

Yiqun Zhang

Address: Department of Computer Science
University of Houston
3551 Cullen Blvd., Room 501
Houston, TX 77204-3010, USA

Cell: +1 (832) 904 1425
E-mail: zhangyiqun9164@gmail.com

ACADEMIC & EDUCATION

Ph.D. candidate in Computer Science, since August 2013

- University of Houston, USA
- **Research interests:** Database Systems, Big Data Analytics, Parallel and scalable algorithms.
 - Integrating statistical and numerical methods with a DBMS;
 - Eliminating RAM limitations from mathematical software like R and Matlab;
 - Query optimization: recursive queries, cubes, join processing, pivoting;
 - Transaction processing.

Master of Science, Computer Science, May 2016

- University of Houston, USA
- Cumulative GPA: 3.8/4.0

B.Sc. Computer Science, July 2013

- Liaoning University, China

SKILLS

- **Programming skills:** C/C++, SQL, R, MPI, OpenMP, Java, Python, C#;
- **Database systems:** HP Vertica, SciDB, MySQL, SQL Server, Oracle;
- **Systems and tools:** Windows, Linux/UNIX, bash, Matlab, Visual Studio;

EXPERIENCE

Software Engineer, VoltDB Inc., October 2016 – present

Software Engineer Intern, VoltDB Inc., May to August in 2015 and 2016

Worked in the SQL team on system development tasks.

- Optimized query execution plans for aggregation queries.
- Optimized refreshing performance of materialized views.
- Building materialized views for supporting complex join queries.

Research Assistant, University of Houston, August 2013 – present

The Gamma matrix operator to summarize dense and sparse data sets in array database system

- Designed and implemented a parallel operator for data summarization in array database (SciDB) using C++;
- Applied the operator to accelerate the process of linear regression and PCA in R;
- Eliminated the RAM limitations of R with the help of the integration of our array operator.

Statistical and graph analysis integrating R and column-based database system

- Unified the data exchange interface between R and column-based database systems (SciDB and Vertica);
- Overloaded basic matrix operators in R to translate matrix operations to SQL queries;
- Achieved equal or even faster analysis speed compared with the R package even when data fit in RAM.

Teaching Assistant, University of Houston, August 2013 – present

- Teaching assistant for core courses: Database Systems, Data Structures and C Programming.
- Over 80% of the programming tasks in these courses are done in C++.

PROJECTS

Coin recognition using digital image processing approach (course project, using Matlab)

- Aimed to distinguish coins from the background and count the total denomination;
- Designed and implemented a ring projection approach to eliminate the effect of coin rotation;
- Trained the neuron network to identify the coin features.

HONORS & AWARDS

- Graduate Assistant Tuition Fellowship at University of Houston, since fall 2013.
- 1st Class Scholarship at Liaoning University from 2009-2012.

SELECTED PUBLICATIONS

“The Gamma Matrix to Summarize Dense and Sparse Data Sets for Big Data Analytics”
IEEE Transactions on Knowledge and Data Engineering (TKDE), 28(7): 1905-1918 (2016)

“Big Data Analytics Integrating a Parallel Columnar DBMS and the R Language”
IEEE CCGrid (Distributed Big Data Management workshop), 2016

“The Gamma Operator for Big Data Summarization on an Array DBMS”
Journal of Machine Learning Research (JMLR): Workshop and Conference Proceedings (BigMine 2014)