

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

In many machine learning applications, the relation between nodes are not predefined but uses distanced-based relation type instead. Currently, the Neo4j Graph Algorithms library supports a distanced-based relation type and similarity search that provides very efficient search time but suffers from a high complexity index build-time and once the relation is build it cannot update itself if a new node is inserted into the graph later on.

Goal

Our goal is to cover several widely used algorithms and offer highly parallelized implementation that works well with large scale graphs. Additionally, inspired by several very recent algorithms designed for distributed and online learning, we aim to build a new dynamic distance-based relation type with efficient online similarity search algorithm.

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

Online similarity search : Mansfield, Philip Andrew, et al. "Links: A high-dimensional online clustering method." arXiv preprint arXiv:1801.10123 (2018).

Plan

我們預計實作以下：

- RelationType
 - Approximate nearest neighbor
 - Online similarity search
- Centrality
 - Group betweenness centrality
 - Group closeness centrality
 - Group degree centrality
 - Reaching centrality
- Katz centrality
- Link prediction
 - Jaccard coefficient
- Community
 - Biconnected components
- Path finding
 - Simple 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