University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Andreas Gruhler

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, October 10, 2014

Vertiefung
Topic: Graph Indexing for Efficient Graph Database Query Processing

In the core of many graph-related applications, there is a common and critical problem: *how to efficient process graph queries and retrieve related graphs.* The classical graph query problem can be described as follows: *Given a graph database* $D = \{g_1, g_2, \cdots, g_n\}$ *and a graph query* q*, find all the graphs in which* q *is a subgraph.* Various indexing methods have been developed to process graph queries, the majority of which can be categorized as *path-based indexing* approach, e.g. *GraphGrep* [1]. Yan et al. [2] proposed a graph indexing method named *gIndex* that, instead of taking *path* as the basic indexing unit, makes use of *frequent subgraphs* as the basic indexing feature. Performance study shows that *gIndex* has smaller index size, but achieves better performance in comparison with *GraphGrep*.

In this project, student should carefully study *gIndex* and *GraphGrep* methods, including two major issues: (1) *index construction* that enumerates and selects frequent subgraphs, and (2) *query processing* that searches and verifies all subgraphs of a query q in database D.

Tasks

1. Study *gIndex* and *GraphGrep* methods proposed in [2, 1], especially on index construction, maintenance, and query processing.
2. Compare the two methods in a written report, including algorithms for index construc-

tion and query processing, discussion on time and space complexity, advantages and disadvantages of the two approaches.

3. Explain the above issues in an oral exam.

Supervisor:	Pei Li (peili@ifi.uzh.ch)
Start date:	13.10.2014
End date:	15.12.2014
Duration:	2 months

University of Zürich
Department of Informatics

Prof. Dr. Michael Böhlen

References

[1] Dennis Shasha, Jason T. L. Wang, and Rosalba Giugno. Algorithmics and applications of tree and graph searching. In *PODS '02*, pages 39–52.

[2] Xifeng Yan, Philip S. Yu, and Jiawei Han. Graph indexing: A frequent structure-based approach. In *SIGMOD '04*, pages 335–346.