DBMS final project proposal

Graph Algorithms Library Extension as Neo4j plugins

R07902043 易大中 R08921089 蘇彥齊 R08944019 王郁婷 R08946007 蔡昀達

Motivation

The Neo4j Graph Algorithms library is developed to enable large scale graph analysis and to support machine learning pipelines. It is supported via online community and are actively developed and maintained. Our goal is to cover several widely used algorithms and offer highly parallelized implementation that works well with large scale graphs.

Related work

Neo4j Graphic Algorithm https://neo4j.com/docs/graph-algorithms/current/labs-algorithms/
The library contains implementations of the following types of algorithms:

- **Path Finding** – find the shortest path or evaluate the availability of routes
- **Centrality** – determines the importance of distinct nodes in a network
- **Community Detection** – these algorithms evaluate how a group is clustered or partitioned, as well as its tendency to strengthen or break apart
- **Similarity** – these algorithms help calculate the similarity of nodes

Plan

我們預計implement以下幾個algorithm:

- Centrality
  - Group betweenness centrality
  - Group closeness centrality
  - Group degree centrality
  - Reaching centrality
  - Katz centrality
- Link prediction
  - Jaccard coefficient
- Community
  - Attracting components
  - Biconnected components
- Path finding
  - Simple path
  - Beam search
  - Cycle finding

Group betweenness centrality of a group of nodes.
Group closeness centrality of a group of nodes.
Group degree centrality of a group of nodes.
Counts the number of nodes each node can reach in k or less steps.
Measures the number of all nodes that can be connected through a path

Timeline

- 11/20 熟悉框架&developer guideline
- 11/28 分工實作
- 12/18 Deploy as extension plugins
- 12/30 Project presentations
- 1/9 Submission Deadline