ZHILONG WANG

LinkedIn | Google Scholar | GitHub

EXPERIENCE

Bytedance, Security Engineer, Security Assurance Team *Jan. 2024 - Current*

- *Security Assurance:* security infrastructure, **software security**.
- *LLM Security*: Research the **attack and defense for LLM applications [8, 5, 1]**.
- *Code Analysis and Vulnerability Discovery:* Increase the **vulnerability discovery** rate through program analysis and AI.
- *Skill:* Golang, Python, Shell, Git, CI/CD, SAST, LLM security, Penetration Testing, Compiler.

Bytedance, Security Engineer Intern, Security Assurance Team *May 2022 - Aug. 2022*

- Adopt static taint analysis to identify **fuzz** targets.
- Optimize the speed of inter-procedural **program analysis**.
- *Skills:* Golang, SSA, Taint Analysis, OSFuzz, libFuzzer

JD.COM American, R & D Intern, Security Infrastructure Team *May 2021 - Aug. 2021*

- Deploy the **Privacy-Preserving Machine Learning cluster** based on Graphene-SGX.
- Develop transparent data encryption with Intel SGX enclave.
- *Skills:* SGX, Graphene-SGX, Machine Learning Cluster, Transparent Data Encryption

Penn State, Graduate Research Assistant *Aug 2019 - Dec. 2023*

- **DL for Security-Oriented Program Analysis.** Leverage GNN, Tree-LSTM, and GPT models to analyze program dependency graphs for detecting silent buffer overflows[4], identifying security-critical variables [3], and assisting ransomware reverse engineering [6,7].
- **LLM Security.** Research the vulnerability in LLM application [1,9].
- *Skills:* C/C++, Python, PyTorch, deep learning for security, imbalanced distribution issue, GNN, Tree-LSTM, BERT, Ransomware, LLVM, Obfuscator-LLVM, AddressSanitizer, AFL, Intel Pin, Dynamic Analysis, Static Analysis, Program Slicing.

Nanjing University Master Student, Penn State Visiting Scholar *Aug 2016 - June. 2019*

- **Software/System Security.** Modify compiler and program runtime to **protect software** against the Blind-ROP attack; **obfuscate binary code** through return-oriented programming; trace and **protect Linux kernel** through ARM hardware feature (i.e., Embedded Trace Macrocell).
- *Skills:* Compiler, Linux Kernel, Intel Pin, x86/ARM Assembly, C/C++.

EDUCATION

Pennsylvania State University, Ph.D. in Informtics, advisor: Peng Liu *Aug. 2019 - Dec. 2023*
Dissertation: Deep Learning for Security-oriented Program Analysis

Nanjing University, M.S., Computer Science *Sept. 2016 - Jun. 2019*
Research: System and Software Security.

Zhengzhou University, B.S., Computer Science *Sept. 2012 - Jul. 2016*
GPA: 3.6/4.0 Ranking: 1/240

AWARDS & HONORS

- National Scholarship, 2014.
- Outstanding Graduates of Nanjing University, 2019.
- First-Class Academic Scholarship of Nanjing University, 2016.
- First-Class Scholarship of Zhengzhou University, 2013 & 2015 & 2016.

- The First Prize of Microsoft Wheeled Micro-Robot Simulation Competition in China Robot Competition, Beijing, 2014.

Publications

1. **Zhilong Wang**, Haizhou Wang, Nanqing Luo, Lan Zhang, Xiaoyan Sun, Yebo Cao, Peng Liu. “Hide Your Malicious Goal Into Benign Narratives: Jailbreak Large Language Models through Carrier Articles" EAI SecureComm, 2026.
2. Neha Nagaraja, Lan Zhang, **Zhilong Wang**, “From Pixels to Prompts: A Systematic Study and Introduction to Image Prompt Injection Attacks." AI Column of IEEE Computer, 2026.
3. **Zhilong Wang***, Haizhou Wang*, Hong Hu, and Peng Liu. “Identifying Non-Control Security-Critical Data in Program Binaries with a Deep Neural Model." accepted by *Transactions on Dependable and Secure Computing (TDSC)* 2025. (* co-first author)
4. **Zhilong Wang**, Chen Cao, Yu Li, Suhang Wang, Xiaoyan Sun, Peng Liu “DeepSanitizer: Combining Heuristic Rules and Deep Learning Models to Spot Silent Buffer Overflows in Binary", accepted by *Transactions on Dependable and Secure Computing (TDSC)* 2025.
5. Neha Nagaraja, Lan Zhang, **Zhilong Wang**, Bo Zhang, Pawan Patil “Image-based Prompt Injection: Hijacking Multimodal LLMs through Visually Embedded Adversarial Instructions", accepted by *The 3rd International Conference on Foundation and Large Language Models (FLLM2025)*.
6. Nanqing Luo*, Haizhou Wang*, **Zhilong Wang***, Lan Zhang, Ping Chen, Peng Liu “Deep Learning Assisted Reverse Engineering: Recognizing Encryption Loops in Ransomware", accepted by *TrustCom* 2025. (*co-first author)
7. Xushu Dai, Nanqing Luo, Haizhou Wang, **Zhilong Wang**, Chen Cao, Peng Liu “DESCG: Data Encoding Scheme Classification with GNN in Binary Analysis", *Automated Software Engineering (ASE)*, Springer 2025.
8. **Zhilong Wang**, Neha Nagaraja, Lan Zhang, Hayretdin Bahsi, Pawan Patil and Peng Liu “To Protect the LLM Agent Against the Prompt Injection Attack with Polymorphic Prompt", on the industry track of *the 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 2025.
9. **Zhilong Wang**, Lan Zhang, Chen Cao, Nanqing Luo, Xinzhi Luo, and Peng Liu “How Does Naming Affect Language Models on Code Analysis Tasks?" Journal of Software Engineering and Applications, 2024.
10. **Zhilong Wang** and Peng Liu. “GPT Conjecture: Understanding the Trade-offs between Granularity, Performance and Timeliness in Control-Flow Integrity." *Cybersecurity*, 2021.
11. Yunlan Du, Zhenyu Ning, Jun Xu, **Zhilong Wang**, Yueh-Hsun Lin, Fengwei Zhang, Xinyu Xing, and Bing Mao. “HART: Hardware-assisted Kernel Module Tracing on Arm." In Proceedings of The *25th European Symposium on Research in Computer Security (ESORICS)*, 2020.
12. Yoon-Ho Choi, Peng Liu, Zitong Shang, Haizhou Wang, **Zhilong Wang**, Lan Zhang, Junwei Zhou and Qingtian Zou. “Using Deep Learning to Solve Computer Security Challenges: A Survey." *Cybersecurity*, 2020. (The authors of this paper are listed in alphabetic order)
13. **Zhilong Wang**, Xuhua Ding, Chengbin Pang, Jian Guo, Jun Zhu and Bing Mao. “To Detect Stack Buffer Overflow With Polymorphic Canaries." In *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2018.
14. Jun Zhu, Weiping Zhou, **Zhilong Wang**, Dongliang Mu, and Bing Mao.“DiffGuard: Obscuring Sensitive Information in Canary Based Protections." *International Conference on Security*

and Privacy in Communication Systems (SecureComm). Springer, Cham, 2017.

15. Dongliang Mu, Jia Guo, Wenbiao Ding, **Zhilong Wang**, Bing Mao, and Lei Shi.“ ROPOB: Obfuscating Binary Code via Return Oriented Programming." *International Conference on Security and Privacy in Communication Systems (SecureComm)*. Springer, Cham, 2017.

Public Services

Reviewer

- The Journal of Computer Security *2021 & 2022*
- The European Conference on Computer Systems (EuroSys) *2022*
- IEEE/IFIP International Conference on Dependable Systems and Networks *2022 & 2023*
- IEEE/IFIP International Conference on Dependable Systems and Networks *2023*
- The Journal of Computer Security *2023*
- Transactions on Information Forensics & Security (TIFS) *2024*
- The International Conference on Security and Cryptography (SECRYPT) *2024*
- ACM Transactions on Internet of Things *2024*
- Journal of Computer Networks *2024*
- The Journal of Supercomputing *2024 & 2025*
- Transactions on Dependable and Secure Computing (TDSC) *2025 & 2026*