**PowerLyra**: Differentiated Graph Computation and Partitioning on Skewed Graphs

Rong Chen, Jiaxin Shi, Yanzhe Chen and Haibo Chen

**Background and Motivation**

Existing graph-parallel systems usually use a "ONE size fits ALL" design that uniformly processes all vertices. However, this leads to high contention and load imbalance.

**Hybrid**

- **Low-degree Vertex**
  - Locality
  - `G` (all in-edges)
- **High-degree Vertex**
  - Parallelism
  - `H` (combined)

**Computation**

- Netflix Movie Recommendation [32]
- Replication Factor

**Graph Partitioning**

- Low-master
  - 4
- Low-mirror
  - 2
- High-master
  - 3
- High-mirror
  - 5

**Evaluation**

(Code and Instruction: [http://ipads.se.sjtu.edu.cn/projects/powerlyra.html](http://ipads.se.sjtu.edu.cn/projects/powerlyra.html))

**Performance**

- PageRank (VM-based 48-node Cluster)
- Replication Factor

**Communication**

- PR + Power-law
- Time (Sec)
- Memory Usage (GB)

**Scalability**

- 6-node Cluster (24 pCore, 64G RAM)
- 48-node Cluster (4 vCore, 12G RAM)

**Others**

- 6-node Cluster (24 pCore, 64G RAM)
- PR (10 iterations, 6-node Cluster)
- ALS (d=50)

**Institute of Parallel and Distributed Systems**

**MLDM**

- Big data processing
- IPADS, China

---

**GraphLab**

- Load Imbalance
- High Contention

**Pregel**

- make resource locally accessible
- evenly parallelize workload

**PowerGraph**

- Excessive Msgs
- Mem. Pressure