Li Zhong

lizhong@ucsd.edu | github.com/floridsleeves | +1 8582144467

Research Area: enhance code quality for developers and AI coding assistants

Education

University of California, San Diego (UCSD) *Sept 2019 - June 2024 (expected)*
PhD Candidacy in Computer Science
University of Science and Technology of China (USTC) *Sept 2015 - June 2019*
Bachelor in Computer Science GPA: 4.0 (Rank 1/157)

Publication

1. **Li Zhong**, Chengcheng Xiang, Haochen Huang, Yuanyuan Zhou. Unused Value in Systems Considered Harmful. (**Eurosys'24**) [paper][code]
2. Haochen Huang, Bingyu Shen, **Li Zhong**, Yuanyuan Zhou. Protecting Data Integrity of Web Applications with Database Constraints Inferred from Application Code. (**ASPLOS'23**) [paper][code]
3. Haochen Huang*, Chengcheng Xiang*, **Li Zhong**, Yuanyuan Zhou. PYLIVE: On-the-Fly Code Change for Python-based Online Services. Proceedings of the 2021 USENIX Annual Technical Conference (**ATC'21**) [paper][code]
4. **Li Zhong**, Zilong Wang. LLM vs StackOverflow: Can LLMs Replace the Online Programming Forum? Under Submission to **AAAI** [paper][code]
5. Chengcheng Xiang, **Li Zhong**, Haochen Huang, Eric Mugnier, Yuanyuan Zhou. Detecting risky access-control changes with automatic configuration testing. Under Submission to **S&P**.
6. ACSym: Detecting Access Control Change with Symbolic Execution. Eric Mugnier, Cheng-hung He, Duolan Ouyang, Bingyu Shen, **Li Zhong**, Tianyi Shan, Yuanyuan Zhou. Under Submission to **SOSP**.
7. **Li Zhong**, Bokai Zhang, Yuanyuan Zhou. Don't Trust Your Code! Detect Vulnerabilities by Code Mutation Analysis. Under Submission to **Usenix Security**

Industry Experience

Large Language Model Inference Weight Update on Neural Engine (Apple)
Host: Cecile Foret *June 2023 - Sept 2023*
- Implement the inference weight update for large language model in the compiler.
- Measure the runtime overhead for weight update
- Support dense and compressed convolution, and linear layers in transformers.

Value Profiling Driven Memory Operation Optimization (Meta)
Host: Hongtao Yu *June 2022 - Sept 2022*
- Extend llvm-profgen to parse perf traces
- Measure workload value characterisitcs on Clang-10 and HHVM
- Improve implementation of memcpy lowering in LLVM IR.

Call Arg Profiling Driven Code Specialization (Google)
Host: Snehasish Kumar, Co-host: Sotiris Apostolakis *June 2021 - Sept 2021*
- Develop an LLVM pass to instrument FFmpeg and Clang to profile arguments
- Identify limitations of current function specialization solution in LLVM

- Improve the function specialization pass in LLVM with profile results. Reduce runtime instruction count on the Transcoder benchmarks up to 5%.

Dynamic Self-tuning Optimizer (Microsoft Research Asia)
Host: Youshan Miao *Nov 2018 - May 2019*

- Investigate different optimizers in machine learning
- Come up with a novel sgd optimizer using adaptive learning rate and prove its convergence
- Achieve a better performance over the Adam optimizer

Research Experience

Unauthorized Access Vulnerability Testing in Web Applications (UCSD)
Advisor: Yuanyuan Zhou *Sept 2021 - now*

- Collect and study the unauthorized access vulnerabilities in web applications
- Develop automatic testing framework for access control in web applications
- Static analysis on the vulnerable endpoints in Web JavaScript with ESLint.
- Detect 26 vulnerabilities in WriteFreely and phpBB.

Detect Unused Definition to Expose Bugs in System (UCSD)
Advisor: Yuanyuan Zhou *Aug 2020 - May 2021*

- Develop an inter-procedural static analysis tool to detect 'unused' value
- Prune false positives in analysis based on cross-scope features and SMT solvers
- Detect 40 existed bugs and 213 new bugs in MySQL, Linux, NFS-ganesha, etc.

Detecting Missing Constraints in Database from Web App Code (UCSD)
Advisor: Yuanyuan Zhou *Sep 2021 - June 2022*

- Summarize code patterns with implicit assumptions on data integrity
- Static analysis on Django apps based on AST
- Detect 210 missing constraints in popular Django apps

Detecting Risky Access-control Changes with Configuration Testing (UCSD)
Advisor: Yuanyuan Zhou *May 2021 - Sep 2021*

- Automatized test creation for access control configuration changes using logs
- Sandbox testing with Docker using copy-on-write and IP remapping
- Designed a trimming method based on static analysis to accelerate the tests up to 98% by skipping the side-effect operations
- Tested real-world Docker images from 72 images and found 168 vulnerabilities

Python Framework for Dynamic Code Change (UCSD)
Advisor: Yuanyuan Zhou *Sep 2019 - May 2020*

- Develop a framework for dynamic code change for Python-based service system
- Take advantage of meta-object protocol, dynamic typing to make it flexible
- Implement light-weight mechanism for byte code manipulation on Python

KV Database on Open-channel SSD (UIUC)
Advisor: Jian Huang *July 2018 - Oct 2018*

- Profile the I/O amplification of popular KV store
- Combine Cuckoo filter with software-defined FTL
- Optimize the implementation based on LevelDB

Professional Service

EuroSys'22 Shadow PC
NeurIPS 2022 Workshop MetaLearn Reviewer
PAKDD'23 Reviewer