

Tianrun Li

E-mail: tli278@wisc.edu

Education Background

Department of Computer Science and Technology, Tsinghua University	09/2012-06/2016
GPA: 91/100, Department Ranking: 9/116	
Department of Computer Science, University of Wisconsin-Madison	08/2016-present
Ph.D. student in Database Group	

Research Experiences

Research Assistant, Database Group, UW-Madison (08/2016-present) *Advisor: Dr. Jignesh Patel*

- Working on cyber threats tracing using graph analysis techniques;
- Working on performance tuning for graph analysis on relational database.

Research Intern, Imperial College London (07/2015-08/2015) *Advisor: Dr. Thomas Heinis*

- Proposed an approximate optimization of DBSCAN clustering algorithm using hashing, which is one order of magnitude faster than state-of-the-art approximate DBSCAN algorithm on high dimensional data [\[code\]](#);
- Proposed an optimization of Locality Sensitive Hashing using Maxeler Dataflow Engine, this project is now available at the [Maxeler Application Gallery](#).

Research Intern, Knowledge Engineering Group, Tsinghua (10/2014- 11/2015) *Advisor: Dr. Jie Tang*

- Similar document search using LDA and q-Gram for both semantic and literal similarity;
- Sentiment analysis for millions of Weibo (Chinese version of Twitter), using emotion icon as initial labeling of sentiment, obtained a good accuracy with no human supervision;
- Developed a knowledge graph generator to show the evolution between research topics using scientific citation network, which is part of the Aminer.org.

Internship

XuetangX, the largest MOOC platform in China (02/2016-06/2016) *Advisor: Dr. Xiu Li*

- Indexed user behavior log data for over 100 courses gathered by XuetangX, analyzed students' learning pattern in different courses;
- Developed two data visualization schemes of user behavior data, deployed them at the backstage course management system for MOOC teachers to monitor their class;
- Used python for data analysis and D3.js for visualization.

Projects

Cache-CPU	A MIPS32 computer architecture design. CPU hardware design in VHDL, a Unix-like OS written in C, a cross-compiler to generate executable code and various applications.
OurSQL	A single user relational DBMS written in C++. Support B+ tree index, integrity constraint and inner join on multiple tables, portable among Win, Mac and Linux.
THUSearch	A front-end to back-end search engine design based on Lucene. BM25 and Pagerank for ranking, word2vec for semantic search.

Awards

-
- Awarded CS Special Scholarship in University of Wisconsin-Madison 2016
 - Awarded the outstanding graduate student in Tsinghua University 2016
 - Awarded the National Scholarship of China (top 3 in CS Department) in 2014

Skills

-
- **Programming language:** C++(highly proficient), python, JAVA
 - **Skills:** database, data mining, machine learning, linux
 - **Github:** <https://github.com/Terranlee>