Chengpeng Wang

Postdoctoral Research Associate in Computer Science, Purdue
(+1) 765-421-0267 · wang6590@purdue.edu · Homepage · Chengpeng Wang

RESEARCH INTEREST

Software Engineering, Machine Learning, and Programming Languages, with a focus on combining symbolic static analysis and large language models to advance code auditing and strengthen software security and performance, particularly in the context of AI-assisted coding.

EDUCATION

Hong Kong University of Science and Technology *Aug 2019 - Jan 2024*
Ph.D. in Computer Science and Engineering
Committees: *Charles Zhang (advisor)*, *Xiangyu Zhang*, *Shing-Chi Cheung*, *Shuai Wang*
Dissertation: *Enhancing Reliability and Performance of Data-Centric Systems with Static Analysis*

Tsinghua University *Aug 2012 - Jul 2019*
B.S and M.S. in Software Engineering

ACADEMIC AND INDUSTRY APPOINTMENTS

Now | **Postdoctoral Research Associate, West Lafayette, IN, USA**
Mar 2024 | LLM-driven code auditing, working with Prof. Xiangyu Zhang

Feb 2024 | **Research Intern at Ant Group, Shenzhen**
Dec 2023 | Task: Large language model-aided static analysis

HONORS AND AWARDS

ASE 2025 Distinguished Reviewer Award	*2025*
NeurIPS 2025 Spotlight	*2025*
Postdoc Travel Award, Purdue	*2024*
ACM SIGARCH Best Paper Award, ASPLOS'24	*2024*
UGC Research Travel Grant, HKUST	*2022, 2023*
Ant Group Outstanding Collaboration Award	*2023*
ACM SIGPLAN Distinguished Paper Award, OOPSLA'22	*2022*
ACM SIGPLAN PAC Award	*2022*
Outstanding Tutor Award, Student Learning and Development Center, THU	*2019*
Future Scholar Award, honored 100 students at Tsinghua University	*2016*
Scholarship for Academic Excellence, Tsinghua University	*2013, 2014*

GRANTS

PI. "*Synthesizing Realistic Benchmarks for LLM-Aided Code Security Auditing*". Researcher Access Program, Funded by ***OpenAI***

Co-Advisor, with Xiangyu Zhang (PI). "*Amazon Trusted AI Challenge: An Agentic Red Teaming Framework for Code LLMs*". Funded by ***Amazon*** (acceptance rate: 11%, 10/90).

PUBLICATIONS

Highlight: I authored 7 first-author papers and 2 co-first-author papers (labeled with #) published in top-tier conferences and journals, including **ICSE**, **FSE**, **ASE**, **TOSEM**, **OOPSLA**, **ECOOP**, **ICML**, **NeurIPS**, and **EMNLP**. The underlined first authors are Ph.D. students under my mentoring.

JOURNAL PUBLICATIONS

[J1] Chao Wang, Li Lin, **Chengpeng Wang**, Jiafeng Huang, Congxia Wu, and Rongxin Wu, ReachCheck: Compositional Library-Aware Call Graph Reachability Analysis in the IDEs, In **TOSEM 2025**: Transactions on Software Engineering and Methodology

[J2] Hao Ling, Heqing Huang, **Chengpeng Wang**, Yuandao Cai, and Charles Zhang, GiantSan: Efficient Operation-Level Memory Sanitization with Segment Folding, In **TOCS 2024**: The ACM Transactions on Computer Systems. (Invited extension of [C11])

[J3] Rongxin Wu, Zhiling Huang, Zige Tian, **Chengpeng Wang**, and Xiangyu Zhang, PackHunter: Recovering Missing Packages for C/C++ Projects, In **TSE 2024**: Transactions on Software Engineering

[J4] Yi Sun, **Chengpeng Wang**, Gang Fan, Qingkai Shi, and Xiangyu Zhang, Fast and Precise Static Null Exception Analysis with Synergistic Preprocessing, In **TSE 2024**: Transactions on Software Engineering

[J5] Wensheng Tang, Dejun Dong, Shijie Li, **Chengpeng Wang**, Peisen Yao, Jinguo Zhou, and Charles Zhang, Octopus: Scaling Value-Flow Analysis via Parallel Collection of Realizable Path Conditions, In **TOSEM 2023**: Transactions on Software Engineering and Methodology

[J6] **Chengpeng Wang**, Wenyang Wang, Peisen Yao, Qingkai Shi, Jinguo Zhou, Xiao Xiao, and and Charles Zhang, Anchor: Fast and Precise Value-Flow Analysis for Containers via Memory Orientation, In **TOSEM 2022**: Transactions on Software Engineering and Methodology

CONFERENCE PUBLICATIONS

[C1] Mingwei Zheng, **Chengpeng Wang**, Xuwei Liu, Jinyao Guo, Shiwei Feng, and Xiangyu Zhang, RFCAudit: AI Agent for Auditing Protocol Implementations Against RFC Specifications, In **ASE 2025**: International Conference on Automated Software Engineering.

[C2] Danning Xie, Mingwei Zheng, Xuwei Liu, Jiannan Wang, **Chengpeng Wang**, Lin Tan, and Xiangyu Zhang, CORE: Benchmarking LLMs' Code Reasoning Capabilities through Static Analysis Tasks, In **NeurIPS 2025** (Spotlight): Annual Conference on Neural Information Processing Systems

[C3] Jinyao Guo[#], **Chengpeng Wang**[#], Xiangzhe Xu, Zian Su, and Xiangyu Zhang. RepoAudit: An Autonomous LLM-Agent for Repository-Level Code Auditing. In **ICML 2025**: International Conference on Machine Learning.

[C4] Yuan Li, Peisen Yao, Kan Yu, **Chengpeng Wang**, Yaoyang Ye, Song Li, Meng Luo, Yepang Liu, and Kui Ren. Understanding Industry Perspectives of Static Application Security Testing Evaluation, In **FSE 2025**: International Conference on the Foundations of Software Engineering.

[C5] Mingwei Zheng, Danning Xie, Qingkai Shi, **Chengpeng Wang**, and Xiangyu Zhang, Validating Network Protocol Parsers with Traceable RFC Document Interpretation, In **ISSTA 2025**: International Symposium on Software Testing and Analysis.

[C6] Wei Chen, Bowen Zhang, **Chengpeng Wang**, Wensheng Tang, and Charles Zhang, Seal: Towards Diverse Specification Inference for Linux Interfaces from Security Patches, In **EuroSys 2025**: European Conference on Computer Systems.

[C7] **Chengpeng Wang**, Wuqi Zhang, Zian Su, Xiangzhe Xu, Xiaoheng Xie, and Xiangyu Zhang, LLMDFA: Analyzing Dataflow in Code with Large Language Models, In **NeurIPS 2024**: Annual Conference on Neural Information Processing Systems

[C8] **Chengpeng Wang**, Wuqi Zhang, Zian Su, Xiangzhe Xu, and Xiangyu Zhang, Sanitizing Large Language Models in Bug Detection with Data-Flow, In **Findings of EMNLP 2024**: Empirical Methods in Natural Language Processing

[C9] Li Lin, Zongyin Hao, **Chengpeng Wang**, Zhuangda Wang, Rongxin Wu, and Gang Fan, SQLess: Dialect-Agnostic SQL Query Simplification, In **ISSTA 2024**: International Symposium on Software Testing and Analysis

[C10] **Chengpeng Wang**, Jipeng Zhang, Rongxin Wu, and Charles Zhang, DAInfer: Inferring API Aliasing Specifications from Library Documentation via Neurosymbolic Optimization, In **FSE 2024**: International Conference on the Foundations of Software Engineering

[C11] Bowen Zhang, Wei Chen, Peisen Yao, **Chengpeng Wang**, Wensheng Tang, and Charles Zhang, SIRO: Empowering Version Compatibility in Intermediate Representations via Program Synthesis, In **ASPLOS 2024**: ACM Conference on Architectural Support for Programming Languages and Operating Systems

[C12] Hao Ling, Heqing Huang, **Chengpeng Wang**, Yuandao Cai, and Charles Zhang, GiantSan: Efficient Memory Sanitization with Segment Folding, In **ASPLOS 2024**: ACM Conference on Architectural Support for Programming Languages and Operating Systems 🏆 **ACM SIGARCH Best Paper Award**

[C13] Rongxin Wu, Yuxuan He, Jiafeng Huang, **Chengpeng Wang**, Wensheng Tang, Qingkai Shi, Xiao Xiao, and Charles Zhang, LibAlchemy: A Two-Layer Persistent Summary Design for Taming Third-Party Libraries in Static Bug-Finding Systems, In **ICSE 2024**: International Conference on Software Engineering

[C14] **Chengpeng Wang**, Peisen Yao, Wensheng Tang, Gang Fan, Charles Zhang, Synthesizing Conjunctive Queries for Code Search, In **ECOOP 2023**: European Conference on Object-Oriented Programming

[C15] Zongyin Hao, Quanfeng Huang, **Chengpeng Wang**, Jianfeng Wang, Yushan Zhang, Rongxin Wu, and Charles Zhang, Pinolo: Detecting Logical Bugs in Database Management Systems with Approximate Query Synthesis, In **ATC 2023**: USENIX Annual Technical Conference

[C16] **Chengpeng Wang**, Gang Fan, Peisen Yao, Fuxiong Pan, and Charles Zhang, Verifying Data Constraint Equivalence in FinTech Systems, In **ICSE 2023**: International Conference on Software Engineering

[C17] Wensheng Tang[#], **Chengpeng Wang**[#], Peisen Yao, Rongxin Wu, Xianjin Fu, Gang Fan, and Charles Zhang, DCLink: Bridging Data Constraint Changes and Implementations in FinTech Systems, In **ASE 2023** : International Conference on Automated Software Engineering

[C18] Rongxin Wu, Minglei Chen, **Chengpeng Wang**, Gang Fan, Jiguang Qiu, and Charles Zhang, Accelerating Build Dependency Error Detection via Virtual Build, In **ASE 2022** : International Conference on Automated Software Engineering

[C19] **Chengpeng Wang**, Peisen Yao, Wensheng Tang, Qingkai Shi, and Charles Zhang: Complexity-Guided Container Replacement Synthesis, In **OOPSLA 2022**: SIGPLAN Conference on Objected Oriented Programming, Systems, Languages and Applications 🏆 **ACM SIGPLAN Distinguished Paper Award**

[C20] Gang Fan, **Chengpeng Wang**, Rongxin Wu, Xiao Xiao, Qingkai Shi, and Charles Zhang: Escaping Dependency Hell: Finding Build Dependency Errors with the Unified Dependency Graph, In **ISSTA 2020**: International Symposium on Software Testing and Analysis

ACADEMIC SERVICES

Program Organizing Committee Member:

International Workshop on Language Models and Programming Languages (LMPL)	*2025*

Program Committee Member:

International Conference on Software Engineering (ICSE)	*2027*
ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)	*2026*
International Conference on the Foundations of Software Engineering (FSE)	*2025, 2026*
International Symposium on Software Testing and Analysis (ISSTA)	*2025, 2026*
International Conference on Automated Software Engineering (ASE)	*2025*
International Conference on Automated Software Engineering (ASE), Industrial Track	*2024*
International Symposium on Software Reliability Engineering (ISSRE)	*2024, 2025*
Annual Conference on Neural Information Processing Systems (NeurIPS)	*2025*
International Conference on AI Foundation Models and Software Engineering (Forge)	*2024, 2025*
SPLASH, Student Research Contest Track	*2024*
International Workshop on Large Language Models for Code (LLM4Code)	*2026*

Journal Reviewer:

Transactions on Software Engineering and Methodology (TOSEM) 2025
Transactions on Software Engineering (TSE) 2024

Artifact Evaluation Committee Member:

International Conference on OOP, Systems, Languages, and Applications (OOPSLA) *2024*
ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI) *2023*
International Conference on the Foundations of Software Engineering (FSE) *2022*
International Symposium on Software Testing and Analysis (ISSTA) *2022*

Volunteer:

International Conference on OOP, Systems, Languages, and Applications (OOPSLA) *2022*
Student Volunteer at International Symposium on Software Testing and Analysis (ISSTA) *2019*
Long-term mentor of SIGPLAN-M *2024, 2025*

TEACHING

CS 510: Software Engineering, Purdue University *Fall 2025*
CS 592: AI and Security, Purdue University *Fall 2024*
COMP 3021: Java Programming, HKUST *Spring/Fall 2022/2023*
COMP 4631: Computer and Communication Security, HKUST *Fall 2021*
COMP 3111/H: Software Engineering, HKUST *Fall 2020*
COMP 2011: Programming with C++, HKUST *Spring 2020*
Haskell: Functional Language Programming, THU *Spring 2019*
Automaton and Formal Logic, THU *Fall 2018*
Drop-in Tutoring for STEM Courses, THU *Spring/Fall 2018/2019*

ADVISING AND MENTORING

Jinyao Guo (Graduate Student, Purdue University) *Nov 2024 - Present*
Currently a Ph.D. student at Purdue University
Working on LLM-driven static analysis and coauthored ICML'25 [C3]

Mingwei Zheng (Graduate Student, Purdue University) *Aug 2024 - Present*
Currently a Ph.D. student at Purdue University
Working on network protocol bug detection and coauthored ISSTA'25 [C5] and ASE'25 [C1].

Dominic DeLuca (Undergraduate Student, Purdue University) *May 2025 - Aug 2025*
Currently a third-year undergraduate student at Purdue University
Working on LLM-assisted program debugging

Danning Xie (Graduate Student, Purdue University) *Oct 2024 - May 2025*
Currently a Ph.D. student at Purdue University
Working on evaluating LLMs in code reasoning tasks and coauthored NeurIPS'25 [C2].

Calix Barrus (Graduate Student, The University of Texas at San Antonio) *June 2025 - Aug 2025*
Currently a visiting Ph.D. student at Purdue University
Working on LLM-assisted call graph analysis for Java

Yifei Gao (Graduate Student, Purdue University) *May 2024 - April 2025*
Currently a Ph.D. student at Purdue University
Working on automated translation from C to safer Rust.

Pengxiang Huang (Graduate Student, Northwestern University) *May 2024 - Present*
Currently a Master student at Northwestern University
Working on automated translation from C to safer Rust.

Bowen Zhang (Graduate Student, HKUST) *Oct 2022 - Dec 2023*
Currently a Ph.D. student at Hong Kong University of Science and Technology
Worked on IR Translator Synthesis and coauthored ASPLOS'24 [C11].

PRESENTATIONS AND INVITED TALKS

Neuro-Symbolic Code Auditing for Software Reliability in the AI Era
National University of Singapore, Singapore Management University — Dec 2025

Static Code Auditing for Software Quality Assurance in the AI Era
Tufts University, Boston — Oct 2025

RepoAudit: Human-like AI Auditor for Code Repositories
RSA Conference 2025, San Francisco — Apr 2025

Advances in AI-powered Code Security: Next-Level Bug Detection
Uber's Programming Systems Team. Virtual — Feb 2025
CodeQL@GitHub. Virtual — Jan 2025

Neuro-Symbolic Static Analysis for Reliable Software Systems
UMass Amherst — Sep 2025
University of British Columbia, Columbia University — Aug 2025
UC Davis, UIUC, UNSW — May 2025
AST Lab, ETH Zurich. Virtual — Mar 2025
Department of Computing, Hong Kong Polytechnic University — Jan 2025
Institute of Data Science, The University of Hong Kong — Dec 2024

Neuro-Symbolic Static Analysis for Reliable and Performant Software Systems
School of Computer Science, Nanjing University. Nanjing, China — Dec 2024

When Static Analysis Meets Large Language Models: A Neuro-Symbolic Approach
Ant Group. Virtual — Dec 2024
Boston Computation Club. Virtual — Nov 2024
AI-CyberSecurity Research Lunch Talk, Texas A&M University (TAMU). Virtual — Oct 2024
School of Software, Tsinghua University. Virtual — Oct 2024

Towards Enhancing Reliability and Performance of Data-Centric Systems with Static Analysis
School of Informatics, Xiamen University. Xiamen, China — Aug 2023

Synthesizing Conjunctive Queries for Code Search
ByteDance. Virtual — Jun 2023

Complexity-Guided Container Replacement Synthesis
AST Lab, ETH Zurich. Virtual — Mar 2023