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

Yi Lu

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

I am interested in the general areas of database and systems. My recent research focuses on distributed graph computing systems.

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), Pages 281-292, 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), Pages 1981-1992, 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), Pages 1821-1832, 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), Pages 721-732, 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

Under Review

Da Yan, James Cheng, Yi Lu, Wilfred Ng. **Effective Techniques for Message Reduction and Load Balancing in Distributed Graph Computation**. Submitted to *International World Wide Web (WWW) Conference*, 2015

Research Experience

July, 2014 - August, 2014 **Big Data Research Intern**, *Taobao Data Lab, Alibaba Group*, Hangzhou, China.
Deployed and applied the Pregel+ system in Taobao to develop large scale graph analytics algorithms. **Remarks:** As one of the original developers of Pregel+, I was sent to Taobao by Prof. Cheng to teach and lead the team to use Pregel+ in Taobao.

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 the design and implementation of systems and algorithms for large-scale graph computing. I am the core developer of the following projects.

Blogel `http://www.cse.cuhk.edu.hk/blogel/` **PVLDB'14, Vol. 7(14)**

Blogel 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 apply the "think like a block" programming paradigm to 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/` **PVLDB'15, Vol. 8(3)**

Pregel+ improves Blogel's messaging model by introducing two effective message reduction techniques: (1)vertex mirroring and (2)a new request-respond. These two techniques address the communication bottleneck and the corresponding imbalanced workload of existing 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 dense graphs.

Teaching Experience

Spring 2015	Advanced Topics in Database Systems
Fall 2014	Introduction to Database Systems
Spring 2013	Data Structures
Fall 2013	Introduction to Discrete Mathematics and Algorithms

External Review

SIGMOD'15, ICDE'15, DASFAA'15, KDD'14, PVLDB'14, DASFAA'14, WAIM'14

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 - 2012	HIT Undergraduate Scholarship

Standardized Tests

TOEFL	Reading: 29, Listening: 28, Speaking: 23, Writing: 25. Total: 105
GRE	Verbal: 152 (54%), Quantitative: 166 (92%), Analytical Writing: 3.5 (38%)

Skills

Languages	Proficient in C/C++, C#, Java, Python, Scala
Frameworks	Skillful in developing applications using Hadoop, Giraph, GraphLab, GraphChi, Spark

Last updated: November 23, 2014