Dixin Tang

Address	1100 E 58th St, Ry177 Chicago, IL 60637	**Homepage** **Email**	people.cs.uchicago.edu/~totemtang 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

Honors & Awards

2016	University Unrestricted (UU) Fellowship - The University of Chicago
2016	CERES 1st year Graduate Research Award - The University of Chicago

Research Projects at UChicago

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

  We consider a scenario where queries are executed on incomplete or flawed datasets and the query results will be updated in response to future intermittent deltas. To efficiently process the deltas, we develop a dynamic programming algorithm to determine which intermediate states of the orignal query plan are kept or discarded given a memory constraint. Our initial results on PostgreSQL show that our algorithm has remarkable improvement over greedy solutions.

- **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 opitmized 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**, 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

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