

CHENGPENG WANG

ADDRESS

Room 3154A, Lawson Computer Science Building (LWSN)
Department of Computer Science, Purdue University
305 N. University Street, West Lafayette, IN, USA
Email: wang6590@purdue.edu, stephenw.wangcp@gmail.com
URL: <https://chengpeng-wang.github.io>
ORCID: 0000-0003-0617-5322

EDUCATION

The Hong Kong University of Science and Technology *August 2019 - Dec 2023*
Ph.D. in Computer Science
Tsinghua University *2012-2019*
Master Degree in Software Engineering, 2019
BA in Software Engineering, Minor in Math, 2016

INDUSTRY EXPERIENCE

FEB 2024	Research Intern at Ant Group, Shenzhen
DEC 2023	TASK: Large language model-aided static analysis
JULY 2023	Research Intern at Veridise, Remote
MAY 2023	TASK: Static bug detection of zero-knowledge proof circuits
AUG 2022	Research Intern at Ant Group, Shenzhen
FEB 2021	TASK: Value flow analysis for programs manipulating containers and database-backed application verification.
AUG 2019	Research Intern at Sourcebrella, Shenzhen
FEB 2019	TASK: Inconsistency detection of build systems

PUBLICATIONS

(**OOPSLA** is a top tier Programming Language conference. **FSE**, **ICSE**, **ASE**, and **ISSTA** are top tier Software Engineering conferences. **TOSEM** and **TSE** are top tier Software Engineering journals. **ASPLOS** and **ATC** are top tier Computer System conferences. **NeurIPS** and **EMNLP** are the top tier conferences of machine learning and natural language processing, respectively.)

JOURNAL PUBLICATIONS

[J1] Rongxin Wu, Zhiling Huang, Zige Tian, **Chengpeng Wang**, Xiangyu Zhang, PackHunter: Recovering Missing Packages for C/C++ Projects, In **TSE: IEEE Transactions on Software Engineering**, Just Accepted.
[J2] Yi Sun, **Chengpeng Wang**, Gang Fan, Qingkai Shi, Xiangyu Zhang, Fast and Precise Static Null Exception Analysis with Synergistic Preprocessing, In **TSE: IEEE Transactions on Software Engineering**, Sept, 2024.

[J3] 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**: ACM Transactions on Software Engineering and Methodology, Oct, 2023.

[J4] **Chengpeng Wang**, Wenyang Wang, Peisen Yao, Qingkai Shi, Jinguo Zhou, Xiao Xiao, and Charles Zhang, Anchor: Fast and Precise Value-Flow Analysis for Containers via Memory Orientation, In **TOSEM**: ACM Transactions on Software Engineering and Methodology, Sept, 2022.

CONFERENCE PUBLICATIONS

[C1] **Chengpeng Wang**, Wuqi Zhang, Zian Su, Xiangzhe Xu, Xiaoheng Xie, Xiangyu Zhang, LLM DFA: Analyzing Dataflow in Code with Large Language Models, In **NeurIPS 2024**: Annual Conference on Neural Information Processing Systems, Dec, 2024.

[C2] **Chengpeng Wang**, Wuqi Zhang, Zian Su, Xiangzhe Xu, Xiangyu Zhang, Sanitizing Large Language Models in Bug Detection with Data-Flow, In **Findings of EMNLP 2024**: Empirical Methods in Natural Language Processing, Nov, 2024.

[C3] Li Lin, Zongyin Hao, **Chengpeng Wang**, Zhuangda Wang, Rongxin Wu, Gang Fan, SQLess: Dialect-Agnostic SQL Query Simplification, In **ISSTA 2024**: The ACM SIGSOFT International Symposium on Software Testing and Analysis, Sept, 2024

[C4] **Chengpeng Wang**, Jipeng Zhang, Rongxin Wu, Charles Zhang, DAInfer: Inferring API Aliasing Specifications from Library Documentation via Neurosymbolic Optimization, In **FSE 2024**: ACM International Conference on the Foundations of Software Engineering, July, 2024

[C5] Bowen Zhang, Wei Chen, Peisen Yao, **Chengpeng Wang**, Wensheng Tang, 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, April, 2024

[C6] Hao Ling, Heqing Huang, **Chengpeng Wang**, Yuandao Cai, Charles Zhang, GiantSan: Efficient Memory Sanitization with Segment Folding, In **ASPLOS 2024**: ACM Conference on Architectural Support for Programming Languages and Operating Systems, April, 2024. (**Best Paper Award**)

[C7] 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**: The IEEE/ACM International Conference on Software Engineering, April, 2024.

[C8] 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** : The 38th IEEE/ACM International Conference on Automated Software Engineering, Sept, 2023. ([#]: Equal Contribution)

[C9] **Chengpeng Wang**, Peisen Yao, Wensheng Tang, Gang Fan, and Charles Zhang, Synthesizing Conjunctive Queries for Code Search, In **ECOOP 2023**: European Conference on Object-Oriented Programming, July, 2023.

[C10] 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, July, 2023.

[C11] **Chengpeng Wang**, Gang Fan, Peisen Yao, Fuxiong Pan, and Charles Zhang, Verifying Data Constraint Equivalence in FinTech Systems, In **ICSE 2023**: The IEEE/ACM International Conference on Software Engineering, May, 2023.

[C12] Rongxin Wu, Minglei Chen, **Chengpeng Wang**, Gang Fan, Jiguang Qiu, and Charles Zhang, Accelerating Build Dependency Error Detection via Virtual Build, In **ASE 2022** : The 37th IEEE/ACM International Conference on Automated Software Engineering, Oct, 2022.

[C13] **Chengpeng Wang**, Peisen Yao, Wensheng Tang, Qingkai Shi, Charles Zhang: Complexity-Guided Container Replacement Synthesis, In **OOPSLA 2022**: The ACM SIGPLAN Conference on Objected Oriented Programming, Systems, Languages and Applications, Dec, 2022. (**ACM SIGPLAN Distinguished Paper Award**)

[C14] Gang Fan, **Chengpeng Wang**, Rongxin Wu, Xiao Xiao, Qingkai Shi, Charles Zhang: Escaping Dependency Hell: Finding Build Dependency Errors with the Unified Dependency Graph, In **ISSTA 2020**: The ACM SIGSOFT International Symposium on Software Testing and Analysis, July, 2020.

WORKSHOP PUBLICATIONS

[W1] **Chengpeng Wang**, CodeSpider: Automatic Code Querying with Multi-modal Conjunctive Query Synthesis, In **SPLASH SRC 2022**: The ACM SIGPLAN conference on Systems, Programming, Languages, and Applications: Software for Humanity, Student Research Competition, Dec, 2022.

[W2] **Chengpeng Wang**, Yixiao Yang, Han Liu, Le Kang: Statistical API Completion Based on Code Relevance Mining, In **MAINT 2019**: International Workshop on Mining and Analyzing Interaction Histories, 2019:7-13.

HONORS AND AWARDS

ACM SIGARCH Best Paper Award, ASPLOS'24	<i>2024</i>
UGC Research Travel Grant, HKUST	<i>2022, 2023</i>
Ant Group Outstanding Collaboration Award	<i>2023</i>
ACM SIGPLAN Distinguished Paper Award, OOPSLA'22	<i>2022</i>
ACM SIGPLAN PAC Award	<i>2022</i>
Future Academician Scholarship, honored 100 students at Tsinghua University	<i>2016</i>
Scholarship for Academic Excellence, Tsinghua University	<i>2013, 2014</i>

PRESENTATIONS AND INVITED TALKS

When Static Analysis Meets Large Language Models: A Neuro-Symbolic Approach, Boston Computation Club, Nov, 2024.

When Static Analysis Meets Large Language Models: A Neuro-Symbolic Approach, AI-CyberSecurity Research Lunch Talk, TAMU, Oct, 2024.

When Static Analysis Meets Large Language Models: A Neuro-Symbolic Approach, School of Software, Tsinghua University, Oct, 2024.

DAInfer: Inferring API Aliasing Specifications from Library Documentation via Neurosymbolic Optimization, In FSE 2024: ACM International Conference on the Foundations of Software Engineering, July, 2024

SIRO: Empowering Version Compatibility in Intermediate Representations via Program Synthesis, In ASPLOS 2024: ACM Conference on Architectural Support for Programming Languages and Operating Systems, April, 2024.

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

Synthesizing Conjunctive Queries for Code Search, In ECOOP 2023: European Conference on Object-Oriented Programming, July, 2023.

Pinolo: Detecting Logical Bugs in Database Management Systems with Approximate Query Synthesis, In ATC 2023: USENIX Annual Technical Conference, July, 2023.

Synthesizing Conjunctive Queries for Code Search, In ByteDance, June, 2023

Verifying Data Constraint Equivalence in FinTech Systems, In ICSE 2023 : The IEEE/ACM International Conference on Software Engineering, May, 2023.

Complexity-Guided Container Replacement Synthesis, In AST lab @ ETH Zurich, March, 2023.

CodeSpider: Automatic Code Querying with Multi-modal Conjunctive Query Synthesis, In SPLASH SRC 2022: The ACM SIGPLAN conference on Systems, Programming, Languages, and Applications: Software for Humanity, Student Research Competition, Dec, 2022.

Complexity-Guided Container Replacement Synthesis, In OOPSLA 2022 : The ACM SIGPLAN Conference on Objected Oriented Programming, Systems, Languages and Applications, Dec, 2022.

PROFESSIONAL SERVICES

Program Committee Member: Forge'25, FSE'25, ISSTA'25, SPLASH'24 SRC, ASE'24 (industrial track), ISSRE'24, Forge'24

Journal Reviewer: IEEE Transactions on Software Engineering (TSE)

Artifact Evaluation Committee: OOPSLA'24, PLDI'23, FSE'22, ISSTA'22

Volunteer: Student Volunteer @ SPLASH 2022, Student Volunteer @ ISSTA 2019

Mentoring: Long-term mentor of SIGPLAN-M

TEACHING EXPERIENCE

CS 592: AI and Security, Purdue University	<i>Fall 2024</i>
COMP 3021: Java Programming, HKUST	<i>Spring/Fall 2022/2023</i>
COMP 4631: Computer and Communication Security, HKUST	<i>Fall 2021</i>
COMP 3111/H: Software Engineering, HKUST	<i>Fall 2020</i>
COMP 2011: Programming with C++, HKUST	<i>Spring 2020</i>
Haskell: Functional Language Programming, THU	<i>Spring 2019</i>
Automaton and Formal Logic, THU	<i>Fall 2019</i>

MENTORING EXPERIENCE

NOW AUG 2024	Wavid Bowman, University of Florida SIGPLAN-M mentee
-----------------	--

NOW JUNE 2024	Jungyeom Kim, Korea University SIGPLAN-M mentee
------------------	---

NOW MAY 2024	Yifei Gao, Purdue University PROJECT: C2Rust Translation
-----------------	--

NOW JUNE 2023	Syed Yusuf Ahmed, Purdue University PROJECT: Customizable Cryptographic API Misuse Detection
------------------	--

DEC 2023 OCT 2022	Bowen Zhang, HKUST PROJECT: IR Translator Synthesis. Published at ASPLOS 2024 [C5]
----------------------	--