Real-time Spatial Task Assignment for Weather Crowdsourcing

Hien To, Liyue Fan, Luan Tran, Cyrus Shahabi
{hto, liyuefan, luantran, shahabi}@usc.edu
University of Southern California

1. Weather Crowdsourcing (WC)

With the ubiquity of smart phones, every person with a mobile phone can report weather conditions, e.g., precipitation, air quality. The reported data, often in real-time, offer a valuable addition to the satellite remote sensing and radar detections technologies.

Applications

- ✓Environmental sensing
- ✓Disaster response
- ✓Transportation decision making

Fig.1) WC apps

Task Assignment

1. Requester issues tasks to server
2. Server assigns tasks to **nearby** workers
3. Server sends tasks to assigned workers

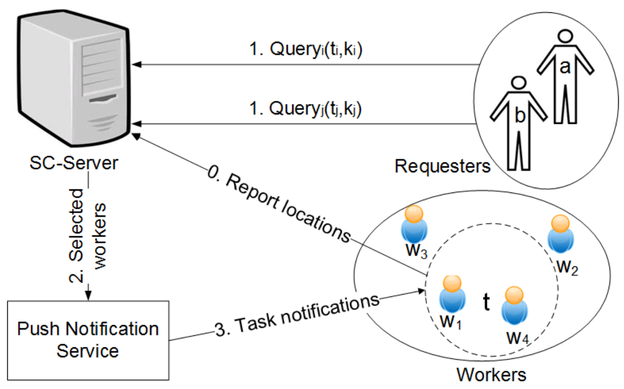


Fig.2) WC framework

2. Challenges

Distinctions from crowdsourcing paradigms:

1. Reported observation is *near* ***(within)*** task's *location* (***duration***)
2. Server maximizes task coverage with worker **budget** constraint
3. **Dynamic** arrivals of tasks/workers

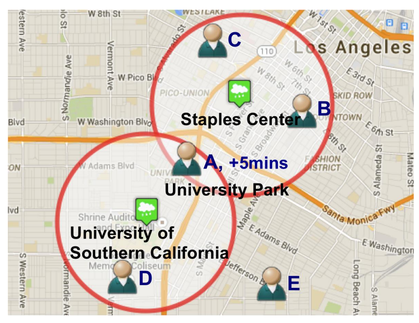
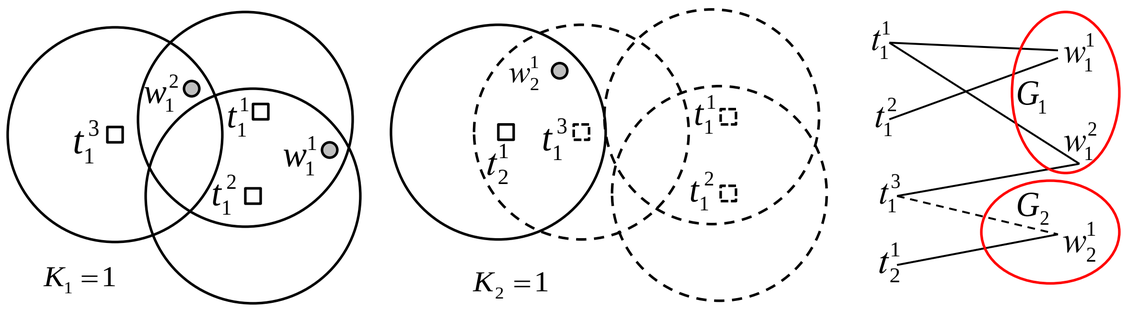


Fig.3) WC example

Select workers that **m**aximize **t**ask **c**overage (MTC)

$$\left| \bigcup_{i=1}^{Q} \bigcup_{w^j{}_i \in L_i} C(w^j{}_i) \right|$$



a) Time instance 1 b) Time instance 2 c) Bipartite graph

Fig. 4) Graphical example of worker-task coverage (2 time instances)

3. Fixed Budget

Off-line Scenario

MTC is **NP-hard** by reduction from the *maximum coverage with group budget constraint problem*

Greedy algorithm gives *0.5-approximation ratio*.

On-line Scenario

Heuristics to MTC time instance: **Basic**, **Spatial**, **Temporal**

4. Dynamic Budget

Off-line Scenario

MTC is **NP-hard** by reduction from the *maximum coverage problem*

Greedy algorithm gives *0.63-approximation ratio*

On-line Scenario

Adapt allocates budget to time instances based on contextual bandits
i.e., balances ***budget status*** and ***coverage gain***

Improve *Adapt* by leveraging workers' activity patterns

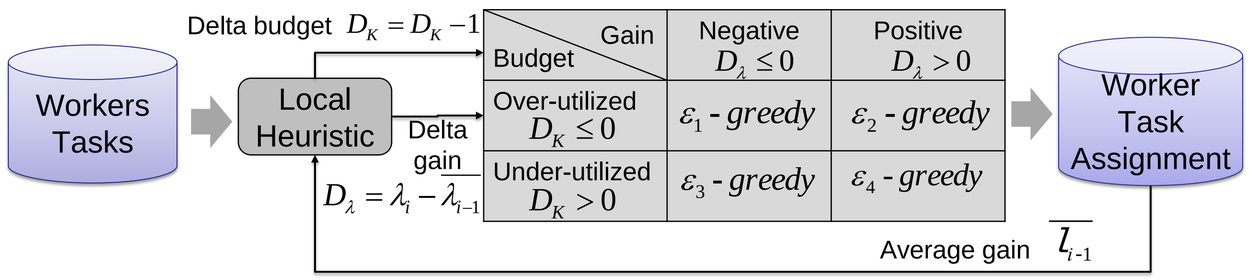
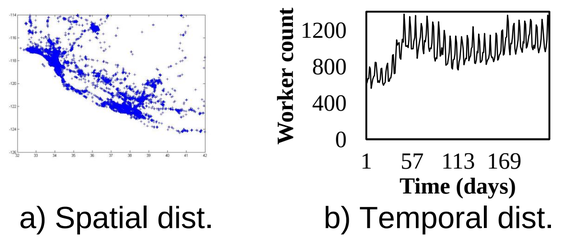


Fig. 5) Work-flow for adaptive budget allocation based on contextual bandits

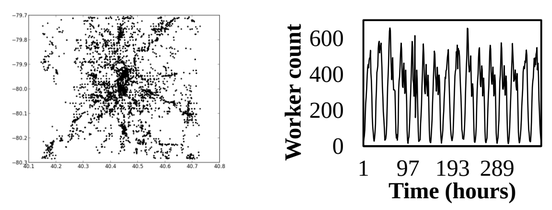
5. Performance Evaluations

Real Datasets



a) Spatial dist. b) Temporal dist.

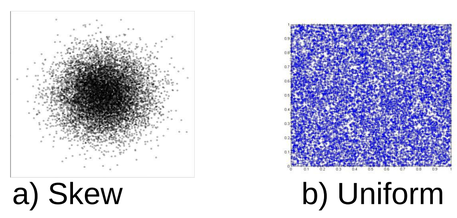
Fig. 6) Gowalla, CA



a) Spatial dist. b) Temporal dist.

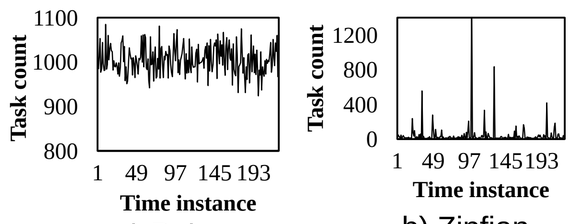
Fig. 7) Foursquare PA

Synthetic Datasets



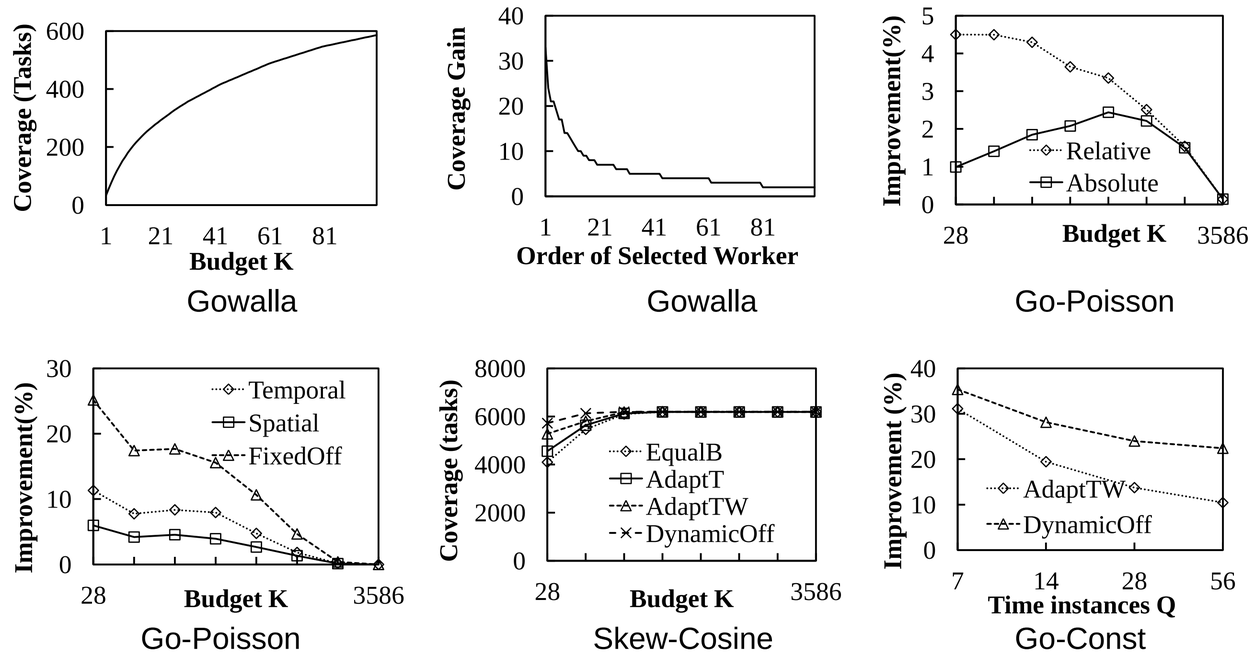
a) Skew b) Uniform

Fig. 8) Spatial distributions



a) Poisson b) Zipfian

Fig. 9) Temporal distributions



6. Discussions

- ✓*Worker overload* to avoid repetitive activations of the same worker
- ✓*Profit* for each task to represent the importance of tasks
- ✓*Utility* of tasks as a function of spatial/temporal distance to worker
- ✓*Activation cost* of workers is not uniform

USC Viterbi School of Engineering

CHRS Center for Hydrometeorology & Remote Sensing UCIrvine

IMSC Integrated Media Systems Center