

## Research Article

# Truthful Mechanism Design for Multiregion Mobile Crowdsensing

Yu Qiao <sup>1</sup>, Jun Wu,<sup>2</sup> Hao Cheng,<sup>1</sup> Zilan Huang,<sup>1</sup> Qiangqiang He,<sup>1</sup> and Chongjun Wang <sup>1</sup>

<sup>1</sup>National Key Laboratory for Novel Software Technology, Nanjing University, China

<sup>2</sup>Jiangsu Provincial Key Laboratory of E-Business, Nanjing University of Finance and Economics, China

Correspondence should be addressed to Chongjun Wang; [chjwang@nju.edu.cn](mailto:chjwang@nju.edu.cn)

Received 11 March 2020; Revised 18 June 2020; Accepted 24 July 2020; Published 19 August 2020

Academic Editor: Wei Yu

Copyright © 2020 Yu Qiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the age of the development of artificial intelligence, we face the challenge on how to obtain high-quality data set for learning systems effectively and efficiently. Crowdsensing is a new powerful tool which will divide tasks between the data contributors to achieve an outcome cumulatively. However, it arouses several new challenges, such as incentivization. Incentive mechanisms are significant to the crowdsensing applications, since a good incentive mechanism will attract more workers to participate. However, existing mechanisms failed to consider situations where the crowdsourcer has to hire capacitated workers or workers from multiregions. We design two objectives for the proposed multiregion scenario, namely, weighted mean and maximin. The proposed mechanisms maximize the utility of services provided by a selected data contributor under both constraints approximately. Also, extensive simulations are conducted to verify the effectiveness of our proposed methods.

## 1. Introduction

With the rapid development of the hand-held mobile devices, mobile crowdsensing [1] has become a new tool for problem-solving, and there has been lots of applications in real life, such as healthcare [2], smart city [3], and localization [4]. Due to the fact that workers' participation will depreciate their devices, such as consumption of CPU and battery, they may take a toll on their participation. Incentive mechanisms are significant to the crowdsensing applications, since a good incentive mechanism will attract more workers to participate. Estimating workers' personal cost is a hurdle. Workers' private cost is subjective, which is associated with various factors. These factors are hard to observe, such as difficulties of tasks and abilities of workers. A reasonable method is to offer an estimated optimal bonus to compensate costs of workers.

With the limited budget, workers may be reluctant to participate, if the offered price is too low. Offering a high price to workers may result in lower outcome. We will use a powerful tool from the algorithmic game to ensure both the efficiency and workers' incentivization.

The mechanism design is a tricky issue for the crowdsourcer with budget constraint, since designing a budget constraint allocating scheme needs understanding its payoff to workers, which is also related to the allocating scheme itself. Most existing works aim at maximizing the efficiency of crowdsensing by hiring workers in a single region. Nevertheless, with some geographical limitations, the crowdsourcer is necessary to procure service from workers in multiple regions to perform tasks and consider the interactions of different regions. Consider a crowdsourcer with a hard budget who wants to estimate the residential information in Eastern Asia which consisted of five countries, China, Japan, South Korea, North Korea, and Mongolia. Due to the geographical limitation, the crowdsourcer has to hire workers in each country to perform tasks separately. Furthermore, the populations or the area of the territories of five countries are different. Thus, the crowdsourcer has to design mechanisms to allocate proper budgets to each region and considers the incentives for workers from different regions.

In this paper, we introduce and study a new scenario, where the crowdsourcer or buyer wants to buy service in a macroregion which is composed of several nonoverlapped

microregions. The crowdsourcer will get a utility for each microregion, respectively, and aggregate results in each microregion to get a final result. We introduce and study two optimization goals when combining results of each microregion, namely, weighted mean and maximin. Under the first model, we define the crowdsourcer's utility as the weighted mean of utility obtained in each microregion. Under the second model, the crowdsourcer's utility is defined as the minimum utility obtained in all microregions. Compared with single region settings, our multiple region setting leads to more useful solutions in practice. By this work, we proposed an incentive mechanism for the weighted mean model firstly, which consists of a task allocation algorithm and a worker compensation algorithm. Then, we extend the proposed mechanism to the maximin model.

The main of our paper can be summarized as follows:

- (i) A major contribution of this work is introducing a new problem, multiregion crowdsensing. Two objectives are introduced to measure the utility of multiregion crowdsensing. Although we study our problem in the context of crowdsensing, the framework and solution proposed in this paper can be applied to a broad range of domains, such as procurement and resource allocation
- (ii) For a multiunit budget-feasible mechanism, we propose a novel method via a proportional share allocation rule instead of the random sampling method applied in [5]. Also, our proposed mechanism can handle the multiregion procurement settings

## 2. Related Work

In recent years, the mobile crowdsensing system has a wide range of application in our daily life [1–3, 6–10]. At the same time, game theoretic models become an effective tool for incentive mechanism in mobile crowdsensing systems because of the strategic behavior of workers [11, 12]. Jin et al. considered integrating information quality into the design of the incentive mechanism for mobile systems [13]. Zhang et al. designed an incentive mechanism by which crowd workers were encouraged to label a set of binary tasks within strict budget constraints [14]. Another Zhang et al. studied three models of cooperation and competition among service providers [15]. Gao et al. devoted to providing a sufficient long-term participation incentive for crowd workers [16]. Feng et al. studied the key dimension of location information in smartphones which are assigned sensing tasks [17]. Qiao et al. studied the task crowdsensing under the background of the cross market [18]. Zhao et al. designed online incentive mechanisms for crowdsensing systems [19]. Liu et al. studied a two-phase group buying based on the auction mechanism for mobile crowdsensing [20]. Xu et al. improved online crowdsourcing auctions via a two-tiered framework [21]. Xu et al. [22] also proposed a mechanism with a strategy-proof discovery phase and a budget-constrained purchasing stage.

Budget-feasible incentive mechanism design, which was initially studied by Singer [23] when he found a novel class of mechanism design problems where the outcomes are limited by payments, becomes a newly emerging branch of mechanism design. Chen et al. designed a stochastic budget-feasible mechanism with a polynomial time approximate ratio and a deterministic budget-feasible mechanism with an exponential time approximate ratio [24]. Bei et al. studied the budget-feasible mechanism through random sampling and designed a constant approximation mechanism for all subadditive functions in Bayesian environment [25]. Anari et al. investigated a model to solve the mechanism design problem in the context of large-scale crowdsensing markets such as Amazon's Mechanical Turk [26]. Then, Singer applied this model to influence maximization in social networks [27] while Horel et al. applied this model to an experiment design problem [28]. Chan and Chen came up with multiunit budget-feasible mechanism design which has been a new setting in budget-feasible incentive mechanism design [5].

In our setting, a crowdsourcer or buyer wants to procure service in a macroregion, which is composed of  $[r] = \{1, \dots, r\}$  nonoverlapping microregions. A set of sellers or workers  $N = \{1, \dots, n\}$  is scattered in each microregion.

Let  $r(\cdot)$  be a function to specify the region. Let  $c_i$  index the private cost of the worker  $i$ . Each worker  $i$  could provide  $m_i$  units of homogeneous service, which is common knowledge. We use  $s_i$  to denote the set of service which could be provided by worker  $i$ . Let  $M = \cup_{i \in N} s_i$  denote the total service that can be provided by all workers in the macroregion, thus the number of total service in the macroregion  $|\mathcal{M}| = \sum_{i \in N} m_i$ .

This paper studies the service procurement with strategic players, and each worker will report his cost strategically in order to maximize his benefit. Let  $\mathbf{b} = (b_1, b_2, \dots, b_n)$  index bids of all workers. In our mechanism, there are two functions, a payment function and an allocation function. The payment scheme will decide the payoff  $p_i$  to the worker  $i$ , and the quota of service  $x_i$  procured from worker  $i$  is the outcome of the allocating scheme. The output of the designed mechanism is two vectors, payment vector  $\vec{\mathbf{p}} = \{p_1, p_2, \dots, p_n\}$  and allocating vector  $\vec{\mathbf{x}} = \{x_1, x_2, \dots, x_n\}$ . The benefit  $u_i(\vec{\mathbf{b}})$  of service provider  $i$  is quasilinear. Formally, we have

$$u_i(\vec{\mathbf{b}}) = \begin{cases} p_i(\vec{\mathbf{b}}) - c_i \cdot x_i(\vec{\mathbf{b}}), & \text{if } x_i(\vec{\mathbf{b}}) \leq m_i, \\ -\infty, & \text{otherwise.} \end{cases} \quad (1)$$

### 2.1. Problem Formulation

**2.1.1. Single-Region Utility Measurement.** Since the utility in each microregion is calculated, respectively, we need to know the quota of procured service in each microregion. We use  $k_j = \sum_{i \in N, r(i)=j} x_i$  to present the quota of procured services in microregion  $j$ . Then, the utility obtained in microregion  $j$  is measured by a symmetrical submodular function  $\mathcal{A}(k_j) = \sum_{i=1}^{k_j} v_i$ , where  $v_i$  is the marginal incremental value by

procuring the  $i$ -th unit service from the workers when the crowdsourcer has brought  $i-1$  service. The symmetrical submodularity means  $v_i$  will decrease as  $i$  grows, and  $v_1 \geq v_2 \geq \dots \geq v_n \geq 0$ .

**2.1.2. Multiregion Utility Measurement.** The crowdsourcer will get a final result by combining the outcomes in each microregion. Two optimization objectives are introduced, namely, *maximin* and *weighted mean*. In the weighted mean scheme, each microregion  $i$  is related with a region weight  $w_i (0 \leq w_i \leq 1, \sum_{i \in [r]} w_i = 1)$ , which measures the relative significance of every microregion. The global utility is calculated by the weighted mean of utility obtained in each microregion. Formally, the weighted mean maximization model can be summarized as follows.

Problem: weighted mean maximization.

Objective: maximize  $\mathcal{A}_{wm} = \sum_{j=1}^r w_j \mathcal{A}(k_j)$ ,

Subject to:

$$\sum_{j=1}^r w_j = 1, \quad 0 \leq w_j \leq 1, \quad (2)$$

$$\sum_{i \in N} p_i \cdot x_i \leq B. \quad (3)$$

In the maximin model, the global benefit of the crowdsourcer is calculated by the minimum of the accuracy rate achieved in all microregions.

Problem: minimum maximization.

Objective: maximize  $\mathcal{A}_{mm} = \min_{j \in \{1, \dots, r\}} \mathcal{A}(k_j)$ ,

Subject to:

$$\sum_{i \in N} p_i \cdot x_i \leq B. \quad (4)$$

We use  $\text{OPT}_{wm}(B, M)$  and  $\text{OPT}_{mm}(B, M)$  to index the optimal results for weighted mean and maximin, respectively.

Due to the fact that the worker  $i$ 's cost  $c_i$  is private information, the problem in our paper belongs to the single parameter mechanism design problem. According to the well-known Myerson's lemma [29], any truthful auctions must have a monotone allocating scheme and a payment scheme with a threshold payment. For the more complex multiunit setting in our paper, we have the following lemma.

**Lemma 1.** *A single-parameter multiunit domain, a normalized auction  $\mathcal{H} = (\mathbf{x}, \mathbf{p})$  is truthful, if and only if,*

(1) *Monotone allocation:*  $x_i(c_i, c_{-i}) > 0, \forall i \in N$ , implied  $x_i(c_i', c_{-i}) > 0$ , if  $c_i' < c_i$

(2) *Threshold payment to the selected workers:*  $p_i(c_i) = \sum_{j=1}^{x_i(c_i)} q_j$ , where  $q_j = \inf \{c_i : x_i(b_i) < j\}$

*Proof.* The first condition is trivial and could be obtained via Myerson's lemma [29]. For the second condition, we use  $w$  to index  $x_i(b_i)$ . Then, we obtain

$$\begin{aligned} p_i(b_i) &= w b_i + \int_{b_i}^{qw} w dz + \int_{qw}^{qw-1} w - 1 dz + \dots + \int_{q_2}^{q_1} 1 dz \\ &= q_w + q_{w-1} + \dots + q_1. \end{aligned} \quad (5)$$

Thus, the lemma holds.

We expect our designed auctions are

- (i) *Truthful:* for any worker  $i$ , his benefit is maximized by reporting his true cost, regardless of the bidding of others:  $\forall b_i, \vec{\mathbf{b}}_{-i} : u_i(b_i, \vec{\mathbf{b}}_{-i}) \geq u_i(b_i', \vec{\mathbf{b}}_{-i})$
- (ii) *Individually rational:* workers are not worse off by participating in the campaign. Formally, it requires  $\forall \vec{\mathbf{b}}_{-i} : u_i(b_i, \vec{\mathbf{b}}_{-i}) \geq 0$
- (iii) *Budget-feasible:* the total payment to workers does not exceed the buyer's budget
- (iv) *Computationally efficient and performance guaranteed:* the results of function  $x$  and  $p$  could be found out in polynomial time, and constant approximations are expected for our proposed mechanisms

### 3. Weighted Mean Maximization

Based on Myerson's well-known lemma [29], strategy-proof auctions must have a monotone allocating scheme and a payment scheme with a threshold payment. For the mechanism design under budget constraint, selecting items for the candidate set according to the benefit-to-cost ratio is a natural fit. The basic idea of our proposed mechanism is to select workers greedily based the marginal of each worker and terminate with an appropriate threshold payment.

First, we would like to introduce a greedy but nonmonotone selecting rule. Then, we modify this rule to make it monotone. Given a set of procured service  $M$ , we define the marginal value of a unit service  $s^i \in \mathcal{M} \setminus M$  for the weighted mean model as

$$\mathcal{U}_{i|M} = \mathcal{A}_{wm}(M \cup s^i) - \mathcal{A}_{wm}(M). \quad (6)$$

Our greedy algorithm selects a group of services one by one. In each stage  $i$ , it adds a service  $s^i$  that maximizes the benefit-to-cost ratio  $\mathcal{U}_{i|M_{i-1}}/c_i$ , in which we use  $M_{i-1}$  to denote first  $i-1$  services selected by our algorithm; without loss of generality, we assume  $M_0 = \emptyset$ . Since the utility is computed separately, the contribution of service  $s^i$  to the global utility equals the incremental value of  $s^i$  in its own microregion. We use  $u_{\{i|M_{i-1}\}}$  to denote the utility increment in

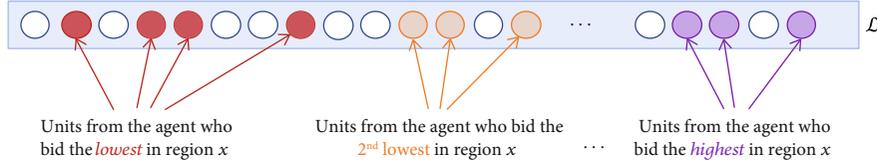


FIGURE 1: For each microregion, each worker's units are placed in  $\mathbb{L}$  as nonoverlapping clusters and ordered according to increasing bid. Total payments do not exceed the budget.

region  $r(s^i)$  with service  $s^i$  and the given set  $M_{i-1}$ . Formally, we have

$$s^i = \arg \max_{s^i \in \mathcal{M} \setminus M_{i-1}} \frac{u_{i|M_{i-1}}}{c_i} = \arg \max_{s^i \in \mathcal{M} \setminus M_{i-1}} \frac{u\{i | M_{i-1}\} \cdot w_{r(i)}}{c_i}. \quad (7)$$

If there are multiple candidate workers, we sort them lexicographically. For each  $s^i$ , selected in stage  $i$ , we use  $u_i^*$ ,  $w_i^*$ , and  $c_i^*$  as abbreviations for  $u_{\{i|M_{i-1}\}}$ ,  $w_{r(i)^*}$ , and  $c_i^*$ , respectively. Through the greedy selecting principle, we have a sorted list  $\mathbb{L}$ :

$$\mathbb{L} : \frac{u_1^* w_1^*}{c_1^*} \geq \frac{u_2^* w_2^*}{c_2^*} \geq \dots \geq \frac{u_M^* w_M^*}{c_M^*}. \quad (8)$$

We choose workers as candidates iteratively until the budget constraint is exceeded. Formally, we have

$$c_i^* \leq \frac{B}{(1 + n \cdot \mathcal{M})} \cdot \frac{w_i^* u_i^*}{\mathcal{A}_{wm}(M_i)}. \quad (9)$$

This allocation rule is a variation of the proportional share rule which is the basis for mechanisms under a budget. Observing that if  $k$  services are selected according to this rule and each selected worker is paid according to their proportional contribution, the one nice feature about this allocation rule is that for any submodular function, it can be proven that the incentive compatible payments do not exceed the workers' proportional contributions. We will further discuss payments after describing the entire mechanism below. Importantly, this allocation rule achieves a bounded approximation ratio.

$$\begin{aligned} & \sum_{i=1}^k B \cdot \left( \frac{\mathcal{A}_{wm}(M_i) - \mathcal{A}_{wm}(M_{i-1})}{\mathcal{A}_{wm}(M_k)} \right) \\ & \leq \frac{B}{\mathcal{A}_{wm}(M_k)} \sum_{i=1}^k (\mathcal{A}_{wm}(M_i) - \mathcal{A}_{wm}(M_{i-1})) \\ & = \frac{B}{\mathcal{A}_{wm}(M_k)} \cdot \mathcal{A}_{wm}(M_k) = B. \end{aligned} \quad (10)$$

**Lemma 2.** Given a sorted list  $\mathbb{L}$  and a service  $s^i \in \mathbb{L}$ , we use  $\mathbb{L}(s^i)$  to denote the position of  $s^i$  in  $\mathbb{L}$ . For any two services from the same region,

- (i) if  $c_i < c_j$ , then  $\mathbb{L}(s^i) < \mathbb{L}(s^j)$
- (ii) if  $c_i = c_j$ , then  $\mathbb{L}(s^i) < \mathbb{L}(s^j)$ , iff  $i < j$

*Proof.*

- (1) With  $c_i < c_j$ , we assume  $\mathbb{L}(s^j) < \mathbb{L}(s^i)$ . Thus, we have

$$\frac{u\{j | M_{j-1}\} \cdot w_{r(j)}}{c_j} \geq \frac{u\{i | M_{i-1}\} \cdot w_{r(i)}}{c_i}. \quad (11)$$

Since  $r(j) = r(i)$ , we have  $u_{\{j|M_{j-1}\}} = u_{\{i|M_{i-1}\}}$ , then  $c_i \geq c_j$ . Thus, we have a contradiction!

- (2) With  $c_i = c_j$ , service  $s^i$  and  $s^j$  are sorted according to the lexicographical order

No two services from the same region intersect with each other in  $\mathbb{L}$  if they are provided by two different workers. For each region, each agent's units are placed in  $\mathbb{L}$  as nonoverlapping clusters and ordered according to increasing bid. An illustration of the services' position is in Figure 1.

**Lemma 3.** The greedy selection principle is monotone.

*Proof.* For each service  $s^l$  chosen as the candidate, the rest of the service in  $M$  can be classified into 4 classes:

- (i)  $T^+$ : all services ranked higher than  $s^l$  in region  $r(l)$
- (ii)  $T^-$ : all services ranked lower than  $s^l$  in region  $r(l)$
- (iii)  $F^+$ : all services not in region  $r(l)$  whose rank is higher than  $s^l$
- (iv)  $F^-$ : all services not in region  $r(l)$  whose rank is lower than  $s^l$

Assume that worker  $l$  declares a cost  $c^l \leq c_l$  and achieves a utility  $u^l$ . Let  $\hat{T}^+$ ,  $\hat{T}^-$ ,  $\hat{F}^+$ , and  $\hat{F}^-$  denote the four updated classes. We can easily obtain that no service  $s^d$  can be moved forward to the front of  $s^l$ . If  $l = d$ , they are owned by the same worker, and we still have  $\mathbb{L}(s^l) < \mathbb{L}(s^d)$ . If  $l \neq d$ , we have  $c_d \geq c_l > c^l$ , and we obtain  $\mathbb{L}(s^l) < \mathbb{L}(s^d)$ . Thus, we obtain  $\hat{T}^+ \subset T^+$  and  $u^l \leq u^*$ . Services in  $F^+ \cup F^-$  are unchanged. No service  $s^d$  in  $F^-$  could be moved before  $s^l$ ; this is because the cost-to-benefit ratio of  $s^d$  is unchanged, and

$$\frac{w_d^* u_d^*}{c_d^*} \leq \frac{w_l^* u_l^*}{c_l^*} \leq \frac{w_l^* u^l}{c^l}. \quad (12)$$

Therefore, we obtain  $\widehat{F}^+ \subset F^+$ . Furthermore,

$$\begin{aligned}
\frac{c'}{w_i^* u_i^*} &< \frac{c_s^*}{w_i^* u_i^*} \\
&\leq \frac{B}{(1 + \ln|\mathcal{M}|) \sum_{i < l} w_i^* u_i^*} \\
&= \frac{B}{(1 + \ln|\mathcal{M}|) \sum_{(s' \in (T^+ \cup F^+ \cup S^+))} w_i^* u_i^*} \\
&= \frac{B}{(1 + \ln|\mathcal{M}|) (w_i^* \mathcal{A}_{wm}(|F^+| + 1) + \sum_{s' \in F^+} w_i^* u_i^*)} \\
&\leq \frac{B}{(1 + \ln|\mathcal{M}|) (w_i^* \mathcal{A}_{wm}(|\widehat{F}^+| + 1) + \sum_{s' \in \widehat{F}^+} w_i^* u_i^*)}. \tag{13}
\end{aligned}$$

Thus, each service will be allocated with a lower cost declaration, and the monotonicity achieved.

Through the greedy selecting principle, we get a candidate set, denoted by  $M_k$ . Unfortunately,  $M_k$  could not provide any lower bound guarantee, if we regard  $M_k$  as the winners. We consider another feasible candidate solution,  $s^*$  which is the single service with the highest marginal utility, i.e.,  $s^* = \arg \max_{s' \in \mathcal{M}} (w_i u_i / c_i) \cdot \{s^*\}$ , is a possible solution. In order to get a bounded approximation ratio, we will compare these two sets,  $M_k$  and  $\{s^*\}$ , and select the one with higher utility. We have the following theorem.

**Theorem 4.** Let  $\text{OPT}_{wm}(B, M)$  denote the optimal solution for the weighted mean model, i.e.,  $\text{OPT}_{wm}(B, \mathcal{M}) \leq (3 + 21\ln|\mathcal{M}|) \times \max \{\mathcal{A}_{wm}(\{s^*\}), \mathcal{A}_{wm}(M_k)\}$ .

Sort workers like  $\mathbb{L}$ . We use  $l$  to present the largest index such that  $\sum_{i=1}^l c_i < B$ . For analysis convenience, we add a new virtual worker  $v$  with a single service  $s^v$ , who reports his cost  $c_v = B - \sum_{i=1}^l c_i$  and his utility is  $u^v = \mathcal{A}_{wm}(s^v) = ((B - \sum_{i=1}^l c_i) / (c_l + 1)) (\mathcal{A}_{wm}(M_{l+1}) - \mathcal{A}_{wm}(M_l))$ . Obviously, the optimal result over all workers in  $\mathcal{M}^+ = \mathcal{M} \cup s^v$  is higher than the optimal result from all workers in  $\mathcal{M}$ . Through greedy selection principle, the first  $l$  workers will be chosen in both  $\mathcal{M}^+$  and  $\mathcal{M}$ . In the rest of the proof, we will index service  $s^v$  as  $s_{l+1}$ .

Due to the optimality, we have

$$\mathcal{A}_{wm}(M_{l+1}) = \sum_{i=1}^l w_i^* u_i^* + w_{l+1}^* u^v \geq \text{OPT}_{wm}(B, M). \tag{14}$$

Next,  $\mathcal{A}_{wm}(M_{l+1})$  will be regarded as the benchmark for analysis. Since all workers are sorted in decreasing order of their benefit-to-cost ratio, for each  $i \in [k+1, \dots, l+1]$ , we get

$$\frac{w_i^* u_i^*}{c_i^*} \leq \frac{w_{k+1}^* u_{k+1}^*}{c_{k+1}^*}. \tag{15}$$

Combining the above inequalities, which imply

$$\begin{aligned}
\frac{c_{k+1}^*}{w_{k+1}^* u_{k+1}^*} \sum_{i=k+1}^{l+1} w_i^* u_i^* &= \frac{c_{k+1}^*}{w_{k+1}^* u_{k+1}^*} (\mathcal{A}_{wm}(M_{l+1}) - \mathcal{A}_{wm}(M_k)) \\
&\leq \sum_{i=k+1}^{l+1} c_i^* \leq B.
\end{aligned} \tag{16}$$

Then, we obtain

$$c_{k+1}^* \leq B \cdot \frac{w_{k+1}^* u_{k+1}^*}{\mathcal{A}_{wm}(M_{l+1}) - \mathcal{A}_{wm}(M_k)}. \tag{17}$$

Based on the concept of notation  $k$ , we obtain

$$c_{k+1}^* > \frac{B}{(1 + \ln|\mathcal{M}|)} \cdot \frac{w_{k+1}^* u_{k+1}^*}{\mathcal{A}_{wm}(M_{k+1})}. \tag{18}$$

From inequalities (17) and (18), we know

$$\mathcal{A}_{wm}(M_{l+1}) - \mathcal{A}_{wm}(M_k) < (1 + \ln|\mathcal{M}|) \mathcal{A}_{wm}(M_{k+1}). \tag{19}$$

Thus, we have

$$\begin{aligned}
\mathcal{A}_{wm}(M_{l+1}) &= \mathcal{A}_{wm}(M_{l+1}) - \mathcal{A}_{wm}(M_k) + \mathcal{A}_{wm}(M_k) \\
&< (1 + \ln|\mathcal{M}|) \mathcal{A}_{wm}(M_{k+1}) + \mathcal{A}_{wm}(M_k) \\
&\leq (1 + \ln|\mathcal{M}|) \mathcal{A}_{wm}(s^*) + (2 + \ln|\mathcal{M}|) \mathcal{A}_{wm}(M_k).
\end{aligned} \tag{20}$$

Finally, we get

$$\begin{aligned}
\text{OPT}_{wm}(B, \mathcal{M}) &\leq \text{OPT}_{wm}(B, \mathcal{M}^+) \\
&\leq \mathcal{A}_{wm}(M_{l+1}) \\
&\leq (3 + 21\ln|\mathcal{M}|) \max \{\mathcal{A}_{wm}(s^*), \mathcal{A}_{wm}(M_k)\}.
\end{aligned} \tag{21}$$

However, one can show that simply selecting  $\max \{\mathcal{A}_{wm}(\{s^*\}), \mathcal{A}_{wm}(M_k)\}$  would violate incentive compatibility. To address this issue, we further design Algorithm 1 to ensure the monotonicity, in which the total cost is bounded by  $B/(1 + \ln|\mathcal{M}|)$ . Therefore, we remove the services with costs larger than  $B/(1 + \ln|\mathcal{M}|)$ , indexed by  $M(B/(1 + \ln|\mathcal{M}|))$ . Different from the previous greedy algorithm, we compare  $\mathcal{A}_{wm}(s^*)$  with the optimal solution for the set  $M_-$  with budget  $B/(1 + \ln|\mathcal{M}|)$ , where  $M_- = \mathcal{M} \setminus (\{s^*\} \cup M_{B/(1 + \ln|\mathcal{M}|)})$ . Algorithm 1 will compare  $\text{OPT}(B/(1 + \ln|\mathcal{M}|), M_-)$  with  $3 \mathcal{A}_{wm}(s^*)$  to determine the final outcome.

**Lemma 5.** Algorithm 1 is monotone.

*Proof.* We prove this lemma through contradiction. With a cost vector  $\vec{c}$ , we assume there exists a winner  $i \in \mathcal{M}$  with a single service such that he is not allocated when

```

1: Input: a set of workers  $N$  from  $r$  microregions with total  $\mathcal{M}$  service available, a budget  $B$  and a bidding profile  $\mathbf{b}$ ;
2: Output: a set of winning worker  $M^*$ 
3:  $M_k \leftarrow 0$ ;
4:  $s^* \leftarrow \arg \max_{s^i \in M} (w_i u_i / c_i)$ ;
5: while  $\mathcal{M} \setminus M_k$  and  $c_{i^*} \leq (B/(1 + \ln |\mathcal{M}|)) \times (w_i^* u_i^* / \mathcal{A}_{wm}(s^{i^*} \cup M_k))$  do
6:    $M_k \leftarrow M_k \cup s^{i^*}$ ;
7:    $s^{i^*} \leftarrow \arg \max_{\mathcal{M} \setminus (M_k^*)} (w_i^* u_i^* / c_i^*)$ 
8: end while
9:  $M^- \leftarrow \mathcal{M} \setminus (\{s^*\} \cup M_{B/(1 + \ln |\mathcal{M}|)})$ ;
10:  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) \leftarrow$  optimal solution for the set  $M_-$  with budget  $B/(1 + \ln |\mathcal{M}|)$ ;
11: if  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) \geq 3\mathcal{A}_{wm}(s^{i^*})$  then
12:    $M^* \leftarrow M_k$ 
13: else
14:    $M^* \leftarrow s^*$ 
15: end if

```

ALGORITHM 1: Task allocation algorithm ( $\mathcal{TA}(B, M)$ ).

declaring  $c_i' \leq c_i$ . The lemma will be analyzed from two conditions:

- (1) If  $M^* = s^*$ , then we have  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) \leq 3\mathcal{A}_{wm}(s^*)$ . Due to the fact that the cost of the owner of  $s^*$  will not change  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-)$  and  $\mathcal{A}_{wm}(s^*)$ ,  $s^*$  will still be allocated if  $s^*$ 's owners reported cost changes
- (2) If  $M^* = M_k$ , then it implies that  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) > 3\mathcal{A}_{wm}(s^{i^*})$ .  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-)$  increases as  $i$ 's reported cost decreases. Furthermore,  $i$  ranks higher in  $\mathbb{L}$ , since  $c_i' < c_i$ . Assume the update index in new  $\mathbb{L}$ ,  $j \leq i$ . According to the symmetrical submodularity of  $\mathcal{A}_{wm}(\cdot)$ , the incremental value of service at  $i$  is lower than that at  $j$ , i.e.,  $w_j^* u_j^* \geq w_i^* u_i^*$ . We obtain  $\mathcal{A}_{wm}(M_{i-1} \cup \{i\}) \geq \mathcal{A}_{wm}(M_{j-1} \cup \{j\})$ . Then, we have

$$\begin{aligned}
c_i' < c_i &\leq \frac{B}{(1 + \ln |\mathcal{M}|)} \times \frac{w_i^* u_i^*}{\mathcal{A}_{wm}(M_{i-1} \cup \{i\})} \\
&\leq \frac{B}{(1 + \ln |\mathcal{M}|)} \times \frac{w_j^* u_j^*}{\mathcal{A}_{wm}(M_{i-1} \cup \{i\})}.
\end{aligned} \tag{22}$$

Thus, it does not violate the budget constraint, and the worker  $i$  will still be in the solution set. The above conclusion implies that the worker  $i$  will still be selected as a winner when he decreases his bid. Thus, the allocating scheme is monotone.

Now, we move to our payment scheme which is presented in Algorithm 2. The basic idea of the payment scheme can be summarized as follows. If  $M^* = s^*$ , we pay  $s^*$ 's owner  $B$ . If  $M^* = M_k$ , we sort workers in region  $r(i)$  in nondecreasing order of their bids. Let  $y$  denote the position of worker  $i$ . Then, remove all service from worker  $i$  and sort the remain-

ing units to the sequence  $\mathbb{L}'$  according to the benefit-to-cost ratio. Then, for each worker  $j$  in the region  $r(i)$ , we index the first and the last of its service in  $\mathbb{L}'$  as  $f_j$  and  $l_j$ . According to Lemma 2, we obtain

$$\begin{aligned}
L' = f_1 < l_1 < f_2 < l_2 < \dots < f_{y-1} < l_{y-1} \\
< f_{y+1} < l_{y+1} < \dots < f_{|h|} < l_{|h|},
\end{aligned} \tag{23}$$

where  $|h|$  denotes the number of workers in region  $h$ . Thus, the threshold could be found between these intervals

$$\left(f_{y-1}, \dots, l_{y+1}\right], \left(f_{y+1}, \dots, l_{y+2}\right], \left(f_{y+2}, \dots, l_{y+3}\right], \dots, \tag{24}$$

and the search for the largest value worker  $i$  could be reported, which could make  $s^i$  be positioned next to one of the unit intervals and be selected. Based on worker  $i$ 's current reported cost information, we will insert unit  $s^i$  to the interval  $(f_{y-1}, \dots, l_{y+1}]$  and be selected. Worker  $i$  must report the cost lower in order to move  $s^i$  to some earlier intervals. So the search could begin at the interval  $(f_{y-1}, \dots, l_{y+1}]$ .

We can show that the payment for the selected workers can be bounded.

**Lemma 6.** *If  $s^i$  is selected,  $c_i \leq B/(1 + \ln |\mathcal{M}|)\mathbb{L}(s^i)$ .*

*Proof.* We assume that the mechanism will select first  $k$  services, i.e.,  $\{s^1, \dots, s^k\}$ , in  $\mathbb{L}$ . We then resort these  $k$  services in nonincreasing order of their benefit-to-cost ratio, as  $\{s_1^+, \dots, s_k^+\}$ . Let  $s^{j^+}$ 's incremental contribution as  $u_j^+$ . We use  $c_{s_j^+}$  and  $w_{r(s_j^+)}$ , as  $c_j^+$  and  $w_j^+$  for short.

```

1: Input: a set of workers  $N$  from  $r$  microregions with total  $M$  service available, a budget  $B$  and a bidding profile  $b$ , and a set of winning worker  $M^*$ ;
2: Output: a payment profile of workers  $\vec{p}$ ;
3:  $\vec{p} \leftarrow 0$ ;
4: if  $M^* = \{s^*\}$  then
5:    $p^* \leftarrow B$ 
6:   return  $\vec{p}$ ;
7: else
8:   while  $i \in M^*$  do
9:     1. Order the workers in region  $h$  as  $\{n_1, n_2, \dots, n_{|h|}\}$  according to increasing cost bid, let  $y$  satisfy  $n_y = i$ ;
10:    2. Order all services except from worker  $i$ , according to decreasing benefit-to-cost ratio, as  $\mathbb{L}' = \{s^1, \dots, s^{|\mathcal{M} \setminus \Phi_i|}\}$ ;
11:    3. For each  $j \in [|\mathcal{M}'|] \setminus \{y\}$ , get the first unit  $f_j$  and last unit  $l_j$ ;
12:    4. Get the last position  $k'$  s.t.  $(c'_k/w'_k, u'_k) \leq B/(1 + \ln |\mathcal{M}'|) \sum_{y \leq k'} w'_y u'_y$ ;
13:     $l_0 = 0$ ;  $k = l_{y-1} + 1$ ;  $z = 1$ ;  $S = \emptyset$ ;
14:    while  $k \leq k' + 1$  do
15:      if  $k = f_{y+z} + 1$  then
16:         $k = l_{y+z}$ ;  $z++$ ;
17:      else
18:         $\Gamma = \sum_{j=1}^{y-1} |s^j| + \sum_{j=2}^z |s^{y+j-1}| + j$ ;
19:         $\delta = \mathcal{A}_{wm}(\Gamma) - \mathcal{A}_{wm}(\Gamma - 1)$ 
20:         $t = \min \{ \delta(c'_k/w'_k, u'_k), (B\delta/(1 + \ln |\mathcal{M}'|))(1/\sum_{i \leq k-1} u'_i + l_{\Gamma-j+1} + \dots + l_\Gamma) \}$ 
21:         $S = S \cup \{t\}$ ;  $k++$ ;
22:      end if
23:    end while
24:    5.  $p_i \leftarrow$  the max value in  $S$ 
25:  end while
26:  return  $\vec{p}$ ;
27: end if

```

ALGORITHM 2: Payment algorithm ( $\mathcal{P}(B, M)$ ).

We assume that  $\exists 1 \leq e \leq k$  satisfying  $c_e^+ > B/(1 + \ln |\mathcal{M}'|)e$ , then

$$\begin{aligned}
\frac{c_k^*}{w_k^* u_k^*} &\leq \frac{B}{(1 + \ln |\mathcal{M}'|) \sum_{j \leq k} w_j^* u_j^*} \\
&= \frac{B}{(1 + \ln |\mathcal{M}'|) \sum_{j \leq k} w_j^+ u_j^+} \\
&\leq \frac{B}{(1 + \ln |\mathcal{M}'|) \sum_{i \leq e} w_i^+ u_i^+} \\
&\leq \frac{B}{(1 + \ln |\mathcal{M}'|) e \cdot w_e^+ u_e^+}.
\end{aligned} \tag{25}$$

Furthermore, we have  $B < (1 + \ln |\mathcal{M}'|)e \cdot c_e^+$  and so

$$\frac{B}{(1 + \ln |\mathcal{M}'|) e \cdot w_e^+ u_e^+} < \frac{c_e^+}{w_e^+ u_e^+}. \tag{26}$$

Combining these two inequalities, we have

$$\frac{c_k^*}{w_k^* u_k^*} < \frac{c_e^+}{w_e^+ u_e^+}. \tag{27}$$

It is a contradiction, because  $\mathbb{L}(s^k) < \mathbb{L}(s^e)$  in  $\mathbb{L}$ , and therefore,  $s^k$  cannot be selected.

With the previous lemma, we still do not know whether worker  $i$  could report  $c' > B/(1 + \ln |\mathcal{M}'|)\mathbb{L}(s^i)$  and get a new position index  $\mathbb{L}'(s^i)$ , and make  $s^i$  win.

**Lemma 7.** *The threshold reported cost  $\epsilon_i$  for a single service  $s^i$  is bounded by  $B/(1 + \ln |\mathcal{M}'|)\mathbb{L}(s^i)$ .*

*Proof.* Suppose  $\exists s^{e+}$ ,  $e \leq k$ , and  $\epsilon_e > B/(1 + \ln |\mathcal{M}'|)e$ . So, worker  $i_e^+$  could report a higher cost  $c'$  to satisfy  $(B/(1 + \ln |\mathcal{M}'|)e) < c' < \epsilon_e$  and make  $s^{e+}$  be selected. Now, assume worker  $i_e^+$  bid  $c'$  and let  $u'$  be the incremental value of unit  $s^{e+}$  in such a case. Obviously, we have  $u' \leq u_e^+$ , and

$$\frac{c'}{w_e^+ u'} \geq \frac{c'}{w_e^+ u_e^+} > \frac{B}{(1 + \ln |\mathcal{M}'|) e \cdot w_e^+ u_e^+} \geq \frac{c_k^*}{w_k^* u_k^*}. \tag{28}$$

For any  $s^{j+}$  where  $j < e$  (the unit used to rank ahead unit  $s^{e+}$  in the greedily sorted list), Let  $u'_j$  index its new incremental contribution as  $u'_j$  and new reported cost as  $c'_j$ . There could be

1: **for** each microregion  $j$  **do**  
 2:     Run task allocation  $\mathcal{TA}(B/r, M_j)$  under budget  $B/s$  and  $M_j$ ;  
 3:     Run worker compensation  $\mathcal{P}(B/r, M_j)$  under budget  $B/s$ .  
 4: **end for**  
 5: Aggregate the outcome from each microregion.

ALGORITHM 3: Maximin mechanism.

3 cases for the relations of the items of  $s^{j+}$  and  $s^{e+}$ , which are as follows:

*Case 1* (belong to the same worker). By the lexicographical rule,  $\mathbb{L}(s^{j+}) < \mathbb{L}(s^{e+})$  still holds in  $\mathbb{L}$ , and thus,  $u'_j > u'$ .

*Case 2* (belong to different workers in the same region). Since  $u_j^+ \geq u_e^+$ ,  $c_j^+ \leq c_e^+$ , and therefore,

$$c'_j = c_j^+ \leq c_e^+ \leq c'. \quad (29)$$

*Case 3* (belong to different regions). It is trivial to have  $u'_j = u_j^+$  and  $c'_j = c_j^+$ , and thus,

$$\frac{c'_j}{u'_j w_j^+} = \frac{c_j^+}{u_j^+ w_j^+} \leq \frac{c_k^*}{u_k^* w_k^*} < \frac{c'}{u' w_e^+}. \quad (30)$$

Moreover, we have  $u'_j = u_j^+ \geq u_e^+ \geq u'$ .

It is obvious to know in all 3 cases, service  $s^{j+}$  will be selected and be ranked before  $s^{e+}$  in  $\mathbb{L}$ . We denote  $e'$  as the unit  $s^{e+}$ 's new position index. We have  $e' \geq e$ , and thus,

$$c' > \frac{B}{(1 + \ln |\mathcal{M}|)e} \geq \frac{B}{(1 + \ln |\mathcal{M}|)e'}. \quad (31)$$

According to Lemma 6,  $s^{e+}$  could not be selected, and we have a contradiction. Based on the previous lemma, we can get the upper bound for total threshold payment of all the selected units directly.

**Lemma 8.** *The payment to all workers in worker compensation algorithm is upper bounded by  $B$ .*

*Proof.* By Lemma 7, the threshold payment is

$$\begin{aligned} \sum_{e=1}^k e_i &\leq \sum_{e=1}^k \frac{B}{(1 + \ln |\mathcal{M}|)e} \leq \sum_{e=1}^{|\mathcal{M}|} \frac{1}{e} \frac{B}{(1 + \ln |\mathcal{M}|)} \\ &\leq (1 + \ln |\mathcal{M}|) \frac{1}{(1 + \ln |\mathcal{M}|)} B. \end{aligned} \quad (32)$$

That is, it is upper bounded by  $B$ .

Now, we have the following theorem.

**Theorem 9.** *The weighted mean mechanism is truthful, budget feasible, individual rational and tractable.*

*Proof. Truthful:* based on Lemma 1, 5, and 6, we conclude that reporting truthful cost is a dominant strategy for each worker.

*Individual rational:* the worker whose cost is larger than  $B$  will be removed directly. Therefore, if  $M^* = s^*$ , the individual rational is obvious. For the case of  $M^* = M_k$ , the payment to each worker is the sum of the threshold bid of his procured services, which is larger than his bid unit cost. *Budget feasible:* from Lemma 8, we can assure the property of being budget feasible. *Computationally efficient:* it is obvious that the task allocation algorithm and payment algorithm can be computed in a polynomial time.

**Theorem 10.** *Weighted mean multiunit multiregion budget-feasible mechanism achieves a constant ratio approximation:  $(4 + 3 \ln |M|)$ .*

*Proof.* Drawing from the proof of Theorem 4, we obtain  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) \leq 3 \times \max \{\mathcal{A}_{wm}(\{s^*\}), \mathcal{A}_{wm}(M_k)\}$ . There are two cases. If  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) \geq 3 \times \mathcal{A}_{wm}(s^*)$ , from above inequality, we have  $\mathcal{A}_{wm}(M_k) \geq \mathcal{A}_{wm}(\{s^*\})$ . Based on Theorem 4, we have  $\text{OPT}_{wm}(B, M) \leq (3 + 2 \ln |M|) \times \mathcal{A}_{wm}(M_k)$ . Then, we have the expected approximation ratio. Otherwise, if  $\text{OPT}(B/(1 + \ln |\mathcal{M}|), M_-) < 3 \times \max \mathcal{A}_{wm}(\{s^*\})$ ; in this case, we get

$$\begin{aligned} \text{OPT}_{wm}(B, \mathcal{M}) &= \text{OPT}_{wm}(B, \mathcal{M}^-) + \mathcal{A}_{wm}(\{s^*\}) \\ &\leq (1 + \ln |\mathcal{M}|) \times \text{OPT}_{wm}\left(\frac{B}{(1 + \ln |\mathcal{M}|)}, M^-\right) \\ &\quad + \mathcal{A}_{wm}(\{s^*\}) \\ &\leq 3(1 + \ln |\mathcal{M}|) \mathcal{A}_{wm}(\{s^*\}) + \mathcal{A}_{wm}(\{s^*\}) \\ &= (4 + 3 \ln |\mathcal{M}|) \times \mathcal{A}_{wm}(\{s^*\}). \end{aligned} \quad (33)$$

#### 4. Minimum Maximization

In the weighted mean model, the buyer will perform a single auction to allocate tasks and compensate workers from all micro-regions. However, in the maximin, the buyer will allocate budget  $B/r$  to each microregion and perform mechanism in each microregion in parallel. Let  $M_j = \cup_{i \in N, r(i)=j} \Phi_i$  index services could be provided in microregion  $j$ . The maximin mechanism is referred to Algorithm 3.

**Theorem 11.** *Maximin mechanism is truthful, budget feasible, individual rational, and computationally efficient.*

Since mechanisms performed in each microregion are truthful, budget feasible, individual rational, computationally efficient, and independent from each other, multiregion maximin mechanism is also truthful, budget feasible, individual rational, and computationally efficient.

**Corollary 12.** *With payment algorithm  $\mathcal{P}(B/r, M_j)$  and task allocation algorithm  $\mathcal{T}\mathcal{A}(B/r, M_j)$ , the mechanism will achieve a constant approximation ratio  $4 + 3\ln M_j$ , in a single microregion  $j$ .*

*Proof. Trivial:* the setting in a microregion  $j$  can be regarded as a special case of weighted multiregion setting with  $r = 1$  and  $w_j = 1, j \in [r]$ .

Use  $\text{OPT}_j(B, M_j)$  to index the optimal result for the  $M_j$  with budget  $B$  in a single microregion  $j$ . Let us make these following observations on the optimal solutions in microregion  $j$ .

**Lemma 13.**  *$\text{OPT}_j(B, M_j)$  is nondecreasing with  $B$ .*

**Lemma 14.**  *$\text{OPT}_j(B, M_j)/B$  is nonincreasing with  $B$  between  $(0, +\infty]$ .*

**Lemma 15.** *For any  $B, M_j, \epsilon \in [0, 1]$ ,  $\text{OPT}_j(\epsilon B, M_j) \geq \epsilon \text{OPT}_j(B, M_j)$ .*

*Proof.* For the set  $M_j$ , if  $B$  increases, more service will be procured. However, the incremental value is nonincreasing under  $B$ . So the payment for unit service will be higher, as  $B$  increases. It implies that  $\text{OPT}_j(B, M_j)/B$  is nonincreasing with  $B$ , and we get

$$\frac{\text{OPT}_j(B, M_j)}{B} \leq \frac{\text{OPT}_j(\epsilon B, M_j)}{\epsilon B} \Rightarrow \epsilon \text{OPT}_j(B, M_j) \leq \text{OPT}_j(\epsilon B, M_j). \quad (34)$$

Let  $B_{\max}$  be the largest amount of the budget that the buyer used in a microregion in  $\text{OPT}_{mm}(B, M)$ . Let  $\beta$  index the  $B/B_{\max}$ . The  $1/\beta$  denotes the largest fraction of total budget allocated for a microregion in  $\text{OPT}_{mm}(B, M)$ , intuitively. We use  $k_j^*$  to denote the total service procured in microregion  $j$  in the optimal result  $\text{OPT}_{mm}(B, \mathcal{M})$ .

**Theorem 16.** *In the maximin model, the mechanism achieves a constant approximation ratio:  $r(4 + 3 \ln |\mathcal{M}|)/\beta$*

*Proof.* Based on the definition of the maximin problem, we get

$$\mathcal{A}_{mm}(B, \mathcal{M}) = \min([\mathcal{A}(k_1), \dots, \mathcal{A}(k_r)], \quad (35)$$

$$\text{OPT}_{mm}(B, \mathcal{M}) = \min(\mathcal{A}(k_1^*), \dots, \mathcal{A}(k_r^*)). \quad (36)$$

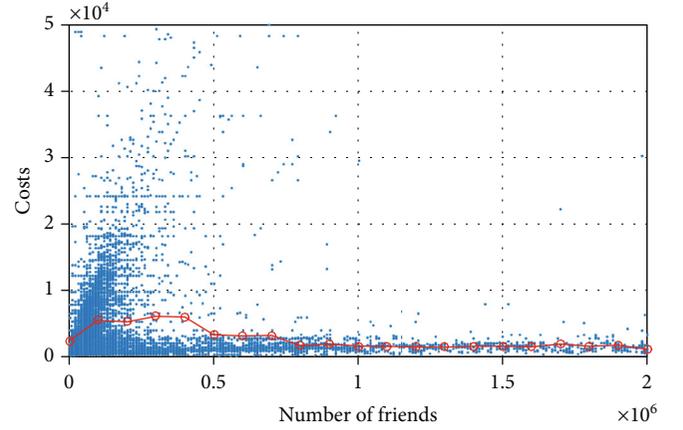


FIGURE 2: Cost information with numbers of friends.

Since the single-region performance is guaranteed in Corollary 12, we obtain

$$\begin{aligned} \mathcal{A}(k_j) &\geq \frac{1}{4 + 3 \ln M_j} \text{OPT}_j\left(\frac{B}{r}, M_j\right) \\ &\geq \frac{1}{4 + 3 \ln |\mathcal{M}|} \text{OPT}_j\left(\frac{B}{r}, M_j\right). \end{aligned} \quad (37)$$

Since we have known

$$B_j \leq \frac{B}{\beta}, \quad (38)$$

and the conclusions of Lemma 13, 14, and 15, we obtain

$$\text{OPT}_j\left(\frac{B}{r}, M_j\right) \geq \beta \text{OPT}_j\left(\frac{B/\beta}{r}, M_j\right) \geq \beta \text{OPT}_j\left(\frac{B_j}{r}, M_j\right). \quad (39)$$

Thus, we have

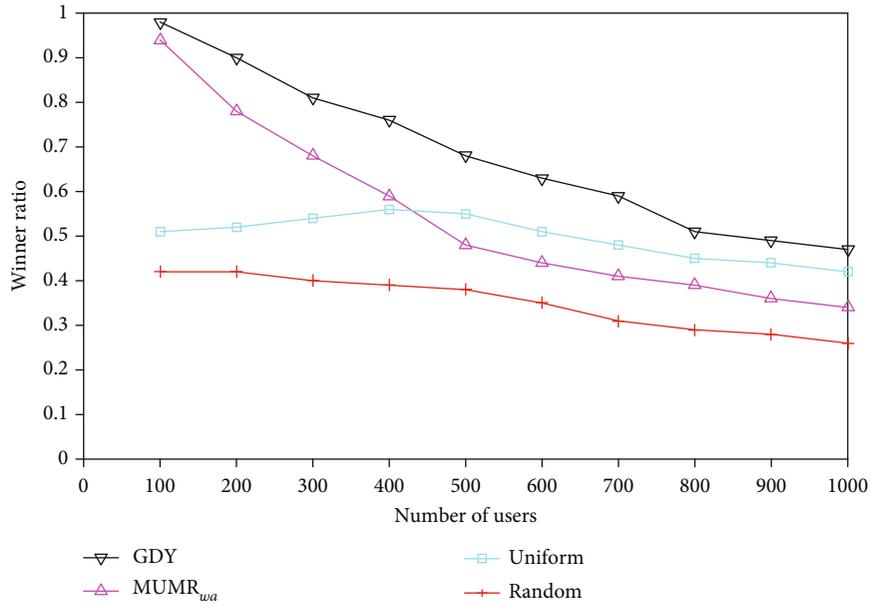
$$\mathbb{E}[\mathcal{A}(k_j)] \geq \frac{1}{4 + 3 \ln |\mathcal{M}|} \beta \text{OPT}_j\left(\frac{B_j, M_j}{r}\right). \quad (40)$$

Then, we have

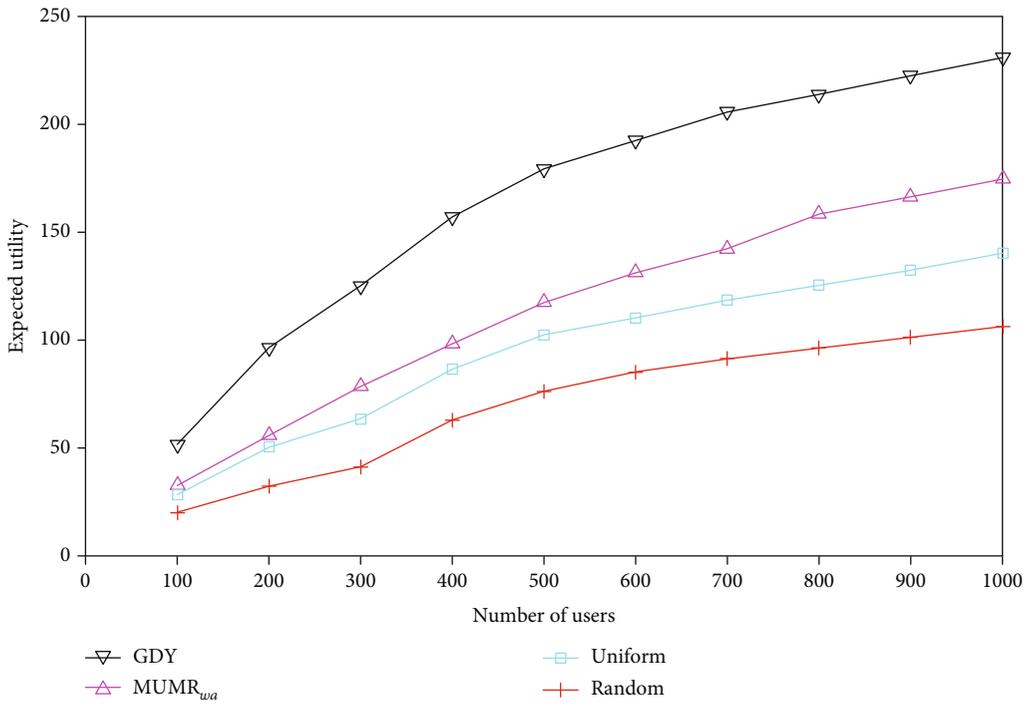
$$\begin{aligned} \mathcal{A}_{mm}(B, \mathcal{M}) &\geq \min_j \left( \frac{1}{4 + 3 \ln |\mathcal{M}|} \beta \text{OPT}_j\left(\frac{B_j, M_j}{r}\right) \right) \\ &\geq \frac{\beta}{(4 + 3 \ln |\mathcal{M}|)r} \min_j (\text{OPT}_j(B_j, M_j)) \\ &\geq \frac{\beta}{(4 + 3 \ln |\mathcal{M}|)r} \text{OPT}_{mm}(B, \mathcal{M}). \end{aligned} \quad (41)$$

## 5. Simulations

**5.1. Simulation Setup.** It is hard to get the bidding data of workers directly; thus, it is necessary for us to infer the cost profile via historical cost information. Drawing from [30],

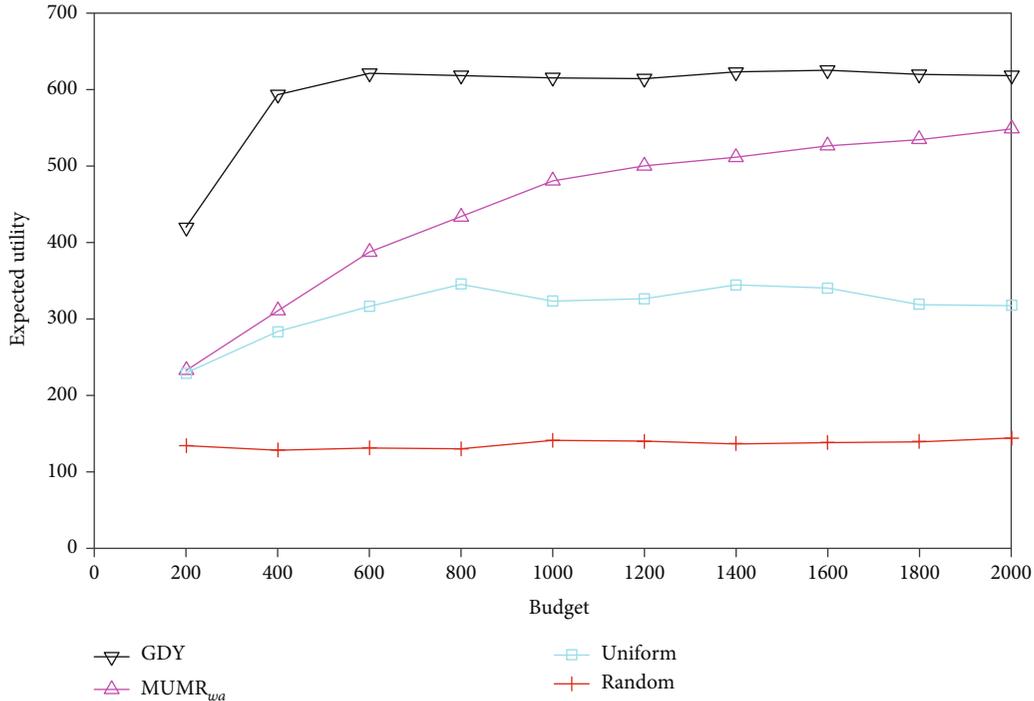


(a) Winner ratio



(b) Utility

FIGURE 3: Continued.



(c) Utility

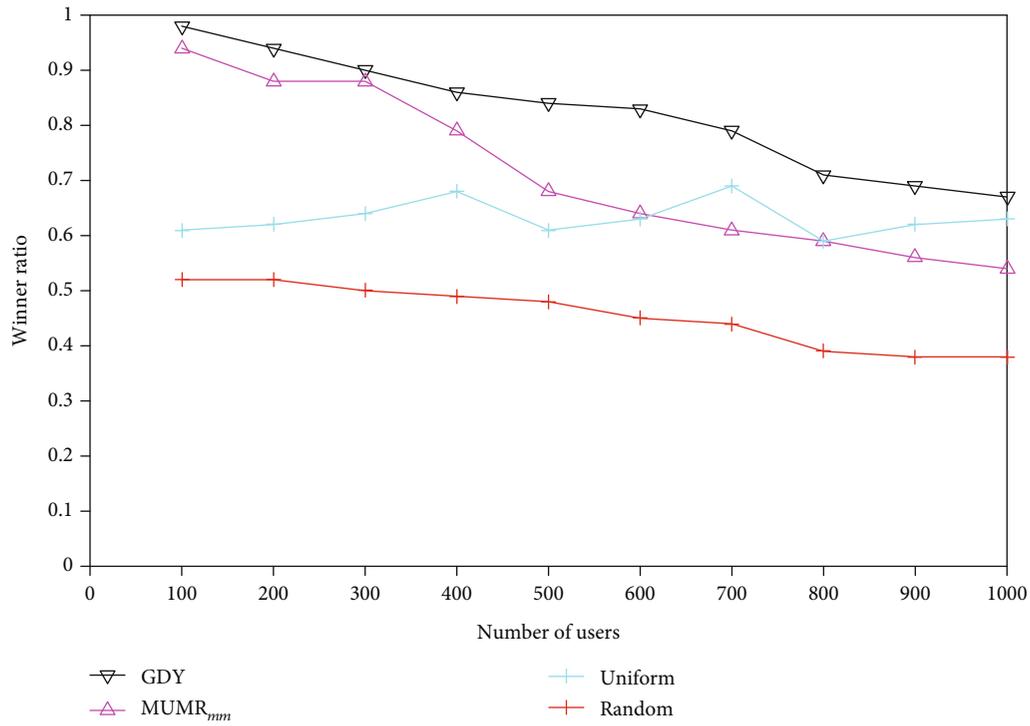
FIGURE 3: Results on weighted mean model.

we will estimate workers' cost information on a website, called weiboyi (<http://www.weiboyi.com>), which is for social advertising. In the website, advertisers are required to post their desired compensation to advertise to fans in the Weibo. From Figure 2, we show the cost information of 15012 advertisers. The range of their reported cost is from 0 to  $5 * 10^4$ , and the range of the number of their friends is from 0 to  $2 * 10^6$ . We use advertisers' number of friends to simulate workers' capacity. The cost information of advertisers will be normalized to (0,5) and the number of friends to [0,2]. The data set generation process is to select an instance in weiboyi data set and use the  $c_a/m_f$  for the worker's cost and the normalized number of friends for workers' capacity iteratively, in which  $c_a$  denotes the advertiser's normalized cost and  $m_f$  indexes the advertiser's normalized number of friends. We set crowdsourcer's utility  $\mathcal{A}(k)$  (in a single region) with three symmetrical submodular functions, where  $\mathcal{A}(k) = k^{0.9}$ ,  $\mathcal{A}(k) = k$ , and  $\mathcal{A}(k) = \sum_{i=1}^k 100/(100+k)$ . We consider the a service subscriber wants to hire workers from 6 microregions. To obtain statistically 330 sound results, we run the simulation 1000 rounds for each parameter setting, and the microregion weight will be randomized in each round of experiments.

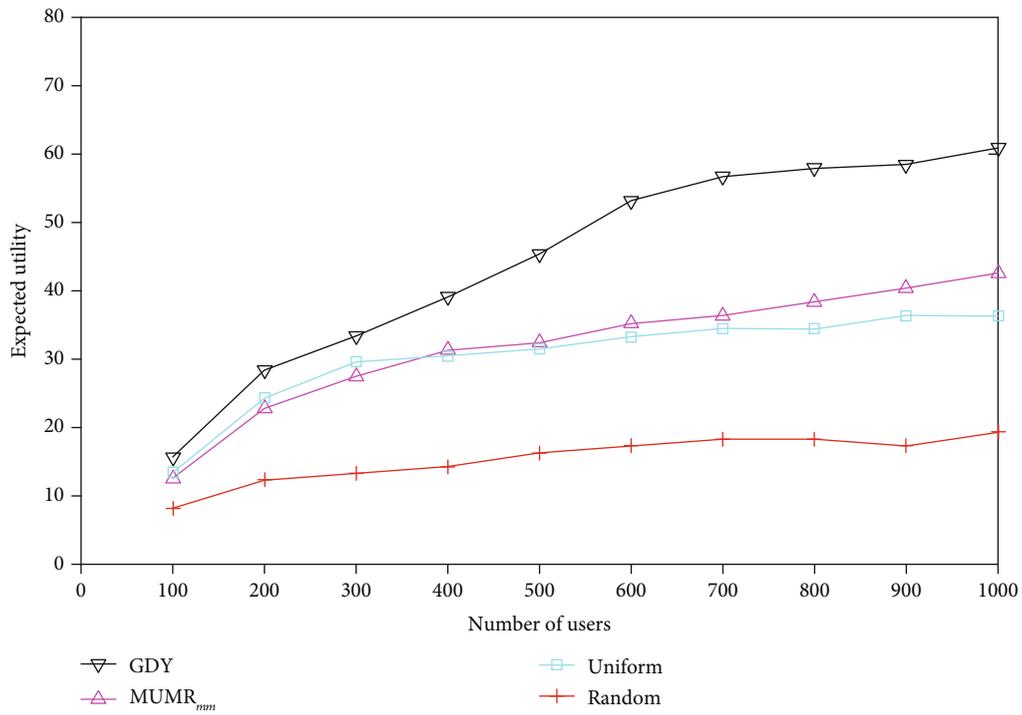
**5.2. Benchmarks.** We compare our designed auctions (MUMR<sub>wm</sub> and MUMR<sub>mm</sub> for two models) with 3 baselines. The first benchmark is the greedy algorithm, which will return the optimal solution. The greedy algorithm is not truthful and has full information of workers' costs. The sec-

ond benchmark is a uniform-price mechanism which is near-optimal, which is introduced in [31]. The last one is a random selecting algorithm. The mechanism chooses workers randomly and payoffs to workers are their declared cost, which are the lower bound performance of our designed problem. For maximin, the above three algorithms will be performed in each microregion in parallel with the same budget. We evaluate two performances of our proposed mechanism, crowdsourcer's valuation and winner ratio. The first one means the utility achieved in the entire region. The second one represents the percentage of selected workers over all workers.

**5.3. Results.** We will present the results of the winner ratio, firstly. Since the simulation results of the three crowdsourcer's valuations are similar, we just present the results of  $\mathcal{A}(k) = k^{0.9}$  here. We set the budget at 1000 and the number of workers from 100 to 1000. The results are shown in Figures 3(a) and 4(a) for both two models. We observe that the winner ratio decreases as the number of workers increases. The reason is that the auctions are more competitive with more workers. For the uniform pricing mechanism, the ratio will be stable, since the uniform pricing mechanism will choose a set of workers as the candidates when the scale of workers is relatively small. According Figures 3(b) and 4(b), we can know that the valuation of the buyer when the budget is 500 and the scale of workers is from 100 to 1000. The buyer's utility grows with the increase of workers in all algorithms in both models. With a larger scale of workers, the buyer will utilize the budget



(a) Winner ratio



(b) Utility

FIGURE 4: Continued.

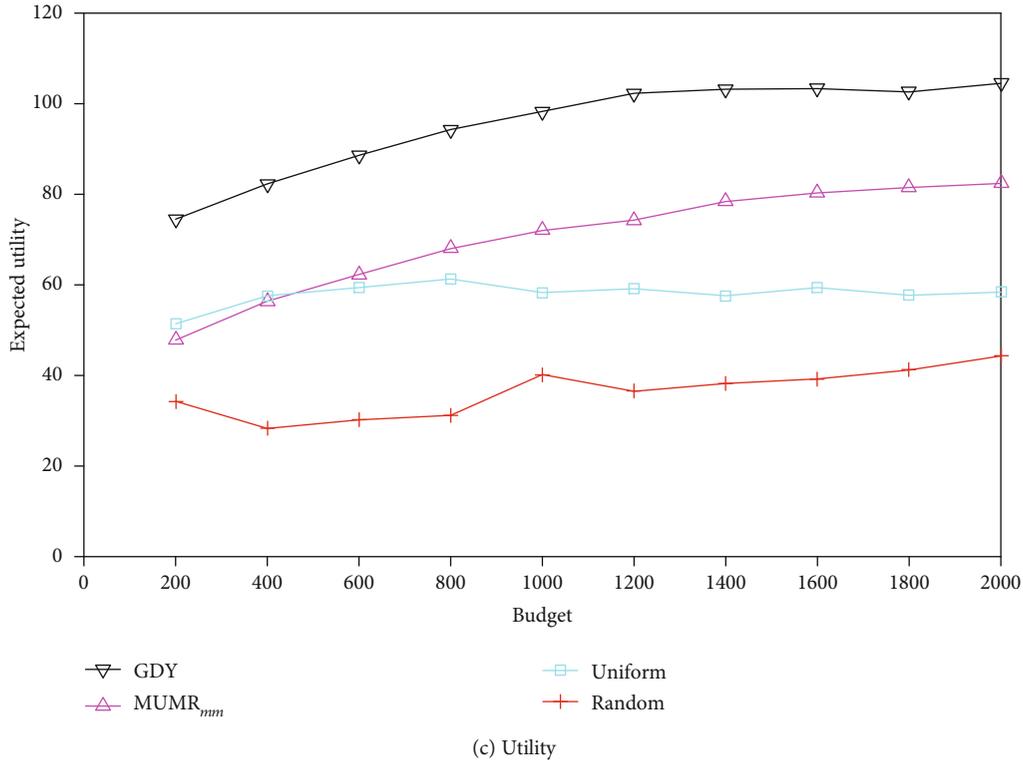


FIGURE 4: Results on maximin model.

more effectively. Our proposed mechanism outperforms random and uniform pricing mechanisms in the weighted mean model. MUMR's performance approaches the performance of the greedy algorithm and uniform pricing mechanism in the maximin model. Figures 3(c) and 4(c) show the service subscriber's valuation when the number of workers is 500 with the budget from 200 to 2000. From the figure, we know that the utility of the buyer increases simultaneously in algorithms MUMR<sub>wm</sub> and MUMR<sub>mmm</sub> and greedy algorithm as the budget increases. The reason is that when the budget becomes larger, more workers can be hired. The greedy-based method performs well in the simulation; however, it is not strategy-proof. The performance of the uniform price mechanism approaches the one of MUMR<sub>wm</sub> and MUMR<sub>mmm</sub> with the small budget. For the uniform price mechanism, the buyer's valuation decreases as the budget increases. Since the mechanism will infer a threshold payoff based on the workers' cost information, the one selected by the mechanism with larger budget may not be the optimal one.

## 6. Conclusion

In this work, we introduce and study the multiregion crowdsensing problem and design two models for the multiregion crowdsensing. We propose a novel multiunit budget-feasible mechanism to solve problems. Our designed mechanisms are budget feasible, truthful, and individual rational and have a constant approximation ratio. Extensive simulations demonstrate the effectiveness of our solution.

## Data Availability

The data adopted to support the findings of this manuscript are available from all authors upon request.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Authors' Contributions

Yu Qiao: the contribution of Yu Qiao is to provide the idea of the paper. Also, he finished the proof and wrote the paper. Jun Wu: the contribution of Jun Wu is to improve the idea and help the first author to finish the proof. Hao Cheng: the contribution of Hao Cheng is to conduct the experiments of the paper. Zilan Huang: the contribution of Zihan Huang is to conduct the experiments of the paper. Qiangqiang He: the contribution of Qiangqiang He is to depict the figures in the editable format in the paper after acceptance. Chongjun Wang: the contribution of Chongjun Wang is to improve the idea and help the first author to finish the proof.

## Acknowledgments

This paper is supported by the National Key Research and Development Program of China (Grant No. 2016YF-B1001102), the National Natural Science Foundation of China (Grant Nos. 61502227 and 61876080), the Fundamental Research Funds for the Central Universities

(No. 020214380040), and the Collaborative Innovation Center of Novel Software Technology and Industrialization at Nanjing University.

## References

- [1] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.
- [2] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones," *Mobile Sensing*, vol. 1, pp. 1–5, 2012.
- [3] S. Hu, L. Su, H. Liu, H. Wang, and T. F. Abdelzaher, "SmartRoad," *ACM Transactions on Sensor Networks*, vol. 11, no. 4, pp. 1–27, 2015.
- [4] R. Gao, M. Zhao, T. Ye et al., "Jig410 saw: indoor floor plan reconstruction via mobile crowdsensing," in *Proceedings of the 20th annual international conference on Mobile computing and networking - MobiCom '14*, pp. 249–260, 2014.
- [5] H. Chan and J. Chen, "Truthful multi-unit procurements with budgets," in *International Conference on Web and Internet Economics*, pp. 89–105, Springer, 2014.
- [6] Y. Zhang, Y. He, J. Wang et al., "Share brings benefits: towards maximizing revenue for crowdsourced mobile network access," in *2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1–9, San Diego, CA, USA, June 2017.
- [7] J. Xu, Z. Rao, L. Xu, D. Yang, and T. Li, "Mobile crowd sensing via online communities: incentive mechanisms for multiple cooperative tasks," in *14th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2017*, pp. 171–179, Orlando, FL, USA, October 2017.
- [8] J. Xu, J. Xiang, and Y. Li, "Incentivize maximum continuous time interval coverage under budget constraint in mobile crowd sensing," *Wireless Networks*, vol. 23, no. 5, pp. 1549–1562, 2017.
- [9] X. Yang, T. Wang, X. Ren, and W. Yu, "Copula-based multi-dimensional crowd-sourced data synthesis and release with local privacy," in *2017 IEEE Global Communications Conference, GLOBECOM 2017*, pp. 1–6, Singapore, December 2017.
- [10] X. Yang, C. Zhao, W. Yu, X. Yao, and X. Fu, "A user incentive-based scheme against dishonest reporting in privacy-preserving mobile crowdsensing systems," in *Wireless Algorithms, Systems, and Applications - 12th International Conference, WASA 2017*, pp. 755–767, Guilin, China, June 2017.
- [11] C. Li, S. Chang, H. Zhu, H. Chen, and T. Lu, "Lotus: evolutionary blind regression over noisy crowdsourced data," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1–9, 2018.
- [12] X. Zhang, Z. Yang, Y. Liu, and S. Tang, "On reliable task assignment for spatial crowdsourcing," *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 1, pp. 174–186, 2019.
- [13] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 167–176, 2015.
- [14] Q. Zhang, Y. Wen, X. Tian, X. Gan, and X. Wang, "Incentivize crowd labeling under budget constraint," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 2812–2820, 2015.
- [15] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 2830–2838, 2015.
- [16] L. Gao, F. Hou, and J. Huang, "Providing long-term participation incentive in participatory sensing," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 2803–2811, 2015.
- [17] Z. Feng, Y. Zhu, Q. Zhang, L. M. Ni, and A. V. Vasilakos, "Trac: truthful auction for location-aware collaborative sensing in mobile crowdsourcing," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 1231–1239, Toronto, ON, Canada, April 2014.
- [18] Y. Qiao, J. Wu, L. Zhang, and C. Wang, "Mechanism design for cross-market task crowdsourcing," in *Proceedings of the International Symposium on Quality of Service*, pp. 1–10, June 2019.
- [19] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsource tasks truthfully without sacrificing utility: online incentive mechanisms with budget constraint," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 1213–1221, Toronto, ON, Canada, April 2014.
- [20] T. Liu, Y. Zhu, and L. Huang, "Tgba: a two-phase group buying based auction mechanism for recruiting workers in mobile crowd sensing," *Computer Networks*, vol. 149, pp. 56–75, 2019.
- [21] J. Xu, C. Guan, H. Wu, D. Yang, L. Xu, and T. Li, "Online incentive mechanism for mobile crowdsourcing based on two-tiered social crowdsourcing architecture," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 55–63, Hong Kong, China, June 2018.
- [22] J. Xu, S. Yang, W. Lu, L. Xu, and D. Yang, "Incentivizing for truth discovery in edge-assisted large-scale mobile crowdsensing," *Sensors*, vol. 20, no. 3, p. 805, 2020.
- [23] Y. Singer, "Budget feasible mechanisms," in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 765–774, Las Vegas, NV, USA, October 2010.
- [24] N. Chen, N. Gravin, and P. Lu, "On the approximability of budget feasible mechanisms," in *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 685–699, San Francisco, California USA, January 2011.
- [25] X. Bei, N. Chen, N. Gravin, and P. Lu, "Budget feasible mechanism design: from prior-free to Bayesian," in *Proceedings of the 44th symposium on Theory of Computing - STOC '12*, pp. 449–458, 2012.
- [26] N. Anari, G. Goel, and A. Nikzad, "Mechanism design for crowdsourcing: an optimal 1-1/e competitive budget-feasible mechanism for large markets," in *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pp. 266–275, Philadelphia, PA, USA, October 2014.
- [27] Y. Singer, "How to win friends and influence people, truthfully: influence maximization mechanisms for social networks," in *Proceedings of the fifth ACM international conference on Web search and data mining - WSDM '12*, pp. 733–742, 2012.
- [28] T. Horel, S. Ioannidis, and S. Muthukrishnan, "Budget feasible mechanisms for experimental design," in *LATIN 2014: Theoretical Informatics*, pp. 719–730, Springer, 2014.

- [29] R. B. Myerson, "Optimal auction design," *Mathematics of Operations Research*, vol. 6, no. 1, pp. 58–73, 1981.
- [30] L. Zhang, H. Chen, J. Wu, C.-J. Wang, and J. Xie, "False-name-proof mechanisms for path auctions in social networks," in *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, pp. 1485–1492, 2016, IOS Press.
- [31] R. Engelbrecht-Wiggans and C. M. Kahn, "Multi-unit auctions with uniform prices," *Economic Theory*, vol. 12, no. 2, pp. 227–258, 1998.