gMission: A General Spatial Crowdsourcing Platform

Zhao Chen Rui Fu Ziyuan Zhao Zheng Liu Leihao Xia Lei Chen
Peng Cheng Caleb Chen Cao Yongxin Tong Chen Jason Zhang
Hong Kong University of Science & Technology, Hong Kong SAR, PR China
{zchenah,rfu,zzhaoab,zliual,lxia,leichen,pchengaa,caochen,yxtong,czhangad}@cse.ust.hk

ABSTRACT

As one of the successful forms of using Wisdom of Crowd, crowdsourcing, has been widely used for many human intrinsic tasks, such as image labeling, natural language understanding, market predication and opinion mining. Meanwhile, with advances in pervasive technology, mobile devices, such as mobile phones and tablets, have become extremely popular. These mobile devices can work as sensors to collect multimedia data(audios, images and videos) and location information. This power makes it possible to implement the new crowdsourcing mode: *spatial crowdsourcing*. In spatial crowdsourcing, a requester can ask for resources related a specific location, the mobile users who would like to take the task will travel to that place and get the data. Due to the rapid growth of mobile device uses, spatial crowdsourcing is likely to become more popular than general crowdsourcing, such as Amazon Turk and Crowdflower. However, to implement such a platform, effective and efficient solutions for worker incentives, task assignment, result aggregation and data quality control must be developed.

In this demo, we will introduce gMission, a general spatial crowdsourcing platform, which features with a collection of novel techniques, including geographic sensing, worker detection, and task recommendation. We introduce the sketch of system architecture and illustrate scenarios via several case analysis.

1. INTRODUCTION

Geographic information is of great importance to people's daily life. For example, a driver need to know the traffic condition along his travel path so that he can effectively arrange his journey, and manufacturing companies need to be aware of the inventory situation of the retailers so as to schedule the production plan. Modern technology has provided people with numerous pathways to meet the need of acquiring such information, like Google map and other web based geographic information providers. Although such methods can help people to some extent, there are still several shortcomings in using them, among which lacking mechanisms for real-timing reporting is one that has to be overcome urgently. This point can be illustrated by the following case. If someone searches on Google for the information corresponding to Time Square, he/she may get its related photos and news, a comprehensive introduction of it through Wikipedia record, or video fragments of it on Youtube. However, he/she may never get the information whether there is a traffic jam in the Columbia Avenue, nor is it possible to obtain the current taxi availability along Broadway.

Proceedings of the VLDB Endowment, Vol. 7, No. 13

Recently, crowdsourcing has become a very efficient way in helping people to acquire information [1, 2, 3, 4, 6]. Unlike traditional methods, it is human workers who play the role of data sourcing and people can obtain the information they need actively by asking specific workers to answer their questions. With the help of crowdsourcing, the process of data collection no longer relies solely on existing information on the Internet, and a much broader spectrum of information that is reachable to people becomes easy to collect with the advance developed of mobile technology. People can use smart phones to capture events or sceneries, answer questions, conduct data analysis, and even find friends with respect to their current locations, we call this type of location-based crowdsourcing as *spatial crowdsourcing*. To address the needs to get information related to geographic locations, it is necessary to develop effective spatial crowdsourcing platforms to well use the human intelligence.

In this demo, we show a general spatial crowdsourcing platform developed by us, called gMission. Generally speaking, our system consists of three fundamental modules: the User Interface module, the Function Manager module, and the Data Manger module. The user interface module provides user with the function of question publishing and answer submitting, with which people can ask questions they are interested in and answer questions by collecting data with their phones. The function manager realizes the functions of geographic sensing, task recommendation, and quality control. The data manager module serves as the data storage and provides API for other module to visit the data. Schemes of multimedia input and output, localization, task recommendation, and quality control are exercised to improve the data quality and system efficiency.

In summary, our work contributes in these aspects:

- The system provides people with the platform to ask and collect geographic information in multimedia manner with ordinary mobile phones.
- The system is equipped with multiple location sensing

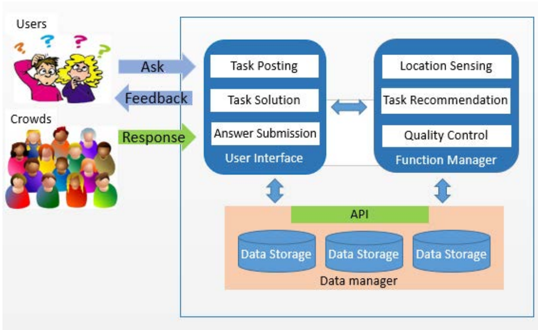


Figure 1: System Architecture of gMission

functions so that worker visualization and geographic based task recommendation are realized.

- We design a set of task recommendation schemes to facilitate the task processing of the system.

The rest of the paper is arranged as follows: Section 2 presents the system architecture and introduces the function of gMission. Section 3 demonstrates the system with two cases: the indoor scenario and the outdoor scenario. Finally, in Section 4 we conclude with our future plan.

2. SYSTEM ARCHITECTURE

gMission consists of four core modules: the user interface module, the location sensing module, the task recommendation module and the quality control module. The user interface runs on the terminal device and provides people with three functions, task posting, task solution and answer submission. The other three modules runs on the server side. The architecture of gMission is depicted in Figure 1. In the rest part of this section, we present the technical details of each module in gMission.

2.1 User Interface

In this subsection, we will introduce the user interface module, which can provide people to ask geographic tasks and collect corresponding answers in multimedia manner. This module consists of the following three functions.

2.1.1 Task Posting

With the task posting function, people may compose their interested questions and publish them to the platform. Before composing the task content, the user can specify the location of it first. The selection work can be done through either clicking the location on map, or choosing from the location category. Therefore, each task is associated with a unique location tagging, and such information is crucial to other modules, like the task recommendation module. After the location specification, the user may further specify the content of his task in multimedia ways. Text, audio and video are all supported as inputs in gMission. In this way, tasks will be easy to interpret. Apart from the composing work, users can also set the required format of the answer and the properties of the question. In gMission, users can ask the answer to be returned in various formats, including text, image and video. The properties of a question include a valid date, the required number of answers, and the rewarding credit. With these properties specified, users can get control of the real-timing and quality of a question.

2.1.2 Task Solution

After a task is posted to gMission, the task solution function will return the corresponding solution via the interaction with crowds. Once sufficient number of answers are arrived, gMission utilizes the task recommendation and quality control modules to aggregate different answers from workers and generate the final solution to the requester.

2.1.3 Answer Submission

The answer submission function provides users with checking the current published questions and submitting their answers to them. In gMission, people get access to their interested question in two ways: one is to traverse the question list, where all the valid questions are listed; the other is to check the news feed list, where the system automatically pushes the questions to each user through the task recommendation function. After picking the interested question, people start the work of answer composing. All the answers have to be submitted within the valid time and in accord to the format requirement. When a question reaches the end of its valid time or terminated by its holder, all the submitted answers will be rated and the user whose answer is qualified will be assigned with the predefined rewarding credit.

2.2 Location Sensing

gMission keeps track of the user's location through the module of location sensing. Currently, two types of localizing techniques are utilized by our system: the WiFi fingerprint localization for indoor environment and the GPS-based localization for outdoor environment. The system adopts a time triggered method to perform the localization. That is to say the user's location is periodically checked by gMission. Once the system is to report the user's location, it first checks whether WiFi fingerprint localization is available. If discernible WiFi fingerprint is detected, the indoor localizing is launch, and both the indoor location and outdoor location of the user are reported. Otherwise, the system will switch to GPS-based localization mode and only the outdoor position of the user is reported. In particular, as discussed below, the location sensing module provides a foundation for other modules, such as the task recommendation module and the quality control module.

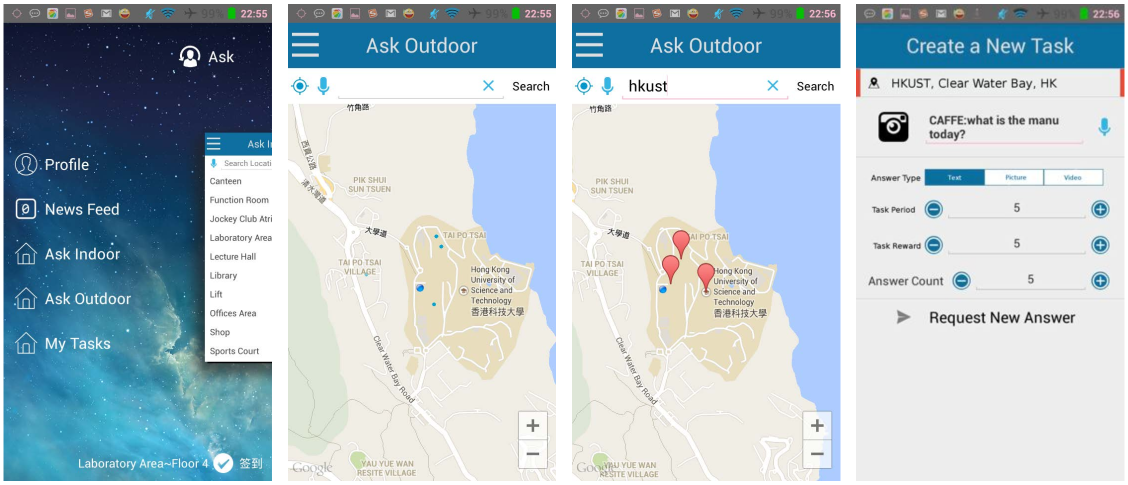
2.3 Task Recommendation

Task recommendation is one of the crucial functions in most crowd-powered systems. To achieve the properties of real-time responding and geo-sensitivity, gMission is equipped with a specially adapted task recommendation module, which consists of the following mechanisms.

Location Based Recommendation. There is a great resemblance between the process of assigning a suitable task to a potential worker and the activity to wake up a sensor in sensor network, where the major difference is that the "human sensor" is prone to move and the task is constrained with a time limit. Therefore, given a task $t(g)$ and a promised budget k, gMission first obtains all the potential workers $W_g = \{w_1, w_2, \ldots, w_n\}$ who are close to the targeted area $g = (x, y, z)$. Then gMission assigns the task $t(g)$ to the first k workers in W_g according to the ranking below:

$$w_i \succ w_j \quad if \quad d(w_i) \cdot |l(w_i) - g| < d(w_j) \cdot |l(w_j) - g|$$

where $d(w_i)$, *duration of worker* w_i, measures the duration that worker w_i has spent in the neighborhood of g, which



(a) Menu of Asking Questions (b) Location Selection (c) Location Recognition (d) Setting of Answer Requirements

Figure 2: Task Posting Scenario

implies that he/she is much more likely to stay than w_j. Simply put, a time-weighted kNN algorithm is adapted to facilitate the location based recommendation.

Whom-to-ask Enhancement. Since users are able to register and report their location information on the gMission, and all historical task records are evaluated in the system, task recommendation may also be enhanced by incorporating the best combination of the workers. Among all n possible workers, only the best k workers, who have the highest corporate task accuracy, will receive the assigned tasks. Paper [1] and [2] are considered as a basic crowdsourcing framework to actively enroll and manage potential workers, where the probability is considered as a primary measure of quality. The addition of geographical information, however, enhances the usability of such a framework in real applications. For more details, please refer to our previous work[1].

Load Balance Enhancement. There are some hotspot locations like metro stations, shopping malls in our daily life. Lots of tasks are about these hotspots and workers are very likely to emerge from there too. This brings the worker load issue. Since each POI has a fixed location, tasks about the same POI always have the same coordinate. Then with the normal kNN mechanism mentioned above, all these tasks will be assigned to the nearest k workers. Even a worker just appears near a hot POI for a short time, his task list will insanely increase while workers who are just a little far away will get relatively quite few tasks. To make the assignments more reasonable for this kind of extreme cases, a dynamic weight of work load is introduced. The weight is calculated from the former assignments of a worker to enhance the current assignment.

2.4 Quality Control

The quality control module is an important component and contribution of gMission. It consists of three strategies: the life cycle control, the location-based verification and the wise-market quality control. We will introduce the three strategies as follows.

Life Cycle Control. For each published task in gMission, the system assigns a life cycle according to the requirement of users. Once a task reaches its deadline, the system will shut it down and submission afterwards will no longer be accepted. Thus, this strategy guarantee the real-time quality of collected answers.

Location-based Verification. By means of the module of location sensing in Section 2.3, the system can track the location of crowd workers when they submitted answers. If the tracked location does not coincide with the location where the task is required, the answer will be recognized as a fake one and rejected afterwards. In this way, gMission guarantees the accuracy of the spatial information.

Wise-Market Quality Control. gMission utilizes a probabilistic model, called Wise-Market model [2], to balance the trade-off between the cost and the confidence of submitted answers. More technical details of Wise-Market model can be found in our previous work [2].

3. DEMONSTRATION OVERVIEW

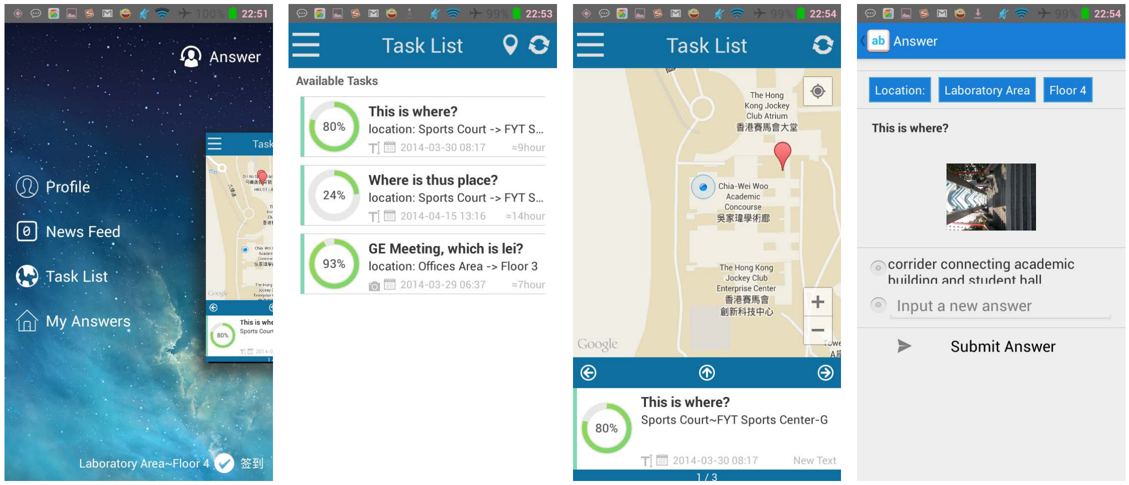
In this section, we describe two kinds of scenarios, task posting and answer submission cases, of gMission system in detail, and explain the aims of our demonstration. Moreover, both of scenarios can support the indoor and outdoor environments. However, for brevity's sake, we illustrate the functionalities of the system in the outdoor case, the indoor case is similar. The detailed process is shown as follows.

3.1 Task Posting Scenario

In this subsection, we exhibit the following four steps of task posting scenario in gMission system to the audiences.

Step 1: Interface of Task Posting. When a user login in the task posting interface, he/she can choose to ask a question outside campus or within campus (Figure 2(a)). Here, the user chose to ask a question outside the campus.

Step 2: Location Selection. After the user makes the choice, a map is presented for him/her to select the location where the user would like to publish the task, in which the amount of available human worker and their location is pointed on the map (Figure 2(b)).



(a) Menu of Answering Questions (b) Task Selection (c) Task Acceptance (d) Answer Submission

Figure 3: Answer Submission Scenario

Step 3: Location Recognition. Suppose that the user intends to post a task about Hong Kong University of Science and Technology, and he/she fills in the word HKUST. Several pointers will then be placed on the map, indicating the recognized locations, of which the user can pick one so as to confirm the location (Figure 2(c)).

Step 4: Task Posting. Then, the user can compose the content of the task. gMission support three kinds of task posting: text posting, image posting and voice posting. After the user tap in or speak out the content of the question, several values are to be set, including the valid time, the rewarding credit, and the format of the answer (Figure 2(d)). Finally, the user can launch this task and wait for answers.

To sum up, through the aforementioned four steps, a user can publish his/her task requirement to gMission system.

3.2 Answer Submission Scenario

In this subsection, we focus on the following four steps regarding the Answer submission scenario in gMission system.

Step 1: Interface of Answer Submission. When a crowd worker login in the interface of answer submission, he/she can get view of the published task through two ways: one is to check the "News Feed", which shows the system recommended tasks, the other is to check the Task List, which shows the complete list of all published tasks (Figure 3(a)). In the remaining part, we assume that the user will get into the News Feed.

Step 2: Task Selection. In the News Feed list, a crowd worker can find three tasks, which are recommended to him. The left side cycle indicates the remaining time of the question and the title of the question shows the location where the question is to ask (Figure 3(b)).

Step 3: Task Acceptance. If the crowd worker picks the first task, he/she will accept this task and is able to check the detailed content of the question (Figure 3(c)).

Step 4: Answer Submission. Finally, the crowd worker finishes this task and submits the answer according to the task requirement (Figure 3(d)). Then the answer is pushed to the question holder after the task is terminated. In particular, by means of the location-based quality control strategy of gMission, fake answers are filtered out and only qualified answer are presented. Therefore, the task publisher will confirm the answers based on the suggestion from gMission, then a crowd worker can get the reward credits if his answer is adopted.

4. CONCLUSION

In this demo, we present a novel spatial crowdsourcing system, gMission, which provides a general platform for conducting various crowdsourcing tasks based on location information. Several novel techniques have been implemented in the system, such as task assignment[1] and answer aggregation [2]. There is a closely related work proposed by Kazemi and Shahabi [5], which mainly focused on task assignment problem. Compared to that work, gMission provides a whole picture of an implemented spatial crowdsourcing platform and proposes several techniques for effective spatial crowdsourcing in addition to task assignment.

5. ACKNOWLEDGMENT

This work is supported in part by the Hong Kong RGC Project N_HKUST637/13, National Grand Fundamental Research 973 Program of China under Grant 2012-CB316200, National Natural Science Foundation of China (NSFC) Grant No. 61232018, Microsoft Research Asia Gift Grant and Google Faculty Award 2013.

6. REFERENCES

[1] C. C. Cao, J. She, Y. Tong, and L. Chen. Whom to ask? jury selection for decision making tasks on micro-blog services. *PVLDB*, 5(11):1495–1506, 2012.

[2] C. C. Cao, Y. Tong, L. Chen, and H. V. Jagadish. Wisemarket: a new paradigm for managing wisdom of online social users. In *KDD*, pages 455–463, 2013.

[3] J. Gao, X. Liu, B. C. Ooi, H. Wang, and G. Chen. An online cost sensitive decision-making method in crowdsourcing systems. In *SIGMOD*, pages 217–228, 2013.

[4] S. Guo, A. G. Parameswaran, and H. Garcia-Molina. So who won?: dynamic max discovery with the crowd. In *SIGMOD*, pages 385–396, 2012.

[5] L. Kazemi and C. Shahabi. Geocrowd: enabling query answering with spatial crowdsourcing. In *SIGSPATIAL/GIS*, pages 189–198, 2012.

[6] C. J. Zhang, L. Chen, H. V. Jagadish, and C. C. Cao. Reducing uncertainty of schema matching via crowdsourcing. *PVLDB*, 6(9):757–768, 2013.