Quality and Budget Aware Task Allocation for Spatial Crowdsourcing

(Extended Abstract)

[1]Han Yu, [1]Chunyan Miao, [1]Zhiqi Shen, and [1,2]Cyril Leung
[1]LILY Research Centre, Nanyang Technological University, Singapore
[2]Department of Electrical and Computer Engineering, the University of British Columbia, Canada
{han.yu, ascymiao, zqshen, cleung}@ntu.edu.sg, cleung@ece.ubc.ca

ABSTRACT

A major research challenge for spatial crowdsourcing is to improve the expected quality of the results. However, existing research in this field mostly focuses on achieving this objective in volunteer-based spatial crowdsourcing. In this paper, we introduce the budget limitations into the above problem and consider realistic cases where workers are paid unequally based on their trustworthiness. We propose a novel quality and budget aware spatial task allocation approach which jointly considers the workers' reputation and proximity to the task locations to maximize the expected quality of the results while staying within a limited budget.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Intelligent agents, Multiagent systems*

Keywords

Reputation, task allocation, spatial crowdsourcing

1. INTRODUCTION

Crowdsourcing refers to the arrangement in which contributions (such as services, observations, contents, etc.) are solicited from a large group of unrelated people [1]. As smart phones become increasingly available, a new type of crowdsourcing - *spatial crowdsourcing* (a.k.a. mobile or location-based crowdsourcing) - has emerged which *requesters* crowdsource results for *spatial tasks* from *workers* who are required to travel to specific locations to complete the tasks.

Task requesters in spatial crowdsourcing face uncertainty about the performance of the workers. As they also have limited budgets for obtaining results, the main research problem is to find an effective trade-off between cost-aware task allocation and the expected quality of the results. In the case of spatial crowdsourcing, workers are required to travel and incur extra costs. This may further negatively affect their willingness to participate.

Reputation aware computing is a promising approach for addressing this problem [4, 6, 7, 8]. In [3], the authors discussed the special characteristics of spatial crowdsourcing, and proposed an efficient heuristic-based greedy approach - *GeoTruCrowd*. However, they focused on self-incentivized spatial crowdsourcing in which people are self-motivated to perform the tasks, and they did not consider the impact of a limited budget. The CrowdBudget approach [5] allocates the budget in advance through analysis of cost and expected quality of the results. Nevertheless, CrowdBudget did not consider the spatial information of tasks and workers.

In this paper, we propose a Budget-aware trustworthy Task Allocation approach for Spatial Crowdsourcing (*Budget-TASC*) to bridge this gap in current research. We focus on the budget allocation problem in which an agent, on behalf of the task requester, allocates the task to a number of workers such that the total cost does not exceed a pre-specified budget. The agent's goal is to find an optimal budget allocation that maximizes the expected quality of the collective result provided by the selected workers.

2. THE PROPOSED APPROACH

We consider the problem of delegating a spatial crowdsourcing task τ_i proposed by a requester i to N candidate workers subject to a budget limit of $B^{\tau_i} \in \mathbb{R}^+$. A spatial task τ_i is represented as a tuple of the form $\langle l^{\tau_i}, R^{\tau_i}, B^{\tau_i}, p_H^{\tau_i}, p_M^{\tau_i} \rangle$. Similar to the model in [3], information related to a spatial task includes l^{τ_i}, which is the location of the task specified by a point in 2D space (e.g., represented by a latitude-longitude coordinate), and R^{τ_i}, which is the radius of the spatial region within which workers are most likely to accept the assigned task.

The workers' reputation information is used by the spatial crowdsourcing system to determine how much payment they should receive for each task. Workers are classified as having *high*, *medium* or *low* reputation standing. $p_H^{\tau_i}$ and $p_M^{\tau_i}$ are the amounts of money the task requester is willing to pay for a result provided by a worker with high reputation and a worker with medium reputation, respectively. Based on the combined consideration of worker reputation and distance to the task location, the expected credibility of a result provided by a worker j for a spatial task τ_i is:

$$c_j^{\tau_i}(t) = r_j(t) \cdot \delta(l_j(t), l^{\tau_i}) \quad (1)$$

where $l_j(t)$ is the worker's location at time step t when the task request is delegated to him. $\delta(l_j(t), l^{\tau_i})$ is a function calculating the discount to the worker's reputation as a result of his proximity to the task location. In this paper, it is defined as:

$$\delta(l_j(t), l^{\tau_i}) = 1 - \max[0, \min[\log_{D_c}(d(l_j(t), l^{\tau_i})), 1]] \quad (2)$$

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey.*

where $d(l_j(t), l^{\tau_i})$ calculates the Euclidean distance between two GPS coordinates in kilometers.

To reduce the complexity of the problem, we identify possible simplifying heuristics. Firstly, for two workers at the same distance from the location of a spatial task, the one with higher reputation has higher credibility. Secondly, for workers with the same reputation, the ones closer to the location of a spatial task have higher credibility. Based on these heuristics, we propose the Budget-TASC approach as shown in Algorithm 1.

Algorithm 1 Budget-TASC

Require: A set of N workers, a spatial task τ_i.
for $j = 1$ to N **do**
  Compute $c_j^{\tau_i}(t)$ according to Eq.(1)
end for
Rank workers in descending order of their $c_j^{\tau_i}(t)$
for $j = 1$ to $\min[\lfloor \frac{B^{\tau_i}}{p_H^{\tau_i}} \rfloor, N]$ **do**
  $J = \min[N - j, \lfloor \frac{B^{\tau_i} - p_H^{\tau_i} j}{p_M^{\tau_i}} \rfloor]$
  Select a set of workers, $W_j \subseteq \{1, ..., j+J\}$, who satisfy $d(l_j(t), l^{\tau_i}) \leq R^{\tau_i}$
  Select a subset, $S_j \subseteq W_j$ workers who satisfy $r_j(t) \geq Th_{ML}$ (subject to actual availability) with ties broken arbitrarily
  $C_{S_j}^{\tau_i} = \sum_{j=1}^{|S_j|} c_j^{\tau_i}(t)$
end for
$P_{\tau_i} = \arg\max_{S_j} C_{S_j}^{\tau_i}$

3. EXPERIMENTAL EVALUATION

We compare the performance of Budget-TASC against two state-of-the-art approaches on a range of spatial crowdsourcing tasks generated from the *Foursquare* dataset.[1] Data for users and venues in the city of Singapore (13,919 users and 430 venues) are extracted from this dataset. The locations of the venues and the users in this subset of data are used to generate the locations for the spatial crowdsourcing tasks and the workers, respectively.

We vary the task radius, R^{τ_i}, which represents the spatial region of each task in the experiments to create different scenarios. The values of $p_H^{\tau_i}$ and $p_M^{\tau_i}$ are set at \$2 and \$1 respectively. We uniformly divide the reputation value range from 0 to 1 into three levels by setting the values of Th_{HM} and Th_{ML} to $\frac{2}{3}$ and $\frac{1}{3}$ respectively.

After the location and reliability values of a population of the workers have been generated at the beginning of each experiment, two additional copies were cloned so that the *BudgetTASC*, *CrowdBudget* [5], and *GeoTruCrowd* [3] approaches can run in parallel for performance comparison. We use the majority voting rule for aggregating the results returned by the workers involved in each task. All approaches in the experiments adopt the method in [2] to calculate their reputation.

Under a budget constraint of \$30, the performance of the three approaches are shown in Figure 1. It can be observed that Budget-TASC achieves significantly lower error rates than GeoTruCrowd and CrowdBudget for all task radius settings studied.

[1]`https://archive.org/details/201309_foursquare_dataset_umn`

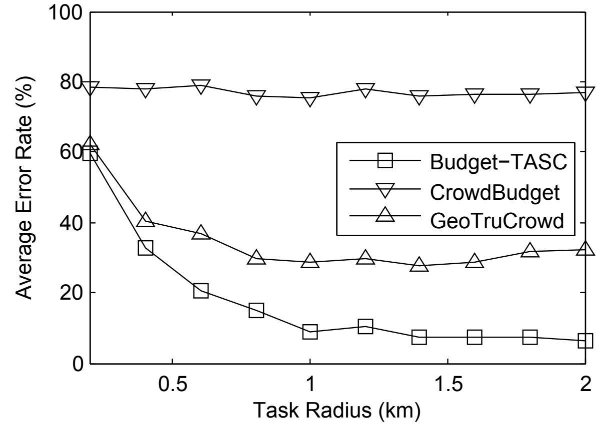

Figure 1: The average error rates.

4. CONCLUSIONS

We studied the issue of reputation aware task allocation in spatial crowdsourcing and introduced a budget constraint into this research problem. We proposed a novel heuristic-based approach - Budget-TASC - to efficiently solve this problem. Through extensive numerical experiments based on real-world data from *Foursquare*, we demonstrated that the proposed approach outperforms state of the art approaches through significant reduction in the average error rates.

Acknowledgements

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative and administered by the Interactive and Digital Media Programme Office.

REFERENCES

[1] A. Doan, R. Ramakrishnan, and A. Y. Halevy. Crowdsourcing systems on the world-wide web. *Communications of the ACM*, 54(4):86–96, 2011.

[2] A. Jøsang and R. Ismail. The beta reputation system. In *Proceedings of 15th Bled Electronic Commerce Conference*, pages 41–55, 2002.

[3] L. Kazemi, C. Shahabi, and L. Chen. GeoTruCrowd: Trustworthy query answering with spatial crowdsourcing. In *ACM SIGSPATIAL'13*, pages 304–313, 2013.

[4] S. Liu, H. Yu, C. Miao, and A. C. Kot. A fuzzy logic based reputation model against unfair ratings. In *AAMAS'13*, pages 821–828, 2013.

[5] L. Tran-Thanh, M. Venanzi, A. Rogers, and N. R. Jennings. Efficient budget allocation with accuracy guarantees for crowdsourcing classification tasks. In *AAMAS'13*, pages 901–908, 2013.

[6] H. Yu, C. Miao, B. An, Z. Shen, and C. Leung. Reputation-aware task allocation for human trustees. In *AAMAS'14*, pages 357–364, 2014.

[7] H. Yu, Z. Shen, C. Miao, and B. An. Challenges and opportunities for trust management in crowdsourcing. In *IAT'12*, pages 486–493, 2012.

[8] H. Yu, Z. Shen, C. Miao, and B. An. A reputation-aware decision-making approach for improving the efficiency of crowdsourcing systems. In *AAMAS'13*, pages 1315–1316, 2013.