Yi Lu

Room 106, SHB, CUHK
(+852) 9608 3790
ylu@cse.cuhk.edu.hk
www.cse.cuhk.edu.hk/ỹlu/

Education

2013-Present **Master of Philosophy**, *Department of Computer Science and Engineering*, the Chinese University of Hong Kong.

Supervised by Prof. James Cheng

2009-2013 **Bachelor of Engineering**, *School of Computer Science and Technology*, Harbin Institute of Technology, China.

Research Interests

Distributed Graph Computing System Design and Implementation

Database and Graph Data Management

Publications

[1] **Yi Lu**, James Cheng, Da Yan, Huanhuan Wu, *Large-Scale Distributed Graph Computing Frameworks: An Experimental Evaluation*, In *Proc. of the VLDB Endowment (PVLDB), Volumn 8(3)*, Kohala, Hawaii, 2015.

[2] Da Yan, James Cheng, **Yi Lu**, Wilfred Ng, *Blogel: A Block-Centric Framework for Distributed Computation on Real-World Graphs*, In *Proc. of the VLDB Endowment (PVLDB), Volumn 7(14)*, Hangzhou, China, 2014.

[3] Da Yan, James Cheng, Kai Xing, **Yi Lu**, Wilfred Ng, Yingyi Bu, *Pregel Algorithms for Graph Connectivity Problems with Performance Guarantees*, In *Proc. of the VLDB Endowment (PVLDB), Volumn 7(14)*, Hangzhou, China, 2014.

[4] Huanhuan Wu, James Cheng, Silu Huang, Yiping Ke, **Yi Lu**, Yanyan Xu, *Path Problems in Temporal Graphs*, In *Proc. of the VLDB Endowment (PVLDB), Volumn 7(9)*, Hangzhou, China, 2014.

[5] Xiaohua Liu, Yitong Li, Haocheng Wu, Ming Zhou, Furu Wei, **Yi Lu**, *Entity linking for tweets*, In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, Sofia, Bulgaria, 2013.

Internship

July, 2014 – August, 2014 **Data Mining Research Intern**, *Alibaba Group*, Hangzhou, China.
Worked on large scale graph analytics problems in the Search and Recommendation Research Group. Pregel+ system was deployed to conduct large scale graph analytics algorithms.

July, 2012 – May, 2013 **Research Intern**, *Microsoft Research Asia*, Beijing, China.
Mentors: Dr. Xiaohua Liu (Researcher) and Dr. Ming Zhou (Principal Researcher)
Worked on people search and entity linking in the Natural Language Computing Group.

Projects

I focus on algorithm design and system implementation for large-scale graph computing. I am the core developer of the following projects.

Blogel `http://www.cse.cuhk.edu.hk/blogel/` Accepted by **PVLDB 14**

It is a block-centric framework, which naturally handles all the three adverse graph characteristics, (1)skewed degree distribution, (2)large diameter, and (3)(relatively) high density. Blogel programmers may think like a block and develop efficient algorithms for various graph problems. Our experiments on large real-world graphs verified that Blogel is able to achieve orders of magnitude performance improvements over the state-of-the-art distributed graph computing systems.

Pregel+ `http://www.cse.cuhk.edu.hk/pregelplus/` Accepted by **PVLDB 15**

Two effective message reduction techniques: (1)vertex mirroring and (2)a new request-respond paradigm are proposed in Pregel+, which can handle the communication bottleneck and the corresponding imbalanced workload of the original Pregel-like systems. Extensive experiments over various large real graphs show that Pregel+ is significantly more efficient than the state-of-the-art graph computing systems, especially for processing power-law graphs and (relatively) dense graphs.

Teaching Experience

2014 - 2015 Introduction to Database Systems, Advanced Topics in Database Systems

2013 - 2014 Data Structures, Introduction to Discrete Mathematics and Algorithms

2010 C Programming Language, Advanced Programming Language in C++

External Review

KDD 2014, PVLDB 2014, ICDE 2015

Awards

2013 – 2015 CUHK Postgraduate Studentship

2012 **First Prize in Province**, China Undergraduate Mathematical Contest in Modeling

2012 **Silver Medal**, ACM-ICPC Asia Jinhua Regional Invitational Contest

2011 **Silver Medal**, ACM-ICPC Asia Chengdu Regional Contest

2011 **Gold Medal**, ACM-ICPC China Northeast Multi-Provincial Programming Contest

2010 **First Class**, People's Scholarship

Standardized Tests

TOEFL Reading: 29, Listening: 28, Speaking: 23, Writing: 25. Total: 105

GRE Verbal: 152 (53%), Quantitative: 166 (93%), Analytical Writing: 3.5 (35%)

Skills

Languages C/C++, C#, Java, Python, Scala

Systems Hadoop, Giraph, GraphLab, GraphChi, Spark

Tools Git, OpenMPI

Last updated: October 27, 2014