SHU WANG

shuwanguc@gmail.com

EDUCATION

The University of Chicago *2015 - 2021*
Ph.D. & M.S. in Computer Science

University of Wisconsin-Madison *2013 - 2015*
M.S. in Computer Engineering

Harbin Institute of Technology *2009 - 2013*
B.E. in Electrical Engineering

EMPLOYMENT

LinkedIn *Feb 2022 - Now*
Senior Software Engineer @ Spark Team

- Led Spark Dependency Cache (Featured on LinkedIn Eng Blog)
  - Developed a scalable, fault-tolerant, and high-performance caching solution for Spark to eliminate redundant dependency uploads
  - Achieved a **5%** reduction in distinct LinkedIn member profile scraping
  - Improved median runtime by **40%** for **all** Spark jobs and by **30%** for ML/AI workflows
  - Eliminated **25 Million** JAR uploads daily, reduced **150 TB** of data transfer, and decreased **10%** of HDFS write operations
- Piloted Spark YARN Client Improvements
  - SPARK-44272: Fixed path inconsistency issue within statCache and reduced RPC calls by 50%
  - SPARK-44306: Proposed to fetch all file statuses at the directory instead of individual RPC calls per file
- Optimized Spark Shuffle Server
  - SPARK-43987: Enhanced Netty pipeline by separating heavy IO requests
  - Revamped P80 fetch delay by 98%, decreased SASL authentication timeout by 40%, cut median Spark runtime by 35%, and achieved a substantial 20% reduction in resource consumption.

RESEARCH & INTERNSHIP EXPERIENCES

Automatic Configuration for Software System *Apr 2016 - Aug 2021*
The University of Chicago *Research Assistant*

- Designed an auto-configuration framework for distributed systems (Mapreduce, HDFS, Hbase, Cassandra).
- Developed a self-adaptive algorithm for auto-configuration.
- Implemented a static analysis tool for inferring configurations' properties.
- Improved both performance and reliability (avoiding OOME crashes) of the system.

Experiment Reproducibility in Chameleon Cloud *Jun 2018 - Sep 2018*
Argonne National Laboratory(ANL) *Research Intern*

- Analyzed RabbitMQ events used in OpenStack-based Cloud Computing Infrastructure.
- Composed an actionable OpenStack command list script for reproducible experiments.

Hardware Transactional Memory Application *Jan 2016 - Aug 2016*
The University of Chicago *Research Assistant*

- Fixed concurrency bugs using Intel Hardware Transactions Memory for MySQL, Apache, and Mozilla.
- Designed an accurate and efficient software instrumentation algorithm.
- Improved the system reliability with less overhead.

Fine-grained Wireless Sensing Application *Aug 2014 - Mar 2015*
University of Wisconsin-Madison *Research Assistant*

- Implemented an eavesdropping system based on the vibration of wireless signal strength.

Stochastic Analysis of Full-duplex Wireless Network *Jan 2014 - Jul 2014*
University of Wisconsin-Madison *Research Assistant*

- Analyzed full-duplex networks capacity using stochastic geometry under different MAC protocols.

PUBLICATIONS

AgileCtrl: A Self-adaptive Framework for Configuration Tuning
Shu Wang, Henry Hoffmann, Shan Lu
ACM Foundations of Software Engineering (**FSE**), 2022
Acceptance ratio: 22%, 99 out of 396 submissions

Statically Inferring Performance Properties of Software Configurations
Chi Li, **Shu Wang**, Henry Hoffmann, Shan Lu
ACM European Conference on Computer Systems (**EuroSys**), 2020
Acceptance ratio: 18%, 43 out of 234 submissions

Applying Transactional Memory for Concurrency-Bug Failure Recovery in Production Runs
Yuxi Chen, **Shu Wang**, Shan Lu, Karthikeyan Sankaralingam
IEEE *Transactions on Parallel and Distributed Systems* (**TPDS**), 2018
Impact Factor: 3.402

Applying Hardware Transactional Memory for Concurrency-Bug Failure Recovery in Production Runs
Yuxi Chen, **Shu Wang**, Shan Lu, Karthikeyan Sankaralingam
USENIX *Annual Technical Conference* (**ATC**), 2018
Acceptance ratio: 20%, 76 out of 378 submissions

Understanding and Auto-Adjusting Performance-Related Configurations
Shu Wang, Chi Li, William Sentosa, Henry Hoffmann, Shan Lu
ACM *International Conference on Architectural Support for Programming Languages and Operating Systems* (**ASPLOS**), 2018
Acceptance ratio: 18%, 56 out of 307 submissions

Reproducibility as Side Effect (Poster)
Shu Wang, Zhuo Zhen, Jason Anderson, Kate Keahey
ACM/IEEE *Supercomputing Conference* (**Supercomputing**), 2018

Fundamental Analysis of Full-duplex Gains in Wireless Networks
Shu Wang, Vignesh Venkateswaran, Xinyu Zhang
IEEE/ACM *Transactions on Networking* (**ToN**), 2017
Impact Factor: 3.597

Acoustic Eavesdropping through Wireless Vibrometry
Teng Wei, **Shu Wang**, Anfu Zhou, Xinyu Zhang
ACM *International Conference on Mobile Computing and Networking* (**MobiCom**), 2015
Acceptance ratio: 18%, 38 out of 207 submissions, one of **top 9** pre-accepted papers

Exploring Full-Duplex Gains in Multi-Cell Wireless Networks:A Spatial Stochastic Framework
Shu Wang, Vignesh Venkateswaran, Xinyu Zhang
IEEE *Conference on Computer Communications* (**INFOCOM**), 2015
Acceptance ratio: 19%, 316 out of 1640 submissions

PATENTS

Wireless Vibometer with Antenna Array
Xinyu Zhang, Teng Wei, **Shu Wang**, Anfu Zhou

SKILLS

- **Programming:** C/C++, Java, Python, Matlab.
- **Software:** Spark, Hadoop, HBase, OpenStack.
- **Hardware:** Intel HTM, Embedded System.
- **Platform:** WARP, Intel MCS-51, TI CC2530.
- **IDE:** Emacs, Eclipse, VS Code, IAR, keil, Latex.
- **Related Courses:** OS, Advanced OS, Algorithms, Database, Wireless and Mobile Networks, Computer Architecture, Advanced Computer Networks, Machine Learning, Deep Learning