, EMSOFT'22, ICSE'24, FSE'24, ISSTA'24, EMSOFT'24, ICSE'25

Skills

- **Technical:** Python, C/C++, Git, Linux, LaTeX, SQL, MATLAB, Markdown, HTML/CSS
- **Language:** English (Proficient, TOEFL 105 - Speaking 24), Mandarin (Native)