

DOCTORAL THESIS

Efficient group queries in location-based social networks

Li, Yafei

Date of Award:
2015

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

Abstract

Nowadays, with the rapid development of GPS-equipped mobile devices, location-based social networks have been emerging to bridge the gap between the physical world and online social networking services. Various types of data, such as personal locations, check-ins, microblogs and social relations, have been available in location-based social networks. Efficiently managing and analyzing such data to meet users' daily query requirements become a challenging task. Among all the existing works in location-based social networks, group query is one of the most important research topics. In this thesis, we investigate query techniques for location-based services in social networking applications. Specifically, considering a location-based social network, we study spatial-aware interest group queries, geo-social k -cover group queries, and social-aware ridesharing group queries.

Firstly, we study the spatial-aware interest group queries in location-based social networks. Recently, most of the location-based social networks release check-in services that allow users to share their visiting locations with their friends. These locations, considered as spatial objects, are usually associated with a few tags that describe the features of those locations. Utilizing such information, we propose a new type of *Spatial-aware Interest Group* (SIG) query that retrieves a user group of size k where each user is interested in the query keywords and the users are close to each other in the Euclidean space. We prove this query problem is NP-complete, and develop two efficient algorithms IOAIR and DOAIR based on the IR-tree for the processing of SIG queries. We also validate the performance efficiency of the proposed query processing algorithms by empirical evaluation.

Secondly, we study the problem of geo-social k -cover group queries for collaborative

spatial computing. In this problem, we propose a novel type of geo-social queries, called *Geo-Social K-Cover Group (GSKCG)* query, which is based on spatial containment and a new modeling of social relationships. Intuitively, given a set of spatial query points and an underlying social network, a GSKCG query finds a minimum user group in which the members satisfy certain social relationship and their associated regions can jointly cover all the query points. Albeit its practical usefulness, the GSKCG query problem is NP-complete. We consequently explore a set of effective pruning strategies to derive an efficient algorithm for finding the optimal solution. Moreover, we design a novel index structure tailored to our problem to further accelerate query processing. Extensive experiments demonstrate that our algorithm achieves desirable performance on real-life datasets.

Thirdly, we study the problem of social-aware ridesharing group queries. With the deep penetration of smartphones and geo-locating devices, ridesharing is envisioned as a promising solution to transportation-related problems such as congestion and air pollution for metropolitan cities. Despite the potential to provide significant societal and environmental benefits, ridesharing has not so far been as popular as expected. Notable barriers include the social discomfort and safety concerns when traveling with strangers. To overcome these barriers, in this thesis, we propose a new type of *Social-aware Ridesharing Group (SaRG)* query which retrieves a group of riders by taking into account their social connections besides traditional spatial proximities. Because the SaRG query problem is NP-hard, we design an efficient algorithm with a set of powerful pruning techniques to tackle this problem. We also present several incremental strategies to accelerate the search speed by reducing the repeated computations. Moreover, we propose a novel index tailored to the proposed problem to further speed up the query processing. Experimental results on real datasets show that our proposed algorithms achieve desirable performance.

The works of this thesis show that the group query processing techniques are effective, which would facilitate the wider deployment of such query services in real applications.

Keywords: Location-based services, Query processing, Group queries, Indexing, Spatial database, Location-based social networks, Social constraints, Ridesharing.

Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	viii
List of Figures	ix
Chapter 1 Introduction	1
1.1 Spatial-aware Interest Group Queries	3
1.2 Geo-Social K-Cover Group Queries	4
1.3 Social-aware Ridesharing Group Queries	6
1.4 Thesis Organization	8
Chapter 2 Related Works	9
2.1 Spatial query processing	9
2.2 Social query processing	10
2.3 Spatial keyword query processing	11
2.4 Geo-Social query processing	13
2.5 Ridesharing query processing	14

Chapter 3	Spatial-aware Interest Group Queries in Location-based Social Networks	17
3.1	Problem Definition	17
3.2	Proposed Approaches	21
3.2.1	Preliminary: IR-Tree	21
3.2.2	Overview	22
3.2.3	Interest Oriented Algorithm	23
3.2.4	Diameter Oriented Algorithm	29
3.3	Performance Evaluation	36
3.3.1	Datasets and Queries	36
3.3.2	Setup	37
3.3.3	Experimental Results	37
3.4	Summary	42
Chapter 4	Geo-Social K-Cover Group Queries for Collaborative Spatial Computing	43
4.1	Problem Formulation	43
4.2	Algorithm Design	47
4.2.1	Basic Algorithm	47
4.2.2	Basic Pruning	49
4.2.3	Diameter Based Pruning	51
4.2.4	Access Order Based Pruning	54
4.3	Hybrid Indexing	55
4.3.1	SaR-tree	55
4.3.2	Enhanced SaR-tree	56
4.3.3	GSKCG Query Processing	62
4.4	Performance Evaluation	63
4.4.1	Datasets and Queries	64
4.4.2	Setup	64
4.4.3	Experimental Results	65

4.5	Summary	69
Chapter 5	Towards Social-aware Ridesharing Group Query Services	70
5.1	Problem Formulation	70
5.2	Algorithm Design	75
5.2.1	RSGExplorer Algorithm	75
5.2.2	Quota Available Strategy	81
5.2.3	Group Diameter Strategy	83
5.2.4	\bar{k} -plex Based Strategy	84
5.2.5	Query Processing	85
5.3	Incremental Strategies	87
5.3.1	Incremental Computation of Core Numbers	87
5.3.2	Social Diameter-based Bounding	88
5.3.3	Neighbor-based Bounding	89
5.4	Hybrid Index	91
5.4.1	SIR-tree	91
5.4.2	Query Processing	92
5.5	Performance Evaluation	93
5.5.1	Experimental Settings	93
5.5.2	Experimental Results	93
5.6	Summary	97
Chapter 6	Conclusions and Future Work	98
6.1	Conclusions	98
6.2	Future Work	99
	Bibliography	101
	Curriculum Vitae	109