Longbin **Lai**

(+86) 18067950576 | longbin.lai@gmail.com

Summary

An enthusiastic and dedicated researcher and engineer with exceptional work ethics. Working and leading experiences in both academia and industry. Over ten years' academic and engineering experiences with Hadoop (MapReduce), Spark and their variants. Proficient in Algorithms. Outstanding academic publications in big graph/data processing. Leadership in developing distributed graph processing system.

Education

The University of New South Wales, Australia (UNSW) *Sydney, Australia*

PhD. in Computer Science *Jul. 2013 - present*

- 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%

Experience

Damo Academy, Alibaba Corporation *Hangzhou, China*

Staff Engineer *Dec. 2019 - Present*

- Research on large-scale data/graph processing system, graph query language, and query optimization.
- Lead a team to build Gaia, an interactive graph query system in the distributed runtime.
- Develop GraphScope (`https://github.com/alibaba/GraphScope`), a large-scale one-stop graph processing system, that incorporates processing the workloads of graph analytics, interactive graph queries, graph pattern matching and graph learning.

School of computer science and software engineering, East Normal China University *Shanghai, China*

Research Assistant *May. 2019 - Dec. 2019*

- 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.

Skills

Programming	Rust, Python, Java, C/C++, 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

Selected Publications

GraphScope: a unified engine for big graph processing

*VLDB 2021, ERA A**

Wenfei Fan, Tao He, Longbin Lai, Xue Li, Yong Li, Zhao Li, Zhengping Qian, Chao Tian, Lei Wang, etc.

PVLDB Volume 14, Issue 12

GAIA: A System for Interactive Analysis on Distributed Graphs Using a High-Level Language

*NSDI 2021, ERA A**

Zhengping Qian, Chenqiang Min, Longbin Lai, Yong Fang, Gaofeng Li, Youyang Yao, Bingqing Lyu, etc.

18th USENIX Symposium on Networked Systems Design and Implementation

A framework to quantify approximate simulation on graph data

*ICDE 2021, ERA A**

Xiaoshuang Chen, Longbin Lai, Lu Qin, Xuemin Lin, Boge Liu

2021 IEEE 37th International Conference on Data Engineering

Efficient structural node similarity computation on billion-scale graphs

*The VLDB Journal, ERA A**

Xiaoshuang Chen, Longbin Lai, Lu Qin, Xuemin Lin

Volume 30, Issue 3

Distributed subgraph matching on timely dataflow

*VLDB 2019, ERA A**

Longbin Lai, Zhu Qing, Zhengyi Yang, Xin Jin, Zhengmin Lai, Ran Wang, etc.

PVLDB Volume 12, Issue 10

Scalable Distributed Subgraph Enumeration

*VLDB 2017, ERA A**

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

PVLDB Volume 10, Issue 3

Scalable Subgraph Enumeration in MapReduce, a cost-oriented approach

*VLDB Journal, ERA A**

Longbin Lai, Lu Qin, Xuemmin Lin, Lijun Chang

The VLDB Journal, Volume 26, Issue 3

Scalable Subgraph Enumeration in MapReduce

*VLDB 2015, ERA A**

Longbin Lai, Lu Qin, Xuemmin 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

2021	**N/A,** HangZhou 521 Project of Talent Introduction	*Hangzhou, China*
2012	**Top 1%,** China's National Scholarship	*SJTU, China*
2011	**Top 4%,** Tencent Academic Scholarship	*SJTU, China*
2010	**Top 10%,** Outstranding Graduate of Shanghai Jiao Tong University	*SJTU, China*
2009	**Top 6%,** Sony Academic Scholarship	*SJTU, China*
07, 08	**Top 15%,** B-Class SJTU Academic Scholarship	*SJTU, China*