A Generic Framework for Spatial Crowdsourcing

Hien To, *University of Southern California*
Giorgos Constantinou, *University of Southern California*
Cyrus Shahabi, *University of Southern California*

Spatial Crowdsourcing

Ubiquity of mobile users

6.5 billion mobile subscriptions by the end of 2013
≡ 93% of the world population

Technology advances on mobile phones (e.g., Cameras)

Network bandwidth improvements

Spatial crowdsourcing (SC) engages individuals, groups, and communities in the act of collecting, analyzing, and disseminating urban, social, and other spatiotemporal information.

Applications

Spatial crowdsourcing (SC) has applications in numerous domains such as journalism, tourism, security, intelligence, disaster response and urban planning.

Fig 1. A Scenario of SC for Journalism

Research Challenges

Three major impediments to the success of SC are ***dynamism, trust*** and ***privacy*** issues.

Dynamism: new tasks and workers become available as tasks are completed (or expired) and workers leave the system

Privacy: adversary with access to individuals' whereabouts can infer sensitive details about a person, e.g., health status

Trust: SC-system cannot evaluate the credibility of the contributed data, rendering it ineffective for replacing the traditional data collection means.

Dynamism of Tasks and Workers

The goal is to assign a set of spatial tasks to a set of available workers and maximize the number of assigned tasks, referred to as ***Maximum Task Assignment*** *(MTA).*

Approaches to MTA

+ Greedy
+ Least Location Entropy Priority
+ Nearest Neighbor Priority

Fig 2. MediaQ Screenshots

[Ref] Leyla Kazemi and Cyrus Shahabi, *GeoCrowd: Enabling Query Answering with Spatial Crowdsourcing*, ACM SIGSPATIAL GIS, Redondo Beach, CA, November 2012

Trustfulness of Workers

Non-spatial metrics
+ Rating (R)
+ # of transactions (TR)
+ Fast response time (FRT)
+ Quality of work (Q)

Spatial metrics
+ ***Distance traveled (DT)***
+ ***Spatial coverage (R)***

e.g. $\mathrm{T(u)} = \frac{\sum_{i=1}^{TR} R(i)+FRT(i)+Q(i)+DT(i)+R(i)}{TR}$

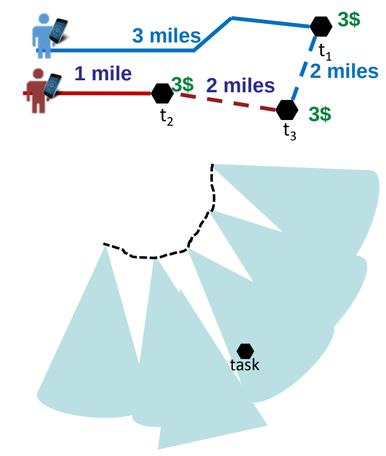

Fig 3. Spatial metrics

Use reputation-based trust:
+ To maximize the ***quality*** of the result in task assignment
+ To direct requesters to "trustful" areas in task posting

Fig 4. Trustfulness Heatmap

Location Privacy of Workers

Cell Service Provider (CSP) releases sanitized location data (PSD) according to ***differential privacy***. When SC-server receives task *t*, it queries the PSD to determine a ***region*** that encloses ***sufficient*** workers. Then, SC-server initializes ***geocast communication*** to disseminate *t* to all workers within the region.

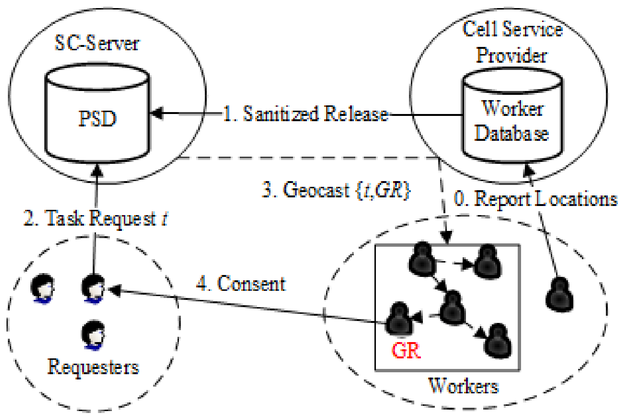

Fig 5. Private SC Framework

~0	~3	~3	~2	
~4	~7	~6	~3	
~4	~5 task	~2	~3	~1
		~2	~0	~0
~0	~6	~3	~0	~0

Fig 6. Sanitized Location Data (PSD)

[Ref] Hien To, Gabriel Ghinita, and Cyrus Shahabi, *A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing*, In Proceedings of the 40th International Conference on Very Large Data Bases (VLDB 2014), Hangzhou, China, September 2014

Generic SC Framework

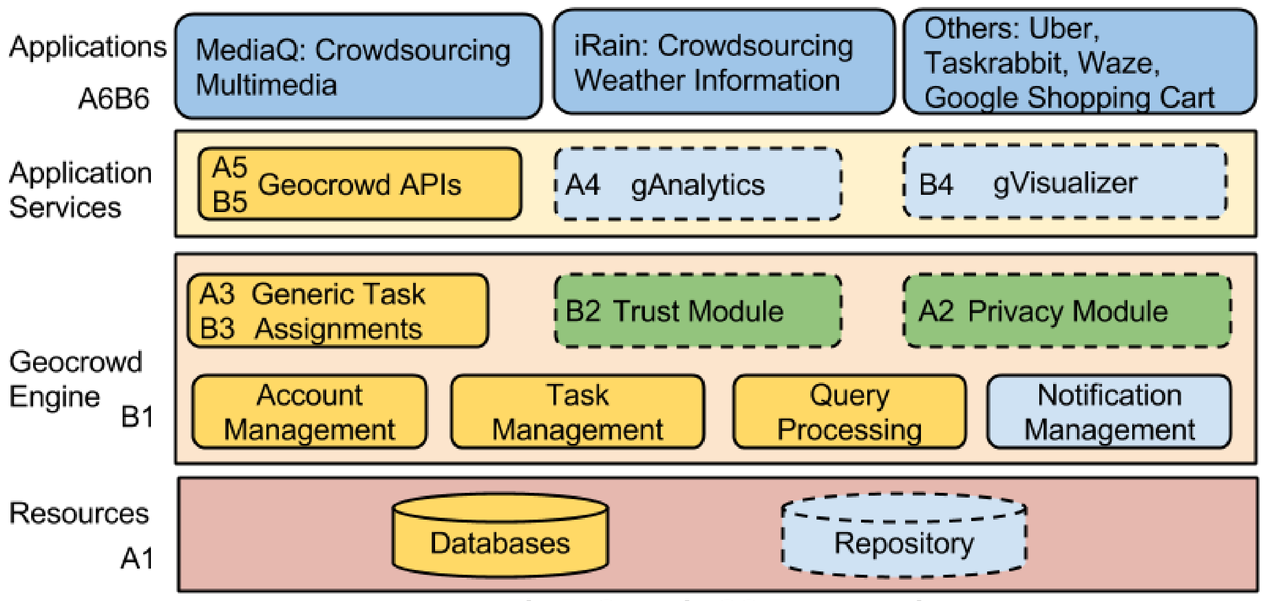

Fig 7. Generic SC Framework

USC Viterbi School of Engineering
Integrated Media Systems Center