

CURRICULUM VITAE

DONG XIE

72 South Central Campus Drive, RM 2780 ◊ Salt Lake City, US 84112

Homepage: <http://www.cs.utah.edu/~dongx> ◊ **Email:** dongx@cs.utah.edu

EDUCATION

University of Utah, USA Aug. 2016 – Now
School of Computing, College of Engineering
Ph.D. in Computer Science

- Advisor: Prof. Feifei Li
- GPA: 4.0/4.0
- Thesis research: Distributed Analytics Engine for Big Spatio-Temporal Data
- Also working on utilizing oblivious RAM to tackle data privacy problems.

Shanghai Jiao Tong University, China Sept. 2011 – Jun. 2015
ACM Honored Class, Zhiyuan College
B.S.E. in Computer Science

- GPA: 3.71/4.3 (87.0/100)
- Got A+ in all project-oriented courses. (Compiler, Operating System and Database)
- Bachelor Thesis: Deterministic Main-Memory Database System. (Rank A.)

RESEARCH INTEREST

My research interests lie in distributed database system and large-scale data management, especially for big spatial-temporal data. I am also interested in main-memory databases, data privacy and security.

HONORS AND AWARDS

Microsoft Research Ph.D Fellow, 2018
Level C Undergraduate Scholarship, Shanghai Jiao Tong University, 2014
Bronze Prize, National Olympiad in Informatics, 2010
Bronze Prize, Asia Pacific Informatics Olympiad, 2010
First Prize, National Olympiad in Informatics (in Province), 2009
First Prize, National Olympiad in Informatics (in Province), 2008

RESEARCH EXPERIENCE

Data Group, University of Utah Sept. 2015 – Present
Research Assistant Salt Lake City, US

- Advised by Prof. Feifei Li
- ***Distributed Analytics Engine for Big Spatio-Temporal Data***
 - Designed and implemented an interactive analytics engine Simba based on Apache Spark, which extended Spark SQL for high-performance spatio-temporal data analysis.
 - Designed new algorithms for different point-based spatial operations.
 - Designed a general distributed framework for trajectory similarity search and implemented it with careful design choices and optimizations in Apache Spark.
- ***Data Privacy based on Oblivious RAM***
 - Did a comprehensive survey on existing Oblivious RAM techniques.
 - Implemented a unified testbed to evaluate practical performance of different ORAM schemes.
 - Exploring possibilities of utilizing ORAM to protect privacy of different queries.

Database Group, Microsoft Research Redmond
Research Intern

May. 2018 – Aug. 2018
Redmond, US

- Mentored by Badrish Chandramouli and Yinan Li
- Working on efficient semi-structured data processing.

DMX Group, Microsoft Research Redmond
Research Intern

May. 2017 – Aug. 2017
Redmond, US

- Mentored by Arnd Christian König, Vivek Narasayya and Manoj Syamala
- ***Multi-tenant Memory Governance across multiple SQL Server instances***
 - Studied the memory governance mechanism within a single SQL Server instance.
 - Proposed and implemented a scheme to evaluate the importance of memory pages across multiple SQL server instances and balance the memory pressure between them.

EPCC Lab, Shanghai Jiao Tong University
Research Assistant

Aug. 2013 – Aug. 2015
Shanghai, China

- Advised by Prof. Bin Yao
- ***Flexible Aggregate Similarity Search***
 - Studied exact and approximate algorithms for flexible aggregate similarity search queries in low and high dimensional spatial databases.
 - Implemented and optimized approximate algorithms for FANN queries.
- ***Practical Private Shortest Path Computation based on Oblivious Storage***
 - Optimized an existing OS scheme by designing a new oblivious shuffle algorithm.
 - Proposed two specific schemes for shortest path queries with a few novel optimizations.
- ***Deterministic Main-Memory Database System***
 - Introduced the concept of Strong Execution Determinism (SED) to main-memory databases.
 - Proposed a concurrency control design paradigm based on stored procedure analyzing.
 - Implemented a prototype main-memory database system and conducted experiments to show the efficiency of our solution.

Database System Group, National University of Singapore
Visiting Researcher

Aug. 2014 – Jan. 2015
Singapore

- Under E2S2 Research Programme (Joint Program between SJTU and NUS)
- Supervised by Prof. Beng Chin Ooi.
- ***Study and Implementation on Main-Memory Databases***
 - Did a comprehensive survey on main-memory and multi-core database in terms of architecture, concurrency control, locking schemes, durability and recovery.

PUBLICATIONS

Towards a Shared-Everything Database on Distributed Log-Structured Storage

by T. Zhu, Z. Zhao, F. Li, W. Qian, A. Zhou, **D. Xie**, R. Stutsman, H. Li, H. Hu
In Proc. of 2018 USENIX Annual Technical Conference (USENIX ATC 2018)

Distributed Trajectory Similarity Search

by **D. Xie**, F. Li, J. Phillips
In Proc. of 43rd International Conference on Very Large Data Bases (VLDB 2017)

Simba: Spatial In-Memory Big Data Analysis (Demo Paper)

by **D. Xie**, F. Li, B. Yao, G. Li, Z. Chen, L. Zhou, M. Guo
In Proc. of 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL 2016)

Simba: Efficient In-memory Spatial Analytics

by **D. Xie**, F. Li, B. Yao, G. Li, L. Zhou, M. Guo

In Proc. of 35th ACM International Conference on Management of Data (SIGMOD 2016)

Practical Private Shortest Path Computation based on Oblivious Storage

by **D. Xie**, G. Li, B. Yao, X. Wei, X. Xiao, Y. Gao, and M. Guo

In Proc. of 32nd IEEE International Conference on Data Engineering (ICDE 2016)

Oblivious RAM: A Dissection and Experimental Evaluation

by Z. Chang, **D. Xie**, F. Li

In Proc. of 42nd International Conference on Very Large Data Base (VLDB 2016)

Exact and Approximate Flexible Aggregate Similarity Search

by F. Li, K. Yi, Y. Tao, B. Yao, Y. Li, **D. Xie**, and M. Wang.

The International Journal on Very Large Databases (VLDBJ), 2016.

TEACHING EXPERIENCE

CS6450 Distributed System

Fall 2017

SJTU MS112 Design and Implementation of Compilers

Fall 2013

OPEN SOURCE PROJECTS

Traj-Sim-Spark: Distributed trajectory similarity search algorithms implemented on Spark

<https://github.com/InitialDLab/traj-sim-spark>

SEAL-ORAM: A unified testbed for evaluating different Oblivious RAM schemes.

<https://github.com/InitialDLab/SEAL-ORAM>

Simba: Distributed In-Memory Spatial Data Analytics System.

<https://github.com/InitialDLab/Simba>

IMDB-D: A prototype of main-memory database with Strong Execution Determinism (SED).

<https://github.com/Skyprophet/IMDB-D>

Compiler: Full-length compiler from simplified C to MIPS.

<https://github.com/Skyprophet/compiler2013>

Nachos: Simple operating system project for understanding OS principals.

<https://github.com/Skyprophet/nachos2013>

Fatworm: Full-length disk-based database management system built from scratch.

<https://github.com/Skyprophet/fatworm>

CPU: A five-level stage pipeline CPU implemented in Verilog HDL.

<https://github.com/Skyprophet/CPU>

Giti.me: Internal git service for ACM Honored Class.

<https://github.com/Skyprophet/Giti.me>

Mini-OJ: Portable online judge system for automatic program grading.

<https://github.com/Skyprophet/Mini-OJ>