

Yifei Liu

✉ yifeliu@cs.stonybrook.edu • 📞 +1 631-710-8377
🌐 github.com/Yifei-Liu • 🌐 https://www.fsl.cs.stonybrook.edu/~yifei/ • 🔗 linkedin.com/in/yifei-liu

Education

Stony Brook University

Stony Brook, NY

- Ph.D. in Computer Science (Advisor: Prof. Erez Zadok) 08/2019 – 05/2025 (expected)
- M.S. in Computer Science (earned en route to Ph.D. program) 08/2019 – 12/2021
- GPA: 3.93 / 4.0

Huazhong University of Science and Technology

Wuhan, China

- M.Eng. in Computer System Architecture (Advisor: Prof. Ke Zhou) 09/2016 – 06/2019

Huazhong Agricultural University

Wuhan, China

- B.Eng. in Computer Science and Technology 09/2012 – 06/2016

Experience

File systems and Storage Lab (FSL), Stony Brook University

Stony Brook, NY

Graduate Research Assistant (C/C++, File Systems, Formal Verification)

05/2020 – Present

- Developed Metis, a file system model-checking framework that identified over 15 bugs in Linux kernel file systems with thorough coverage
- Developed RefFS, a reliable user-space file system with state-saving and restoration features, achieving 3-28x better performance than other file systems

Samsung Semiconductor, Inc.

San Jose, CA

Storage Systems Architect Intern (C++, Databases, Storage)

05/2022 – 08/2022

- Implemented the custom PostgreSQL plans and paths to offload aggregate operations to SmartSSD, for enhancing the performance of queries

Wuhan National Laboratory for Optoelectronics

Wuhan, China

Research Assistant (Python, Deep Learning, Cloud Storage)

09/2016 – 06/2019

- Designed and implemented a storage system that uses deep learning hashing and a graph database to integrate high-precision, low-latency content-based semantic queries

Tencent Cloud

Shenzhen, China

Backend Developer Intern (C++, Machine Learning, Storage)

12/2015 – 08/2016

- Developed infrastructure to collect long-term disk S.M.A.R.T. data from over 10,000 servers, utilizing machine learning to predict disk failures with high precision and recall

Selected Publications

Google Scholar Profile: scholar.google.com/citations?user=WNu87vQAAAAJ

Journal Articles

- [1] Tyler Estro, Mário Antunes, Pranav Bhandari, Anshul Gandhi, Geoff Kuenning, Yifei Liu, Carl Waldspurger, Avani Wildani and Erez Zadok. “Accelerating Multi-Tier Storage Cache Simulations Using Knee Detection.” *Performance Evaluation*, 2024.
- [2] Ke Zhou, Yangtao Wang, Yu Liu, Yujuan Yang, Yifei Liu, Guoliang Li, Lianli Gao, and Zhili Xiao. “A Framework for Image Dark Data Assessment.” *World Wide Web*, 2020.
- [3] Yu Liu, Yangtao Wang, Ke Zhou, Yujuan Yang, and Yifei Liu. “Semantic-aware Data Quality Assessment for Image Big Data.” *Future Generation Computer Systems*, 2020.

Conference and Workshop Papers

- [1] Yifei Liu, Manish Adkar, Gerard Holzmann, Geoff Kuenning, Pei Liu, Scott Smolka, Wei Su and Erez Zadok. “Metis: File System Model Checking via Versatile Input and State Exploration.” In *the 22nd USENIX Conference on File and Storage Technologies (FAST)*, 2024. **(Received all Artifact Evaluation badges: Available, Functional, and Reproduced)**
- [2] Tyler Estro, Mário Antunes, Pranav Bhandari, Anshul Gandhi, Geoff Kuenning, Yifei Liu, Carl Waldspurger, Avani Wildani and Erez Zadok. “Guiding Simulations of Multi-Tier Storage Caches Using Knee Detection.” In *the 31st International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2023.

- [3] Yifei Liu, Gautam Ahuja, Geoff Kuenning, Scott Smolka, and Erez Zadok. “Input and Output Coverage Needed in File System Testing.” In *the 15th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage)*, 2023.
- [4] Wei Su, Yifei Liu, Gomathi Ganesan, Gerard Holzmann, Scott Smolka, Erez Zadok, and Geoff Kuenning. “Model-Checking Support for File System Development.” In *the 13th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage)*, 2021.
- [5] Yu Liu, Hong Jiang, Yangtao Wang, Ke Zhou, Yifei Liu, and Li Liu. “Content Sifting Storage: Achieving Fast Read for Large-scale Image Dataset Analysis.” In *the 57th Design Automation Conference (DAC)*, 2020.
- [6] Yu Liu, Yangtao Wang, Ke Zhou, Yujuan Yang, Yifei Liu, Jingkuan Song, and Zhili Xiao. “A Framework for Image Dark Data Assessment.” In *the 3rd APWeb-WAIM joint conference on Web and Big Data (APWeb-WAIM)*, 2019. **(Best Paper Runner-Up)**
- [7] Yangtao Wang, Yu Liu, Yifei Liu, Ke Zhou, Yujuan Yang, Jiangfeng Zeng, Xiaodong Xu, and Zhili Xiao. “Analysis and Management to Hash-Based Graph and Rank.” In *the 3rd APWeb-WAIM joint conference on Web and Big Data (APWeb-WAIM)*, 2019.

Patents

- [1] Ke Zhou, Yu Liu, Yujuan Yang, Hua Wang, Chunhua Li, Yangtao Wang, Yifei Liu. Method for valuation of image dark data based on similarity hashing. U.S. Patent US11,138,479B2, Granted: 10/05/2021.
- [2] Ke Zhou, Yifei Liu, Yu Liu, Yangtao Wang, Yujuan Yang. Image query method and system based on content semantic metadata. Chinese Patent CN110413807B, Granted: 04/20/2021.

Skills




Programming Languages

- **Fluent ($\geq 7,000$ LoC):** C, C++, Python, Bash
- **Intermediate ($\geq 2,000$ LoC):** SQL, Java, MATLAB, Cypher, JavaScript, Promela, Prolog

Technologies

- **Databases:** MySQL, Neo4j, PostgreSQL, HBase
- **File and Storage:** Linux VFS and kernel file systems, OpenStack Swift, HDFS
- **Other Tools:** Git, GNU make, CMake, GDB, Elasticsearch, Hadoop, Spark, \LaTeX

Projects

- Metis : A Versatile Framework for File System Model Checking (C/C++)** 2020 – 2024
 - A framework for thoroughly checking Linux file systems with minimal constraints
- RefFS : A Fast and Reliable File System for Checking Reference (C++)** 2020 – 2024
 - An in-memory FUSE file system capable of independently saving and restoring its entire state
- IOCov : Input and Output Coverage for File System Testing (Python)** 2022 – 2024
 - A framework for computing syscall input and output coverage in file system test suites

Talks

- **Metis: File System Model Checking via Versatile Input and State Exploration**
- USENIX FAST 2024, Graduate Research Day 2024
- **Input and Output Coverage Needed in File System Testing**
- ACM HotStorage 2023
- **Model-Checking Support for File System Development**
- ACM HotStorage 2021, Dutch Model Checking Day 2022

Service

Journal Reviewer: ACM Trans. on Architecture and Code Optimization (TACO), IEEE Access
Artifact Evaluation Committee: USENIX OSDI '23, USENIX ATC '23

Teaching

Teaching Assistant for CSE376 Advanced Systems Programming in Unix/C S '20, S '21
 Teaching Assistant for CSE306 Operating Systems F '19

Contest Awards

- Finalist, Interdisciplinary Contest in Modeling (MCM/ICM), USA, 2015.
- First Prize, National Postgraduate Mathematic Contest in Modeling, China, 2014.
- First Prize, Contemporary Undergraduate Mathematical Contest in Modeling, China, 2014.