RGC Ref.: N_HKUST637/13 NSFC Ref. : 61361166009
(please insert ref. above)

The Research Grants Council of Hong Kong
NSFC/RGC Joint Research Scheme
Joint Completion Report

(Please attach a copy of the completion report submitted to the NSFC by the Mainland researcher)

Part A: The Project and Investigator(s)

1. Project Title

Effective Location-based Spatial Crowdsourcing

2. Investigator(s) and Academic Department/Units Involved

	Hong Kong Team	Mainland Team
Name of Principal Investigator *(with title)*	Prof. Lei Chen	Prof. Yunhao Liu
Post	Associate Professor	Professor
Unit / Department / Institution	Department of Computer Science and Engineering, HKUST	School of Software, Tsinghua University
Co-investigator(s) *(with title)*	Prof. Cyrus Shahabi (Dept. of CS, University of Southern California)	

3. Project Duration

	Original	Revised	Date of RGC/ Institution Approval *(must be quoted)*
Project Start date	01/01/2014		
Project Completion date	31/12/2017		
Duration *(in month)*	48		
Deadline for submission of Joint Completion Report	31/12/2017		

Part B: The Completion Report

5. Project Objectives

5.1 Objectives as per original application

1. Design effective incentive mechanisms to encourage mobile device users to participate in crowdsourcing tasks.
2. Develop tools to conduct automatic user profile mining and design user latency models.
3. Design near-optimal solutions to assign tasks to registered workers.
4. Propose an effective model to integrate data received from different mobile workers.
5. Develop novel movement pattern-based data quality control mechanisms.
6. Implement a prototype system for general spatial crowdsourcing tasks.

5.2 Revised Objectives

Date of approval from the RGC: _____

Reasons for the change: _____

_____

1.
2.
3. ....

6. Research Outcome

Major findings and research outcome
(maximum 1 page; please make reference to Part C where necessary)

During this reporting period, we mainly focused on finding effective mechanisms to select the correct questions to find high quality mapping and find high quality data sources in crowdsourcing. We have proposed a market strategy to find the high quality mappings, the work was published in KDD 2014. We have design a mechanism to select the proper questions to ask in a crowd-aided routing planning system, the work was published in ICDE 2014. We have proposed a general framework to answer queries via human and machine collaboration, which was published in CIKM 2014. Moreover, based on collected crowdsourced data, we have developed streaming topic detection method, which was published in KDD 2014. In ICDE 2015, we have proposed a solution to select the critical questions to ask the crowd to clean the noisy data. Based on the theoretical findings, we have built a preliminary spatial crowdsourcing platform, called gMission. We have demonstrated

gMission in VLDB 2014 and won the excellent demonstration award. Recently, based on the gMission platform, we have developed solutions to fulfill the task assignment, ask routing questions and aggregated diverse answers, three works were all published in VLDB 2015. With consideration of the skills of workers to form a team accomplishing a specific task, we have developed task assignment methods to assign groups of workers to satisfy the required skill sets of tasks, and one work was published in TKDE 2016 and one work was published in DSE 2017. To study the task assignment methods under online settings, we have developed methods to assign a suitable worker to a newly coming task to maximizing the global optimization goals, and we have 4 papers published in VLDB 2016, TKDE 2016, ICDE 2016 and ICDE 2017. In addition, we have proposed frameworks to manage the budget during task assignment such that the objective goals can be maximized with the minimized budget, and one work was published in BigData 2016 and one work was published in ICDE 2017. To further improve the results, we also studied methods to predict the future worker supply and demand in areas, which was published in KDD 2017. With the prediction methods, we also developed corresponding prediction-based task assignment frameworks to improve the final results, which was published in ICDE 2017. Furthermore, we also considered to benefit both the workers and task requesters during task assignment processes and proposed mutual benefit aware task assignment methods, which was published in ICDE 2016. When the recommended tasks can be rejected by workers, we studied how to utilize the rejection information to maximize the acceptance rate during task assignment, which was published in TKDE 2017.

7. The Layman's Summary

(describe <u>in layman's language</u> the nature, significance and value of the research project, in no more than 200 words)

In this project, we have investigated solutions for three challenges in spatial crowdsourcing, which are finding reliable sources, investigating correct mapping and conducting data cleaning. The results can be applied to any spatial crowdsourcing platform to improve its effectiveness and efficiency.

<u>Part C: Research Output</u>

8. Peer-reviewed journal publication(s) arising <u>directly</u> from this research project

(Please attach a copy of each publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				Author(s)	Title and	Submitted	Attached	Acknowledged
Year of publication	Year of Acceptance *(For paper accepted but not yet published)*	Under Review	Under Preparation *(optional)*	***bold** the authors belonging to the project teams <u>and</u> denote the corresponding author with an asterisk*)*	Journal/Book *(with the volume, pages and other necessary publishing details specified)*	to RGC *(indicate the year ending of the relevant progress report)*	to this report *(Yes or No)*	the support of this Joint Research Scheme *(Yes or No)*

2017				Dawei Gao, Yongxin Tong*, Jieying She, Tianshu Song, **Lei Chen** and Ke Xu	Top-k team recommendation and its variants in spatial crowdsourcing. Data Science and Engineering, 2(2): 136-150 (2017)	2018	Yes	Yes
2017				Rui Meng, **Lei Chen***, Yongxin Tong, and Chen Zhang	Knowledge Base Semantic Integration Using Crowdsourcing. IEEE Trans. Knowl. Data Eng., 29(5): 1087-1100 (2017)	2018	Yes	Yes
2017				Libin Zheng and **Lei Chen***	Maximizing acceptance in rejection-aware spatial crowdsourcing. IEEE Trans. Knowl. Data Eng., 29(9): 1943-1956 (2017)	2018	Yes	Yes
2016				Xiang Lian, **Lei Chen***, and Guoren Wang	Quality-aware subgraph matching over inconsistent probabilistic graph databases. IEEE Trans. Knowl. Data Eng., 28(6): 1560-1574 (2016)	2018	Yes	Yes
2016				Peng Cheng*, Xiang Lian, **Lei Chen**, Jinsong Han and Jizhong Zhao	Task assignment on multi-skill oriented spatial crowdsourcing. IEEE Trans. Knowl. Data Eng., 28(8): 2201-2215 (2016)	2018	Yes	Yes

2016				Jieying She, Yongxin Tong*, **Lei Chen*** and Caleb Chen Cao	Conflict-aware event-participant arrangement and its variant for online setting. IEEE Trans. Knowl. Data Eng., 28(9): 2281-2295 (2016)	2018	Yes	Yes
2016				Anand Inasu Chittilappilly, **Lei Chen*** and Sihem Amer-Yahia	A survey of general-purpose crowdsourcing techniques. IEEE Trans. Knowl. Data Eng., 28(9): 2246-2266 (2016)	2018	Yes	Yes
2016				Wei Gong, Haoxiang Liu, **Lei Chen**, Kebin Liu, and **Yunhao Liu***	Fast composite counting in RFID systems. IEEE/ACM Transactions on Networking, 24(5): 2756-2767 (2016)	2018	Yes	Yes
2015				Xiang Lian, **Lei Chen***, Zi Huang	Keyword Search Over Probabilistic RDF Graphs. IEEE Trans. Knowl. Data Eng. 27(5): 1246-1260 (2015)	2015	No	Yes
2015				Xiping Liu, **Lei Chen***, Changxuan Wan	LINQ: A Framework for Location-Aware Indexing and Query Processing. IEEE Trans. Knowl. Data Eng. 27(5): 1288-1300 (2015)	2015	No	Yes

2015				Ning Xu, Bin Cui, **Lei Chen***, Zi Huang, Yingxia Shao	Heterogeneous Environment Aware Streaming Graph Partitioning. IEEE Trans. Knowl. Data Eng. 27(6): 1560-1572 (2015)	2015	No	Yes
2015				Ye Yuan, Guoren Wang**, Lei Chen***, Haixun Wang	Graph similarity search on large uncertain graph databases. VLDB J. 24(2): 271-296 (2015)	2015	No	Yes
2015				Ye Yuan, Guoren Wang, Jeffery Yu Xu, **Lei Chen***	Efficient distributed subgraph similarity matching. VLDB J. 24(3): 369-394 (2015)	2015	No	Yes
2014				Xiang Lian, **Lei Chen***	Trip Planner Over Probabilistic Time-Dependent Road Networks. IEEE Trans. Knowl. Data Eng. 26(8): 2058-2071 (2014)	2015	No	Yes

9. Recognized international conference(s) in which paper(s) related to this research project was/were delivered *(Please attach a copy of each delivered paper. All listed papers must acknowledge RGC's funding support by quoting the specific grant reference.)*

Month/Year/ Place	Title	Conference Name	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this report *(Yes or No)*	Acknowledged the support of this Joint Research Scheme *(Yes or No)*
04/2014/Chicago	CrowdPlanner: A crowd-based route recommendation system.	ICDE, 2014	2015	No	Yes
08/2014/New York	TCS: efficient topic discovery over crowd-oriented service data	KDD, 2014	2015	No	Yes

08/2014/New York	From labor to trader: opinion elicitation via online crowds as a market	KDD, 2014	2015	No	Yes
06/2014/Snowbird	Efficient cohesive subgraphs detection in parallel	SIGMOD, 2014	2015	No	Yes
06/2014/Snowbird	Parallel subgraph listing in a large-scale graph	SIGMOD, 2014	2015	No	Yes
11/2014/Shanghai	MaC: A Probabilistic Framework for Query Answering with Machine-Crowd Collaboration	CIKM,2014	2015	No	Yes
11/2014/Shanghai	Efficient Probabilistic Supergraph Search Over Large Uncertain Graphs	CIKM,2014	2015	No	Yes
09/2014/Hangzhou	gMission: A General Spatial Crowdsourcing Platform	VLDB,2014	2015	No	Yes
09/2014/Hangzhou	Repairing Vertex Labels under Neighborhood Constraints	VLDB,2014	2015	No	Yes
09/2014/Hangzhou	On Concise Set of Relative Candidate Keys.	VLDB,2014	2015	No	Yes
06/2015/Melbourne	Utility-Aware Social Event-Participant Planning	SIGMOD, 2015	2015	No	Yes
09/2015/Hawaii	Where To: Crowd-Aided Path Selection	VLDB, 2015	2015	No	Yes
09/2015/Hawaii	Hear the Whole Story: Towards the Diversity of Opinion in Crowdsourcing Markets	VLDB, 2015	2015	No	Yes
09/2015/Hawaii	An Efficient Similarity Search Framework for SimRank over Large Dynamic Graphs.	VLDB, 2015	2015	No	Yes
09/2015/Hawaii	Reliable Diversity-Based Spatial Crowdsourcing by Moving Workers	VLDB, 2015	2015	No	Yes
09/2015/Hawaii	Bonding Vertex Sets Over Distributed Graph: A Betweenness Aware Approach	VLDB, 2015	2015	No	Yes
09/2015/Hawaii	Differential Privacy in Telco Big Data Platform	VLDB, 2015	2015	No	Yes
04/2015/Seoul	Cleaning uncertain data with a noisy crowd.	ICDE, 2015	2015	No	Yes
04/2015/Seoul	Conflict-aware event-participant arrangement	ICDE, 2015	2015	No	Yes
12/2016/Washington D.C.	Object Identification with Pay-As-You-Go Crowdsourcing	BigData, 2016	2018	Yes	Yes

03/2016/Bordeaux	Query Workload-based RDF Graph Fragmentation and Allocation	EDBT, 2016	2018	Yes	Yes
05/2016/Helsinki	Online Mobile Micro-Task Allocation in Spatial Crowdsourcing	ICDE, 2016	2018	Yes	Yes
05/2016/Helsinki	Mutual Benefit Aware Task Assignment in a Bipartite Labor Market	ICDE, 2016	2018	Yes	Yes
09/2016/New Delhi	Online Minimum Matching in Real-Time Spatial Data: Experiments and Analysis	VLDB, 2016	2018	Yes	Yes
04/2017/San Diego	Tuning Crowdsourced Human Computation	ICDE, 2017	2018	Yes	Yes
04/2017/San Diego	Prediction-Based Task Assignment in Spatial Crowdsourcing	ICDE, 2017	2018	Yes	Yes
04/2017/San Diego	CrowdFusion: A Crowdsourced Approach on Data Fusion Refinement	ICDE, 2017	2018	Yes	Yes
04/2017/San Diego	Trichromatic Online Matching in Real-time Spatial Crowdsourcing	ICDE, 2017	2018	Yes	Yes
08/2017/Halifax	The Simpler The Better: A Unified Approach to Predicting Original Taxi Demands based on Large-Scale Online Platforms	KDD, 2017	2018	Yes	Yes

10. Student(s) trained

Name	Degree registered for	Date of registration	Date of thesis submission/ graduation
Leihao Xia	Mphil	Fall 2013	Fall 2015
Ziyuan Zhao	Mphil	Fall 2013	Fall 2015
Rui Fu	Mphil	Fall 2013	Fall 2015
Chen Zhang	Ph.D.	Fall 2011	Fall 2015
Peng Cheng	Ph.D.	Fall 2012	Fall 2016

11. Other impact *(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, etc.)*