Computer Networks 132 (2018) 1–14

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

A truthful incentive mechanism for mobile crowd sensing with location-Sensitive weighted tasks

Hui Cai, Yanmin Zhu*, Zhenni Feng

Shanghai Jiao Tong University, China

ARTICLE INFO

Article history:
Received 10 June 2017
Revised 27 October 2017
Accepted 24 December 2017
Available online 25 December 2017

Keywords:
Mobile crowd sensing
Location-sensitive
Incentive mechanism
Truthfulness
Auction

ABSTRACT

Mobile crowd sensing has emerged as an appealing paradigm to provide sensing data for its efficient economy. A large number of incentive mechanisms has been proposed for stimulating smartphone users to participate in mobile crowd sensing applications. Different from existing work, in addition to sensing tasks with diverse weights, we uniquely take into consideration the *crucial dimension of location information* when performing sensing tasks allocation. However, the location-sensitive weighted tasks are more vulnerable to the real life where each sensing task has the evident distinction. Meanwhile, the location sensitiveness leads to the increase of theoretical and computational complexity. In this paper, we investigate a truthful incentive mechanism which consists of two main components, winning bids determination algorithm and critical payment scheme. Since optimally determining the winning bids is *NP hard*, a near-optimal algorithm with polynomial-time computation complexity is proposed, which further approximates the optimal solution within a factor of $1 + \ln(n)$, where n is the maximum number of sensing tasks that a smartphone can accommodate. To guarantee the truthfulness, a critical payment scheme is proposed to induce smartphones to disclose their real costs. Through both rigid theoretical analysis and extensive simulations, we demonstrate that the proposed mechanism achieves truthfulness, individual rationality and high computation efficiency.

1. Introduction

Mobile crowd sensing with smartphones [1–4] has gradually bloomed into an appealing paradigm to collect various distributed sensing data for purposes [5]. The main application can be associated with the development and promotion of mobile crowd sensing systems such as the noise map calculation [6], real-time traffic delay prediction [7], citizen emergency monitoring [8] and so forth. Embedded with a variety of sensors, like GPS(Global Position System), microphone, camera, a smartphone can easily collect the essential data for various applications. Especially, smartphones gather ubiquitous data but only claim the little money, probably leading to enormous economic as well as the improvement of life quality [9].

A mobile crowd sensing system typically consists of a platform residing in the cloud, mobile smartphones and the platform users who consume sensing data. An example is illustrated in Fig. 1. The associated sensing tasks with diverse weights are released by the platform once it receives new arriving sensing requests from the platform users. Then the platform determines the appropriate set of smartphones to provide sensing services for new sensing tasks. Once receiving the hiring decision from the platform, the chosen smartphone starts to collect the required sensing data. Later, it submits the collected data to the platform, which aggregates the data to the platform user. Finally, the platform pays for the data. This demonstrates that a mobile crowd sensing system with geographically distributed smartphones can support a wide range of large-scale monitoring applications [8,10].

Motivation: *Stimulating smartphone users to participate in mobile crowd sensing system is fairly significant to the success of mobile crowd sensing with smartphones*. As we know, it incurs some non-negligible cost (*e.g.,* power consumption, bandwidth occupation) in consideration of limited resources when a smartphone provides sensing service for various applications [11,12]. Specifically, for our case, smartphones allocated to sensing tasks with higher weight are paid for more money, in consideration of more spent cost(*e.g.,* time cost). Furthermore, smartphone users may suffer the risk of privacy breach when providing sensing service related to their current location. Thus, smartphone users are usually reluctant to join a mobile crowd sensing system without sufficient incentives as compensation. However, the hypothetical sensing applications fail when no enough smartphone users provide the desired sensing

* Corresponding author.
E-mail addresses: carolinecai@sjtu.edu.cn (H. Cai), yzhu@sjtu.edu.cn, yzhu@cs.sjtu.edu.cn (Y. Zhu), zhennifeng@sjtu.edu.cn (Z. Feng).

https://doi.org/10.1016/j.comnet.2017.12.012
1389-1286/

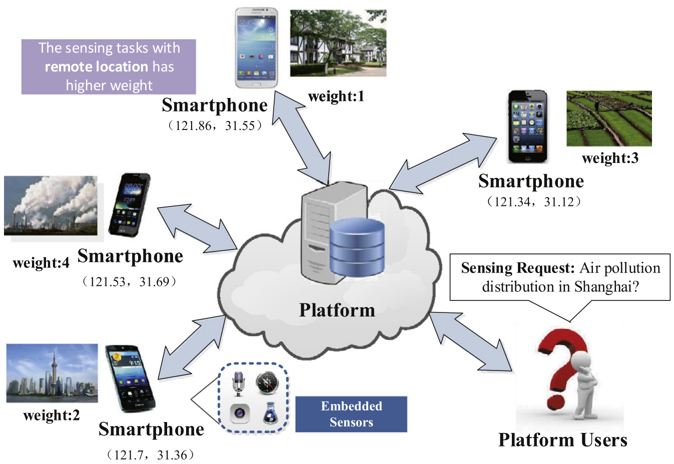

Fig. 1. An example of mobile crowd sensing applications. Smartphones are distributed over a large city. The location of each smartphone is described by a 2-tuple, (*longitude, latitude*). Each sensing task is to collect air pollution data. One with higher weight indicates the location to collect data is more distant. A platform user advertises sensing tasks through the platform residing on the cloud, and smartphones can contribute to the sensing tasks by returning their sensed data (*e.g.*, a photo of its surroundings) to the platform.

service. Unfortunately, although a large number of mobile crowd sensing applications [8,13–15] have been proposed, most of them have assumed that smartphones voluntarily contribute to the mobile crowd sensing system, which is not impractical in the real world.

The problem of stimulating smartphone users to participate in mobile crowd sensing applications is highly complicated because of smartphone user's strategic behaviors. In general, strategic users are selfish and self-interested. Thus, a smartphone user may misreport his real cost for maximizing his utility regardless of others' utility. The mobile crowd sensing system may bear the enormous economic loss in the long term. Furthermore, the cost information is private and unknown to the platform, which has no access to reveal the real value. Thus, designing a truthful incentive mechanism is non-trivial to induce smartphones to disclose their real cost.

There have been several research efforts on developing incentive mechanisms for mobile crowd sensing applications, which can generally be divided into three categories. One category of existing work [16,17] tends to adopt auctions for inducing cooperation from smartphones. In [16], Yang et al. design two incentive mechanisms to maximize the platform utility. Zheng et al. [18] propose a single parameter auction mechanism for the data procurement procedure under the known cost distribution, aiming to minimize the expected payment. The second category of existing work [17,19] utilizes the invisible indicator to stimulate cooperative behaviors of selfish smartphone users. Zhang et al. in [17] provide an incentive mechanism based on the repeated game to model the user's reputation. The final category of the existing work [9,20] designs an incentive mechanism based on the estimations of quality of information (QOI) submitted by smartphone users to avoid the strategy behaviors of smartphones.

In addition to sensing tasks with diverse weights, another significant observation is illustrated that location sensitiveness is central to most mobile crowd sensing application. A sensing task typically specifies the location where the sensing task should be performed. This is because that the desired sensing data are closely related to the specific location. A sensing data collected at an irrelevant location is meaningless or even invalid. This practical consideration on location sensitiveness caters to *more meaning and accurate matching between demands and supplies of sensing services*. Another significant observation is that most smartphone users are only willing to provide sensing service in a crowded area, thus, reluctant to spend more effort for sensing tasks with a remote location. Furthermore, we emphasize that the sensing task in a remote location has higher weight, indicating that it is more significant to incentivize smartphones to participate in such sensing task. Unfortunately, most of existing designs of incentive mechanism [21,22] have neglected this important dimension of location information in their designs.

In this paper, we introduce a practical *reverse auction framework*, in which the platform announces sensing tasks each of which has a location and weight attribute, and smartphones can submit multiple bids for a set of tasks within their service coverage according to the interest. Meanwhile, the corresponding claimed cost is disclosed with each submitted bid. For minimizing the social cost, the platform determines the set of winning bids, allocating all sensing tasks to the associated winning smartphone users. All winning users are paid for the rewards according to their contributions.

To make this reverse auction framework actually work for mobile crowd sensing with location-sensitive weighted tasks, we aim at designing a truthful incentive mechanism by which each smartphone would truthfully disclose its real cost. For our known combinatorial auction problem, *two critical problems* have to be addressed: 1) A task allocation algorithm to determine the cost-efficient bids has to be proposed, and 2) A payment decision scheme to fight against the strategic behavior of smartphones has to be proposed. Unfortunately, we prove that the optimal winning bids determination problem is *NP hard*. Thus, in this paper we design a truthful incentive mechanism which consists of *two main components*, the winning bids determination algorithm and critical payment scheme. The first algorithm approximately determines the set of winning bids, while the second algorithm determines the critical payment for each winning bid. Furthermore, we prove that the near-optimal algorithm can approximate the optimal solution within a factor of $1 + \ln(n)$, where n is the maximum number of sensing tasks that a smartphone can accommodate. As an additional part, we theoretically prove that the number of winning bids from the approximation solution has an upper bound α compared to that of the optimal solution when all smartphone users have the same claimed bid price. After rigorous theoretical proof and extensive simulations, the results demonstrate that our mechanism achieves truthfulness, individual rationality, high computational efficiency and modest overpayment ratio.

We highlight the main intellectual contributions as follows.

- We consider location sensitiveness as well as sensing tasks with diverse weights in the design of a truthful incentive mechanism for mobile crowd sensing with location-sensitive weighted tasks. Especially, the consideration of location information essentially increases the problem complexity of combinatorial auction design.
- We design an algorithmic mechanism for mobile crowd sensing with location-sensitive weighted tasks. We prove that optimally determining the winning bids with location sensitiveness is *NP hard*. The proposed mechanism consists of a *polynomial time and near-optimal task allocation algorithm* and a novel payment scheme that guarantees the truthfulness of the proposed mechanism.
- Through both rigid theoretical analysis and extensive simulations with real trace data sets, we demonstrate that the proposed mechanism achieves the desired properties of truthfulness, individual rationality and high computation efficiency. In addition, we theoretically prove that the number of winning bids from the approximation solution has an upper bound α compared to that of the optimal solution when all smartphone users have the same claimed bid price.

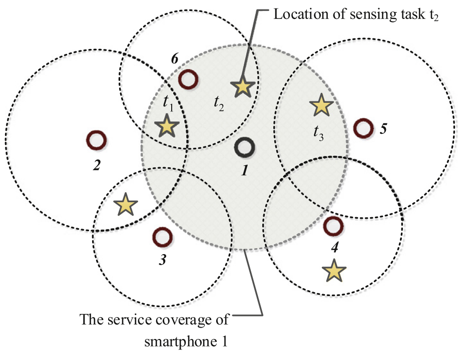

Fig. 2. An example of service coverage of smartphones. The filled area denotes the service coverage of smartphone 1. Three sensing requests fall within the coverage of smartphone 1. Thus, smartphone 1 can provide sensing service to each of the three tasks t_1, t_2, t_3. The sensing coverage of different smartphones can be different.

Some preliminary results of this work were reported in the work [23]. The remainder of the paper is organized as follows. In Section 2, we first present the system model, the reverse auction framework and the mathematical problem formulation. Then, the detailed design of our mechanism is described in Section 3. The theoretical analysis of the proposed mechanism is presented in Section 4. Next, we evaluate the performance of proposed mechanism in Section 5. Then, we review related work in Section 6. Finally, we conclude the paper and discuss future research directions in Section 7.

2. System model and problem formulation

2.1. System model

We consider a mobile crowd sensing system with smartphones consisting of a *platform, platform user*, and many *smartphone users*. The platform resides in the cloud. The platform accepts sensing requests from *platform users* who connect to the platform via the cloud. The platform periodically publicizes sensing tasks to be performed by smartphones. Let T denote the set of sensing tasks, $T = \{t_1, t_2, \cdots, t_m\}$. A sensing task t_i specifies the desired sensing service, the corresponding location where the sensing data should be collected and the task weight. Let $p(t_i)$ and $w(t_i)$ denote the location of the sensing task and the task weight respectively. Different sensing tasks have a respective weight for *platform users*, indicating the corresponding significance. Especially, the higher weight means that the sensing location is more distant for noise mapping application [6]. By carrying out the market techniques like survey, We divide all sensing tasks into K categories. The corresponding weight vector is denoted as $W = (w_1, w_2, \ldots, w_K)$, where $w_i < w_k$, for all $1 \le i < k \le K$. Note that a sensing task is atomic, meaning that it is either entirely performed or it is not completed. The difference from our prior work [23] lies in our distinction with the task weight, leading to more cost-effective task allocation. Thus, we expand the earlier work to the generalized version, which is more consistent with the realistic condition.

There are n smartphones which are interested in performing sensing tasks and the set of smartphones is denoted by $N = \{1, 2, \cdots, n\}$. Each smartphone i is aware of its own location l_i, through Global Positioning Systems (GPS) or other localization schemes [13]. Each smartphone i is intrinsically associated with a geographical service coverage, denoted by ϑ_i (as illustrated in Fig. 2). Only those sensing tasks within the service coverage may potentially receive sensing service of smartphone i. The key notations of this paper are listed in Table 1.

The service coverage can be different from smartphone to smartphone, and is dependent on the smartphone user and the associated preferences or restrictions. It is practical to assume that the service coverage ϑ_i of smartphone i is dependent on the current location l_i of the smartphone. Thus, given the current location l_i of smartphone i, one is able to determine its service coverage ϑ_i. It is worth noting that each smartphone can have a different function mapping its current location to its service coverage ϑ_i. We assume that each smartphone in the system would share such information with the platform. With ϑ_i, the platform is able to determine the subset T_i of sensing tasks, $T_i \subseteq T$, that smartphone i is able to provide sensing service. Each smartphone should not misreport its own service coverage. Misreporting may result in failure of completing a sensing task and a serious penalty would be reinforced. Protection of location privacy of smartphones and is beyond the scope of this paper and subject to future research.

Table 1
Key notations.

Notation	Description
T	Set of sensing tasks
N	Set of smartphones
B	Set of all submitted bids
S	Set of winning bids
w_i	Task weight of category i
B_i	Set of k_i bids submitted from smartphone i
β_i^k	The kth tasks-bid pair from smartphone i
T_i	Set of sensing tasks within geographical service coverage of smartphone i
r_i	Maximum number of winning bids for smartphone i
Q_i^k	The batch of sensing tasks from β_i^k
b_i^k	The claimed cost of the subset of Q_i^k
c_i^k	The true cost of the subset of Q_i^k
p_i^k	The payment to winning bid β_i^k
$u_i(\beta_i^k)$	The payoffs of user i for bid β_i^k

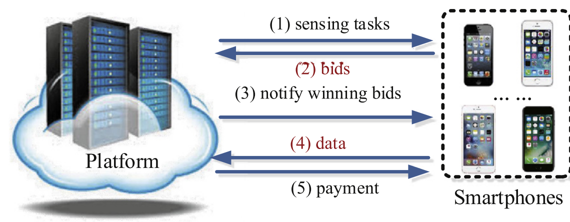

Fig. 3. The interactions between the platform and smartphones within the reverse auction model.

2.2. Reverse auction framework

In the mobile crowd sensing system, smartphones compete for opportunities to provide sensing services. We introduce a *reverse auction framework* for modeling the interactions between the platform and the smartphones, in which smartphones are the *sellers* and the platform is the *buyer* (buying sensing services). The framework is such called as it is a type of auction in which the roles of buyer and seller are reversed.

With the framework, the interactions between the platform and smartphones are described as follows, which is also illustrated in Fig. 3.

1. The platform advertises a set T of sensing tasks to all the smartphones in the mobile crowd sensing system.
2. Each user i replies with a set B_i of k_i bids, each of which is a tasks-bid pair $\beta_i^k = (Q_i^k, b_i^k)$, in which Q_i^k is a subset of sensing tasks that are within its geographical service coverage, $Q_i^k \subset T_i$, and b_i^k is called *claimed cost* of the subset of tasks Q_i^k which is *reserved price* that user i wants to charge for the service. In addition, each user i submits a number r_i which is the max-

imum number of bids that it can accommodate. Note that a smartphone user could not submit two bids for the same set of sensing tasks.

3. The platform determines whether a bid is winning or not, *i.e.*, it selects a subset S from all submitted bids, $S \subseteq \bigcup_{i \in N} B_i$, in which $\beta_i^k \in S$ indicates that smartphone i would perform the set of sensing tasks Q_i^k in its bid β_i^k.
4. Each smartphone i performs the sensing tasks in its winning bids which are $S \bigcap B_i$ and sends the sensed data back to the platform.
5. Each smartphone i is paid an amount of money p_i^k for its winning bid β_i^k, for each $\beta_i^k \in S \bigcap B_i$.

Due to the distributed nature of mobile crowd sensing, the real cost c_i^k of performing the set Q_i^k of the sensing tasks is *private* and *unknown* to others. Each smartphone is owned by a selfish individual who always tries to maximize its own payoff. Thus, it is possible for a smartphone to manipulate the claimed cost for its own good, *i.e.*, b_i^k may not be equal to c_i^k. This kind of behavior is typical *strategic behavior*. The existence of strategic behaviors makes it difficult for the platform to *hire* those smartphones with lower costs.

Smartphones are strategic and hence each smartphone i claims cost b_i^k for bid β_i^k that may be different from the real cost c_i^k for maximizing its own payoff (or benefit). We define the payoff of a smartphone as follows.

Definition 1 (Smartphone payoff)**.** The payoff of a smartphone is the sum of payoffs of all its winning bids. The payoff of a winning bid is the difference between the payment it receives and its real cost. The payoff of the smartphone can be computed as follows,

$$u_i = \sum_{\beta_i^k \in S \bigcap B_i} (p_i^k - c_i^k). \tag{1}$$

2.3. Problem formulation

We next give the mathematical formulation of the mechanism design problem. In this work, the platform determines the winning bids and allocates sensing tasks to the corresponding smartphones.

Definition 2 (Winning bids determination problem (WBDP))**.** For all submitted bids S of bids, the winning bids determination problem is defined as follows:

$$min \sum_{\beta_i^k \in S} c_i^k \tag{2}$$

$$s.t. \ |S \bigcap B_i| \le r_i \tag{3}$$

$$\bigcup_{\beta_i^k \in S} Q_i^k = T. \tag{4}$$

Remarks. The definition of WBDP shows the objective of the platform selecting the winning bids is to minimize the *social cost* which is the sum of the real costs of smartphones completing all the sensing tasks. The constraint (3) indicates that the number of winning bids of user i could not exceed its maximum number r_i. The constraint (4) shows that the platform guarantees that each sensing task is finished unless a sensing task is within less than one submitted bids. Notice that if no bid that covers a sensing task t_j and then it is obvious that t_j could not be finished. And we also exclude the situation where only one bid covers the sensing task t_j in order to prevent the monopoly. Thus, we assume that there are enough smartphones within the system and more than one smartphones compete for each task.

Definition 3 (Truthful mechanism design problem (TMDP))**.** For each bid of smartphone i, let $\tilde{\beta}_i^k = (Q_i^k, c_i^k)$ denote the truthful bid and $\beta_i^k = (Q_i^k, b_i^k)$ denote the untruthful bid. The payoffs of user i for the truthful bid and the untruthful bid is $u_i(\tilde{\beta_i^k}) = p_i(\tilde{\beta_i^k}) - c_i^k$ and $u_i(\beta_i^k) = p_i(\beta_i^k) - c_i^k$, respectively. The TMDP problem is to design a *payment scheme* such that

$$u_i(\tilde{\beta_i^k}) \ge u_i(\beta_i^k). \tag{5}$$

Remarks. A payment scheme resulting from the TMDP can guarantee that smartphones declare their costs truthfully.

The goal of our work is to design a truthful mechanism that solves the two problems defined above. This mechanism should also have the following desired properties.

Definition 4 (Individual rationality)**.** The payoff of each bid of user i is nonnegative, $p_i^k \ge c_i^k$.

Remarks. To stimulate smartphones to participate in the mobile crowd sensing applications, each smartphone has non-negative utility at least.

Definition 5 (Computational efficiency)**.** An algorithm has the property of computational efficiency if it terminates in polynomial time.

Remarks. The computational efficiency of the algorithms of solving WBDP and TMDP is of great importance in realistic scenarios. Any optimal algorithm with high complexity is useless in reality.

Definition 6 (Cost-efficient rank criterion)**.** For any bid $\beta_i^k = (Q_i^k, b_i^k)$ of each user i, the weight of any sensing task t_i is $w(t_i)$ for any $t_i \in Q_i^k$. Then the cost-efficient rank criterion $r(\beta_i^k)$ of the bid β_i^k is defined as

$$r(\beta_i^k) = \frac{b_i^k}{\sum_{t_i \in Q_i^k} w(t_i)}, \tag{6}$$

Remarks. The cost-efficient rank criterion is considered as the only recruiting rule for winning bids determination procedure.

3. Design of our mechanism

Our mechanism consists of two components: The first component solves WBDP to determine the winning bids, and the second component is a payment scheme for solving TMDP. Before describing the algorithm for the first component, we first analyze the complexity of solving the WBDP, *i.e.*, determining the winning bids to minimize the social cost. We rigorously prove that WBDP is NP hard. Then, we propose an algorithm that obtains a near-optimal solution with low computational complexity, which is different from the traditional truthful mechanism. Finally, we propose the payment scheme for solving TMDP to induce smartphones to disclose their real costs truthfully.

3.1. Complexity analysis of WBDP

It is very important to solve WBDP with a time efficient algorithm. Unfortunately, as we are going to prove next, WBDP is NP hard.

Theorem 1. *The WBDP is NP hard.*

Proof. To prove WBDP is NP-hard, we can prove that its decision version is NP-complete. For the decision problem, we should demonstrate that it belongs to NP, and then find another known NP-complete problem that could be reduced to the decision version of WBDP in polynomial time.

The decision version of our problem is a modified minimum weighted set cover (MWSC) problem in which some pairs of subsets are mutually exclusive, *i.e.,* they could not win simultaneously. The decision problem belongs to NP as checking whether a solution is correct or not could end up in polynomial time.

Next, we use the decision version of minimized weighted set cover as the known NP-complete problem. An instance of the known problem is defined as follows.

Definition 7 (An instance of MWSC)**.** For a universe set U and a set $S=\{s_1, s_2, \cdots, s_n\}$, each s_i satisfies $s_i \subseteq U$ and its weight is $w(s_i)$. The question is whether exists a set $Q \subseteq S$, the union $\bigcup_{s_i \in Q} s_i$ of members from Q covers all elements of the universal set U, and further $\sum_{s_i \in Q} w(s_i) \leq K$. We regard the instance as A in later discussion.

Next, we change the instance of MWSC to an instance of our problem. Suppose that set T of the sensing tasks is the universal set U, and the subset of tasks Q_i^k for each submitted bid is considered as the element s_i of set S. Meanwhile, the claimed cost b_i^k of each bid is the corresponding weight $w(s_i)$ of element s_i. Thus, we construct a set $\tilde{S}=\{s_1, s_2, \cdots, s_n, z\}, z \subset U,\ w(z) > K$ and z could not be chosen together with each s_i. Thus, we get an instance of our problem, which is denoted by B.

Then, we can simply see that q is a solution of A if and only if q is a solution of B. Moreover, the reduction from A to B ends in polynomial time. □

Remarks. The previous theorem shows that WBDP is NP hard even though each smartphone honestly declares their costs. This essentially rules out the possibility of exploiting the traditionally VCG mechanism [24] for our problem. The VCG mechanism requires that the optimal set of winning bids must be selected, which is impossible since there are no computationally efficient algorithms for solving NP hard problems. Moreover, an approximation algorithm with the VCG mechanism could not guarantee truthfulness.

Consequently, to compose a truthful and computationally efficient mechanism, we have to propose a non-VCG mechanism. We next present the design of an approximate algorithm for solving WBDP.

In fact, many existing approximation algorithms have been developed for solving the MWSC. They can be categorized into two classes. The first class is exact approaches. However, most exact algorithms require substantial computation time [25], failing to efficiently solve the location-sensitive weighted task assignment problem especially when the problem scale is large. For example, Chudak et al. [26] regard MWSC as an integer programming (IP) problem, and adopt the primal-dual algorithm to transfer the primal IP problem to the linear programming (LP) problem, and obtain the final integer solution through handling the fractional solution. Furthermore, the integrality gap acts as the approximation ratio. As another example, Li et al. [27] achieve a comparable constant competitive ratio according to the primal-dual algorithm.

The other class is heuristic approaches [25], generally used for finding a near-optimal solution with reasonable time complexity. The natural heuristic method is the greedy method because of its simplicity and convenient implementation. Similar to our work, the existing work [28] also adopts the greedy strategy to solve the known MWSC. However, we apply the greedy algorithm into solving the practical location-sensitive weighted task assignment problem, and further achieve a constant competitive ratio. Some other works [29,30] have tried to improve the greedy method by introducing the perturbed weight. The theoretical proof demonstrates that the randomized greedy method achieves a better theoretical competitive ratio than the pure greedy approach. Nevertheless, the randomness in the modified greedy method may fail in the realistic scenario, leading to the poor performance. Therefore, we adopt a natural greedy method to solve WBDP. The details of the approximation method will be shown in next section.

Bid Set

ID	Bid
1	([1,2,3,4], 10)
2	([1,2,3], 8)
3	([4,5,6], 6)
4	([5,6], 2)

Task Set

ID	weight
1	1
2	1
3	1
4	2
5	2
6	2

Fig. 4. The details of all submitted bids and sensing tasks to illustrate a simple example with approximation algorithm for WBDP.

3.2. Approximate method to solve WBDP

To achieve the desired property of computational efficiency, we propose an approximate algorithm to solve WBDP. To make it understood easily, we first assume that each smartphone reports its bids truthfully, and then demonstrate that each smartphone would obey the rule of truthfulness in the next section.

The algorithm adopts a greedy strategy to solve the problem. The main idea is to pick the next most cost-efficient bid that makes the towards finishing all sensing tasks until the all sensing tasks are assigned. Considering that each category i has its own task weight w_i, we redefine the ranking criterion in definition 2.7. As is different from our prior work, Feng et al. [23] regard each sensing task as the same for location-aware collaborative sensing. However, for our case, the bid $\beta_i^k=(Q_i^k, b_i^k)$ with high claimed cost b_i^k but high total task weights $w(Q_i^k)=\sum_{t_i \in Q_i^k} w(t_i)$ where each task t_i is mapped to some category i and $w(t_i)=w_i$ may win. Thus, the bid with the minimal cost per weight would win in each selection. More specifically, for each bid $\beta_i^k=(Q_i^k, b_i^k)$, its progress towards the goal is the ranking criterion $r(\beta_i^k)$ and is computed as follows:

$$r(\beta_i^k)=\frac{b_i^k}{w(Q_i^k-\tilde{Q})}, \tag{7}$$

$$w(Q_i^k-\tilde{Q})=\sum_{t_i \in Q_s^t-\tilde{Q}} w(t_i), \tag{8}$$

where $\tilde{Q}$ denotes the task coverage of all winning bids, *i.e.,* $\tilde{Q}=\bigcup Q_t^s$ and $\tilde{Q}=\emptyset$ initially, β_t^s denotes each bid that has won before β_i^k. $w(Q_i^k-\tilde{Q})$ denotes the total weights of sensing tasks each of which belongs to Q_i^k but not $\tilde{Q}$. In each iteration, $\tilde{Q}$ and $r(\beta_i^k)$ is updated and all bids that could not be selected together with existing winning bids are deleted. The pseudo-code is shown in Algorithm 1.

Illustrating example. Fig. 4 gives the settings of a simple example. There are *6* sensing tasks, each weight of which is demonstrated. Meanwhile, *4* bids are submitted in total on the platform. For smartphone *1*, it bids for task *1, 2, 3, 4* and claims a cost of *10*. According to Algorithm 1, smartphone *4* wins as the first winning bid because it has the lowest rank $r_4=\frac{2}{2+2}=0.5$, however, the ranks of other smartphones are $r_1=2, r_2=\frac{8}{3}, r_3=1$ respectively. Similarly, we select the next winning bid, smartphone 1 with rank $r_1=2$. Then the approximation algorithm covers all sensing tasks and stops after 2 iterations.

Algorithm 1: Approximate Algorithm for WBDP.

Input: set $T = \{t_1, t_2, \cdots, t_m\}$ of sensing tasks, set $B = \bigcup_{i \in N} B_i$ of all submitted bids, maximum number r_i of winning bids for user i, weight $w(t_i)$ of each sensing task.
Output: set S of winning bids, social cost ω, set W of total weight of cost-efficient subset of tasks corresponding to winning bids.

// Initialization
$S \leftarrow \emptyset$, $\omega \leftarrow 0$, $\tilde{Q} \leftarrow \emptyset$;
while $\tilde{Q} \neq T$ **do**
  for all β_i^k in B **do**
    if $Q_i^k \subseteq \tilde{Q}$ **then**
      Remove β_i^k from the set B;
    else
      $r(\beta_i^k) = \frac{c_i^k}{w(Q_i^k - \tilde{Q})}$;
    end if
  end for
  Sort $r(\beta_i^k)$ for all $\beta_i^k \in B$ in the nondecreasing order and the list is denoted by $\mathcal{R}$;
  //Add a bid into the set of winning bids
  β_s^t denotes the head of $\mathcal{R}$;
  $w(\beta_s^t) = \sum_{t_i \in Q_s^t - \tilde{Q}} w(t_i)$ denotes the total weight corresponding to cost-efficient subset of tasks of $Q_s^t - \tilde{Q}$ for winning bid β_s^t;
  $S \leftarrow S \bigcup \beta_s^t$, $\omega \leftarrow \omega + c_s^t$, $\tilde{Q} \leftarrow \tilde{Q} \bigcup Q_s^t$, $W \leftarrow W \bigcup w(\beta_s^t)$;
  Remove β_s^t element from B;
  //Delete bids that conflict with β_s^t
  for all β_i^k in B **do**
    if β_i^k conflicts with β_s^t **then**
      Delete β_i^k;
    end if
  end for
end while
return S, ω, W;

3.3. Critical payment scheme

The payment to each smartphone should be such determined that it is guaranteed that each smartphone honestly reports its real cost. The rule of *critical payment* introduced in [24] is used to determine the payment to each smartphone.

Each smartphone i is paid an amount of monetary reward for each winning bid β_i^k. The amount is determined according to a *critical bid* $c(\beta_i^k)$, which is determined as follows: if β_i^k satisfies $r(\beta_i^k) \le r(c(\beta_i^k))$ bid β_i^k wins, while the bid loses if $r(\beta_i^k) > r(c(\beta_i^k))$. The critical bid of each bid β_i^k is the first bid that makes β_i^k fail. The bid β_i^k fails when no progress it could make towards completing all sensing tasks, *i.e.*, $Q_i^k \subseteq \tilde{Q}$. The payment to β_i^k would be related to the claimed cost of its critical bid $c(\beta_i^k)$ and the payment is called critical payment.

The critical bid of a bid β_x^y is the first bid β_i^k which makes β_x^y useless any longer, *i.e.*, all existing winning bids could do all that β_x^y could do. The basic idea of finding the critical bid is deleting β_x^y and greedily selecting other bids as shown in Algorithm 1 until β_x^y is useless ($Q_x^y - \tilde{Q} = \emptyset$). We assume that β_i^k is the critical bid and it wins in the q-th iteration. Then, the winning bid denoted by β_x^y is paid an amount of money which is proportional to the ranking criterion of β_i^k in that iteration. If we denote the ranking criterion of β_i^k in the q-th iteration as $r^q(\beta_i^k)$, the critical payment is

$$p_c(\beta_x^y) = w(\beta_x^y) \cdot r^q(\beta_i^k), \tag{9}$$

where $w(\beta_x^y)$ denotes the total weight corresponding to its cost-efficient subset of tasks $Q_x^y - \tilde{Q}$ when bid β_x^y wins in Algorithm 1. The algorithm is shown in 2.

Algorithm 2: Critical payment scheme for TMDP.

Require: bid β_x^y, total weight $w(\beta_x^y)$ of cost-efficient subset of tasks corresponding to bid β_x^y, other submitted bid $B_{-(x,y)} = \{\beta_s^t | s \neq x \bigwedge t \neq y\}$.
Ensure: critical bid $c(\beta_x^y)$, critical payment $p_c(\beta_x^y)$.

$p_c(\beta_x^y) \leftarrow 0$, $\tilde{Q} \leftarrow \emptyset$;
while $\tilde{Q} \neq T$ **do**
  for all β_i^k in $B_{-(x,y)}$ **do**
    if $Q_i^k \subseteq \tilde{Q}$ **then**
      Remove β_i^k from the set $B_{-(x,y)}$;
    else
      $r(\beta_i^k) = \frac{c_i^k}{w(Q_i^k - \tilde{Q})}$;
    end if
  end for
  Sort $r(\beta_i^k)$ for all $\beta_i^k \in B_{-(x,y)}$ in the nondecreasing order and the list is denoted by $\mathcal{R}$;
  β_s^t denotes the head of $\mathcal{R}$;
  if β_s^t conflicts with existing winning bids **then**
    CONTINUE;
  end if
  if $Q_x^y \subseteq \tilde{Q} \bigcup Q_s^t$ **then**
    $c(\beta_x^y) \leftarrow \beta_s^t$, $p_c(\beta_x^y) \leftarrow w(\beta_x^y) \cdot r(\beta_s^t)$.
    RETURN $c(\beta_x^y)$, $p_c(\beta_x^y)$;
  end if
  $\tilde{Q} \leftarrow \tilde{Q} \bigcup Q_s^t$;
  Remove β_s^t from $B_{-(x,y)}$;
end while

Thus, the payoff of each bid is derived as follows.

$$u(\beta_x^y) = \begin{cases} w(\beta_x^y) \cdot r^q(\beta_i^k) - c_x^y & \beta_x^y \in S \\ 0 & \beta_x^y \notin S \end{cases}, \tag{10}$$

where $\beta_x^y \notin S, u(\beta_x^y) = 0$ means that losing bids are associated with no payment. The payoff of smartphone x is the sum of all its winning bids, $u_x = \sum_{\beta_x^y \in S} u(\beta_x^y)$.

Illustrating example. we have obtained the solution for winning bids determination problem from Fig. 4, *i.e.*, smartphone *4*, *1* win. Furthermore, we illustrate the payment scheme for all the winning bids. For the payment of smartphone *4*, we excludes it from the bidding list. Then we run Algorithm 1 to pick the current winning bid, smartphone *3* with rank $r_3 = 1$. Fortunately, the subset of tasks according to its bid covers all the sensing tasks smartphone *4* bids for. Then the payment of smartphone *4* is $p_c(4) = 4 * 1 = 4$. Similarly, for another winning bid, the payment for smartphone 1 is $p_c(1) = 5 * 3 = 15$.

4. Theoretical analysis

In the section we present theoretical analysis, demonstrating that our mechanism achieves the desired properties of truthfulness, individual rationality and computational efficiency.

4.1. Individual rationality and truthfulness

To demonstrate that our mechanism is truthful, we should reveal that each smartphone will honestly disclose its real costs when the strategies of other smartphones are fixed. According to [24], our proposed mechanism is truthful if and only if the following two conditions hold: (1) the winning bids determination algorithm for WBDP is *monotonic*, and (2) each winning bid is paid the *critical value*.

Before showing our mechanism satisfies the two conditions, we first define the two conditions of monotonicity and critical value.

Definition 8 (Monotonicity)**.** For each bid β_i^k, if $\beta_i^k = (Q_i^k, c_i^k)$ wins, then bid $\tilde{\beta}_i^k$ also wins, $\tilde{\beta}_i^k = (Q_i^k, c_i^k - \delta)$ and $\delta > 0$.

Definition 9 (Critical value)**.** For each bid β_i^k, there is a critical value γ_i^k. If bid β_i^k declares a cost that is lower than or equal to γ_i^k, it must win; otherwise; it will not win.

Next, we prove our mechanism is truthful by showing that it satisfies both the two conditions.

Lemma 1. *Algorithm 1 is monotonic.*

Proof. Suppose bid β_i^k wins in the q-th iteration. In the previous iterations, a number of winning bids have been determined. Let a sorted list $L = (\xi^1, \xi^2, \cdots)$ store these winning bids in the order that they have been determined. Suppose β_i^k is in the q-th place in the list. Assume bid β_i^k was replaced by another bid $\tilde{\beta}_i^k = (Q_i^k, \tilde{b}_i^k)$, where $\tilde{b}_i^k = b_i^k - \delta$. According to the rule of determining a winning bid in Algorithm 1, we have the conclusion as follows:

$$r(\tilde{\beta}_i^k) = \frac{\tilde{b}_i^k}{w(Q_i^k - \tilde{Q})} \le r(\beta_i^k), \tag{11}$$

where two ranking criterions have the same denominator, the total weight of cost-efficient subset of sensing task $Q_i^k - \tilde{Q}$. Because we pick the cost-efficient bid in each iteration, bid $\tilde{\beta}_i^k$ must have won in the q-th or an even earlier iteration. This proves that Algorithm 1 is monotonic. □

Lemma 2. *Each winning bid is paid the critical value.*

Proof. Assume that a winning bid β_x^y is $(Q_x^y, p_c(\beta_x^y))$ with the total weight $w(\beta_x^y)$ corresponding to the subset of its cost-efficient tasks, and its critical bid β_s^t wins in the q-th iteration, according to the payment rule from Algorithm 2, we have $p_c(\beta_x^y) = w(\beta_x^y) \cdot r^q(\beta_s^t)$. It is obvious that a bid $\tilde{\beta}_x^y = (Q_x^y, p_c(\beta_x^y) + \delta), \delta > 0$ whose claimed cost is higher than β_x^y would lose, because

$$r^q(\tilde{\beta}_x^y) = \frac{p_c(\beta_x^y) + \delta}{w(\beta_x^y)} > r^q(\beta_s^t) \tag{12}$$

which means that its critical bid β_s^t will be picked according to Algorithm 1. Furthermore, bid $\tilde{\beta}_x^y$ could not win in the following iterations because $Q_x^y \subseteq \tilde{Q}^q$. On the contrary, a bid $\tilde{\beta}_x^y = (Q_x^y, p_c(\beta_x^y) - \delta), \delta > 0$ whose claimed cost is lower than β_x^y would still win, because in the q-th iteration $r^q(\tilde{\beta}_x^y) < r^q(\beta_s^t)$. The reason is that our algorithm still selects the bid $\tilde{\beta}_x^y$ with the lower ranking criterion and declines its critical bid β_s^t. This demonstrates that $p_c(\beta_x^y)$ is the critical value. □

Theorem 2. *The proposed auction mechanism is truthful.*

According to [24], this theorem easily follows from Lemma 1 and Lemma 2.

Theorem 3. *The proposed auction mechanism has the property of individual rationality.*

Proof. For a smartphone x that has no winning bids, its payoff is zero, *i.e.*, $u_x = 0$. For a smartphone x with winning bids, its payoff is $u_x = \sum_{\beta_x^y \in S} u(\beta_x^y)$. Next, we show that each $u(\beta_x^y)$ is nonnegative, *e.g.*, $u(\beta_x^y) \ge 0$. In consideration of $u(\beta_x^y) = p_c(\beta_x^y) - c_x^y$, meanwhile, we have $p_c(\beta_x^y) = w(\beta_x^y) \cdot r^q(\beta_s^t)$ and $b_x^y = w(\beta_x^y) \cdot r^q(\beta_x^y)$ according to the payment rule and the allocation result from Algorithm 2 and Algorithm 1 respectively. Because $r^q(\beta_x^y) \le r^q(\beta_s^t)$, furthermore, we prove that our proposed mechanism is truthful, we have $c_x^y = b_x^y$. Thus, for each winning bid β_x^y, we have its critical value must be larger than its claimed cost, *i.e.*, $p_c(\beta_x^y) \ge c_x^y$. Then for each winning bid β_x^y that truthfully reports its real cost, its payoff is $u_x = p_c(\beta_x^y) - c_x^y \ge 0$. □

4.2. Computational efficiency

We next analyze the computation complexity of the two algorithms proposed by our mechanism.

Lemma 3. *Algorithm 1 for winning bids determination has polynomial-time computation complexity.*

Proof. Since the computation complexity of the total weight according to line 12 is $O(1)$, it has no impact on the total complexity. We neglect it in the next proof. The complexity of the first for loop (line 3–9) and the second for loop (line 15–19) in Algorithm 1 is $O(|B|)$. The operation of sorting in line 10 is $O(|B| \cdot \log|B|)$. Thus, the aggregate complexity of a single iteration of the outer while is $O(|B| \cdot \log|B|)$. Since the outer loop is run at most $|T|$ times, it is easy to compute the total computation complexity of Algorithm 1 which is $O(|T| \cdot |B| \cdot \log|B|)$. □

Lemma 4. *Algorithm 2 for critical payment determination has polynomial-time computation complexity.*

Proof. The outer while loop is run at most $|B|$ times because in some steps $\tilde{Q}$ does not expand. In each iteration of the outer while loop, there are two for loops and an operation of sorting. The computation complexity of the first for loop is $O(|B|)$. The computation complexity of the sorting operation is $O(|B|log|B|)$. The computation complexity of the second for loop (removing conflicting bids in line 12) is $O(|T|)$. Thus, the total computation complexity is $O(|B| \cdot (|B|log|B| + |T|))$. □

4.3. Approximation ratio analysis

Next, we analyze the approximation ratio achieved by Algorithm 1.

Theorem 4. *Algorithm 1 can approximate the optimal solution within a factor of $H(m)$, where m is the maximum number of sensing tasks that a smartphone can accommodate,* i.e., $m = max_{\beta_i^k \in B}\{|Q_i^k|\} \le |T|$ *and* $H(m) = \sum_{i=1}^{m} \frac{1}{i} \approx \ln m$.

Proof. When Algorithm 1 chooses a bid β_i^k, imagine that each element in the set $Q_i^k - \tilde{Q}$ introduces a part of the social cost. Then, the total social cost of all winning bids selected by Algorithm 1 equals the amount of the sum of social cost introduces in all iterations.

Consider an arbitrary winning bid $\beta_i^k = (Q_i^k, c_i^k)$, $Q_i^k = \{t_n, \cdots, t_1\}$, $\beta_i^k \in I$. Suppose that the elements of Q_i^k is covered in the order of $t_n, t_{(n-1)}, \cdots$. At the beginning of iteration in which Algorithm covers t_j of Q_i^k, at least j elements of Q_i^k is uncovered, *i.e.*, in $Q_i^k - \tilde{Q}$. Furthermore, any one x of j elements can be mapped to some category denoted by k_x where $k_x \in [1, 2, \ldots, K]$, indicating that the element x belongs to the category k_x with the corresponding weight $w(x), x = 1, 2, \ldots, j$. Thus, if the algorithm chooses β_i^k in that iteration, element t_j introduces at most $\frac{c_i^k}{\sum_{x=1}^{j} w(x)}$. Different from Feng et al.'s work [23], each element t_j has its own weight $w(t_j)$, introducing the diverse social cost $\frac{c_i^k}{\sum_{x=1}^{j} w(x)}$. However, for Feng et al.'s work [23], the social cost introduced by element t_j is $\frac{c_i^k}{j}$ in corresponding to each task with weight 1. For the symbolic simplification, let the notation $w(x) = w_x$, for any $x = 1, 2, \ldots, j$. Let W denote $\sum_{x=1}^{j} w(x)$. Then, the social cost introduced by all elements in Q_i^k is computed as follows:

$$\sum_{j=1}^{n} \frac{c_i^k}{W} = \frac{c_i^k}{w_1} + \frac{c_i^k}{w_1 + w_2} + \cdots + \frac{c_i^k}{w_1 + \cdots + w_n} \tag{13}$$

$$\leq \frac{c_i^k}{1} + \frac{c_i^k}{2} + \cdots + \frac{c_i^k}{n} \tag{14}$$

$$= c_i^k \cdot \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}\right) \tag{15}$$

$$\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \approx ln(n) \tag{16}$$

$$\sum_{j=1}^{n} \frac{c_i^k}{W} \leq c_i^k ln(n) = H(n) c_i^k \tag{17}$$

In consideration of $w_x \geq 1$, the derivation from (13) to (14) holds. Summing over each $\beta_i^k \in I$, the social cost obtained by Algorithm 1 is $\tilde{\omega} \leq \sum_{\beta_i^k \in I} H(n) c_i^k \leq H(m) \sum_{\beta_i^k \in I} c_i^k$ and $m = max\{|Q_i^k|\}$. The optimal social cost ω^* is $\sum_{\beta_i^k \in I} c_i^k$, and then, $\tilde{\omega} \leq H(m) \cdot \omega^*$. □

4.4. The upper bound of sets analysis

Finally, the maximum number of winning bids by the approximation algorithm is discussed, inducing an upper bound α compared to that of the optimal solution.

Theorem 5. *The number $|S|$ of winning bids by Algorithm 1 has achieved an upper bound $\alpha = \ln(|\Delta|)$ compared to the number $|O|$ of the optimal solution, where $\Delta = \sum_{i=1}^{m} w(t_i)$ is the total weights of all sensing tasks T for any $t_i \in T$, i.e., $|S| \leq ln(|\Delta|)|O|$, when all the submitted bids have the same claimed bid price.*

Proof. Compared to Feng et al.'s work [23], we highlight the difference as the sensing task with diverse task weight. Therefore, we have to convert different tasks to the virtual tasks with the same weight 1. Consider all sensing tasks are $T = \{t_1, t_2, \cdots, t_m\}$ with the corresponding weight $w(t_i)$ for any $t_i \in T$. Suppose that task t_j has weight $w(t_j) = 1$, then $t_{j'}$ with weight $w(t_{j'}) = k$ has k elements t_j respectably. Thus, sets of the original sensing tasks can be transferred into a universal set with the number $|\Delta| = \sum_{i=1}^{m} w(t_i)$ of elements with the same weight 1. When all the submitted bids have the same claimed cost, according to Algorithm 1, we have the most cost-efficient bid is one with the most uncovered sensing tasks. Furthermore, assume that the optimal solution produces the number $|O|$ of winning bids to cover all elements. Firstly, we observe that the first set picked by Algorithm 1 covers at least $\frac{|\Delta|}{|O|}$ elements. Otherwise, the number of winning bids obtained by the optimal solution would exceed $|O|$. Then we have the remaining uncovered elements $|\Delta_1|$ as follows:

$$|\Delta_1| \leq |\Delta| - \frac{|\Delta|}{|O|} \tag{18}$$

Next, one of the remaining winning sets has to cover at least $\frac{|\Delta_1|}{|O|-1}$ elements because of the limit of the number of given optimal bids $|O|$. Thus, after Algorithm 1 chooses the most cost-efficient bid to cover the maximum uncovered elements, we still have the uncovered elements as follows:

$$|\Delta_2| \leq |\Delta_1| - \frac{|\Delta_1|}{|O| - 1} = |\Delta_1| \cdot \left(1 - \frac{1}{|O| - 1}\right) \tag{19}$$

$$\leq |\Delta_1| \cdot \left(1 - \frac{1}{|O|}\right) \tag{20}$$

$$\leq |\Delta| \cdot \left(1 - \frac{1}{|O|}\right)^2 \tag{21}$$

As a consequence, for any iteration i, the derivative result is shown as follows:

$$|\Delta_{i+1}| \leq |\Delta_i| \cdot \left(1 - \frac{1}{|O|}\right) \leq |\Delta| \cdot \left(1 - \frac{1}{|O|}\right)^{i+1} \tag{22}$$

Furthermore, suppose that after k iterations, our approximation algorithm covers all elements of the universal set. Therefore, the number $|S|$ of winning bids obtained by Algorithm 1 is equal to k. According to the derivation from (22), we have uncovered elements after k iterations as follows:

$$|\Delta_k| \leq |\Delta| \cdot \left(1 - \frac{1}{|O|}\right)^k \tag{23}$$

Finally, when the upper bound from (23) is less than 1, then Algorithm 1 achieves all elements covered. The proof is derived as follows:

$$|\Delta| \cdot \left(1 - \frac{1}{|O|}\right)^k \leq 1 \tag{24}$$

$$\left(1 - \frac{1}{|O|}\right)^k \leq \frac{1}{|\Delta|} \tag{25}$$

$$\left(1 - \frac{1}{|O|}\right)^{|O| \cdot \frac{k}{|O|}} \leq \frac{1}{|\Delta|} \tag{26}$$

$$e^{-\frac{k}{|O|}} \leq \frac{1}{|\Delta|} \tag{27}$$

$$ln(|\Delta|) \geq \frac{k}{|O|} \tag{28}$$

Considering that we have known knowledge as follows:

$$(1 - x)^{\frac{1}{x}} \approx \frac{1}{e} \tag{29}$$

the derivation from (26) to (27) holds. Thus, we have $k \leq ln(|\Delta|)|O|$. Then we prove that the number of winning bids achieved by our approximation algorithm has $\alpha = ln(|\Delta|)$ upper bound compared to the optimal number $|O|$. □

5. Performance evaluation

5.1. Methodology and simulation settings

We evaluate the performance of the proposed mechanisms with extensive simulations based on a real data set of location traces. The real location traces were collected from around 2600 taxis in Shanghai, as used in prior studies [31,32]. For each taxi, its GPS coordinate (longitude and latitude) and the corresponding ID were recorded every 30 to 60 s. The taxis operate in Shanghai, the largest city in China which covers an area of 6340 km^2. In a simulation, we take the locations of a subset of the taxis at a certain time snapshot. For different simulations, we take different snapshots. We assume that a smartphone is carried by the passenger or the driver of a selected taxi.

To evaluate the performance of our mechanism, we use the following metrics: *social cost, overpayment ratio, individual rationality, running time, approximation ratio and α-sets approximation ratio.* The overpayment is the difference between the total payment to

Table 2
Summary of default settings.

Parameter name	Default value
Number of smartphones n	500
Number of sensing tasks m	40
Cost range R	[0,50]
Average cost μ	25
Maximum number r_i of bids a user could receive	3

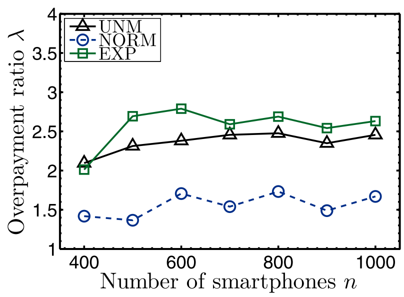

Fig. 5. Overpayment ratio vs. Number of smartphones n.

all contributing smartphones and the social cost. The overpayment measures the cost paid by the platform (or smartphone sensing applications) to induce truthfulness of all smartphones. We define the overpayment ratio to measure the overpayment.

Definition 10 (Overpayment ratio)**.** The overpayment ratio is the ratio of overpayment to the social cost. It is computed as

$$\lambda = \frac{P - \omega}{\omega}, \tag{30}$$

where P denotes the total payment.

In a simulation, location attributes of sensing tasks are uniformly distributed in the whole area of the Shanghai. We generate real costs of bids according to three distributions, *i.e.*, uniform distribution (UNM), normal distribution (NORM) and exponential distribution (EXP). Each experiment is conducted with each of all three distributions. In a simulation, we vary the mean μ of real costs from 15 to 25. The normal distribution sets such a standard deviation σ that 99.73% of samples falling within $[\mu - \sigma, \mu + \sigma]$, *i.e.* $\sigma = \frac{50-\mu}{3}$. Meanwhile, the weight of each sensing task is varied with [1, 5] for simplicity. For another significant parameter, the maximum number of sensing tasks that a smartphone can accommodate, in general, we stimulate it with 3. It is worth noting that we assume that more than 2 smartphones bid for the same task for reducing the uncertainty of duopoly. The default settings of other parameters are summarized in Table 2. Each data point is the average of 20 independent runs under the same setting.

5.2. Evaluation of overpayment ratio

Fig. 5 plots the overpayment ratio when the number of smartphones changes from 400 to 1,000. We can see that the overpayment ratio keeps low when the number of smartphones increases. The overpayment ratio is always lower than 3 for all three kinds of distributions. The overpayment ratio of the exponential distribution is larger than those of the other two distributions. This is because, with the exponential distribution, if the real cost of its critical bid locates in the tail of the exponential distribution, the overpayment is relatively large. The consequence occurs occasionally when the number of bids corresponding to some sensing task is near 2. The overpayment ratio of normal distribution is lower than those of the uniform distribution and exponential distribution. With the normal distribution in our default settings, the real

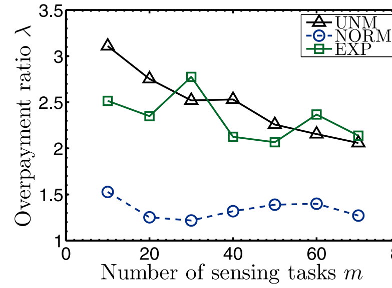

Fig. 6. Overpayment ratio vs. Number of sensing tasks m.

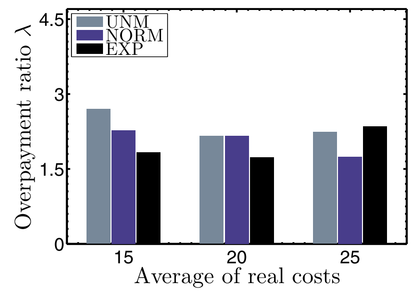

Fig. 7. Overpayment ratio vs. Average cost μ.

cost of the critical bid of a smartphone is closer to μ. Meanwhile, the smartphone has a real cost near to μ. However, the real costs of a smartphone vary within cost range R for the uniform distribution. Thus, The real cost of the critical bid for the smartphone is possibly much larger than his claimed cost, inducing the high overpayment ratio.

Fig. 6 shows that with the increasing number of sensing tasks, the overpayment ratio decreases and smoothes gradually close to each stationary point for all three distributions. The overpayment ratio of the normal distribution still holds down, however, that of the exponential distribution fluctuates around the curve of the overpayment ratio of the uniform distribution, indicating those of two distributions show the similar property. Meanwhile, we observe that the overpayment ratio under three distributions remains small as the number of sensing tasks becomes larger.

Then, in Fig. 7, the overpayment ratio is evaluated when the average of real costs increases. The mean of real costs varies from 15 to 25. Under the normal distribution, the overpayment ratio decreases slightly with the average of real costs increases. This is because, lower average of real costs produces higher standard deviation σ according to $\sigma = \frac{50-\mu}{3}$. Furthermore, larger σ produces real costs of bids with large deviation, thus inducing high overpayment ratio. However, the overpayment of the exponential distribution increases considerably because the claimed cost of winning bids is possibly much smaller than that of the critical bid with average $\mu = 25$. Finally, from the figure, we can see that the overpayment ratio does not exceed 3 under three distributions.

From Fig. 8, we can observe that the overpayment ratio decreases with multiple changes of generative real costs diminish. In particularly, the overpayment ratio under cost range R with lower bound 0 is higher because our approximation algorithm tends to pick the most cost-efficient bid with the real cost near to 0. Furthermore, the much higher real cost of critical bid leads to higher overpayment ratio. Meanwhile, we can see for the cost range R with lower bound 10, the overpayment ratio under the three distributions has a minor distinction because the generated real costs of smartphones have little multiple difference.

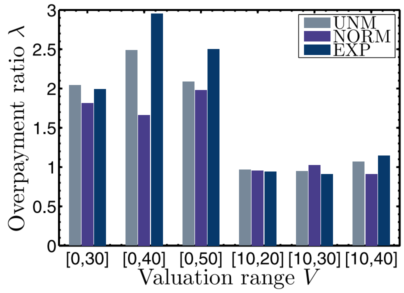

Fig. 8. Overpayment ratio vs. Cost range R.

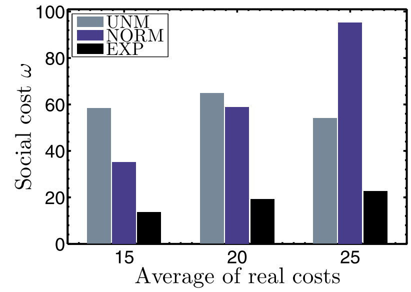

Fig. 11. Social cost vs. Average cost μ.

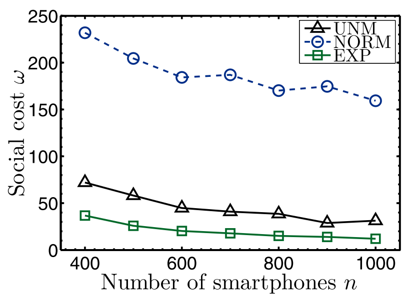

Fig. 9. Social cost vs. Number of smartphones n.

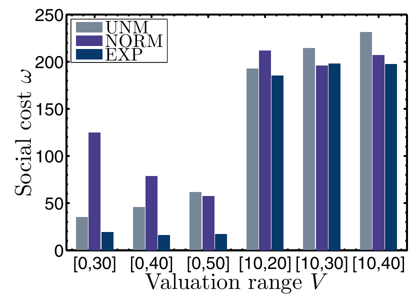

Fig. 12. Social cost vs. Cost range R.

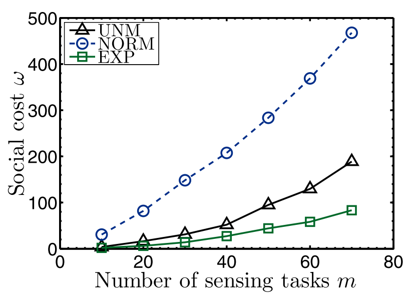

Fig. 10. Social cost vs. Number of sensing tasks m.

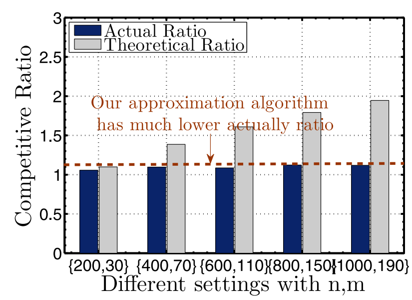

Fig. 13. approximation ratio vs. number of smartphones and sensing tasks.

5.3. Evaluation of social cost

Fig. 9 depicts the performance of social cost with the number of smartphones being varied from 400 to 1,000. The social cost decreases when the number of smartphones increases. This is because when there are more smartphones, the platform can find more cheap smartphones to perform the sensing tasks. The social cost of the exponential distribution is lower than those of the other two distributions since there exists a larger percent of smartphones of low real costs. The social cost of the normal distribution is much higher than those of other two distributions because the real cost of most smartphones falls with $[\mu-\sigma, \mu+\sigma]$, thus there are fewer smartphones with low real costs. Meanwhile, for the exponential distribution and uniform distribution, the smooth of two curves since the number of smartphones $n > 800$ is caused by a large percent of smartphones with low real costs produced as winning bids.

Fig. 10 plots the social cost when the total number of sensing tasks varies from 10 to 50. With more sensing tasks, the platform must employ more smartphones. Thus, more resources are consumed, incurring a higher social cost. The social cost of the exponential distribution is smaller than those of the other two distributions again because the exponential distribution produces more smartphones with low costs. When the number of sensing tasks m is larger, the social cost of the normal distribution is much larger than those of other two distributions. The reason lies in a large percent of bids with high costs produced by the normal distribution. Meanwhile, we observe that the social cost of the uniform distribution is slightly larger than that of the exponential distribution because of the real costs produced by the uniform distribution possibly with slightly less low costs.

In Fig. 11, the social cost rises with the average real cost increases. When the average of real costs goes up, the mobile sensing applications are less likely to find cheap smartphones, thus incurring a higher social cost. The increase of the exponential distribution is rather smaller since there are more cheap smartphones than that of the other two distributions.

From Fig. 12, we can observe that the social cost under cost range R with lower bound 0 is lower than that under R with lower bound 10 because our approximation algorithm tends to pick the most cost-efficient bid with the real cost near to 0, inducing much lower social cost.

5.4. Evaluation of approximation ratio

We adopt the real costs subjecting to the uniform distribution to conduct a series of experiments related to approximation ratio. Fig. 13 plots approximation ratios under different settings. The bar of the theoretical approximation ratio $ln(m)$ rises with $m = max_{\beta_i^k \in B}\{|Q_i^k|\}$, the maximum number of each sensing task that a smartphone can accommodate increases. For all five settings, m is increased one by one. From Fig. 13, we observe that each actual

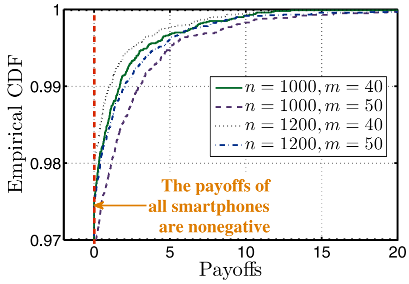

Fig. 14. Empirical CDF of payoffs for all smartphones.

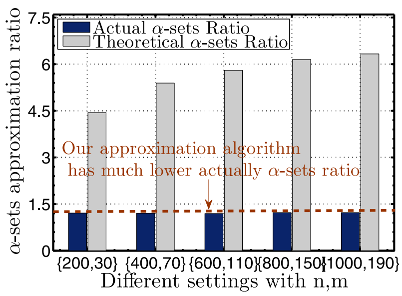

Fig. 15. α-sets approximation ratio vs. number of smartphones and sensing tasks.

approximation ratio is lower than the theoretical value, indicating the correctness of our theoretical approximation ratio. We can see that the approximation ratio is slightly larger than 1, meaning the validity of our approximation algorithm.

With the real location traces, the actual maximum number of sensing tasks that a smartphone can accommodate is usually small and thus the upper bound of the approximation ratio is close to 1.

5.5. Evaluation of individual rationality

The basic property of an incentive mechanism is to guarantee that each smartphone is individually rational. In Fig. 14, we plot the empirical CDF of the payoffs for all smartphones. We can observe that the payoff of over 97% smartphones is equal to zero, indicating that only a few are selected as the winning bids. However, there is a vast majority of smartphones competing for the same task so as to avoid the monopoly, which is consistent with our case. Finally, from the figure, we can see that no smartphone has a negative payoff, thus demonstrating that our mechanism achieves the property of individual rationality.

5.6. Evaluation of α-sets approximation ratio

To prove the correctness of our analysis for the upper bound of sets α, we utilize Fig. 15 to demonstrate α-sets Approximation Ratio computed as $\alpha = ln(|\Delta|)$, meaning that the upper bound of sets of our approximation algorithm divided by that of the optimal solution. According to Fig. 15, we observe that each actual α-sets approximation ratio is lower than the theoretical value, indicating the correctness of our theoretical α-sets approximation ratio. Thus, our simulation shows that our approximation algorithm only picks $\alpha \cdot |O|$ winning bids at most, where $|O|$ is the sets of the optimal solution.

5.7. Evaluation of computation efficiency

We compare the running time of our mechanism with the optimal VCG mechanism (denoted by *OPT*). Since the problem is *NP hard*, we can only obtain the optimal solution when the problem

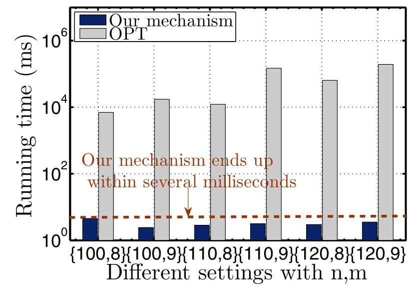

Fig. 16. Evaluation of computation efficiency (y-axis in log scale).

scale is small (*e.g.*, $n \le 120$, $m \le 9$). The *OPT* employs a backtracking approach to find the optimal solution and a VCG-style payment scheme to guarantee truthfulness. In Fig. 16, we show the running time in various settings. We can clearly see that our algorithm uses significantly shorter time than the optimal algorithm. For example, in the 4th set of bars, our mechanism uses only 3 milliseconds, but *OPT* uses more than 2 min..

6. Related work

We review related work from two parts, deterministic allocation and mechanism design. The deterministic allocation only emphasizes on the task allocation and neglects the payment scheme. However, mechanism design contains both sides. In particular, we make a detailed introduction from three aspects, including incentive schemes with strategic users, incentive schemes with cooperative users, and privacy preservation. Other works like [33,34] exploit the auction to employ the cloud application.

6.1. Deterministic allocation

There have been many existing work [35–39] in the deterministic allocation. However, they only consider the task allocation, neglecting the payment determination problem or only making a simple payment scheme. The consideration is unrealistic and fails to motivate users to work better for the platform. He et al. [37] offer a near-optimal allocation approach in consideration of user's time budget and sensing task with a specific location, aiming to maximize the reward of the platform. Although the crucial dimension of location information is taken into consideration, the approximation algorithm *LRBA* [37] proposed by them cannot be applied to our case.

6.2. Mechanism design

Incentives mechanism with strategic behaviors: In [20], Wen et al. provide a novel incentive mechanism in consideration of quality of information (QOI) returned by smartphones, and tend to exploit quality-driven auction against the strategic behaviors of users. Zhang et al. in [17] adopt repeated game to match smartphone users with platform users, taking the maintenance of smartphone users' reputation model into consideration. In [16], two incentive mechanisms are designed for a user-centric model and a platform-centric model, respectively. In the platform-centric model, the platform first announces the total amount of money that the platform is willing to pay, and then each smartphone decides the time it would like to provide the sensing service. In the user-centric model, the objective is system-wide. If a sensing task is finished, it would receive some benefit. The goal is to select a subset of smartphones, maximizing the overall gain.

However, they only consider the single preference of a user, which means a user only submits a single bid and it would win for performing all sensing tasks in the bid or not at all. Such an assumption is not realistic for mobile crowd sensing with location-sensitive weighted tasks, failing to make the best use of the sensing capabilities of all smartphones. Similarly, they assume that only a single bid is submitted by each user in other related work [27,40]. Meanwhile, in [27], Li et al. provide a randomized auction but only achieve the approximate-truthfulness.

In [22], the authors design a truthful incentive mechanism for three models of mobile crowd sensing applications. They assume that only the collective efforts from winning service providers promote the success of service requester's tasks. Each service provider submits the interested tasks and claimed cost. The platform as auctioneer decides the allocation algorithm and relevant payment scheme. The allocation algorithm fails to take the location into consideration and could not be applied to our problem because of the NP-hardness of winning bids determination problem.

Incentives schemes with cooperative users: Some related work proposing to provide monetary rewards to generally cooperative users, such as [41–44]. Users are supposed to be well motivated because they receive some monetary rewards. In [44], the author proposes a cooperative scheme to induce the partners to negotiate the payment. Meanwhile, the source determines the best rewards to buy the service the partners provide. In [41], a subset of users is greedily selected according to their locations and the total budget. The algorithm aims to cover the largest area. In [42], the users sell sensed data to a service provider and a dynamic pricing scheme is designed to stimulate more participants, thus achieving a better quality of service. In [43], it considers that a selfish user has the demand of consuming data, and how much service it could consume depends on how much she or he contributes to the participatory sensing system. Thus, they have to consider how to satisfy all users fairly and how to achieve a desirable result (maximized social welfare) for the whole system. These existing studies do not consider that cost information is private, and that users may misreport their real costs in order to maximize their own payoffs. As a result, these incentive schemes are not truthful.

Privacy preserving schemes: Privacy preserving is crucial to mobile crowd sensing with smartphones. It has attracted many research efforts, such as [19,45,46]. One smartphone is supposed to report its sensed data, but it is reluctant to disclose its private information, such as location, and identity. In [47] the author adopts an efficient encryption technique to achieve the desired data aggregation without the leakage of mobile user's privacy. Some other studies [48,49] add noise or perturbation to original sensory data for the purpose of anonymity. These papers usually utilize k-anonymity or entropy to measure the leakage of privacy. Our work, which has focused on designing incentive mechanisms for mobile crowd sensing with location-sensitive weighted tasks, can benefit from these existing schemes for protecting location privacy.

7. Conclusion

In this paper, we have investigated a truthful incentive mechanism for stimulating smartphone users to participate in mobile crowd sensing applications with smartphones. In addition to sensing tasks with diverse weights, the *crucial dimension of location information* is uniquely taken into consideration in our design of incentive mechanism, which is more consistent with the real condition. Based on the reverse auction framework, we have designed a truthful incentive mechanism which consists of winning bids determination algorithm and a critical payment scheme. With polynomial-time computation complexity, the near-optimal algorithm for determining the cost-efficient winning bids can approximate the optimal solution within a factor of $1 + \ln(n)$, where n is the maximum number of sensing tasks that a smartphone can accommodate. Meanwhile, despite the suboptimal approximation algorithm, the proposed critical payment scheme can guarantee truthfulness. As an additional part, after the rigorous analysis, we prove that the number of winning bids from the approximation solution has an upper bound α compared to that of the optimal solution when all smartphone users have the same claimed bid price. After rigorous theoretical proof and extensive simulations, the results demonstrate that our mechanism achieves truthfulness, individual rationality and high computational efficiency.

8. Acknowledgements

This work is supported in part by National Natural Science Foundation of China (No.61772341, 61472254). This work is also supported by the Program for Changjiang Young Scholars in University of China, the Program for Shanghai Top Young Talents, and the National Top Young Talents Support Program.

References

[1] B. Guo, Z. Yu, From participatory sensing to mobile crowd sensing, in: IEEE International Conference on Pervasive Computing and Communication Workshops, Budapest, Hungary, March 24–28, IEEE, 2014, pp. 593–598.
[2] B. Guo, Mobile crowd sensing and computing: the review of an emerging human-powered sensing paradigm, ACM Comput. Surv. (CSUR) 48 (2015) 7.
[3] H. Ma, Opportunities in mobile crowd sensing, IEEE Commun. Mag. 52 (2014) 29–35.
[4] R. Pryss, M. Reichert, Mobile crowd sensing in clinical and psychological trials - a case study, in: 28th IEEE International Symposium on Computer-Based Medical Systems, Sao Carlos, Brazil, June 22–25, IEEE, 2015, pp. 23–24.
[5] Sensorly, 2010, (http://www.sensorly.com/).
[6] R.K. Rana, C.T. Chou, Ear-phone: an end-to-end participatory urban noise mapping system, in: Proceedings of the 9th International Conference on Information Processing in Sensor Networks, Stockholm, Sweden, April 12–16, ACM, 2010, pp. 105–116.
[7] A. Thiagarajan, L. Ravindranath, Vtrack: accurate, energy-aware road traffic delay estimation using mobile phones, in: Proceedings of the 7th International Conference on Embedded Networked Sensor Systems, Berkeley, California, USA, November 4–6, ACM, 2009, pp. 85–98.
[8] T. Ludwig, C. Reuter, Crowdmonitor: mobile crowd sensing for assessing physical and digital activities of citizens during emergencies, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, April 18–23, ACM, 2015, pp. 4083–4092.
[9] H. Jin, L. Su, 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, Hangzhou, China, June 22–25, ACM, 2015, pp. 167–176.
[10] R.K. Ganti, Mobile crowdsensing: current state and future challenges, IEEE Commun. Mag. 49 (2011) 32–39.
[11] L.G. Jaimes, A survey of incentive techniques for mobile crowd sensing, IEEE Internet Things J. 2 (2015) 370–380.
[12] L. Pournajaf, L. Xiong, Spatial task assignment for crowd sensing with cloaked locations, in: IEEE 15th International Conference on Mobile Data Management, Brisbane, Australia, July 14–18, IEEE, 2014, pp. 73–82.
[13] A. Rai, K.K. Chintalapudi, Zee: zero-effort crowdsourcing for indoor localization, in: The 18th Annual International Conference on Mobile Computing and Networking, Istanbul, Turkey, August 22–26, IEEE, 2012, pp. 293–304.
[14] X. Hu, Vita: a crowdsensing-oriented mobile cyber-physical system, IEEE Trans. Emerg. Top. Comput. 1 (2013) 148–165.
[15] N.D. Lane, Y. Chon, Piggyback crowdsensing (PCS): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities, in: The 11th ACM Conference on Embedded Network Sensor Systems, Roma, Italy, November 11–15, ACM, 2013, pp. 1–14.
[16] D. Yang, G. Xue, Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing, in: The 18th Annual International Conference on Mobile Computing and Networking, Istanbul, Turkey, August 22–26, IEEE, 2012, pp. 173–184.
[17] Y. Zhang, M. van der Schaar, Reputation-based incentive protocols in crowdsourcing applications, in: Proceedings of the IEEE INFOCOM, Orlando, FL, USA, March 25–30, IEEE, 2012, pp. 2140–2148.
[18] Z. Zheng, Trading data in the crowd: profit-driven data acquisition for mobile crowdsensing, IEEE J. Sel. Areas Commun. 35 (2017) 486–501.
[19] A.C. Weaver, J.P. Boyle, Applications and trust issues when crowdsourcing a crisis, in: The 21st International Conference on Computer Communications and Networks, Munich, Germany, July 30-August 2, IEEE, 2012, pp. 1–5.
[20] Y. Wen, Quality-driven auction-based incentive mechanism for mobile crowd sensing, IEEE Trans. Veh. Technol. 64 (2015) 4203–4214.
[21] D. Zhao, X. Li, How to crowdsource tasks truthfully without sacrificing utility: online incentive mechanisms with budget constraint, in: IEEE Conference on Computer Communications, Toronto, Canada, April 27, - May 2, IEEE, 2014, pp. 1213–1221.

[22] X. Zhang, G. Xue, Truthful incentive mechanisms for crowdsourcing, in: Conference on Computer Communications, Kowloon, Hong Kong, April 26, - May 1, IEEE, 2015, pp. 2830–2838.
[23] Z. Feng, Y. Zhu, TRAC: truthful auction for location-aware collaborative sensing in mobile crowdsourcing, in: IEEE Conference on Computer Communications, Toronto, Canada, April 27, - May 2, IEEE, 2014, pp. 1231–1239.
[24] N. Nisan, Algorithmic Game Theory, Cambridge University Press, 2007.
[25] G. Lan, An effective and simple heuristic for the set covering problem, Eur. J. Oper. Res. 176 (2007) 1387–1403.
[26] F.A. Chudak, A primal–dual interpretation of two 2-approximation algorithms for the feedback vertex set problem in undirected graphs, Oper. Res. Lett. 22 (1998) 111–118.
[27] J. Li, Crowdsourcing sensing to smartphones: a randomized auction approach, IEEE Trans. Mob. Comput. 1 (2017) 1–15.
[28] F. Akhter, A heuristic approach for minimum set cover problem, Int. J. Adv. Res. Artif. Intell. 4 (2015) 40–45.
[29] M. Haouari, A probabilistic greedy search algorithm for combinatorial optimisation with application to the set covering problem, J. Oper. Res. Soc. 53 (2002) 792–799.
[30] T.A. Feo, A probabilistic heuristic for a computationally difficult set covering problem, Oper. Res. Lett. 8 (1989) 67–71.
[31] Y. Wu, Y. Zhu, Infrastructure-assisted routing in vehicular networks, in: Proceedings of the IEEE INFOCOM, Orlando, FL, USA, March 25–30, IEEE, 2012, pp. 1485–1493.
[32] Y. Wu, Y. Zhu, CCR: capacity-constrained replication for data delivery in vehicular networks, in: Proceedings of the IEEE INFOCOM, Turin, Italy, April 14–19, IEEE, 2013, pp. 2580–2588.
[33] J. Ding, Auction-based cloud service differentiation with service level objectives, Comput. Networks 94 (2016) 231–249.
[34] W. Shi, An online auction mechanism for dynamic virtual cluster provisioning in geo-distributed clouds, IEEE Trans. Parallel Distrib. Syst. 28 (2017) 677–688.
[35] Y. Liu, B. Guo, Taskme: multi-task allocation in mobile crowd sensing, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, September 12–16, ACM, 2016, pp. 403–414.
[36] L. de Alfaro, V. Polychronopoulos, Reliable aggregation of boolean crowdsourced tasks, in: Proceedings of the Third AAAI Conference on Human Computation and Crowdsourcing, San Diego, California, November 8–11, IEEE, 2015, pp. 42–51.
[37] S. He, D. Shin, Toward optimal allocation of location dependent tasks in crowdsensing, in: IEEE Conference on Computer Communications, Toronto, Canada, April 27-May 2, IEEE, 2014, pp. 745–753.
[38] I. Boutsis, V. Kalogeraki, On task assignment for real-time reliable crowdsourcing, in: IEEE 34th International Conference on Distributed Computing Systems, Madrid, Spain, June 30-July 3, IEEE, 2014, pp. 1–10.
[39] H. Xiong, icrowd: near-optimal task allocation for piggyback crowdsensing, IEEE Trans. Mob. Comput. 15 (2016) 2010–2022.
[40] J. Li, Y. Zhu, Towards redundancy-aware data utility maximization in crowdsourced sensing with smartphones, in: 44th International Conference on Parallel Processing, Beijing, China, September 1–4, IEEE, 2015, pp. 899–908.
[41] L. G.Jaimes, Vergara-Laurens, A location-based incentive mechanism for participatory sensing systems with budget constraints, in: International Conference on Pervasive Computing and Communications, Lugano, Switzerland, March 19–23, IEEE, 2012, pp. 103–108.
[42] J. Lee, B. Hoh, Sell your experiences: a market mechanism based incentive for participatory sensing, in: Eigth Annual IEEE International Conference on Pervasive Computing and Communications, Mannheim, Germany, March 29, - April 2, IEEE, 2010, pp. 60–68.
[43] A.J. Mashhadi, Fair content dissemination in participatory dtns, Ad Hoc Netw. 10 (2012) 1633–1645.
[44] N. Zhang, N. Cheng, A cooperative scheme for secure communications with partner selection and incentive mechanism, in: Sixth International Conference on Wireless Communications and Signal Processing, Hefei, China, October 23–25, IEEE, 2014, pp. 1–6.
[45] X.O. Wang, W. Cheng, Artsense: anonymous reputation and trust in participatory sensing, in: Proceedings of the IEEE INFOCOM, Turin, Italy, April 14–19, IEEE, 2013, pp. 2517–2525.
[46] D. He, User privacy and data trustworthiness in mobile crowd sensing, IEEE Wireless Commun. 22 (2015) 28–34.
[47] Q. Li, Efficient and privacy-aware data aggregation in mobile sensing, IEEE Trans. Dependable Secure Comput. 11 (2014) 115–129.
[48] F. Zhang, L. He, Data perturbation with state-dependent noise for participatory sensing, in: Proceedings of the IEEE INFOCOM, Orlando, FL, USA, March 25–30, IEEE, 2012, pp. 2246–2254.
[49] K. Vu, R. Zheng, Efficient algorithms for k-anonymous location privacy in participatory sensing, in: Proceedings of the IEEE INFOCOM, Orlando, FL, USA, March 25–30, IEEE, 2012, pp. 2399–2407.

Hui Cai is a Ph.D. student in the Department of Computer Science and Engineering at Shanghai Jiao Tong University, China. She received her B.Eng. Degree in Computer Science and Technology from Sichuan University, Chengdu, China, in 2015. Her research interests include incentive mechanism design, mobile crowd sensing and game theory.

Yanmin Zhu received the B.Eng. degree in computer science from the XiAn Jiao Tong University, China, in 2002 and the Ph.D. degree in computer science from Hong Kong University of Science and Technology, Hong Kong, China, in 2007. He is a Professor in the Department of Computer Science and Engineering at Shanghai Jiao Tong University, Shanghai, China. His research interests include crowd sensing, big data analytics and systems, and cloud computing.

Zhenni Feng received her B.Eng. degree in computer science from Huazhong University of Science and Technology, Wuhan, China, in 2012. She is now a Ph.D. student in the Department of Computer Science and Engineering at Shanghai Jiao Tong University. Her research interests include machine learning, spatiotemporal data mining and big data analysis.