Penghui Li

Room 489, CS Building
500 West 120 Street
New York, NY 10027

☎: +1 (646) 294 0890
✉: pl2689@columbia.edu
🏠: https://peng-hui.github.io

Research Interests

My research lies in **software system security** and its intersection with *software engineering* and *machine learning*. Most recently, I have been building agentic systems to secure software at scale. I leverage rigorous program analysis and autonomous LLM agents to tackle critical security tasks, including vulnerability detection, exploit generation, and runtime defense.

Education

The Chinese University of Hong Kong Aug. 2019 – Jul. 2023
Doctor of Philosophy, Computer Science and Engineering
Advisor: Prof. Wei Meng
GPA: 3.97/4

University of Chinese Academy of Sciences Aug. 2015 – Jul. 2019
Bachelor of Engineering, Computer Science and Technology
GPA: 3.87/4

Research Experience

Columbia University Sep. 2024 – Present
Postdoctoral Research Scientist
Host: Prof. Junfeng Yang

Zhongguancun Laboratory Sep. 2023 – Aug. 2024
Security Researcher

Tsinghua University Feb. 2022 – Sep. 2022
Visiting Student
Host: Prof. Chao Zhang

Institute of Information Engineering, CAS Oct. 2018 – Jun. 2019
Research Intern
Host: Prof. Kai Chen

Awards and Honors

ACM CCS Top Reviewer Award Oct. 2025

ACM CCS Distinguished Artifact Award	Oct. 2025
ACM CCS Distinguished Paper Award	Oct. 2024
USENIX Security Distinguished Artifact Reviewer Award	Aug. 2024
ACM CCS Best Paper Honorable Mention	Nov. 2022
HKSAR Reaching Out Award	Apr. 2022
The Web Conference Student Scholarship	Mar. 2021

Grant Experience

Generating CodeQL Queries with LLMs for Privilege Escalation Detection in Microservices
Co-PI, with Prof. Junfeng Yang (PI) and Prof. Yinzhi Cao (Co-PI)
Submitted to *Google YouTube Security and ISE Static Analysis Teams*

Detecting Memory-Safety Vulnerabilities in Multilingual Software
Proposal contributor, with Prof. Wei Meng (PI)
Funded by *General Research Fund of HK RGC*

Publication

Summary:

10 papers in software security (S&P, Security, CCS, NDSS)
3 papers in software engineering (ICSE, FSE, ASE)
2 papers in web security (WWW)
1 paper in database systems (VLDB)

Preprints

[1] **Detecting Privilege Escalation in Polyglot Microservices via Agentic Program Analysis**
Penghui Li, Hong Yau Chong, Yinzhi Cao, and Junfeng Yang
Under Review.

[2] **Automated Static Vulnerability Detection via a Holistic Neuro-Symbolic Approach**
Penghui Li, Songchen Yao, Josef Sarfati Korich, Changhua Luo, Jianjia Yu, Yinzhi Cao, and Junfeng Yang
Under Review, https://arxiv.org/abs/2504.16057.

[3] **A Systematic Investigation of Security Threats in the PHP Supply Chain Ecosystem**
Changhua Luo, Zejun Feng, Minghang Shen, Penghui Li, Mingxue Zhang, and Qian Wang
Under Review.

[4] **Reasoning under Vision: Understanding Visual-Spatial Cognition in Vision-Language Models for CAPTCHA**
Jincen Song, Luke Tenyi Wang, Yun-yun Tsai, Penghui Li, and Junfeng Yang
Under Review, https://arxiv.org/abs/2510.06067.
*** Supervised the project**.

[5] **Chasing Cookies to the Source: In-Browser Data Flow Backtracking for Web Compliance Analysis**
Yi Yang, Mingxue Zhang, Yuxiang Ma, Cong Zhang, Penghui Li, Changhua Luo, and Weina Niu
Under Review.

[6] **Minnie: User Privacy Leak Detection for WeChat Miniapps via Holistic Dynamic Taint Analysis with Concolic Execution**
Jianjia Yu, Zhengyu Liu, Zhihan Xia, Penghui Li, Zifeng Kang, Junfeng Yang, and Yinzhi Cao
Under Review.

[7] **Explainer-Guided Targeted Adversarial Attacks against Binary Code Similarity Detection Models**
Tiancheng Zhu, Mingjie Chen, Mingxue Zhang, Yiling He, Minghao Lin, Penghui Li, and Kui Ren
Under Review, https://arxiv.org/abs/2506.05430.

Referred Papers

[8] **Fuzzing JavaScript Engines by Fusing JavaScript and WebAssembly**
Jiayi Lin, Changhua Luo, Mingxue Zhang, Lanteng Lin, Penghui Li, and Chenxiong Qian
In *Proceedings of the 48th International Conference on Software Engineering (ICSE)*. Apr. 2026.

[9] **PickleBall: Secure Deserialization of Pickle-Based Machine Learning Models**
Andreas D. Kellas, Neophytos Christou, Wenxin Jiang, Penghui Li, Laurent Simon, Yaniv David, Vasileios P. Kemerlis, James C. Davis, and Junfeng Yang
In *Proceedings of the 32nd ACM Conference on Computer and Communications Security (CCS)*. Oct. 2025.
Distinguished Artifact Award.

[10] **Predator: Directed Web Application Fuzzing for Efficient Vulnerability Validation**
Chenlin Wang, Wei Meng, Changhua Luo, and Penghui Li
In *Proceedings of the 46th IEEE Symposium on Security and Privacy (S&P)*. May 2025.

[11] **VulShield: Protecting Vulnerable Code Before Deploying Patches**
Yuan Li, Chao Zhang, Jinhao Zhu, Penghui Li, Chenyang Li, Songtao Yang, and Wende Tan
In *Proceedings of the 31st Annual Network and Distributed System Security Symposium (NDSS)*. Feb. 2025.

[12] **FuzzCache: Optimizing Web Application Fuzzing Through Software-Based Data Cache**
Penghui Li and Mingxue Zhang
In *Proceedings of the 31st ACM Conference on Computer and Communications Security (CCS)*. Oct. 2024.
Distinguished Paper Award.

[13] **Test Suites Guided Vulnerability Validation for Node.js Applications**
Changhua Luo, Penghui Li*, Wei Meng, and Chao Zhang
In *Proceedings of the 31st ACM Conference on Computer and Communications Security (CCS)*. Oct. 2024.
*** Supervised the project**.

[14] **SDFuzz: Target States Driven Directed Fuzzing**
Penghui Li, Wei Meng, and Chao Zhang
In *Proceedings of the 33rd USENIX Security Symposium (Security)*. Aug. 2024.

[15] **Testing Graph Database Systems via Graph-Aware Metamorphic Relations**
Zeyang Zhuang, Penghui Li, Pingchuan Ma, Wei Meng, and Shuai Wang
In *Proceedings of the 50th International Conference on Very Large Data Bases (VLDB)*. Aug. 2024.

[16] **Holistic Concolic Execution for Dynamic Web Applications via Symbolic Interpreter Analysis**
Penghui Li, Wei Meng, Mingxue Zhang, Chenlin Wang, and Changhua Luo
In *Proceedings of the 45th IEEE Symposium on Security and Privacy (S&P)*. May 2024.

[17] **DDRace: Finding Concurrency UAF Vulnerabilities in Linux Drivers with Directed Fuzzing**
Ming Yuan, Bodong Zhao, Penghui Li, Jiashuo Liang, Xinhui Han, Xiapu Luo, and Chao Zhang
In *Proceedings of the 32nd USENIX Security Symposium (Security)*. Aug. 2023.

[18] **SelectFuzz: Efficient Directed Fuzzing with Selective Path Exploration**
Changhua Luo, Wei Meng, and Penghui Li
In *Proceedings of the 44th IEEE Symposium on Security and Privacy (S&P)*. May 2023.

[19] **SEDiff: Scope-Aware Differential Fuzzing to Test Internal Function Models in Symbolic Execution**
Penghui Li, Wei Meng, and Kangjie Lu
In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*. Nov. 2022.

[20] **TChecker: Precise Static Inter-Procedural Analysis for Detecting Taint-Style Vulnerabilities in PHP Applications**
Changhua Luo, Penghui Li, and Wei Meng
In *Proceedings of the 29th ACM Conference on Computer and Communications Security (CCS)*. Nov. 2022.
Best Paper Honorable Mention.

[21] **Understanding and Detecting Performance Bugs in Markdown Compilers**
Penghui Li, Yinxi Liu, and Wei Meng
In *Proceedings of the 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. Nov. 2021.

[22] **LChecker: Detecting Loose Comparison Bugs in PHP**
Penghui Li and Wei Meng
In *Proceedings of the Web Conference (WWW)*. Apr. 2021.

[23] **On the Feasibility of Automated Built-in Function Modeling for PHP Symbolic Execution**
Penghui Li, Wei Meng, Kangjie Lu, and Changhua Luo
In *Proceedings of the Web Conference (WWW)*. Apr. 2021.

Services

Organizer

Columbia Agentic AI Security & Privacy Seminar Series	Fall 2025

Program Committee Member

ACM Conference on Computer and Communications Security	2025 – 2026
USENIX Security Symposium	2026
International Workshop on Large Language Models for Code	2026
Workshop on Measurements, Attacks, and Defenses for the Web	2024 – 2025
European Conference on Computer Systems, Shadow PC	2024

USENIX Security Symposium, Artifact Evaluation Committee	2024
ACM Conference on Computer and Communications Security, Artifact Evaluation Committee	2023

Journal Reviewer

IEEE Transactions on Dependable and Secure Computing	2024 – 2025
IEEE Transactions on Information Forensics and Security	2025
ACM Transactions on Software Engineering and Methodology	2024 – 2025
IEEE Transactions on Software Engineering	2025

External Reviewer

ACM SIGSOFT International Symposium on Software Testing and Analysis	2024
IEEE Symposium on Security and Privacy	2023 – 2024
The Annual Computer Security Applications Conference	2023
ACM Conference on Computer and Communications Security	2021 – 2022
The Web Conference	2020 – 2022
ACM ASIA Conference on Computer and Communications Security	2021 – 2022

Teaching

Guest Lecturer

Agentic Program Analysis, W4152: Engineering Software-as-a-Service, Columbia	Fall 2025
Web Security, EIE553: Security in Data Communication, HK PolyU	Spring 2025

Teaching Assistant

Introduction to Database Systems, CUHK	Fall 2021
Building Web Applications, CUHK	Spring 2021
Introduction to Cyber Security, CUHK	Fall 2019, Fall 2020
Linear Algebra for Engineers, CUHK	Spring 2020

Mentoring

Hong Yau Chong Sep. 2025 – Present
Undergraduate at Columbia, working on LLM-aided static analysis

Chunyi Wang Sep. 2025 – Present
Master’s at Columbia, working on LLM-aided static analysis

Yunfei Ke Jun. 2025 – Present
Master’s at Columbia, working on LLM-aided static analysis

Luke Chang Jan. 2025 – Sep. 2025

Master's at Columbia, working on LLM-based CAPTCHA solving [4]

Sophia Yao Jan. 2025 – Jun. 2025

Master's at Columbia, working on neuro-symbolic static analysis [2]

Josef Sarfati Korich Jan. 2025 – May 2025

Undergraduate at Columbia, working on neuro-symbolic static analysis [2]

Zeyang Zhuang Jan. 2023 – Jul. 2023

Ph.D. student at CUHK, worked graph database system testing [15]

Changhua Luo Nov. 2019 – Jul. 2022

Ph.D. student at CUHK, worked on PHP static analysis [20] and Node.js testing [13]

Yanting Chi Oct. 2021 – May 2022

Undergraduate student from SJTU, worked on symbolic execution

Chiho Cheng Oct. 2018 – Apr. 2019

Undergraduate student from CUHK, worked on PHP taint analysis

Hoihim Chan Oct. 2018 – Apr. 2019

Undergraduate student from CUHK, worked on PHP taint analysis