

Longbin Lai

☎ (+61) 420 600 521 | ✉ longbin.lai@gmail.com

Summary

An enthusiastic and dedicated programmer with exceptional work ethics. Interned in Google Infrastructure and Alibaba Cloud Computing. Over seven years' academic and engineering experiences with Hadoop (MapReduce), Spark and their variants. Proficient in Algorithms. Outstanding academic publications in big graph processing. Leadership in developing distributed graph processing system.

Skills

Programming	Python, Java, C/C++, Rust, SQL, Scala
Big Data	Hadoop, Spark, Timely dataflow system, Flink, AWS Infrastructure
Big Graph	Giraph/Pregel, GraphX, Gelly, Neo4J, GraphLab
Machine Learning	Tensorflow, Scikit-learn, Libsvm, Spark ML
Others	Recommendation Systems, Yii2 on Php, Django on Python

Experience

School of Computer Science and Software Engineering, East Normal China University

Shanghai, China

RESEARCH ASSISTANT

May. 2019 - Present

- Lead a team to develop graph pattern matching system.
- Research on large-scale graph database, especially distributed query processing and storage.

School of Computer Science and Engineering, UNSW

Sydney, Australia

RESEARCH ASSISTANT

May. 2017 - May. 2019

- Design and implement big graph processing primitives and languages.
- Lead a team to develop graph pattern matching system.

Google Inc.

Mountain View, CA, USA

TECH INTERN

Jan. 2017 - Apr. 2017

- Designed and implemented an emulator that simulates the Google backbone network and the routing strategies for testing, debugging and routing validation.

School of Computer Science and Engineering, UNSW

Sydney, Australia

PHD CANDIDATE, INDEPENDENT RESEARCH PROJECT

Jul. 2013 - May. 2017

- **(TwinTwigJoin)** Increased the performance of subgraph enumeration by up to an order of magnitude compared to the state-of-the-art by applying a decomposition-and-join framework in MapReduce.
- **(SEED)** Further improved the TwinTwigJoin by more than one order of magnitude by using a more advanced graph data storage mechanism (extending the traditional adjacency list) and an optimal join structure.

Department of Advertising and Searching, Alibaba Cloud Computing Corporation

HangZhou, China

RESEARCH INTERN, TEAM PROJECT

Jan. 2012 - Sep. 2012

- Designed and implemented a web recommendation system based on Alibaba cloud computing system (MapReduce-like system), which serves over 1000 top websites in China.
- Improved the throughput of the recommendation system to over 2 billion records per hour via a well-designed MapReduce data flow.
- Implemented a prototype of web classification algorithm that is twice faster than existing algorithm by solely using the url of the web page.

IBM Share-With-University Project

Shanghai, China

RESEARCH ASSISTANT, PROJECT LEADER

Oct. 2009 - Oct. 2010

- Saved the storage overhead of Hadoop File System (HDFS) by up to 30% without compromising the storage reliability by replacing the full replication mechanism with erasure coding.
- Improved the performance of Hadoop streaming utility (allowing coding with languages other than Java) by over 60% by replacing the synchronized inter-process communication module in Linux with desynchronized single-read-single-write queue.

Education

The University of New South Wales, Australia (UNSW)

Sydney, Australia

PH.D. IN COMPUTER SCIENCE

Jul. 2013 - May. 2017

- All courses Highly Distinguished

Shanghai Jiao Tong University (SJTU)

Shanghai, China

M.S. IN COMPUTER TECHNOLOGY

Sep. 2010 - Mar. 2013

- GPA 3.8 / 4.0, China's National scholarship, Top 2%

Shanghai Jiao Tong University (SJTU)

Shanghai, China

B.S. IN INFORMATION SECURITY

Sep. 2006 - Jun. 2010

- GPA 3.6 / 4.0, twice B-class SJTU academic scholarship, Top 15%

Publications

Scalable Distributed Subgraph Enumeration

VLDB 2017, ERA A*

Longbin Lai, Lu Qin, Xuemin Lin, Ying Zhang, Lijun Chang

To appear

Scalable Subgraph Enumeration in MapReduce, An Extension

VLDB Journal, ERA A*

Longbin Lai, Lu Qin, Xuemin Lin, Lijun Chang

Under Revision

Scalable Subgraph Enumeration in MapReduce

VLDB 2015, ERA A*

Longbin Lai, Lu Qin, Xuemin Lin, Lijun Chang

PVLDB Volume 8, Issue 10

ShmStreaming: A shared memory approach for improving Hadoop streaming performance

AINA 2013, ERA B

Longbin Lai, Jingyu Zhou, Long Zheng, Huakang Li, Yanchao Lu

Honors & Awards

2012	Top 1% , China's National Scholarship	<i>SJTU, China</i>
2011	Top 4% , Tencent Academic Scholarship	<i>SJTU, China</i>
2010	Top 10% , Outstanding Graduate of Shanghai Jiao Tong University	<i>SJTU, China</i>
2009	Top 6% , Sony Academic Scholarship	<i>SJTU, China</i>
07, 08	Top 15% , B-Class SJTU Academic Scholarship	<i>SJTU, China</i>