Multilayer networks in GraphLab

An open source project study

Mariana Marasoiu, R212

GraphLab

Execution engine: PowerGraph
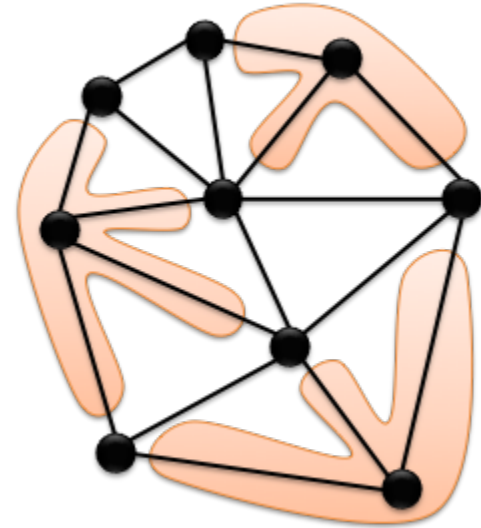
Used for graph processing

    vertex-centric functions

     distributed data graph

     vertex-cut partitioning

     vertex scopes

Good performance

Similar systems

Pregel, Giraph, GPS

GraphX → dataflow, on top of Spark

GraphLab	244s		GraphLab	249s
GraphX	251s		GraphX	419s
Giraph	200s		Giraph	596s

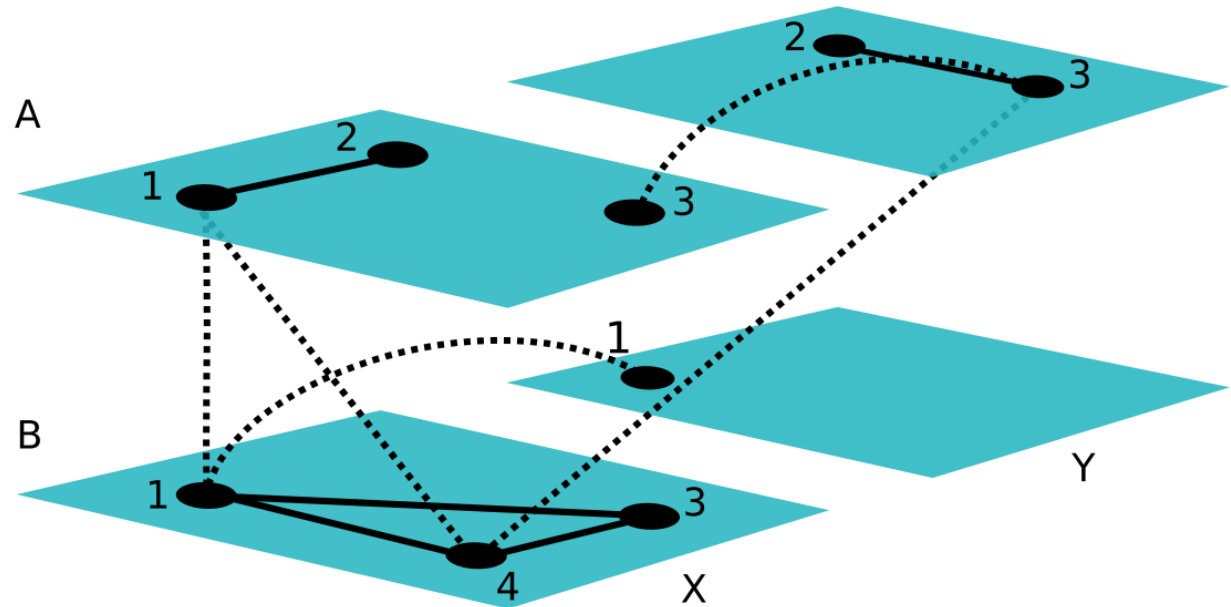twitter-2010, Connected Components   twitter-2010, PageRank, 20 iterations
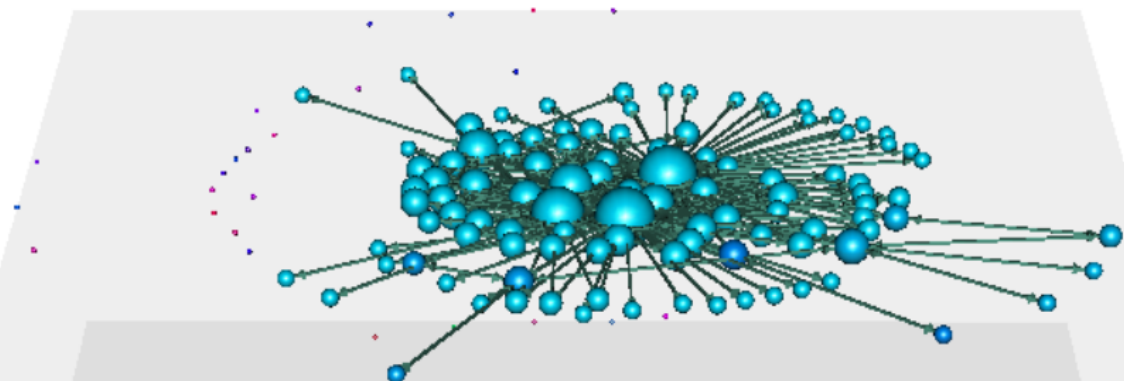
Multilayer networks

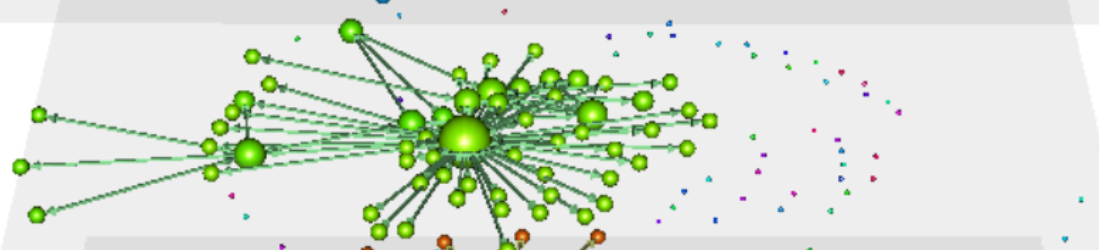social networks → multiplex networks
protein interaction networks
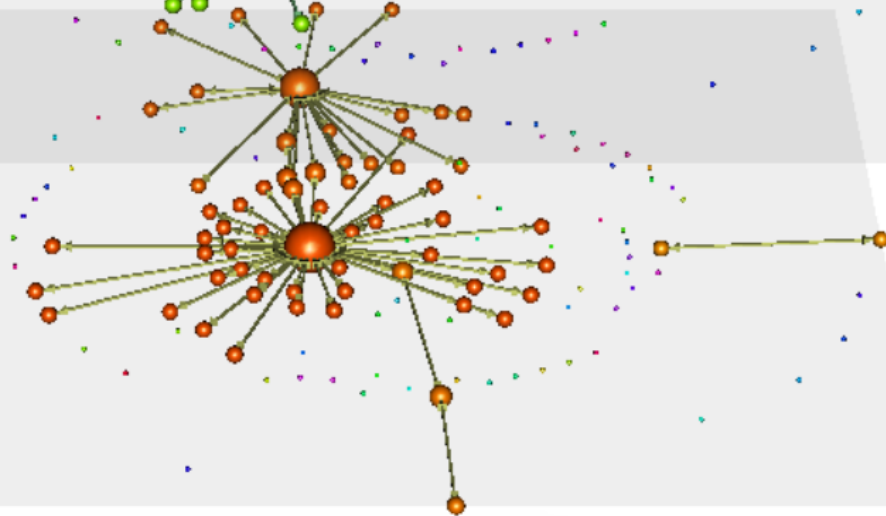transportation

Lufthansa

Air-France

British Airways

Multilayer Networks in GraphLab

What is the best representation of multilayer networks in GraphLab/PowerGraph?

- flat graph with coloured edges
- different vertices for each layer

Test both implementations with PageRank

Experiment and Evaluation

Experiments run on an Amazon EC2 cluster

Evaluate performance measuring

    runtime

    memory consumption

    network IO

Project steps

Acquire multilayer datasets ← working on it

Setup GraphLab on Amazon EC2

Preprocess datasets to fit GraphLab input

Implement PageRank

Run multiple experiments on EC2

Analyse results

Write report ☺