

Dixin Tang

Address 5730 S Ellis Ave, JCL 299
Chicago, IL 60637

Homepage people.cs.uchicago.edu/~totemtang
Email totemtang@uchicago.edu

Research Areas

Query Processing, Adaptable Database, Transaction Processing

Education

- 2015-present Ph.D. Candidate in Computer Science - The University of Chicago
Advisor: Aaron Elmore
- 2011-2014 M.S. in Computer Science - Institute of Computing Technology, Chinese Academy of Sciences
Advisor: Wei Li
- 2007-2011 B.S. in Software Engineering - Huazhong University of Science & Technology

Industrial Experience

- **Internship at Microsoft Research** June. 2018-Sep. 2018
Project: Benchmarking Socrates Mentor: Umar Farooq Minhas

Socrates is a new cloud-native database that decouples computation from storage. My internship job is to test the new database architecture of Socrates in the industrial setting, understand its performance bottlenecks, and propose optimization opportunities.

Research Projects at UChicago

- **Intermittent Query Processing** Dec. 2017-Present

Our project studies how to process a standing query when the query is not necessary to be always active throughout its life time. We propose a new query processing paradigm Intermittent Query Processing, which sits between batch processing and continuous query processing. The new query processing paradigm raises many new problems about query execution engines, memory management, query optimization, and multi-query scheduling.
- **Adaptive Concurrency Control for Main-memory Database** Sep. 2015-Nov. 2017

We build a main-memory database that supports adaptively mixing multiple forms of concurrency control with minimal overhead. Our system can decompose the workload into partitions and selects a concurrency control protocol for each partition of workload that the protocol is optimized for, and during workload changes adaptively reconfigure the protocols online.

Earlier Projects

- **Structured Data Shuffling for Big Data Analytical Stacks** Nov. 2013-Jan. 2015

We build a structured data shuffling procedure that can leverage the semantics of SQL queries to apply efficient compression algorithms and discard unnecessary data during data shuffling.
- **A Fast and Space-efficient Join Method for Log Processing in MapReduce** Sep. 2012-Nov. 2013

We design a join method that achieves high query performance with a small extra storage cost for log processing. It shuffles the log table to avoid huge storage consumption and optimizes the shuffle procedure to achieve high query performance.

Publications

- **Dixin Tang**, Zechao Shang, Aaron J. Elmore, Sanjay Krishnan, Michael J. Franklin:
Intermittent Query Processing
PVLDB, 12(11): 1427-1441, 2019.
- Panagiotis Antonopoulos, Alex Budovski, Cristian Diaconu, Alejandro Hernandez Saenz, Jack Hu, Hanuma Kodavalla, Donald Kossmann, Umar Farooq Minhas, Naveen Prakash, Hugh Qu, Chaitanya Sreenivas Ravela, Krystyna Reisteter, Sheetal Shrotri, **Dixin Tang**, Vikram Wakade:
Socrates: The New SQL Server in the Cloud
SIGMOD Conference 2019: 1743-1756
- **Dixin Tang**, Aaron J. Elmore:
Toward Coordination-free and Reconfigurable Mixed Concurrency Control.
USENIX Annual Technical Conference 2018: 809-822
- **Dixin Tang**, Hao Jiang, Aaron J. Elmore:
Adaptive Concurrency Control: Despite the Looking Glass, One Concurrency Control Does Not Fit All.
CIDR 2017
- **Dixin Tang**, Taoying Liu, Rubao Lee, Hong Liu, Wei Li:
A Case Study of Optimizing Big Data Analytical Stacks Using Structured Data Shuffling.
BigData Congress 2016: 91-100
- Wenjuan Wang, Taoying Liu, **Dixin Tang**, Hong Liu, Wei Li, Rubao Lee:
SparkArray: An Array-Based Scientific Data Management System Built on Apache Spark.
NAS 2016: 1-10
- **Dixin Tang**, Taoying Liu, Rubao Lee, Hong Liu, Wei Li
A Case Study of Optimizing Big Data Analytical Stacks Using Structured Data Shuffling.
CLUSTER 2015: 70-73
- **Dixin Tang**, Taoying Liu, Hong Liu, Wei Li:
RHJoin: A Fast and Space-efficient Join Method for Log Processing in MapReduce.
ICPADS 2014: 975-980
- Liang Li, **Dixin Tang**, Taoying Liu, Hong Liu, Wei Li, Chenzhou Cui:
Optimizing the Join Operation on Hive to Accelerate Cross-Matching in Astronomy.
IPDPS Workshops 2014: 1735-1745

Honors & Awards

2018	USENIX ATC'18 Student Travel Grant
2016	University Unrestricted (UU) Fellowship - The University of Chicago
2016	CERES 1st year Graduate Research Award - The University of Chicago

Teaching Assistant

Fall 2015	MPCS 51040 - C programming
Spring 2016	MPCS 52040 - Distributed Systems
Winter 2017	CMSC 23500 - Introduction to Database
Winter 2018	CMSC 23500 - Introduction to Database

Referees

Name Aaron Elmore
Affiliate The University of Chicago
Position Assistant Professor
Contact aelmore@cs.uchicago.edu

Name Wei Li
Affiliate Institute of Computing Technology
Position Associate Professor
Contact liwei@ict.ac.cn

Name Umar Farooq Minhas
Affiliate Microsoft Research
Position Researcher
Contact ufminhas@microsoft.com