# Towards a Generic Framework for Trustworthy Spatial Crowdsourcing

Cyrus Shahabi
Integrated Media System Center
University of Southern California
shahabi@usc.edu

## ABSTRACT

Many studies foresee significant future growth in the number of mobile smart phone users, the phone's hardware and software features, and the broadband bandwidth. Therefore, a transformative area of research is to fully utilize this new platform for various tasks, among which the most promising is *spatial crowdsourcing*. Spatial crowdsourcing (SC) engages individuals, groups, and communities in the act of collecting, analyzing, and disseminating urban, social, and other spatiotemporal information. This new paradigm of data collection has shown to be useful when traditional means fail (e.g., due to disaster), are censored or do not scale in time and space.

Two major impediments to the success of spatial crowdsourcing in real-world applications are scalability and trust issues. Without scale considerations, it is impossible to develop a generic multi-campaign spatial crowdsourcing system (SC-system) that can efficiently and in real-time match many requesters' tasks to numerous workers. Without trust, the SC-system cannot evaluate the credibility of the contributed data, rendering it ineffective for replacing the traditional data collection means. In this position paper, we propose to study both issues of scale and trust in spatial crowdsourcing.

## 1. INTRODUCTION

Smartphones are ubiquitous: we are witnessing an astonishing growth in mobile phone subscription that has surpassed 4.35 billion worldwide at the end of 2009 3, reached 6 billion in 2011, which is 87% of the world population [2], and by 2013, Gartner predicts that mobile phones will overtake PCs as the most common Web access device worldwide. Meanwhile, the mobile phone's bandwidth is constantly increasing: from 2.5G (up to 384Kbps) to 3G (up to 14.7Mbps) and recently 4G (up to 100 Mbps) [3]. Hence, the multiplication of the above two factors plus the constant progress and increase of smartphone's sensors (e.g., video cameras) suggest an exponential growth in data collection and sharing by smart phones. That is, every person with a mobile phone can now act as a multi-modal sensor collecting and sharing various types of high-fidelity spatiotemporal data instantaneously (e.g., picture, video, audio, location, time, speed, direction, acceleration).

Exploiting this large volume of potential users and their movability, a new mechanism for efficient and scalable data collection has emerged: *"spatial crowdsourcing"* [4]. With spatial crowdsourcing, the goal is to crowdsource a set of spatial *tasks* (i.e., tasks related to a location) to a set of *workers*, which requires the workers to perform the spatial tasks by physically traveling to those locations. Spatial crowdsourcing has applications in numerous domains such as journalism, tourism, intelligence, disaster response and urban planning. To illustrate, consider a disaster-response scenario, where Red Cross (i.e., *requester)* is interested in collecting pictures and videos of disaster areas from various locations of a city. With spatial crowdsourcing, the requester issues a query to a spatial crowdsourcing server (*SC-server*). Consequently, the SC-server crowdsources the query among the available workers in the vicinity of the events. Once the workers document their events with their mobile phones, the results are sent back to the requester.

Towards this end, we are designing and developing a proof-of-concept spatial crowdsourcing system, dubbed *GeoCrowd*. GeoCrowd consists of two main components: a webserver and a mobile client. The webserver would enable the public to post their spatial crowdsourcing requests (utilizing a map-based

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
*MobiDE '13*, June 23, 2013 New York, New York, USA
Copyright © 2013 ACM 978-1-4503-2197-6/13/06...$15.00

interface), which is then translated by the system to a series of spatial tasks. The mobile client (developed for smartphones and tablets, e.g., on Android OS) can be downloaded by public on their phones in order for them to act as workers, performing the spatial tasks. Figure 1 shows our spatial crowdsourcing framework.

## 2. FOCUS STUDY

Within the GeoCrowd framework, we study both issues of scale and trust in spatial crowdsourcing. For scale, we argue that task assignment is the main bottleneck of the system as it needs to be performed frequently and in real-time as new tasks and workers become available or as tasks get completed (or expired) and workers leave the system. Although several variations of the task-assignment problem have been studied in different problem domains, we show that the spatial aspects of the tasks, e.g., the fact that the worker needs to physically travel to a specific location to perform the task, renders the problem unique. In our prior work [4], we show the significant complexity of the problem when considering various spatial and temporal constraints of the workers. Moreover, we plan to study a cloud-based distributed approach to the problem for better scale-out.

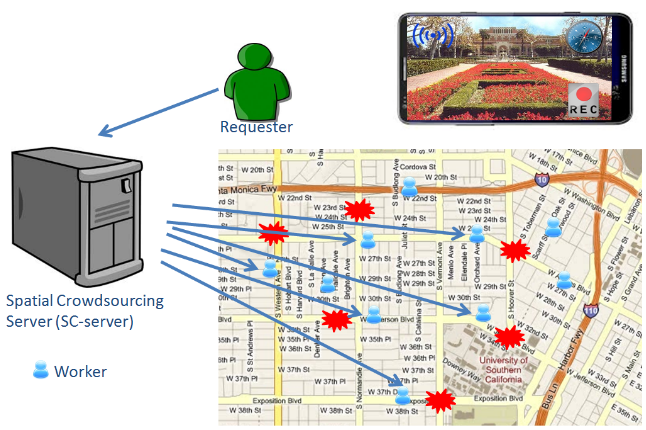


**Figure 1: The GeoCrowd Framework**

For trust, we argue that while trustworthy computing has been successful in developing techniques to avoid any malicious software to manipulate the sensed data before sending it to the server, the issue of trusting the data contributor himself has not been studied thoroughly. To tackle this problem, we extend our framework to maintain a reputation score per worker and a confidence level for every spatial task. Thus, every spatial task should now be assigned to enough number of workers such that their aggregate reputation satisfies the confidence of the task. Again, a unique aspect of our problem is that the optimal assignment of tasks heavily depends on the geographical locations of workers and tasks. An exhaustive solution to this problem is to compute the aggregate reputation (using a typical decision fusion aggregation mechanism, such as voting) for all possible subsets of the workers, which renders the problem complex (we believe it is NP-hard). Our preliminary studies show that this optimization problem cannot be trivially reduced to existing problems such as the classical resource allocation problem. Hence, our goal is to prove this problem is NP-hard and then solve it by utilizing a set of heuristics, which exploit the spatial properties of the problem space.

## 3. REFERENCES

[1] ABI Research: Mobile broadband subscriptions surpassed 271 million in 2009, March 27, 2010, URL: *http://www.intomobile.com/2010/03/27/abi-research-mobile-broadband-subscriptions-surpassed-271-million-in-2009/.*

[2] *http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats.*

[3] M. Ergen, Mobile Broadband - Including WiMAX and LTE, Springer, 1st edition, 2009.

[4] L. Kazemi and C. Shahabi, GeoCrowd: Enabling Query Answering with Spatial Crowdsourcing. In *ACM SIGSPATIAL GIS, 2012*.