

# A Streaming Graph Partitioning Approach on Imbalance Cluster

Yang Cao\*, Ruonan Rao\*\*

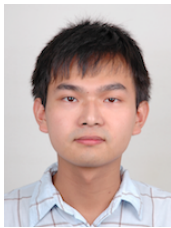
\* School of Software, SJTU, 800 Dong Chuan Road, Minhang, Shanghai, China, 200240

\*\* School of Software, SJTU, 800 Dong Chuan Road, Minhang, Shanghai, China, 200240

alex.nantong@gmail.com, rao-ruonan@cs.sjtu.edu.cn

**Abstract**—Distributed graph computing refers to extract knowledge by performing computations on large graphs. If the data source is continuously input like stream, the system is called streaming graph computing. When computing large graphs, a basic and significant step is to distribute the graph over a cluster of nodes, which is called ‘partition’. If the graph isn’t partitioned properly, the communication will quickly become a limiting factor in scaling the system, especially in streaming graph computing. And inside some cluster, the CPU speed and memory size of different nodes differs from each other. Observing that in this kind of cluster, nodes those has less resource limit the computing speed, we ask if the partition algorithm could be improved. We propose a simple heuristics to do partition in such cluster and compare the performance of some classic algorithms. It makes less cost of communication more efficient, and make better use of nodes those have more resources. Finally, we evaluate the performance gains in imbalance clusters by using our graph partition method to solve standard PageRank computing on a large real-world World-Wide-Web link graph. It shows that in such circumstance, our heuristics are a significant improvement.

**Keyword**—Graph Partition, Streaming, Imbalance Cluster, Distributed Graphs



**Yang Cao** (M’76–SM’81–F’87) This author was born in Nantong, China on 1990. On 2013, he graduated from Nanjing University with his bachelor degree of Software Engineering in Nanjing, China. He will get his master degree of Software Engineering on 2016 in Shanghai, China. His major field of study is distributed computing.



**Mr. Rao.** He is an associate professor in Shanghai Jiao Tong University in Shanghai, China. Before he came to Shanghai Jiao Tong University, he was a visiting scholar in Steven Institute of Technology. His research areas include distributed computing, cloud computing, service computing and middleware of Internet of Things.