BidGuard: A Framework for Privacy-Preserving Crowdsensing Incentive Mechanisms

Jian Lin, Dejun Yang, Ming Li, Jia Xu, Guoliang Xue

Abstract—**With the rapid growth of smartphones, crowdsensing emerges as a new paradigm which takes advantage of the pervasive sensor-embedded smartphones to collect data efficiently. Auction has been widely used to design mechanisms to stimulate smartphone users to participate in the crowdsensing applications and systems. Many auction-based incentive mechanisms have been proposed for crowdsensing. However, none of them has taken into consideration both the bid privacy of smartphone users and the social cost. To the best of our knowledge, we are the first to study the design of privacy-preserving incentive mechanisms that also achieve approximate social cost minimization. In this paper, we design BidGuard, a general privacy-preserving framework for incentivizing crowdsensing. This framework works with different score functions for selecting users. In particular, we propose two score functions, linear and log functions, to realize the framework. We rigorously prove that BidGuard achieves computational efficiency, individual rationality, truthfulness, differential privacy and approximate social cost minimization. In addition, the BidGuard with log score function is asymptotically optimal in terms of the social cost. Extensive simulations evaluate the performance and validate the desired properties of BidGuard.**

I. Introduction

Nowadays, the proliferation of smartphones is changing people's daily lives. With the advance of high-speed 3G/4G networks and more powerful embedded sensors (e.g., camera, accelerometer, compass, etc.), crowdsensing emerges as a new paradigm which takes advantage of the pervasive sensor-embedded smartphones to collect data efficiently.

A typical crowdsensing system consists of a cloud-based platform and a large number of smartphone users. The platform works as a sensing service buyer who posts the required sensing information and recruits a set of smartphone users to provide sensing services. Once selected by the platform, a smartphone user starts to collect the required data and sends it back to the platform. The potential effectiveness of crowdsensing, especially with geographically distributed smartphone users, enables numerous crowdsensing applications [25, 33, 38]. However, most of them assume that the smartphone users contribute to the platform voluntarily. In reality, smartphone users consume their own resources such as battery and sensing time while completing the sensing tasks. In addition, they might suffer from the potential privacy disclosure by sharing their sensed data with personal information (e.g., location tags and bid price). Therefore, smartphone users may be reluctant to participate in a crowdsensing system and application, unless they are paid some rewards to compensate their resource consumption or potential privacy leaks. Since the number of participating smartphone users has a significant impact on the performance of the crowdsensing systems, it is necessary to stimulate users to join the systems.

Auction is an efficient method to design incentive mechanisms. Many auction-based incentive mechanisms have been proposed for crowdsensing [33, 35, 36]. They are essentially reverse auctions in which the platform is the service buyer and the smartphone users are the bidders selling sensing services. In these mechanisms, the service buyer selects bidders according to their submitted task-bid pairs (elaborated in Section III-A). The objectives of these mechanisms focus on either maximizing the total value gained by the platform or minimizing the total payment to the selected users. However, none of them takes users' privacy into consideration.

In most of the proposed truthful auction-based incentive mechanisms, bidders are stimulated to bid their true costs, which are private information of smartphone users. For transparency, the platform will publish the outcome of the auction mechanism, which consists of winning bidders and their payments. Ensuring transparency in the procurement procedure is essential to efficiency, as it enhances the competitiveness of public procurement [27]. However, once the true cost of a smartphone user is reported to the platform, other bidders might infer this private information based on the published outcome. This is known as *inference attack* [14] (we give an example in Section III). Inference attack has been analyzed in many areas, e.g., multilevel secure databases [16], data mining [5], web-based applications [29] and mobile devices [22]. In this paper, we focus on designing truthful auction-based mechanisms to protect users' bid privacy.

To formalize the notion of users' bid privacy, we employ the concept of *differential privacy* [8]. Intuitively, a mechanism provides differential privacy if the change of one user's bid has limited impact on the outcome. We also leverage the *exponential mechanism* [24], a technique to design differentially private mechanisms, to preserve users' bid privacy.

In this paper, we study the problem of designing truthful mechanisms, which achieve computational efficiency, individual rationality, differential privacy, and approximate social cost minimization. We consider the scenario where there is one

This research was supported in part by NSF grants 1444059, 1461886, NSFC grant 61472193, NSF of Jiangsu Province BK20141429, and CCF-Tencent RAGR20150107. The information reported here does not reflect the position or the policy of the funding agencies.. J. Lin, D. Yang and M. Li are with Colorado School of Mines, Golden, CO 80401 USA (e-mail: {jilin, djyang, mili}@mines.edu). J. Xu is with the School of Computer Science and Technology, Institute of Computer technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: xujia@njupt.edu.cn). G. Xue is with Arizona State University, Tempe, AZ 85287 USA (e-mail: xue@asu.edu).

978-1-5090-3065-1/16/$31.00 ©2016 IEEE

buyer and multiple sellers. Smartphone users act as bidders and submit their bids to compete for the chance of being selected to perform the corresponding tasks. Besides, smartphone users do not want others to know their own bid information. We first propose BidGuard, a general differentially private truthful auction-based framework for incentivizing crowdsensing. As an important component of this framework, we design two score functions, which will determine the selection of users.

The main contributions of this paper are as follows:

- To the best of our knowledge, we are the first to propose a framework, BidGuard, for privacy-preserving crowdsensing incentive mechanisms, which achieves computational efficiency, individual rationality, truthfulness, differential privacy, and approximate social cost minimization. Specifically, we design two different score functions, linear score function and log score function, to realize BidGuard.
- With linear score function, BidGuard achieves $(\epsilon(e-1)/e, \delta)$-differential privacy and the social cost is at most $H_K\mathcal{OPT} + gK \cdot O(\ln n)$ with the probability of at least $1 - 1/n^{O(1)}$, where $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$ are two constants, e is the base of the natural logarithm, $H_K = \sum_{j=1}^{K} 1/j$, K is the size of the largest user task set, $\mathcal{OPT}$ is the optimal social cost, g is the size of the optimal user set, and n is the number of the users.
- With log score function, BidGuard achieves $(\epsilon(e-1)/e, \delta)$-differential privacy and the social cost is at most $2^t H_K\mathcal{OPT}$ with the probability of at least $1 - e^{-t}$ for any constant $t > 0$. In this case, BidGuard is proved to be asymptotically optimal.
- We evaluate the performance of BidGuard through simulations based on a real data set. Extensive numerical results show that BidGuard has desired properties.

The remainder of this paper is organized as follows. In Section II, we briefly review the related work. In Section III, we introduce the system model and the objectives. In Section IV, we present our framework in detail and prove its properties. We evaluate the performance of our framework in Section V. We conclude this paper in Section VI.

II. Related Work

In recent years, incentive mechanisms in crowdsensing have been widely studied. As one of the pioneering works on designing incentive mechanisms for crowdsensing, Yang *et al.* [34, 35] proposed two incentive mechanisms for both user-centric and platform-centric models using auction and Stackelberg game, respectively. The objectives of most of the state-of-art incentive mechanisms are either maximizing the total utility/value of the platform under a certain constraint (e.g., budget) [37] or minimizing the total payment of the platform [23]. Feng *et al.* [11] proposed a mechanism called TRAC, which takes into consideration the importance of location information when assigning sensing tasks.

Many pieces of works have explored the privacy-preserving mechanisms in mobile crowdsourcing. Most of them [12, 19] apply the spacial and temporal cloaking techniques like K-anonymity to blur users' locations in a cloaked area or cloaked time interval to preserve users' privacy. PEPSI [7] and AnonySense [28] focus on anonymous data collection, which could protect users' identities when they submit the tasks.

Some efforts have been specially made to protect users' privacy in crowdsensing [10, 20, 21, 26, 30]. Although providing good performance in privacy preservation, the mechanisms in [10, 20, 21, 26] are based on cryptography techniques and do not take into consideration the users' strategic behaviors. Sun *et al.* [30] proposed an auction-based incentive mechanism which encrypts users' bids by oblivious transfer. But it does not solve the issue of inference attack because one user still can infer others' bids from the received payment. Jin *et al.* proposed a privacy-preserving approximately truthful incentive mechanism [17], which minimizes the total payment, and a privacy-preserving framework [18] for data aggregation. However, none of the above works has a performance guarantee on social cost minimization, which is the objective in this paper.

Differential privacy was firstly introduced by Dwork *et al.* [8]. The first differentially private auction mechanism was proposed by McSherry *et al.* [24]. They also incorporate exponential mechanism and mechanism design to achieve differential privacy with different objectives. General methods to design truthful mechanisms while still preserving differential privacy have been studied in [3, 15, 32]. However, our objective is different from above works. Recently, differential privacy has been used in other applications, e.g., location-based systems [1] and spatial crowdsoucing [31]. Zhu *et al.* [39] proposed the first differentially private spectrum auction mechanism, which achieves strategy-proofness and approximate revenue maximization. Note that our objective is to minimize the social cost, which differs from that in [39].

III. Model and Problem Formulation

In this section, we present an overview of our crowdsensing system, model it as a reverse auction, describe the threat models, and give our design objective.

A. System Model

Similar to most crowdsensing systems [11, 33–36], we consider a crowdsensing system consisting of a platform and multiple smartphone users who are interested in performing sensing tasks. The platform first publicizes a set $\mathcal{T} = \{t_1, t_2, \ldots, t_m\}$ of m sensing tasks. Assume there is a set $\mathcal{U} = \{1, 2, \ldots, n\}$ of $n \geqslant 2$ smartphone users. Each user i has a task set $\Gamma_i \subseteq \mathcal{T}$, which she can perform. Each Γ_i is associated with a cost c_i, which is a private information of user i. The platform selects a subset of users $\mathcal{S} \subseteq \mathcal{U}$ to complete all the sensing tasks in $\mathcal{T}$. At last, the platform calculates the payment p_i for each selected user $i \in \mathcal{S}$. Let $\overrightarrow{p} = (p_1, p_2, \ldots, p_n)$ denote the payment profile. The utility of user $i \in \mathcal{U}$ is

$$u_i = \begin{cases} p_i - c_i, & \text{if } i \in \mathcal{S}; \\ 0, & \text{otherwise.} \end{cases}$$

In this paper, we model the interactive process between the platform and the users as a sealed-bid reverse auction, where

$User$ β_i	1	2	3	4	5
Γ_i	t_1, t_2	t_1	t_1, t_3	t_1, t_2	t_1, t_3
b_i	\$3	\$1	\$4	\$5	\$5

TABLE I
EXAMPLE SHOWING THE INFERENCE ATTACK

the platform buys sensing service and the users are bidders who sell sensing service. In order to prevent the monopoly and guarantee the quality of sensing task, we assume each task in $\mathcal{T}$ can be completed by more than one user in $\mathcal{U}$. This assumption is reasonable for crowdsensing as made in [11]. If a task in $\mathcal{T}$ can only be completed by at most one user in $\mathcal{U}$, we simply remove it from $\mathcal{T}$.

At the beginning of this auction, each user $i \in \mathcal{U}$ submits a task-bid pair $\beta_i = (\Gamma_i, b_i)$ to the platform, where b_i is user i's bid, representing the minimum price user i wants to sell her sensing service for. Note that in a truthful auction-based incentive mechanism, users are stimulated to bid their true costs, i.e., $b_i = c_i$. Without loss of generality, we assume that for each user, the bid is bounded by $[b_{min}, b_{max}]$, where b_{min} is normalized to 1 and the difference between b_{max} and b_{min} is denoted by Δ. Let $\overrightarrow{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)$ denote the task-bid profile. Given the task-bid profile $\overrightarrow{\beta}$, the platform determines the outcome of the auction, which consists of selected winning users $\mathcal{S}$ and the payment profile $\overrightarrow{p}$.

B. Threat Models

Threats to Incentive: We assume that users are selfish but rational. Hence user i could report a bid b_i differs from her true cost c_i to maximize her own utility. We also assume that user i does not misreport her task set Γ_i as in [11, 33–36][1]. Other threats to incentive (e.g., collusion among bidders) are out of the scope of this paper.

Threats to Privacy: As mentioned earlier, bidders are stimulated to bid their true costs in a truthful auction-based incentive mechanism, i.e., $b_i = c_i$. However, one bidder could infer other bidders' bid according to the outcome of the mechanism. This inference attack can be seen from the following example. Suppose there are 5 users in the system and their task-bid pairs $\beta_i = (\Gamma_i, b_i), i \in [1, 5]$ are shown in Table I. The platform publicizes a set of 3 sensing tasks $\mathcal{T} = \{t_1, t_2, t_3\}$. According to the proposed truthful mechanism in TRAC [11], the winning users $\mathcal{S} = \{2, 1, 3\}$. Suppose $user$ 5 is a bidder who want to infer other bidders' bid, and she changes her bid b_5 from \$5 to \$3 in the next auction while the other four bidders do not change their task-bid pairs. The winning users of the new auction is $\mathcal{S} = \{2, 1, 5\}$. Since the platform publishes the outcome of the mechanism for transparency, $user$ 5 could know the results and infer that $user$ 3's bid is between \$3 and \$5 by the fact that if she bids \$5 she will be replaced by $user$ 3 and if she bid \$3 she will take $user$ 3's place. We can see that, after many rounds of auction, $user$ 5 might narrow down $user$ 3's bid range, and even infer the exact value in some cases. This inference attack is practical in most crowdsensing applications, e.g., [25, 38], where tasks are publicized periodically for collecting dynamic sensing data.

C. Desired Properties

We consider the following important properties.

- **Computational Efficiency:** A mechanism is computationally efficient if it terminates in polynomial time.
- **Individual Rationality:** A mechanism is individually rational if each user will have a non-negative utility when bidding her true cost.
- **Truthfulness:** A mechanism is truthful if any user's utility is maximized when bidding her true cost.
- **Social Cost Minimization:** A mechanism achieves social cost minimization if the total cost of the users in $\mathcal{S}$ is minimized subject to certain constraints on $\mathcal{S}$.

In addition, we consider users' bid privacy preservation.

Definition 3.1: (Differential Privacy [8]). A randomized function M has ϵ-differential privacy if for any two input sets A and B with a single input difference, and for any set of outcomes $\mathcal{O} \subseteq Range(M)$,

$$Pr[M(A) \in \mathcal{O}] \leq \exp(\epsilon) \times Pr[M(B) \in \mathcal{O}].$$

In this paper, the randomized function M is corresponding to our framework, and $Range(M)$ is the outcome space of the framework. One relaxation of differential privacy is as follows.

Definition 3.2: (Approximate Differential Privacy [9]). A randomized function M gives (ϵ, δ)-differential privacy if for any two input sets A and B with a single data difference, and for any set of outcomes $\mathcal{O} \subseteq Range(M)$,

$$Pr[M(A) \in \mathcal{O}] \leq \exp(\epsilon) \times Pr[M(B) \in \mathcal{O}] + \delta.$$

The truthfulness of an auction mechanism is guaranteed by the following theorem.

Theorem 1: [39] Let $Pr_i(z)$ denote the probability that bidder i is selected when her bid is z. A mechanism with bids $\overrightarrow{b}$ and payments $\overrightarrow{p}$ is truthful in expectation if and only if, for any bidder i,

1) $Pr_i(z)$ is monotonically non-increasing in b_i.

2) $\int_0^\infty Pr_i(z)dz < \infty$.

3) The expected payment satisfies $E[p_i] = b_i Pr_i(b_i) + \int_{b_i}^\infty Pr_i(z)dz$.

Next, we introduce the concept of the exponential mechanism and its properties. In the literature of differential privacy, the exponential mechanism is often used to design privacy-preserving mechanisms. A key component of the exponential mechanism is the score function $q(A, o)$, which maps the input set A and an outcome $o \in \mathcal{O}$ to a real-valued score. The score represents how good the outcome o is for the input set A compared with the optimal outcome.

[1] If user i reports Γ_i' containing tasks not in Γ_i, i.e., $\Gamma_i' \backslash \Gamma_i \neq \emptyset$, she cannot finish Γ_i' when selected. If user i reports $\Gamma_i' \subset \Gamma_i$ with c_i, the probability of user i being selected will not increase according to our mechanism. The case where user i misreports both Γ_i and c_i is challenging, because calculating the true cost of $\Gamma_i' \subset \Gamma_i$ is still an open question.

Given an outcome space $\mathcal{O}$, an input set A, a score function q and a small constant ϵ, the exponential mechanism $\epsilon_q^\epsilon(A)$ chooses an outcome $o \in \mathcal{O}$ with probability

$$Pr\left[\epsilon_q^\epsilon(A) = o\right] \propto \exp\left(\epsilon q(A, o)\right).$$

The exponential mechanism has the following properties.

Theorem 2: [24] The exponential mechanism gives $2\epsilon\Delta$-differential privacy, where Δ is an upper-bound of the difference of any two input sets.

Theorem 3: [13] The exponential mechanism, when used to select an output $o \in \mathcal{O}$, $\epsilon_q^\epsilon(A)$ yields $2\epsilon\Delta$-differential privacy, letting $\mathcal{O}_{OPT}$ be the subset of $\mathcal{O}$ achieving $q(A, o) = \max\limits_{r} q(A, o)$, ensures that

$$Pr\ \left[q\left(A, \epsilon_q^\epsilon(A)\right) < \max_{o} q(A, o) - \ln(|\mathcal{O}|/|\mathcal{O}_{OPT}|)/\epsilon - t/\epsilon\right]$$
$$\leq\ \exp(-t).$$

D. Design Objective

The goal of our framework design is to minimize the social cost while achieving computational efficiency, individual rationality, truthfulness and differential privacy. Solving the minimization problem itself, referred to as the social cost minimization (SCM) problem, is challenging because SCM is NP-hard (proved by Theorem 4), let alone combining with the other three properties. Next, we give the formal formulation of the SCM problem.

SCM problem: Given a task set $\mathcal{T}$ and a user set $\mathcal{U}$, the goal of the SCM problem is to find a subset $\mathcal{S} \subseteq \mathcal{U}$, such that

$$\min C(\mathcal{S}) = \sum_{i \in \mathcal{S}} c_i, \qquad s.t. \bigcup_{i \in \mathcal{S}} \Gamma_i = \mathcal{T}.$$

Theorem 4: The SCM problem is NP-hard.

Proof: We prove the NP-hardness of the SCM problem by a polynomial time reduction from the minimum weighted set cover (MWSC) problem, which is NP-hard [6].

The MWSC problem is defined as follows: Given a universe set U and a set $S = \{s_1, s_2, \ldots, s_n\}$ of subsets of U, i.e., $s_i \subseteq U$ and the weight of each s_i is $w(s_i)$, find the minimum weight subset of S whose union is U.

Next, we construct an instance of the SCM problem from an instance of the MWSC problem in polynomial time. We create a task in the task set $\mathcal{T}$ for each element in U. There is a user in the user set $\mathcal{U}$ corresponding to each element in $\mathcal{S}$, where Γ_i consists of tasks corresponding to s_i and $c_i = w(s_i)$. It is straightforward to see that there is a solution to the MWSC problem if and only if there is a solution to the SCM problem. Therefore, the theorem SCM problem is NP-hard. ■

Since the SCM problem is NP-hard, we aim to find an approximate solution.

IV. BidGuard: Differentially Private Auction Framework

In this section, we design and analyze BidGuard, a differentially private auction framework.

A. Design Rationale

BidGuard integrates the exponential mechanism with the reverse auction to achieve computational efficiency, individual rationality, truthfulness, differential privacy and approximate social cost minimization. In our framework, users are selected iteratively. In each iteration, redundant users are eliminated and each remaining user is assigned a probability of being selected. The framework then selects one of them as the winner according to the probability distribution. Specifically, the probability of a user to be selected is set according to a specific criterion. The above processes repeats until all the sensing tasks can be completed by the selected users. Finally, the framework computes the payment to each winner.

B. Design of BidGuard

In this section, we will describe BidGuard in detail. As illustrated in Algorithm 1, BidGuard consists of three phases: user screening, winner selection, and payment determination. It executes these three phases iteratively until all the sensing tasks can be completed by the selected users.

```
Algorithm 1: BidGuard
  Input  : A set of sensing tasks $\mathcal{T}$, a set of users $\mathcal{U}$,
           submitted task-bid profile $\overrightarrow{\beta}$, and differential
           privacy parameters $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$.
  Output: A set of winners $\mathcal{S}$ and a payment profile $\overrightarrow{p}$.
$\mathcal{S} \leftarrow \emptyset$, $\mathcal{T}_c \leftarrow \emptyset$, $\mathcal{R} \leftarrow \mathcal{U}$;
foreach $i \in \mathcal{U}$ do $p_i \leftarrow 0$;
while $\mathcal{T}_c \neq \mathcal{T}$ do
    foreach $i \in \mathcal{R}$ do
        if $\Gamma_i \subseteq \mathcal{T}_c$ then $\mathcal{R} \leftarrow \mathcal{R} \setminus \{i\}$ ;
    end
    foreach $i \in \mathcal{R}$ do
        Calculate the probability $Pr_i(b_i)$ of each user
        being selected according to the score function;
    end
    Select one user randomly, denoted by $i'$, according to
    the computed probability distribution;
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{i'\}$, $\mathcal{T}_c \leftarrow \mathcal{T}_c \cup \Gamma_{i'}$, $\mathcal{R} \leftarrow \mathcal{R} \setminus \{i'\}$ ;
end
foreach $i \in \mathcal{S}$ do $p_i \leftarrow b_i + \frac{\int_{b_i}^{b_{max}} Pr_i(z)dz}{Pr_i(b_i)}$ ;
return $\mathcal{S}$ and $\overrightarrow{p}$.
```

1) User Screening Phase

BidGuard will eliminate all the redundant users, whose task set can be completed by the currently selected users. The set of remaining users is denoted by $\mathcal{R}$.

2) Winner Selection Phase

BidGuard will assign each user $i \in \mathcal{R}$ a probability of being selected as follows. It first computes a criterion $r(\beta_i)$, which is the bid divided by the number of tasks that cannot be completed by the currently selected users, i.e.,

$$r(\beta_i) = \frac{b_i}{|\Gamma_i - \mathcal{T}_c|}, \tag{1}$$

where $\mathcal{T}_c$ is the set of tasks that can be completed by the currently selected users. BidGuard selects the user with the lowest $r(\beta_i)$ in each iteration. To apply the exponential mechanism, we need to design a score function, which is a non-increasing function of $r(\beta_i)$. The probability of each user of being selected is set to the value of the score function.

3) Payment Determination Phase

Let $Pr_i(z)$ denote the probability of user i being selected with bid z. According to Theorem 1, we calculate the payment to winner i is

$$p_i = b_i + \frac{\int_{b_i}^{b_{max}} Pr_i(z)dz}{Pr_i(b_i)}.$$

C. Design of Score Functions

To apply the exponential mechanism, we need to design a score function. Specifically, we design two score functions, *linear score function* and *log score function*. We will show that they have different theoretical bounds on the social cost (Section IV-D) and performance in simulations (Section V).

Linear score function: $f_{LIN}(x) = 1 - x$. For any bidder $i \in \mathcal{R}$, the probability of being selected in each iteration is

$$Pr_i(b_i) \propto \begin{cases} \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right), & \text{if } i \in \mathcal{R}; \\ 0, & \text{otherwise,} \end{cases}$$

where $\epsilon' = \epsilon/(e\Delta \ln(e/\delta))$. Note that in order to guarantee the value of the score function is nonnegative, we normalize $r(\beta_i)$, i.e., $\frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|}$. Then the probability is

$$Pr_i(b_i) = \begin{cases} \frac{\exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}{\sum_{j \in \mathcal{R}} \exp\left(\epsilon'(1 - \frac{b_j}{b_{max}|\Gamma_j - \mathcal{T}_c|})\right)}, & \text{if } i \in \mathcal{R}; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Log score function: $f_{LOG}(x) = \log_{1/2} x$. For any bidder $i \in \mathcal{R}$, the probability of being selected in each iteration is

$$Pr_i(b_i) \propto \begin{cases} \exp\left(\epsilon' \log_{1/2} \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|}\right), & \text{if } i \in \mathcal{R}; \\ 0, & \text{otherwise,} \end{cases}$$

where $\epsilon' = \epsilon/(e \ln(e/\delta) \log_{1/2}(1/(1+\Delta)))$. We also normalize the $r(\beta_i)$, i.e., $\frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|}$ to guarantee the value of the score function is nonnegative. Then the probability is

$$Pr_i(b_i) = \begin{cases} \frac{\exp\left(\epsilon' \log_{1/2} \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|}\right)}{\sum_{j \in \mathcal{R}} \exp\left(\epsilon' \log_{1/2} \frac{b_j}{b_{max}|\Gamma_j - \mathcal{T}_c|}\right)}, & \text{if } i \in \mathcal{R}; \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Throughout the rest of this paper, we denote the BidGuard with linear score function f_{LIN} and log score function f_{LOG} by LIN and LOG, respectively.

D. Analysis of BidGuard

In this section, we first analyze the properties of LIN.

Theorem 5: LIN achieves computational efficiency, individual rationality, truthfulness, and $(\epsilon(e-1)/e, \delta)$-differential privacy, where $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$ are constants, e is the base of the natural logarithm. In addition, it has social cost at most $H_K \mathcal{OPT} + gK \cdot O(\ln n)$ with probability at least $1 - 1/n^{O(1)}$, where $H_K = \sum_{j=1}^{K} 1/j$, K is the size of the largest user task set, $\mathcal{OPT}$ is the optimal social cost, g is the size of the optimal user set, and n is the number of users.

Proof: We first prove the computational efficiency. The outer while-loop (Lines 3-12) will run at most m iterations since there are m tasks. Meanwhile, the two inner for-loops (Lines 4-6) and (Lines 7-9) will run at most n iterations since there are n users. Therefore, the total computational complexity of LIN is $O(mn)$. The individual rationality is guaranteed by the fact that the payment to each winner i is $p_i = b_i + \frac{\int_{b_i}^{b_{max}} Pr_i(z)dz}{Pr_i(b_i)} \geq b_i$. In order to prove the rest of this theorem, we prove the following lemmas. ■

Lemma 4.1: LIN is truthful.

Proof: According to (2) and (3), the probability $Pr_i(b_i)$ of user i being selected in BidGuard is monotonically non-increasing in her bid b_i. In addition, no bid is greater than b_{max} in our model. Thus we have $\int_0^\infty Pr_i(z)dz = \int_0^{b_{max}} Pr_i(z)dz < \infty$. Furthermore, we have

$$\begin{aligned} & E[p_i] \\ = \; & (1 - Pr_i(b_i)) \times 0 + Pr_i(b_i) \times (b_i + \frac{\int_{b_i}^{b_{max}} Pr_i(z)dz}{Pr_i(b_i)}) \\ = \; & b_i Pr_i(b_i) + \int_{b_i}^{\infty} Pr_i(z)dz. \end{aligned}$$

Then, according to Theorem 1, the lemma holds. ■

Lemma 4.2: For any constants $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$, LIN achieves $(\epsilon(e-1)/e, \delta)$-differential privacy, where e is the base of the natural logarithm.

Proof: Let $\overrightarrow{\beta}$ and $\overrightarrow{\beta'}$ be two input task-bid profiles that differ in any user d's bid, respectively. Let $M(\overrightarrow{\beta})$ and $M(\overrightarrow{\beta'})$ denote the sequences of users selected by LIN with inputs $\overrightarrow{\beta}$ and $\overrightarrow{\beta'}$, respectively. We show that LIN, even revealing the order in which the users are chosen, achieves differential privacy for an arbitrary sequence of users $\mathbb{I} = i_1, i_2, \ldots, i_l$ of arbitrary length l. We consider the relative probability of LIN for given task-bid inputs $\overrightarrow{\beta}$ and $\overrightarrow{\beta'}$:

$$\begin{aligned} \frac{Pr\left[M(\overrightarrow{\beta}) = \mathbb{I}\right]}{Pr\left[M(\overrightarrow{\beta'}) = \mathbb{I}\right]} &= \prod_{j=1}^{l} \frac{\frac{\exp\left(\epsilon'(1 - \frac{b_{i_j}}{b_{max}|\Gamma_{i_j} - \mathcal{T}_c|})\right)}{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}}{\frac{\exp\left(\epsilon'(1 - \frac{b'_{i_j}}{b_{max}|\Gamma_{i_j} - \mathcal{T}_c|})\right)}{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b'_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}} \\ &= \prod_{j=1}^{l} \frac{\exp\left(\epsilon'(1 - \frac{b_{i_j}}{b_{max}|\Gamma_{i_j} - \mathcal{T}_c|})\right)}{\exp\left(\epsilon'(1 - \frac{b'_{i_j}}{b_{max}|\Gamma_{i_j} - \mathcal{T}_c|})\right)} \\ &\quad \times \prod_{j=1}^{l} \frac{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b'_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}, \end{aligned}$$

where $\mathcal{U}_j = \mathcal{U} \setminus \{i_1, i_2, \ldots, i_{j-1}\}$ and the first equation is based on (2). We then prove this lemma by cases. When $b_d > b'_d$, the second product is at most 1 because the factor for any

$j \in [1, l]$ is less than 1 if $d \in \mathcal{U}_j$ and equal to 1 otherwise. Therefore, we have

$$\begin{aligned}
\frac{Pr\left[M(\overrightarrow{\beta}) = i_1, i_2, \ldots, i_l\right]}{Pr\left[M(\overrightarrow{\beta'}) = i_1, i_2, \ldots, i_l\right]} &\leq \frac{\exp\left(\epsilon'(1 - \frac{b_d}{b_{max}|\Gamma_d - \mathcal{T}_c|})\right)}{\exp\left(\epsilon'(1 - \frac{b'_d}{b_{max}|\Gamma_d - \mathcal{T}_c|})\right)} \\
&= \exp\left(\epsilon' \frac{b'_d - b_d}{b_{max}|\Gamma_d - \mathcal{T}_c|}\right) \\
&\leqslant \exp\left(\epsilon'(b'_d - b_d)\right) \\
&\leqslant \exp(\epsilon'\Delta).
\end{aligned}$$

When $b_d \leqslant b'_d$, the first product is at most 1 because the factor for any $j \in [1, l]$ is less than 1 if $i_j = d$ and equal to 1 otherwise. In the remainder of the proof, we focus on this case. Therefore, we have

$$\begin{aligned}
&\frac{Pr\left[M(\overrightarrow{\beta}) = i_1, i_2, \ldots, i_l\right]}{Pr\left[M(\overrightarrow{\beta'}) = i_1, i_2, \ldots, i_l\right]} \\
\leqslant\ & \prod_{j=1}^{l} \frac{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b'_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)} \\
=\ & \prod_{j=1}^{l} \frac{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon' \frac{\theta_i}{|\Gamma_i - \mathcal{T}_c|}\right) \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)}{\sum_{i \in \mathcal{U}_j} \exp\left(\epsilon'(1 - \frac{b_i}{b_{max}|\Gamma_i - \mathcal{T}_c|})\right)} \\
=\ & \prod_{j=1}^{l} E_{i \in \mathcal{U}_j}\left[\exp\left(\epsilon' \frac{\theta_i}{|\Gamma_i - \mathcal{T}_c|}\right)\right] \\
\leqslant\ & \prod_{j=1}^{l} E_{i \in \mathcal{U}_j}\left[\exp(\epsilon'\theta_i)\right],
\end{aligned}$$

where $\theta_i = b'_i - b_i$. For all $x \leqslant 1$, $e^x \leqslant 1 + (e-1) \cdot x$. Therefore, for all $\epsilon' \leqslant 1$, we have

$$\begin{aligned}
\prod_{j=1}^{l} E_{i \in \mathcal{U}_j}[\exp(\epsilon'\theta_i)] &\leqslant \prod_{j=1}^{l} E_{i \in \mathcal{U}_j}[1 + (e-1)\epsilon'\theta_i] \\
&\leqslant \exp\left((e-1)\epsilon' \sum_{j=1}^{l} E_{i \in \mathcal{U}_j}[\theta_i]\right).
\end{aligned}$$

Lemma B.2 in [13] implies that $Pr[\sum_{j=1}^{l} E_{i \in \mathcal{U}_j}[\theta_i] > \Delta \ln(e/\delta)] \leqslant \delta$. Let $\mathcal{O}$ denote the outcome space, where each $o \in \mathcal{O}$ is a sequence of users $i_1, i_2, \cdots, i_l$. We split $\mathcal{O}$ into two sets $\mathcal{O}'$ and $\mathcal{O}''$, where $\mathcal{O}' = \{o \in \mathcal{O} | \sum_{j=1}^{l} E_{i \in \mathcal{U}_j}[\theta_i] \leq \Delta \ln(e/\delta)\}$ and $\mathcal{O}'' = \mathcal{O} \setminus \mathcal{O}'$. Thus we have

$$\begin{aligned}
& Pr\left[M(\overrightarrow{\beta}) \in \mathcal{O}\right] \\
=\ & \sum_{o \in \mathcal{O}} Pr\left[M(\overrightarrow{\beta}) = o\right] \\
=\ & \sum_{o \in \mathcal{O}'} Pr\left[M(\overrightarrow{\beta}) = o\right] + \sum_{o \in \mathcal{O}''} Pr\left[M(\overrightarrow{\beta}) = o\right] \\
\leq\ & \sum_{o \in \mathcal{O}'} \exp\left((e-1)\epsilon'\Delta \ln(e/\delta)\right) Pr\left[M(\overrightarrow{\beta'}) = o\right] + \delta \\
\leq\ & \exp((e-1)\epsilon'\Delta \ln(e/\delta) Pr\left[M(\overrightarrow{\beta'}) \in \mathcal{O}\right] + \delta \\
=\ & \exp(\epsilon(e-1)/e) Pr\left[M(\overrightarrow{\beta'}) \in \mathcal{O}\right] + \delta.
\end{aligned}$$

The lemma holds. ■

Lemma 4.3: With probability at least $1 - 1/n^{O(1)}$, LIN has social cost at most $H_K\mathcal{OPT} + gK \cdot O(\ln n)$, where $H_K = \sum_{j=1}^{K} 1/j$, K is the size of the largest user task set, i.e., $K = \max_{i \in \mathcal{U}} |\Gamma_i|$, $\mathcal{OPT}$ is the optimal social cost, g is the size of the optimal user set, and n is the number of users.

Proof: Consider any task-bid pair $\beta_i = (\Gamma_i, b_i)$ of user i in the optimal solution $\mathcal{S}^*$, where $\Gamma_i = \{t_j, t_{j-1}, \ldots, t_1\}$. The social cost of the optimal solution is $\mathcal{OPT} = \sum_{i \in \mathcal{S}^*} b_i$. For truthful mechanisms, we have $b_i = c_i$. Without loss of generality, suppose the task in Γ_i is completed in the order of $t_j, t_{j-1}, \ldots, t_1$. At the start of the iteration in which the algorithm completes task t_k of Γ_i, at least k tasks of Γ_i remain uncovered. Thus, if user i is selected in this iteration, the cost per task is at most $\frac{b_i}{k}$. According to Theorem 3, by taking $t = O(\ln n)$, we have the cost per task of our framework is at most $\frac{b_i}{k} + O(\ln n)$ with a probability of at least $1 - 1/n^{O(1)}$. Summing over j, the total amount of social cost of Γ_i is at most $b_i H_K + K \cdot O(\ln n)$, with a probability of at least $1 - 1/n^{O(1)}$. Summing over $i \in \mathcal{S}^*$, the social cost is at most $\sum_{i \in \mathcal{S}^*}(b_i H_K + K \cdot O(\ln n)) = H_K\mathcal{OPT} + |\mathcal{S}^*|K \cdot O(\ln n)$, with a probability of at least $1 - 1/n^{O(1)}$. ■

For LOG we have the following properties. The proofs are similar to those for LIN, and thus omitted.

Theorem 6: LOG achieves computational efficiency, individual rationality, truthfulness, and $(\epsilon(e-1)/e, \delta)$-differential privacy, where $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$ are two constants, e is the base of the natural logarithm. In addition, it has social cost at most $2^t H_K\mathcal{OPT}$ with probability at least $1 - e^{-t}$, for any constant $t > 0$ and $H_K = \sum_{j=1}^{K} 1/j$, where K is the size of the largest user task set, and $\mathcal{OPT}$ is the optimal social cost.

Remarks: We have demonstrated in Theorem 4 that the minimum weighted set cover problem can be reduced to the SCM problem. It is well known that the best-possible polynomial time approximation algorithm is an H_K-approximation algorithm for the weighted set cover problem [4], where H_K is the K-th harmonic number. LOG has social cost at most $2^t H_K\mathcal{OPT}$, where t is a constant, and thus it is asymptotically optimal. Even though LIN cannot be proved to be asymptotically optimal in terms of the social cost, we will show in Section V that it achieves better privacy protection than LOG.

V. Performance Evaluation

In this section, we evaluate the performance of BidGuard and compare it with TRAC [11], which is closest to our work in terms of the design objective, but does not protect users' bid privacy.

A. Simulation Setup

All the evaluation results are based on a real data set of taxi traces. The dataset consists of the traces of 320 taxi drivers, who work in the center of Rome [2]. Each taxi driver has a tablet that periodically (every 7s) retrieves the GPS locations (latitude and longitude) and sends it with the corresponding driver ID to a central server.

We consider a crowdsensing system where the task is to measure the cellular signal strength at specific locations. Each user can sense the cellular signal strength within the area centered at the user's location with a radius of 30m. Tasks are represented by GPS locations reported by taxis. We assume that the driver of each taxi is a user. We preprocess the tasks such that each task can be sensed by at least two users according to our system model.

We use three metrics to evaluate the performance of BidGuard: *social cost*, *total payment* and *privacy leakage*. The social cost, as defined in Section III, refers to the total cost of all selected users. The total payment measures the payment paid by the platform to all selected users. We first compare the social cost and total payment with TRAC. Then we compare the social cost of BidGuard with the optimal social cost. We define *privacy leakage* to quantitatively measure the differential privacy performance of BidGuard.

Privacy Leakage: Given a mechanism M, let $\overrightarrow{\beta}$ and $\overrightarrow{\beta}'$ be two task-bid profiles, which only differ in one user's bid. Let $M(\overrightarrow{\beta})$ and $M(\overrightarrow{\beta}')$ denote the outcome of M with input $\overrightarrow{\beta}$ and $\overrightarrow{\beta}'$, respectively. The privacy leakage, denoted by PL, is defined as the Kullback-Leibler divergence of the two outcome probability distributions based on $\overrightarrow{\beta}$ and $\overrightarrow{\beta}'$, i.e.,

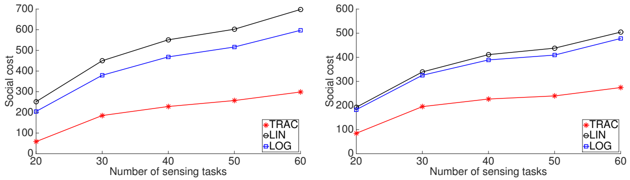
$$PL = \sum_{o \in \mathcal{O}} Pr\left[M(\overrightarrow{\beta}) = o\right] \ln \left(\frac{Pr\left[M(\overrightarrow{\beta}) = o\right]}{Pr\left[M(\overrightarrow{\beta}') = o\right]} \right). \quad (4)$$

Note that the smaller the PL value is, the harder it is to distinguish the two task-bid profiles, and thus the better the privacy preserving performance is achieved.

In our simulations, we randomly select locations as the sensing tasks according to the settings. We assume the bids of users are randomly distributed over $[1, 50]$. We generate users' bids according to two different distributions, i.e., uniform distribution and normal distribution. To evaluate the impact of the number of sensing tasks on the performance metrics, we set the number of users to 200 and vary the number of sensing tasks from 20 to 60 with a step of 10. To evaluate the impact of the number of users on the performance metrics, we set the number of sensing tasks to 150 and vary the number of users from 100 to 300 with a step of 50. For the differential privacy parameters, we set $\epsilon = 0.1$ and $\delta = 0.25$. All the results are averaged over 1000 independent runs for each setting.

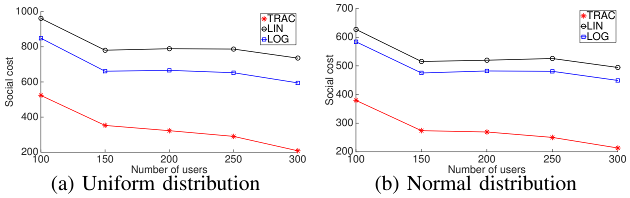
B. Evaluation of Social Cost

We first compare the social cost of BidGuard with that of TRAC. The impact of the number of sensing tasks on the social cost under uniform distribution and normal distribution is shown in Fig. 1(a) and Fig. 1(b), respectively. For both distributions, the social cost of TRAC and that of BidGuard both increase when the number of sensing tasks grows. This is because with more sensing tasks, the platform may select more users incurring a higher social cost. We can also see that the social cost of TRAC is smaller than that of BidGuard. This is because TRAC is determinate to select the user with lowest criterion value (defined in (1)) in each iteration. In contrast, since BidGuard is randomized, it cannot always guarantee to select the user with the lowest criterion value in each iteration. Besides, the social cost of LOG is smaller than that of LIN for both uniform distribution and normal distribution. This is because LOG prefers to select users with low bid, as the log score function will give more probability of being selected to low-bid users.



(a) Uniform distribution (b) Normal distribution

Fig. 1. Impact of the number of sensing tasks on the social cost

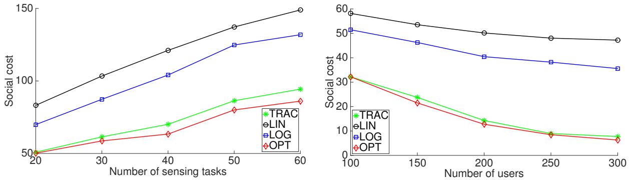


(a) Uniform distribution (b) Normal distribution

Fig. 2. Impact of the number of users on the social cost

Fig. 2(a) and Fig. 2(b) depict the impact of the number of users on the social cost under uniform distribution and normal distribution, respectively. We can see that, no matter what the distribution is, the social cost decreases slightly when the number of users increases for both TRAC and BidGuard. This is because, with more users, the platform can find more low-cost users to complete the sensing tasks. The social cost of TRAC is smaller than that of BidGuard. The reason is same as explained for Fig. 1. Meanwhile, the social cost of LOG is smaller than that of LIN for the same reason as above.

In Fig. 3, we compare the social cost of BidGuard and TRAC with the optimal solution, which is denoted by OPT. In this case, we only consider the uniform distribution, because the results will have similar pattern for the normal distribution according to Fig. 1 and Fig. 2. Since finding the optimal solution takes exponential time, we set the number of the users to 10 for Fig. 3(a), and set the number of sensing tasks to 4 for Fig. 3(b). We can see that the social cost in Fig. 3(a) and Fig. 3(b) have the same pattern as shown in Fig. 1(a) and Fig. 2(a), respectively. The reason is similar to those explained for Fig. 1(a) and Fig. 2(a). Furthermore, we can observe that BidGuard sacrifices the social cost for the users' bid privacy, compared to TRAC and the optimal solution. Note that in Fig. 3(b), the social cost of TRAC is very close to that of OPT. This is because TRAC is an H_K-approximation algorithm, where $H_k \approx 2.34$ in this figure.

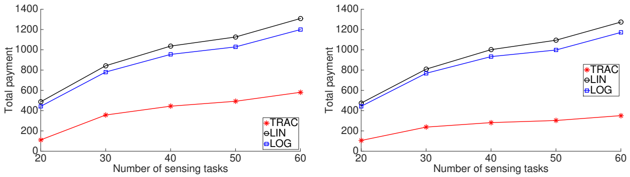


(a) Impact of the number of sensing tasks (b) Impact of the number of users

Fig. 3. Comparison of BidGuard, TRAC and OPT

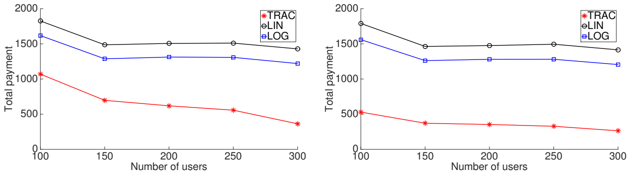
C. Evaluation of Total Payment

In Fig. 4 and Fig. 5, we plot the impact of the number of sensing tasks and the impact of the number of users on the total payment under two distributions, respectively. The results show that the total payment of both TRAC and BidGuard follow the same pattern as the social cost. In addition, the LOG has smaller total payment than that of LIN because the log score function could select users with lower social cost as shown in Fig. 1 and Fig. 2.



(a) Uniform distribution (b) Normal distribution

Fig. 4. Impact of the number of sensing tasks on the total payment



(a) Uniform distribution (b) Normal distribution

Fig. 5. Impact of the number of users on the total payment

D. Evaluation of Privacy Leakage

Next, we evaluate BidGuard in terms of privacy leakage. Since TRAC is deterministic, the privacy leakage is undefined according to (4). We only consider the uniform distribution in this case because the normal distribution has similar patterns. Fig. 6(a) and Fig. 6(b) plot the impact of the number of sensing tasks and the number of users on the privacy leakage, respectively. We observe that the privacy leakage values in both figures are very small which indicates that BidGuard achieves a good differential privacy.

Fig. 6(a) shows the privacy leakage of BidGuard when the number of sensing tasks varies. We see that the privacy leakage of LIN is always smaller than that of LOG, which indicates that LIN has better privacy protection performance than LOG. This is because the linear score function treats the probability of every outcome uniformly, however, the log score function gives more probability to the outcome with low social cost. We do not observe a pattern of the privacy leakage when the number of tasks increases. The reason is, according to the definition of privacy leakage, the difference between the probabilities of two outcomes to be selected is independent of the number of sensing tasks.

In Fig. 6(b), we can see the impact of the number of users on the privacy leakage of BidGuard. Note that the privacy leakage value decreases when the number of users increases for both LIN and LOG. This is because the probability of each outcome decreases as the number of users increases. Specifically, the more users in the system, the more possible outcomes of BidGuard, the less difference between the probabilities of two outcomes to be selected, and thus the better differential privacy performance. We can also see the privacy protection performance of LIN is better than that of LOG. The reason for this is similar to that discussed above.

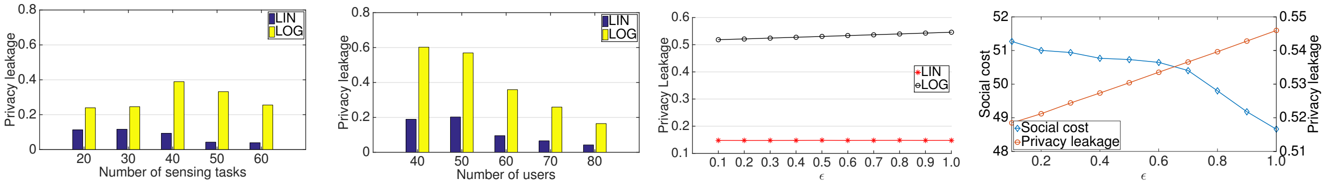
Fig. 6(c) shows the impact of the differential privacy parameter ϵ on the privacy leakage. The results show that the value of ϵ has more impact on the privacy leakage for LOG than that of LIN. This is because the log score function is more sensitive than the linear score function. For LOG, the privacy leakage increases slightly when the value of ϵ grows. This is because, theoretically, the lower the ϵ is, the better the differential privacy is achieved, and thus the lower privacy leakage. Meanwhile, it is easy to observe that the privacy leakage of LIN is smaller than that of LOG. This can also be explained by the same reason for Fig. 6(a).

Fig. 6(d) illustrates the tradeoff between the social cost and the privacy leakage of LOG. We observe that the privacy leakage decreases as the decreasing of ϵ. The reason is similar to that discussed for Fig. 6(c). However, this improvement in privacy comes at a cost of the increased social cost.

Remarks: Compared with TRAC, which does not protect users' bid privacy, BidGuard sacrifices the social cost and payment for the users' bid privacy. Besides, LIN outperforms LOG in terms of privacy protection, while LOG has lower social cost and payment.

VI. Conclusion

In this paper, we have proposed BidGuard, the first general framework for privacy-preserving crowdsensing incentive mechanisms, which achieves computational efficiency, individual rationality, truthfulness, approximate social cost minimization, and differential privacy. We designed two score



(a) Impact of the number of sensing tasks

(b) Impact of the number of users

(c) Impact of ϵ

(d) Social cost v.s. privacy leakage

Fig. 6. Evaluation of privacy leakage

functions, linear and log, to realize the framework. Specifically, the BidGuard with log function is asymptotically optimal in terms of the social cost. Extensive simulations evaluate the performance and validate the desired properties of BidGuard.

References

[1] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. CCS*, 2013, pp. 901–914.

[2] L. Bracciale, M. Bonola, P. Loreti, G. Bianchi, R. Amici, and A. Rabuffi, "CRAWDAD data set roma/taxi (v. 2014-07-17)," Downloaded from http://crawdad.org/roma/taxi/, Jul. 2014.

[3] Y. Chen, S. Chong, I. A. Kash, T. Moran, and S. Vadhan, "Truthful mechanisms for agents that value privacy," in *Proc. EC*, 2013, pp. 215–232.

[4] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of operations research*, vol. 4, no. 3, pp. 233–235, 1979.

[5] C. Clifton, "Using sample size to limit exposure to data mining," *Journal of Computer Security*, vol. 8, no. 4, pp. 281–307, 2000.

[6] T. H. Cormen, *Introduction to algorithms*. MIT press, 2009.

[7] E. De Cristofaro and C. Soriente, "Short paper: Pepsi—privacy-enhanced participatory sensing infrastructure," in *Proc. WiSec*, 2011, pp. 23–28.

[8] C. Dwork, "Differential privacy," in *Encyclopedia of Cryptography and Security*, 2011, pp. 338–340.

[9] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of cryptography*, 2006, pp. 265–284.

[10] J. Fan, Q. Li, and G. Cao, "Privacy-aware and trustworthy data aggregation in mobile sensing," in *Proc. CNS*, 2015, pp. 31–39.

[11] 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 *Proc. INFOCOM*, 2014, pp. 1231–1239.

[12] B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: Architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 1–18, 2008.

[13] A. Gupta, K. Ligett, F. McSherry, A. Roth, and K. Talwar, "Differentially private combinatorial optimization," in *Proc. SODA*, 2010, pp. 1106–1125.

[14] T. H. Hinke, H. S. Delugach, and R. P. Wolf, "Protecting databases from inference attacks," *Computers & Security*, vol. 16, no. 8, pp. 687–708, 1997.

[15] Z. Huang and S. Kannan, "The exponential mechanism for social welfare: Private, truthful, and nearly optimal," in *Proc. FOCS*, 2012, pp. 140–149.

[16] S. Jajodia and C. Meadows, "Inference problems in multilevel secure database management systems," *Information Security: An integrated collection of essays*, vol. 1, pp. 570–584, 1995.

[17] H. Jin, L. Su, B. Ding, K. Nahrstedt, and N. Borisov, "Enabling privacy-preserving incentives for mobile crowd sensing systems," in *Proc. ICDCS*, 2016.

[18] H. Jin, L. Su, H. Xiao, and K. Nahrstedt, "INCEPTION: incentivizing privacy-preserving data aggregation for mobile crowd sensing systems," in *Proc. MobiHoc*, 2016, pp. 341–350.

[19] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preventing location-based identity inference in anonymous spatial queries," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 12, pp. 1719–1733, 2007.

[20] I. Krontiris and T. Dimitriou, "A platform for privacy protection of data requesters and data providers in mobile sensing," *Computer Communications*, 2015.

[21] Q. Li and G. Cao, "Providing efficient privacy-aware incentives for mobile sensing," in *Proc. ICDCS*, 2014, pp. 208–217.

[22] X. Liu, Z. Zhou, W. Diao, Z. Li, and K. Zhang, "When good becomes evil: Keystroke inference with smartwatch," in *Proc. CCS*, 2015, pp. 1273–1285.

[23] T. Luo, H.-P. Tan, and L. Xia, "Profit-maximizing incentive for participatory sensing," in *Proc. INFOCOM*, 2014, pp. 127–135.

[24] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proc. FOCS*, 2007, pp. 94–103.

[25] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proc. SenSys*, 2008, pp. 323–336.

[26] X. Niu, M. Li, Q. Chen, Q. Cao, and H. Wang, "EPPI: An e-cent-based privacy-preserving incentive mechanism for participatory sensing systems," in *Proc. IPCCC*, 2014, pp. 1–8.

[27] H. Ohashi, "Effects of transparency in procurement practices on government expenditure: A case study of municipal public works," *Review of Industrial Organization*, vol. 34, no. 3, pp. 267–285, 2009.

[28] M. Shin, C. Cornelius, D. Peebles, A. Kapadia, D. Kotz, and N. Triandopoulos, "Anonysense: A system for anonymous opportunistic sensing," *Pervasive and Mobile Computing*, vol. 7, no. 1, pp. 16–30, 2011.

[29] A. Stoica and C. Farkas, "Secure xml views," *Research Directions in Data and Applications Security*, pp. 133–146, 2003.

[30] J. Sun and H. Ma, "Privacy-preserving verifiable incentive mechanism for online crowdsourcing markets," in *Proc. ICCCN*, 2014, pp. 1–8.

[31] H. To, G. Ghinita, and C. Shahabi, "A framework for protecting worker location privacy in spatial crowdsourcing," in *Proc. VLDB*, vol. 7, no. 10, 2014, pp. 919–930.

[32] D. Xiao, "Is privacy compatible with truthfulness?" in *Proc. ITCS*, 2013, pp. 67–86.

[33] J. Xu, J. Xiang, and D. Yang, "Incentive mechanisms for time window dependent tasks in mobile crowdsensing," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2015.

[34] D. Yang, G. Xue, G. Fang, and J. Tang, "Incentive mechanisms for crowdsensing: Crowdsourcing with smartphones," *IEEE/ACM Trans. Netw.*, vol. PP, no. 99, pp. 1–13, 2015.

[35] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing," in *Proc. MobiCom*, 2012, pp. 173–184.

[36] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. INFOCOM*, 2015, pp. 2830–2838.

[37] X. Zhang, Z. Yang, Z. Zhou, H. Cai, L. Chen, and X. Li, "Free market of crowdsourcing: Incentive mechanism design for mobile sensing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3190–3200, 2014.

[38] P. Zhou, Y. Zheng, and M. Li, "How long to wait?: predicting bus arrival time with mobile phone based participatory sensing," in *Proc. MobiSys*, 2012, pp. 379–392.

[39] R. Zhu and K. G. Shin, "Differentially private and strategy-proof spectrum auction with approximate revenue maximization," in *Proc. INFOCOM*, 2015, pp. 918–926.