香港浸會大學理學院
HKBU Faculty of Science

DEPARTMENT OF COMPUTER SCIENCE

PhD Degree Oral Presentation

PhD Candidate:	Mr. Jiaxin JIANG
Date	August 21, 2020 (Friday)
Time:	10:00 am – 12:00 pm (35 mins presentation and 15 mins Q & A)
Venue:	Zoom ID: 99340522517 (The password and direct link will only be provided to registrants)
Registration:	https://bit.ly/sem-zm (Deadline: 5:00pm, 20 August 2020)

Efficient Frameworks for Keyword Search on Massive Graphs

Abstract

Due to the unstructuredness and the lack of schema information of knowledge graphs, social networks and RDF graphs, keyword search has been proposed for querying such graphs/networks. Recently, various keyword search semantics have been designed. However, these keyword search semantics and algorithms encounter efficiency or scalability issues. In this thesis, we propose new three generic frameworks or index techniques to address these issues. The thesis results show that the keyword search on massive graphs under different scenarios can be effective and efficient, which would facilitate keyword search services on graphs in the real world.

First, we study the keyword search on massive knowledge graphs. In particular, we propose a generic ontology-based indexing framework for keyword search, called Bisimulation of Generalized Graph Index (BiG-index), to enhance the search performance. The novelties of BiG-index reside in using an ontology graph GOnt to summarize and index a data graph G iteratively, to form a hierarchical index structure G. Regarding query evaluation, we transform a keyword search q into Q according to GOnt in runtime. The transformed query is searched on the summary graphs in G. The efficiency is due to the small sizes of the summary graphs and the early pruning of semantically irrelevant subgraphs. To illustrate BiG-index's applicability, we show popular indexing techniques for keyword search can be easily implemented on top of BiG-index. Our extensive experiments show that BiG-index clearly reduced the runtimes of popular keyword search algorithms.

Second, we study the problem of keyword search on public-private graph. In many applications (e.g., social networks), users may prefer to hide parts or all of her/his data graphs (e.g., private friendships) from the public. This leads to a recent graph model, namely the public-private network model, in which each user has his/her own network. While there have been studies on public-private network analysis, keyword search on public-private networks has not yet been studied. Hence, we propose a new keyword search framework, called public-private keyword search (PPKWS). PPKWS consists of three major steps: partial evaluation, answer refinement, and answer completion. We select three representative ones and show that they can be implemented on the model with minor modifications. We propose indexes and optimizations for PPKWS. We have verified through experiments that, on average, the algorithms implemented on top of PPKWS run 113 times faster than the original algorithms directly running on the public network attached to the private network for retrieving answers that span through them.

Third, we study the keyword search in distributed graph evaluation systems. In the recent research on query evaluation, parallel evaluation has attracted much interest. However, the study on keyword search on distributed graphs has still been limited. We propose a novel distributed keyword search framework called DKWS. We propose a notify-push paradigm which can exchange the upper bounds of answers across all the workers asynchronously. In particular, the workers notify the coordinator when the local upper bound is refined. The coordinator pushes the refined global upper bound to all the workers. Moreover, we propose an efficient and generic keyword search algorithm for the workers. We have implemented DKWS on top of GRAPE, a distributed graph process system from our previous research collaboration. Extensive experimental results show that DKWS outperforms current-state-of-art techniques.

***** ALL INTERESTED ARE WELCOME *****