Keyword-Centric Community Search

PI: Dr. ZHANG Zhiwei

OBJECTIVES
1. To formally analyze the differences and benefits of a new framework based on graph contraction compared to existing frameworks.
2. To design new algorithms for basic graph operations, including node-centric graph operations and edge-centric graph operations.
3. To efficiently integrate these developed basic graph algorithms for graph applications.

HIGHLIGHTS

Problem Statement
- Community Search Query with Keywords in Large Graphs
  - ACQ query aim to compute the number of keywords in the query shared by all vertices.
  - ATC query aim to minimize the distance from each node to the query vertex.
  - Existing solutions consider the keyword relationships and graph structure separately.
  - The time cost is high.

Keyword-Centric Community Search
- Keyword-Structure Measurement
  - Given a graph \( G \), a set of keywords \( W \) and an integer \( k \):
    - The community to find is a \( k \)-core,
    - **Keyword closeness**: there does not exist any subgraph \( H' \), such that \( H \) is minimal among all the subgraphs \( H' \not\subseteq H \), and \( KwdC(W, H') < KwdC(W, H) \).
    - \( KwdC(W, H) \): largest value of the shortest distance from each vertex in \( H \) to each keyword in \( W \).

Our Approach
- **Idea**: iteratively update \( cd^k(u) \) according to the keyword-closeness and structural cohesiveness conditions.
  - Given \( d' = cd^k(u) \): \( u \) is only possible belong to \( G = KC^d \) with \( d \geq d' \).
  - Theorem: if for each node with \( cd^k(u) \leq d \), \( cd^k(u) \) cannot be updated, then \( G = KC^d \) is the induced graph of the nodes with \( cd^k \leq d \).
- Algorithm
  - (recursively) compute the \( k \)-core of the induced graph of the node with \( cd^k \leq d \).
  - update \( cd^k \)
  - return the induced subgraph of the nodes with the minimum \( d \) at the condition that the nodes with \( cd^k \leq d \) will not change their \( cd^k \) value.

Performance Studies
- **Effectiveness**
- **Efficiency**

SELECTED PUBLICATIONS