

Yuzhe (Richard) Tang

Addr: 329142, Atlanta, GA 30332, USA
Phone: +1-678-793-2706

Email: yztang@gatech.edu
Web: <http://www.cc.gatech.edu/~ytang36/>

Education

Georgia Institute of Technology

Aug. 2009 – May 2014 (expected)

Ph.D. in Computer Science

Adviser: **Ling Liu**

Thesis: Scalable and Secure Cloud Services in Big Data Systems

Fudan University (in China)

Sep. 2006 – Jun. 2009

M.Sc. in Computer Science

Adviser: Shuigeng Zhou

Fudan University (in China)

Sep. 2002 – Jun. 2006

B.Sc. in Computer Science

Work Experiences

Georgia Tech, Atlanta, GA USA

Aug/2009–present

Lead Research Assistant

This NSF-funded project, entitled “Privacy Preserving Index on Access-controlled Documents”, aims at developing a privacy preserving information networks, where access-controlled documents distributed cross multiple mutually-untrusted domains can be efficiently searched without sacrificing privacy. One result of this on-going project is a research paper published in CIKM’11.

- Proposed a secret-sharing based mechanism to securely compute the privacy preserving index (i.e. PPI).
- Proposed various computation models for different privacy concerns, including multi-keyword search and differentiated sensitivity.
- Implemented the proposed computation models based on a multi-party computation platform, FairplayMP.

IBM Research T.J. Watson, Yorktown

May/2013–Aug/2013

Heights, NY USA

Research Intern

The project, entitled “Authenticating Big Data Streams”, is to authenticate data stream in an outsourced-database scenario. One result of this on-going project is a system demo paper published in ICDE’14.

- Proposed a stream signing approach based on Bloom filters and Merkle tree.
- Implemented a prototype system for outsourcing stream based on HBase and Netty.

IBM Research T.J. Watson, Hawthorne, NY USA

May/2012–Aug/2012

Lead Research Intern

The project, entitled “Index Support on NoSQL”, is to enrich NoSQL store (e.g. HBase) with secondary index support. One result of this on-going project is a software integrated in the IBM’s BigInsights product and a paper published in EDBT’14.

- Proposed a real-time indexing framework that lends itself to the system of key-value stores. The idea is to take advantage of the fast-write-slow-read characteristic of the HBase alike key-value stores.
- Implemented a complete and functioning prototype system based on HBase’s CoProcessor Interface.
- Conducted experiments with a Cloud benchmark tool, YCSB and in a Cloud platform, Emulab. The experiments are fully automated.

IBM Research T.J. Watson, Hawthorne, NY USA

May/2011–Sep/2011

Research Intern

The project, entitled “System Optimization for Stream Processing”, is to optimize the performance of data stream processing dynamically on a multi-core system. The result of the project is a research paper published in TPDS’13.

- Identified the streaming system bottleneck to be a CPU utilization problem.
- Proposed a dynamic optimization technique by automatic pipelining the data stream to avoid the bottleneck scenario.
- Implemented the proposed algorithm with full integration into the IBM’s streaming platform, System S. The implementation requires Linux system programming.

NEC Labs America, Cupertino, CA USA

May/2010–Aug/2010

Research Intern

The project, entitled “Top- k Aggregation in Cloud”, is to support interactive query processing for top- k aggregations in the key-value store in cloud.

- Proposed an adaptive algorithm to maintain materialized views of top- k aggregations.

Fudan University, Shanghai China

Sep/2006–Sep/2009

Lead Research Assistant

The project, entitled “P2P Data Management”, aims at scaling out database over the distributed hash tables (i.e. DHT) in P2P networks. The results of the project are research papers published in TPDS’11, TKDE’10, ICDCS’09 and ICDCS’08.

- Proposed an original mapping mechanism for low-maintenance indexing over DHT without sacrificing query performance.
- Applied the proposed mapping mechanism to various data models including multi-dimensional data and range and k -NN queries, and proposed algorithms for query processing.

Microsoft Research Asia, Beijing China

Feb/2009–Jun/2009

Research Intern

The project, entitled “Log Analysis in Data Centers”, aims at automatic performance diagnosis and debugging of distributed systems in a large-scaled data center.

- Analyzed the system log and modeled the distributed system behavior by the acyclic function-call graphs.
- Designed various analysis operators (e.g. behavior-diff and clustering) by applying graph mining techniques (e.g., graph edit distances and metric space clustering).

IBM Development Lab, Shanghai China

Apr/2006–Sep/2006

Software Engineering Intern

The project, entitled “A Tool for Dynamic Code Review”, is to implement an automatic code review tool; it can help programmers to check their code against predefined programming rules by performing profile run of the program.

- Implemented an Eclipse plugin; used Eclipse’s TPTP and RCP to instrument the target client JVM process with the rule class.

Courses (at Georgia Tech)

Applications : Software Analysis and Testing, Computer System Security, Internet Computing and Application Development, Databases

Systems : Advanced Operating Systems, High Performance Computer Architecture

Theory : Randomized Algorithms

Business(Minor) : Principle Management for Engineers, Financial & Managerial Accounting I, Legal Issues-Technology Transfer

Selected Course Projects

- Intercepted system call using PTrace to enforce additional access controls in a FTP server.
- Implemented using secret sharing to securely store misspelling patterns, for tolerating password case misspelling.
- Administrated Hadoop over virtual machines for performance evaluation.
- Implemented Linux kernel module to capture read/write() system call.
- Implemented pointer analysis.

Skills

System: HBase/Hadoop/YCSB/Cassandra, MySQL/JDBC

Tool: *nix/POSIX , Emulab, Netty/ProtoBuf, L^AT_EX

Programming: Java/Ant, C/C++, Bash/Awk/Expect, Perl/Python

Honors and Awards

- Chinese Government Award for Outstanding Self-financed Students Abroad, 2012
- Best paper award, 5th International Conference on Cloud Computing, 2012 (co-recipient).
- Outstanding Master Thesis of Shanghai, Shanghai Government, 2010
- Tung's Oriental Scholarship, Tung's Oriental, 2008
- HP Distinguished Chinese Student Scholarship, Hewlett-Packard, 2008
- ICDCS Student Travel Grant, TCDP (IEEE Computer Society), 2008
- Graduate Student Fellowship of Fudan University, 2007-2008 (2 times)
- Outstanding Graduated Student of Fudan University, 2006
- Excellence Award, Tencent Innovation Contest, 2006
- The People's Scholarship of Fudan University, 2002–2006 (4 times)
- Chinese Physics Olympiads, First Prize in Hunan Province, 2001

References

Prof. **Ling Liu**, Professor, Adviser
College of Computing
Georgia Institute of Technology
Email: lingliu@cc.gatech.edu
Phone: +1-404-385-1139

Prof. **Mustaque Ahamad**, Professor
College of Computing
Georgia Institute of Technology
Email: mustaq@cc.gatech.edu
Phone: +1-404-894-2593

Prof. **Calton Pu**, Professor and John P. Imlay,
Jr. Chair in Software
College of Computing
Georgia Institute of Technology
Email: calton.pu@cc.gatech.edu
Phone: +1-404-385-1106

Prof. **Edward Omiecinska**, Associate Professor
College of Computing
Georgia Institute of Technology
Email: edwardo@cc.gatech.edu