Privacy-Preserving Data Aggregation for Mobile Crowdsensing With Externality: An Auction Approach

Mengyuan Zhang, Lei Yang, *Senior Member, IEEE*, Shibo He, *Senior Member, IEEE*, Ming Li, *Member, IEEE, ACM*, and Junshan Zhang, *Fellow, IEEE*

***Abstract*— We develop an auction framework for privacy-preserving data aggregation in mobile crowdsensing, where the platform plays the role as an auctioneer to recruit workers for sensing tasks. The workers are allowed to report noisy versions of their data for privacy protection; and the platform selects workers by taking into account their sensing capabilities to ensure the accuracy level of the aggregated result. Observe that when moving the control of data privacy from the data aggregator to the workers, the data aggregator has limited market power in the sense that it can only partially control the noise by judiciously choosing a subset of workers based on workers' privacy preferences. This introduces *externalities* because the privacy of each worker depends on the total noise in the aggregated result that in turn relies on which workers are selected. Specifically, we first consider a privacy-passive scenario where workers participate if their privacy loss can be adequately compensated by the rewards. We explicitly characterize the externalities and the hidden monotonicity property of the problem, making it possible to design a truthful, individually rational and computationally efficient incentive mechanism. We then extend the results to a privacy-proactive scenario where workers have individual requirements for their perceivable data privacy levels. Our proposed mechanisms for both scenarios can select a subset of workers to (nearly) minimize the cost of purchasing their private sensing data subject to the accuracy requirement of the aggregated result. We validate the proposed scheme through theoretical analysis as well as extensive simulations.**

***Index Terms*— Crowd sensing, incentive mechanism, privacy-preserving, data aggregation.**

Manuscript received June 22, 2018; revised August 29, 2019 and December 20, 2020; accepted January 19, 2021; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Shakkottai. This work is supported by National Key R&D Program of China(No. 2018YFB1702100), and in part by NSFC under Grant 61731004, and in part by U.S. National Science Foundation under Grants CNS-1950485, IIS-1838024, and EEC-1801727. A preliminary version of this article was presented in ACM MobiHoc 2018 conference with the title of "Crowd-Empowered Privacy-Preserving Data Aggregation for Mobile Crowdsensing." *(Corresponding author: Shibo He.)*

Mengyuan Zhang and Shibo He are with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: zhang418@zju.edu.cn; s18he@zju.edu.cn).

Lei Yang is with the Department of Computer Science and Engineering, University of Nevada, Reno, NV 89557 USA (e-mail: leiy@unr.edu).

Ming Li is with the Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX 76019 USA (e-mail: ming.li@uta.edu).

Junshan Zhang is with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: junshan.zhang@asu.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNET.2021.3056490, provided by the authors.

Digital Object Identifier 10.1109/TNET.2021.3056490

I. INTRODUCTION

A. Motivation

MOBILE crowdsensing arises as a promising sensing paradigm that leverages the sensing capability of human-carried mobile devices to perform various sensing tasks (e.g., healthcare, environment monitoring, indoor localization, and smart transportation) [2]. By outsourcing the sensing tasks to the public crowd, mobile crowdsensing systems can collect fine-grained information effectively and efficiently. However, any individual involved in a sensing task inevitably authorizes the task agent a certain level of privilege to access her sensing data which may be sensitive, thereby giving rise to the privacy leakage when being released to an untrusted party. This becomes a key challenge hindering individuals (workers) from participation, other than the consumption of the limited system resources (e.g., battery and computing power) of their mobile devices. Therefore, the success of mobile crowdsensing hinges closely upon the design of efficient incentive mechanisms to stimulate workers' participation.

Most of incentive mechanisms developed for mobile crowdsensing systems (e.g., [3]–[17]) take into account only workers' sensing costs. Only a few recent works consider workers' privacy costs. However, in these works, either workers have no control of their data privacy (e.g., the platform is assumed to be trustworthy and fully responsible for protecting workers' private data [10]), or the platform interacts with workers via game-theoretic models (e.g., [18]), which may lead to an inefficient equilibrium, i.e., the platform may not achieve a desirable accuracy level of the aggregated result. To address these issues, it is of paramount importance to develop novel data aggregation schemes for mobile crowdsensing that not only allow the platform to selectively recruit workers based on their sensing quality,[1] but also allow workers to report their locally perturbed sensing data to the untrusted platform for privacy protection. And a key question here is how to achieve a good balance between workers' data privacy and the aggregation accuracy by the design of an incentive mechanism.

[1]The reliability of the sensing results depends on the total noise added by the workers and the sensor quality of their mobile devices [10].

Due to the existence of multiple Nash Equilibria (e.g., [18]), game-theoretic models cannot guarantee a desirable accuracy level of the data aggregation. Therefore, in this paper we take an auction approach that includes the accuracy requirement when designing the incentive mechanism. However, using an auction-based approach to select privacy-sensitive workers and collect their noisy sensing data has to tackle four major challenges for the design of effective incentive mechanisms:

- *Strategic Behavior.* As workers are allowed to perturb their data locally, if the noise is fully specified by the workers themselves, it is possible that they would play strategically by adding more noise into their sensing data to enhance their data privacy. Moreover, workers may manipulate their bids to maximize their own benefits, leading to a higher cost of achieving a desirable aggregation accuracy. Therefore, a truthful incentive mechanism is required, which integrates a carefully designed data aggregation scheme that endues the platform certain control over the workers' data perturbation.
- *Externalities.* Compared with the existing works (e.g., [10]), where the platform adds noises into workers' sensing data and workers' data privacy only depends on the noise added by the platform, the data privacy of each worker in this paper depends on which workers are selected to fulfill the task and how much noise the selected workers generate (see Section II-C), which introduces *externalities*. This makes the design of incentive mechanism in this paper more challenging.
- *Rational Behavior.* In the crowdsensing modeling, workers aim to maximize the difference between their rewards from the platform and their data privacy loss. Conventionally, a worker will opt into the system as long as her privacy loss is fully compensated by the received reward, which is called *privacy-passive* case in this paper. However, in some cases workers' behaviors can be more proactive in the sense that they might have intrinsic preferences on their data privacy levels. In such a *privacy-proactive* case, a worker would refuse to participate if the noise level determined by the mechanism is below a certain customized threshold, regardless of how much reward she could receive. Novel incentive mechanisms are required to deal with workers with different kinds of rational behaviors.
- *Computational Complexity.* To achieve a desirable accuracy level of the aggregated result in a cost-effective manner, the platform needs to find an optimal subset of workers to fulfill the sensing task. Because different workers have different valuations of their data privacy and workers' data privacy is interdependent due to externalities, it is of combinatorial nature to find an optimal subset of workers to minimize the system cost while achieving the desirable accuracy level. Therefore, a computationally efficient mechanism is needed.

B. Summary of Main Contributions

In this paper, we develop an auction framework for privacy-preserving data aggregation in mobile crowdsensing, where the workers submit their bids to the platform and the platform plays the role as an auctioneer to recruit workers for a sensing task. When aggregating noisy data from workers, the platform aims to minimize the cost of purchasing the *private* sensing data, while achieving a desirable accuracy level of the aggregated result. The externalities introduced by the coupling of users' *differential privacy* levels induce great challenges to the mechanism design problem. The consideration of an extended scenario with intrinsic workers' privacy requirements further differentiate our solution from others in the literature. Our main contributions are summarized as follows:

- *Differentially Private Data Aggregation.* To tackle the challenge due to workers' strategic behaviors, we propose a differentially private data aggregation scheme by leveraging the celebrated concept of differential privacy. The key idea is to carefully design the noise distribution for each worker based on the divisible property of Laplace distribution, such that each worker can report a privacy-preserving version of their data based on the noise distribution suggested by the platform, who can guarantee the differential privacy of each worker's data. By using this scheme, the platform can have certain control over the aggregated noise level without knowing workers' true sensing data.
- *Externalities*. Under the proposed differentially private data aggregation scheme, for different sets of workers, different noise distributions will be designed for the workers. In other words, the privacy of a worker would change if the platform chooses different workers, which introduces externalities. For the Laplace noise distribution, we explicitly characterize the externalities among workers and the impact of each worker's participation on the privacy of other workers, which is accounted in the incentive mechanism design.
- *Privacy-Accuracy Tradeoff.* To maintain the accuracy of the aggregated result, the platform would reward workers more if the reported data is of higher accuracy (i.e., less noise is added). Clearly, there is a tradeoff between the (privacy) cost and the accuracy. We characterize the tradeoff between workers' data privacy and the accuracy of the aggregated result based on the concept of differential privacy. The accuracy of the aggregated result is characterized in terms of the distortion, due to the noise added by workers.
- *Differentially Private Data Auction.* Based on the proposed differentially private data aggregation, the design of the incentive mechanism boils down to solving a privacy auction of allocating the sensing task to a set of workers that can minimize the total payment to the workers, subject to the accuracy constraint of the aggregated result. We show that it is NP-hard to find the optimal solution to this problem. By exploring the problem structure, we discover the *hidden monotonicity* property of the problem and determine the critical bid of workers. Based on these findings, we propose a computationally efficient differentially private data auction scheme despite the combinatorial nature of the problem. Moreover, we show that the proposed differentially private data auction scheme

is truthful, individually rational and close to the optimal solution. The performance of the proposed scheme is evaluated via extensive simulations.

- *Intrinsic Privacy Requirements*. We make an extension of the basic differentially private data auction to deal with the case with privacy-proactive workers. Specifically, we incorporate workers' intrinsic privacy preference on their data privacy levels to the incentive mechanism. Each worker would report her lowest acceptable data privacy level together with her unit privacy cost to the platform. An amended truthful auction mechanism is provided in Section IV, which takes the two-dimensional bids of workers as input.

C. Related Work

Incentive mechanism design for mobile crowdsensing systems has recently garnered much attention (e.g., [3]–[13], [15], [16]). Different models (e.g., auction [3]–[11] and game-theoretic models [12], [13], [15], [16]) have been introduced to design incentive mechanisms with different objectives, including social welfare maximization (e.g., [9], [15], [19]), cost or payment minimization (e.g., [4], [10]), and platform's profit maximization (e.g., [11], [13]). Most of the existing works (e.g., [3]–[9]) consider only the sensing costs of the participants.

Recently, there has been much attention paid to data privacy (e.g., [10], [11], [17], [18], [20]–[22]). Most of these works (e.g., [10], [11], [20]–[22]) assume that the platform (i.e., the data collector) is trustworthy and the true data is reported to the platform, where workers have no control of their data privacy. Very recent works [18], [23] allow the workers to protect their data privacy by reporting noisy data and study how to trade private data in game-theoretic models, which, however, may result in an inefficient equilibrium, i.e., the accuracy of the aggregated result cannot be guaranteed. To address these issues, this paper proposes a novel auction framework for mobile crowdsensing, where the workers can protect their data privacy by adding noise based on the noise distributions determined through the proposed data aggregation scheme. Specifically, we consider *frugal* mechanism design [24], [25] which aims to minimize the total payment of the buyer (the platform) for procuring a feasible set of workers whose aggregated data achieves a desirable accuracy level. We caution that, however, the threshold-based mechanism for frugal mechanism [24], [26] cannot be applied directly to the problem under consideration due to the effect of externalities, which makes the mechanism design more challenging. [2]

Although the control of data privacy is moved from the platform side to the workers side by allowing workers to conduct local noise injection, the platform still has the power to specify the noise injection level for each selected worker [1]. In order to grant workers more control of their data privacy levels, we further extend the discussion to the scenario where privacy-proactive workers can impose restrictions on their least acceptable noise levels. We devise an efficient truthful incentive mechanism for this scenario where workers' bids are of two-dimension, including their bids for the unit privacy cost and their intrinsic requirements on their privacy levels. Pre-processing for selecting feasible worker set is also needed at the beginning stage for winner determination. These differentiates our approach from the existing mechanism designs on private data aggregation in mobile crowdsensing [9]–[11].

[2]Some of our preliminary results have been presented in [1].

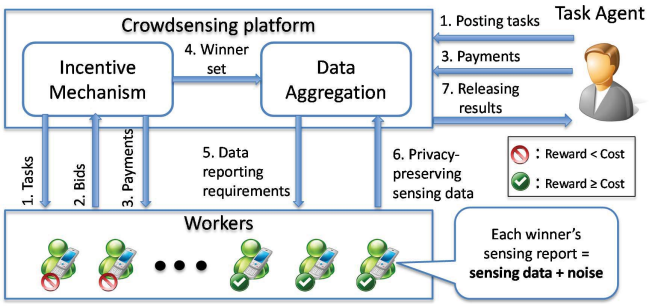


Fig. 1. An auction framework for privacy-preserving data aggregation.

The rest of the paper is organized as follows. In Section II, we describe the privacy-preserving data aggregation framework for mobile crowdsensing systems. In Section III, we propose the incentive mechanism and analyze its properties for the privacy-passive scenario. In Section IV, we extend the study to the privacy-proactive scenario where workers impose intrinsic requirement on their data privacy levels. In Section V, we evaluate the performance of the proposed incentive mechanism. The paper is concluded in Section VI.

II. Privacy-Preserving Data Aggregation for Mobile Crowdsensing

A. System Overview

Consider a mobile crowdsensing system consisting of a centralized platform $\mathcal{A}$, a task agent $\mathcal{T}$ and a set of participating workers $\mathcal{N} \triangleq \{1, \cdots, N\}$, as illustrated in Fig. 1. The task requires workers to report to the platform their local sensing data of a specific object or phenomenon (e.g., spectrum sensing and environmental monitoring).[3] To enhance the reliability of the result, the platform will aggregate the sensing data, as the reliability of each worker's sensing data may be different due to different sensor qualities [10]. *Different from the existing works on auctions in mobile crowdsensing systems (e.g., [3]–[13], [15]–[17]), we allow each individual to report a privacy-preserving version of her data to protect her own data privacy [18].*

Specifically, the workflow (see Fig. 1) of the proposed privacy-preserving data aggregation is as follows:

- First, the task agent posts a task in the crowdsensing platform, which then announces the task to a set of N workers, denoted as $\mathcal{N}$ (step 1).
- **Incentive Mechanism.** Then the platform runs an auction to recruit workers. The workers first submit their bids to the platform (step 2), where the bids reflect the personal information of workers such as the valuation of privacy loss and the lowest acceptable data privacy protection level of each worker (see Section II-B). Based on the

[3]In this work, we assume that all the N workers are in connection with the platform and active in the crowdsensing system all the time.

collected bids, the platform determines the winners (i.e., the workers to fulfill the task), and the corresponding payments to the winners (steps 3 and 4).
- **Data Aggregation.** Next, the platform sends the data reporting requirements to the winners and allows the winners to report a privacy-preserving version of their sensing data (steps 5 and 6).
- Finally, the platform releases the aggregated result to the task agent (step 7).

B. Crowdsensing Auction Model

In the crowdsensing system, the platform plays the role as an auctioneer who recruits workers to complete the sensing task and then aggregates the sensing data. As the bidders, workers provide their private sensing data to the platform in return for payments that compensate their privacy loss. In the following, we introduce the privacy cost model, the workers model and the platform model, followed by the design objectives.

1) Platform Model: At the beginning of the auction, the platform (auctioneer) would elicit bids (defined in Section II-B.2) from the workers. By running a carefully designed winner determination procedure and a payment determination procedure, the platform outputs an allocation result $(\mathbf{x}, \mathbf{p})$, in which $\mathbf{x} = (x_1, \cdots, x_N)$ indicates the participants and $\mathbf{p} = (p_1, \cdots, p_N)$ indicates the amount of payments to the participants. Specifically, $x_i \in \{0, 1\}$ denotes if worker i is selected to execute the task: $x_i = 1$ means that worker i is selected (i.e., winner) and $x_i = 0$, otherwise. Accordingly, we define $\mathcal{S}$ as the winner set with S workers. For each worker $i \in \mathcal{N}$, the platform will pay $p_i \geq 0$ amount of reward to collect her private data, and use the data in a differentially private manner after the data aggregation (see Section II-C). The total payment that the platform spent can be expressed as $\sum_{i \in \mathcal{N}} p_i$. We denote the data aggregation accuracy requirement for the sensing task as Δ, which will be defined later in Section III.

2) Worker Model: Next, we introduce the privacy cost model, the bidding model, and the utility model for the workers.

Privacy Cost. In our model, workers incur privacy cost when providing their private sensing data to the platform. Such privacy cost is quantified using differential privacy [27]. We let $v_i > 0$ denote worker i's valuation of unit privacy cost. Intuitively, a larger value of v_i indicates that worker i has a higher intrinsic valuation of privacy loss by revealing her sensing data. We assume that all unit privacy costs are unknown to the platform or to the other workers. We let ϵ_i denote worker i's data privacy level (see formal definition in Section II-C), which is specified by the platform and closely related to the noise level for data perturbation. Specifically, the smaller the value of ϵ_i, the larger the noise worker i is allowed to add on her data, and thus the higher the data privacy level. The privacy cost c_i of worker i can be given by[4]

$$c_i = v_i \epsilon_i(\mathbf{x}). \tag{1}$$

[4]According to the utility theoretic characterization of differential privacy [22], the privacy cost can be modeled as the difference between the utility with true data vector and the utility with perturbed data vector, which is a linear function of worker's privacy ϵ_i.

Note that this cost function has been used in many existing works (e.g., [10], [11], [21], [22]). However, the worker's privacy ϵ_i in (1) in this paper is a function of $\mathbf{x}$ and depends on not only the noise added by herself but also the total noise in the aggregated result, which introduces externalities (see Section II-C). This is a key difference between this work and other related works in mobile crowdsensing (e.g., [10]), where the privacy cost of a worker purely depends on her own participation.

Bidding Model. We assume that each worker's unit privacy cost is independent to her private data, so that she would not reveal private information during the bidding process. Nevertheless, a worker may not report the true value of her unit privacy cost to gain more benefit. We differentiate the bidding model for the *privacy-passive* scenario and *privacy-proactive* scenario due to the different roles workers play during determination of their data privacy levels. In privacy-passive scenario, each worker $i \in \mathcal{N}$ simply report her unit privacy cost as bid b_i, which could be different from the true value v_i. Let $\mathbf{b} = (b_1, \cdots, b_N)$ denote the vector of bids submitted by the workers and $\mathbf{b}_{-i}$ denote the bid vector without worker i's bid. The platform runs the auction with the outcome specifying the data privacy level ϵ_i of each worker i. Worker i would passively accept the data privacy level and conduct local noise injection accordingly (see Section II-C).

In the privacy-proactive scenario, we assume each worker $i \in \mathcal{N}$ possesses an intrinsic requirement on her data privacy level such that she would drop out if ϵ_i assigned by the platform is greater than a customized threshold E_i. To impose such a constraint, worker i would report a bid tuple (b_i, g_i) for her unit privacy cost v_i and her requirement for the data privacy level E_i, respectively.

Worker's Utility. In our crowdsensing framework, each worker reports the noisy data to the platform in return for the payment p_i that compensates her privacy cost c_i. Workers are assumed to be selfish and strategic, in order to maximize their own utilities. Based on the privacy cost (1), the utility u_i of a privacy-passive worker i can be given as,

$$u_i(b_i, \mathbf{b}_{-i}) = p_i(b_i, \mathbf{b}_{-i}) - c_i = p_i(b_i, \mathbf{b}_{-i}) - v_i \epsilon_i(\mathbf{x}), \tag{2}$$

For a privacy-proactive worker, her utility is,

$$u_i(b_i, g_i, \mathbf{b}_{-i}, \mathbf{d}_{-i}) = \begin{cases} p_i(b_i, g_i, \mathbf{b}_{-i}, \mathbf{g}_{-i}) - v_i \epsilon_i(\mathbf{x}), & \text{if } \epsilon_i(\mathbf{x}) \leq g_i, \\ -\infty, & \text{Otherwise.} \end{cases} \tag{3}$$

where u_i and p_i are functions of the bid vector, given ϵ_i and $\mathbf{x}$. Here it holds that for a non-participant $i \in \mathcal{N}$ (i.e., $x_i = \epsilon_i = p_i = 0$), her utility turns out to be zero. Notice that we do not explicitly include the sensing cost of carrying out the task into the utility function (2) in order to ease the presentation. Meanwhile, our results in this paper can be easily extended to incorporate the sensing cost as in [16], [28]. For example, similar to [16], letting s_i denote the sensing cost of user i, we can modify the individual utility of user i as $u_i = p_i - s_i - \epsilon_i v_i$, and define $p_i' = p_i - s_i$ to incorporate the sensing cost in the reward. Therefore, our results can be extended to this case.

3) Design Objectives: We aim to design an auction based allocation mechanism that minimizes the total payment to the workers with satisfactory data aggregation accuracy, by designing an incentive mechanism with the following desirable properties:

- **Truthfulness:** Each worker i can maximize her utility by truthfully bidding her privacy valuation, i.e., $u_i(v_i, \mathbf{b}_{-i}) \geq u_i(b_i, \mathbf{b}_{-i})$ for any $\mathbf{b}$.
- **Individual Rationality:** Each worker $i \in \mathcal{N}$ can obtain a non-negative utility. According to (2) and (3), it implies that the payment p_i and the privacy level $\epsilon_i(\mathbf{x})$ are determined such that $u_i = p_i - c_i \geq 0$; moreover, for a privacy-proactive worker i, the constraint $\epsilon_i(\mathbf{x}) \leq g_i$ needs to be further satisfied.
- **Cost Minimization:** The mechanism can minimize the total payment to the workers.
- **Computational Efficiency:** The solution $(\mathbf{x}, \mathbf{p})$ can be computed in polynomial time.

C. Differentially Private Data Aggregation

In both the privacy-passive case and the privacy-proactive case, to protect the data privacy, each winner i will report a privacy-preserving version $\hat{d}_i$ of her data d_i by adding random noise n_i. Without loss of generality, we assume that all the sensing data d_i are normalized values within the range $[0, 1]$. In this paper, we consider a weighted aggregation operation f to calculate the aggregated result r based on workers' data. Let $\mathbf{d}$ be the vector of workers' data. The aggregated result r can be written as

$$r = f(\mathbf{d}) = \sum_{i \in \mathcal{N}} w_i(d_i + n_i)x_i = \sum_{i \in \mathcal{S}} w_i(d_i + n_i), \quad (4)$$

where $w_i > 0$ is the normalized weight of worker i such that the sum of these weights is equal to 1. Similar to [10], [29], [30], the weighted aggregation is to capture the effect of workers' diverse skill levels on the calculation of the aggregated results. Intuitively, higher weights will be assigned to workers whose sensing data are more likely to be close to the ground truths. This makes the aggregated results closer to the data provided by more reliable workers, which have been used by many state-of-the-art data aggregation methods [10], [29], [30]. The choice of weights can be based on workers' skill levels as in [10], which is *a priori* known to the platform and the workers.

In this paper, we quantify the privacy loss incurred in data aggregation based on the celebrated concept of differential privacy [27], and the proposed differentially private data aggregation is defined as follows.

Definition 1 (Differentially Private Data Aggregation): An aggregation operation $f : [0, 1]^S \rightarrow \mathbb{R}$ is ϵ_i-differentially private with respect to worker i, if for any pair of neighboring vectors $\mathbf{d}$ and $\mathbf{d}_{(i)}$ differing only in the i^{th} worker's data and any set of aggregation results $O \subseteq Range(f)$, the following inequality holds:

$$Pr[f(\mathbf{d}) \in O] \leq \exp(\epsilon_i) Pr[f(\mathbf{d}_{(i)}) \in O], \quad (5)$$

with ϵ_i being a positive parameter quantifying the data privacy level of worker i.

It follows that worker i's data is used in an ϵ_i-differentially private manner under operation f. This definition differs slightly from the definition in [27], which is stated in terms of the worst-case privacy (i.e., ϵ-differentially private, where $\epsilon = \sup_i \epsilon_i$).

Given the aggregation operation f, a well-known method to provide differential privacy is to add random noise drawn from a Laplace distribution to this function [27]. As we allow each worker to add noise by themselves, we need to carefully design the noise distribution for each worker such that the sum of the noises is equivalent to the random noise drawn from a Laplace distribution, i.e., the aggregated noise $n = \sum_{i \in \mathcal{S}} w_i n_i$ follows the Laplace distribution.

Proposition 1: For the aggregation operation f in (4), define $\epsilon_i = s_i(f)/\sigma$, where $s_i(f) = \max_{\mathbf{d}, \mathbf{d}_{(i)} \in [0,1]^S} |f(\mathbf{d}) - f(\mathbf{d}_{(i)})|$ is the sensitivity of f to the i^{th} entry d_i and σ is the parameter of the Laplace distribution. The aggregation operation f is ϵ_i-differentially private with respect to worker i, if $n_i = G_1(S, \sigma/w_i) - G_2(S, \sigma/w_i)$ for all $i \in \mathcal{S}$ are independent, where $G_1(S, \sigma/w_i)$ and $G_2(S, \sigma/w_i)$ are i.i.d. random variables following gamma distribution with pdf $g(x; S, \sigma/w_i) = \frac{1}{\Gamma(1/S)} (\frac{w_i}{\sigma})^{\frac{1}{S}} x^{\frac{1}{S}-1} e^{\frac{w_i x}{\sigma}}$.

Proof: To show Proposition 1, it suffices to show that the aggregated noise follows the Laplace distribution. Based on the divisible property of Laplace distribution [31], the Laplace distribution is divisible and can be constructed as the sum of *i.i.d.* gamma distributions. Based on the scaling law of gamma distribution, $w_i n_i = G_1(S, \sigma) - G_2(S, \sigma)$. Therefore, we have

$$\sum_{i \in \mathcal{S}} w_i n_i = \sum_{i \in \mathcal{S}} (G_1(S, \sigma) - G_2(S, \sigma)) = L(\sigma), \quad (6)$$

where the second equality follows from the divisible property of Laplace distribution [31], which concludes the proof. □

Based on Proposition 1, if the noise distribution of each worker is carefully designed, the aggregation operation f in (4) is ϵ_i-differentially private with respect to worker i. Therefore, we propose the data aggregation mechanism in Algorithm 1. In Algorithm 1, the platform only informs the workers the values of S and σ/w_i, based on which each worker generates a random noise and reports $\hat{d}_i$ back to the platform. Specifically, the quantity σ/w_i characterizes the data privacy level of worker i (see Proposition 2 in Section II-D).

Remarks:

Algorithm 1 Differentially Private Data Aggregation

1: **Input:** Worker set $\mathcal{S}$, Number of workers S, weight of each worker $w_i, \forall i \in \mathcal{S}$, Laplace distribution parameter σ
2: **Output:** Aggregated result r.
3: For each worker $i \in \mathcal{S}$, the platform informs the values of parameters S and $\frac{\sigma}{w_i}$.
4: Each worker generates a random noise n_i based on the distribution of $G_1(S, \sigma/w_i) - G_2(S, \sigma/w_i)$, and then reports $\hat{d}_i = d_i + n_i$ to the platform.
5: The platform aggregates the data from the workers using (4) and releases the aggregated result r to the task agent.

- Note that in the privacy-passive case, workers will passively inject random noise parameterized by the data privacy level informed by the platform. In contrast, in the privacy-proactive case, workers have customized requirements on their data privacy levels, and the noise parameters informed by the platform will satisfy their data privacy requirements (see detailed discussion in Section IV).
- Note that in the proposed data aggregation algorithm, the platform does not know the true value of worker's data, but a privacy-preserving version of her data, which is generated using the noise distributions that the workers agree with in the proposed auction framework. By doing so, other than protecting workers' data privacy, the proposed algorithm can also prevent workers from adding arbitrarily large noise into their sensing data, in which case the aggregated result becomes useless. For the platform, it is easy to check if the distribution of each worker's reports follows the assigned noise distribution, based on which reputation management techniques can applied to recognize the dishonest workers [32]. Also, the client-side APP can be designed to enforce the locally generated noise following the distribution regularized by the platform, to address the *moral hazard* issue.
- Also note that for different sets of winners, different noise distributions will be assigned to the winners. In other words, the privacy of each winner depends on the selection of the winner set, which introduces the *externalities*. This makes the design of incentive mechanism in this paper different from the existing works on auctions in mobile crowdsensing systems.

D. Privacy Versus Accuracy

When allowing workers to report noisy data, the noise added into the aggregated result would inevitably reduce the accuracy of the result. From Proposition 1, we observe that ϵ_i depends on the value of σ. The higher the value of σ, the smaller ϵ_i, and hence, the better the privacy guarantee. However, the higher the value of σ, the lower the accuracy of the aggregated result. *Clearly, there is a natural trade-off between workers' data privacy and the accuracy of the aggregated result.*

To characterize the accuracy, we introduce the notion of *distortion* between two aggregation functions: one using all the workers' data with no noise and the other using the selected workers' data with noise (i.e., the aggregated result r in (4)). As the platform needs to pay for the workers' data, it would be costly to get all workers' data and workers would also add noise to protect their data privacy. Therefore, we can treat the aggregation of all the workers' data with no noise as the benchmark.

Definition 2 (Distortion): Given the vector $\mathbf{x}$*, the distortion* $\delta(\mathbf{x})$ *is defined as*

$$\delta(\mathbf{x}) = \max_{\mathbf{d}\in[0,1]^N} \mathbb{E}[(\sum_{i\in\mathcal{N}} w_i d_i - \sum_{i\in\mathcal{N}} w_i(d_i+n_i)x_i)^2]. \quad (7)$$

In Definition 2, the distortion is defined as the maximum of expected deviation from the true result for any sensing data reported by the workers. It is clear that the distortion depends on the set of workers fulfilling the task and the noise added into the data. Their dependence is quantified by the following proposition.

Proposition 2 (Privacy versus Distortion): Given x_i *and* w_i *for all the workers, under the aggregation function (4), the privacy of each worker and the distortion of the aggregated result can be given as*

$$\epsilon_i = \frac{w_i x_i}{\sigma}, \quad \forall i \in \mathcal{N} \quad (8)$$

$$\delta(\mathbf{x}) = (\sum_{i\in\mathcal{N}} w_i(1-x_i))^2 + 2\sigma^2. \quad (9)$$

Proof: Given x_i and w_i under the aggregation function (4), we have

$$s_i(f) = \max_{\mathbf{d},\mathbf{d}_{(i)}\in[0,1]^S} |w_i(d_i - d_i')x_i| = w_i x_i.$$

Therefore, we have $\epsilon_i = \frac{s_i(f)}{\sigma} = \frac{w_i x_i}{\sigma}$. For the distortion, we have

$$\begin{aligned}
\delta(\mathbf{x}) &= \max_{\mathbf{d}\in[0,1]^N} \mathbb{E}[(\sum_{i\in\mathcal{N}} w_i d_i - \sum_{i\in\mathcal{N}} w_i(d_i+n_i)x_i)^2] \\
&\overset{(a)}{=} \max_{\mathbf{d}\in[0,1]^N} \mathbb{E}[(\sum_{i\in\mathcal{N}} w_i d_i(1-x_i) - \sum_{i\in\mathcal{S}} w_i n_i)^2] \\
&\overset{(b)}{=} \max_{\mathbf{d}\in[0,1]^N} (\sum_{i\in\mathcal{N}} w_i d_i(1-x_i))^2 + 2\sigma^2 \\
&= (\sum_{i\in\mathcal{N}} w_i(1-x_i))^2 + 2\sigma^2,
\end{aligned}$$

where (a) follows from equation (4) that $\sum_{i\in\mathcal{N}} w_i n_i x_i = \sum_{i\in\mathcal{S}} w_i n_i$, and (b) follows from Proposition 1 that $\sum_{i\in\mathcal{S}} w_i n_i$ is a Laplace random variable with zero mean and $2\sigma^2$ variance. □

From Proposition 2, it is clear that given σ, the more workers fulfilling the task, the less the distortion; given the set of selected workers $\mathcal{S}$, the higher the value of σ, the smaller ϵ_i (i.e., better privacy) and the worse the distortion. Following [21], we call the aggregated operation f in (4) *canonical* if the Laplace noise added by workers has a parameter of the following form

$$\sigma = \sigma(\mathbf{x}) = \sum_{i\in\mathcal{N}} w_i(1-x_i). \quad (10)$$

Based on (10), the privacy of each worker and the distortion of the aggregated result can be given as

$$\epsilon_i(\mathbf{x}) = \frac{w_i x_i}{\sum_{i\in\mathcal{N}} w_i(1-x_i)}, \quad \forall i \in \mathcal{N} \quad (11)$$

$$\delta(\mathbf{x}) = 3(\sum_{i\in\mathcal{N}} w_i(1-x_i))^2. \quad (12)$$

Eqs. (11) and (12) introduce *externalities* among the workers such that the data privacy of worker i depends on other workers' participations. Specifically, the more participants, the less the distortion but the larger ϵ_i (i.e., worse privacy). Intuitively, as the same sensing task is fulfilled by all the workers, the more participants, the more easily the true data can be figured out (i.e., the more privacy loss). Moreover, we need to carefully choose the workers as they have different

skill levels (i.e., w_i) that may contribute differently to the distortion. Further, the costs of choosing different workers are different. Therefore, it is a challenging task to find a suitable set of workers to fulfill the sensing task.

III. Incentive Mechanism: The Case With Privacy-Passive Workers

In this section, we study the incentive mechanism design for data crowdsensing in the privacy-passive scenario where workers have no intrinsic requirements on their data privacy level. In other words, they would passively participate as long as their privacy loss are compensated by the reward from the platform.

A. Mathematical Formulation

The goal of crowdsensing platform is to minimize the total payment to the workers such that the accuracy of the aggregated result is above certain predetermined threshold (i.e., the distortion is below a threshold Δ). Specifically, this problem can be formulated as

$$\begin{aligned} \text{minimize } & \sum_{i\in\mathcal{N}} p_i \\ \text{subject to } & p_i \geq b_i\epsilon_i(\mathbf{x}), \quad \forall i \in \mathcal{N}, \text{ (Individual rationality)} \\ & \delta(\mathbf{x}) \leq \Delta, \text{ (Accuracy requirement)} \\ & x_i \in \{0,1\}, \quad \forall i \in \mathcal{N}. \end{aligned} \tag{13}$$

In problem (13), the decision variables are $\{x_i\}_{i\in\mathcal{N}}$ and $\{p_i\}_{i\in\mathcal{N}}$, and the individual rationality constraints ensure that each worker can obtain non-negative utility. For the accuracy requirement constraint, the threshold will generally determine the total payment and the data privacy levels of the workers. With a low threshold (i.e., high accuracy), the platform would pay more to the workers to obtain less noisy data (i.e., worse privacy for the workers). Note that different from most works on crowdsensing, problem (13) considers the externalities among workers such that workers' data privacy depends on each other, which has been discussed in Sections II-C and II-D. Due to the externalities, designing an incentive mechanism to solve (13) is a challenging task. Theorem 1 shows that problem (13) is NP-hard.

Theorem 1: The crowdsensing auction problem (13) is NP-hard.

To show Theorem 1, we first establish the equivalence between problem (13) and the following problem:

$$\begin{aligned} \text{minimize } & \sum_{i\in\mathcal{N}} b_i\epsilon_i(\mathbf{x}) \\ \text{subject to } & \sum_{i\in\mathcal{N}} w_i x_i \geq W, \\ & x_i \in \{0,1\}, \quad \forall i \in \mathcal{N}, \end{aligned} \tag{14}$$

where $W = \sum_{i\in\mathcal{N}} w_i - (\Delta/3)^{1/2}$.

Lemma 1: The optimal allocation $\mathbf{x}^*$ for problem (13) is the same as that for problem (14).

Proof: Observe that to minimize (13), p_i is always equal to $b_i\epsilon_i(\mathbf{x}^*)$. Therefore, the inequalities for individual rationality are tight. In other words, minimizing $\sum_{i\in\mathcal{N}} p_i$ is equivalent to minimizing $\sum_{i\in\mathcal{N}} b_i\epsilon_i(\mathbf{x})$. Next, we can rewrite the constraint $\delta(\mathbf{x}) \leq \Delta$ as $\sum_{i\in\mathcal{N}} w_i x_i \geq W$ after some algebra, which concludes the proof. □

Algorithm 2 Differentially Private Data Auction: Winner Determination

Input: worker set $\mathcal{N}$, weight of each worker $w_i, \forall i \in \mathcal{N}$, bid of each worker $b_i, \forall i \in \mathcal{N}$.
Output: winner set $\mathcal{S}$.
Sort the set of workers in the increasing order of $w_i b_i$.
Find the target cost C by solving problem (16).
Let $k = 1$, $x_1 = 1$ and $x_i = 0$, $\forall i = 2, \ldots, N$.
Set $\mathcal{S} = \{1\}$ and compute $C' = b_1\epsilon_1(\mathbf{x})$.
while $C' < C$ **do** \\ Find the set of winners
  $k = k + 1$.
  Set $x_k = 1$ and $\mathcal{S} = \mathcal{S} \cup \{k\}$.
  $C' = \sum_{i=1}^{k} b_i\epsilon_i(\mathbf{x})$.
end while
return $\mathcal{S}$.

Algorithm 3 Differentially Private Data Auction: Payment Determination

Input: worker set $\mathcal{N}$, weight of each worker $w_i, \forall i \in \mathcal{N}$, bid of each worker $b_i, \forall i \in \mathcal{N}$, winner set $\mathcal{S}$.
Output: payments $\mathbf{p}$.
Set $\mathbf{p} = (0, \ldots, 0)$ and $b_c = b_{k+1}$, where k is the worker's index in $\mathcal{S}$ with the largest bid.
for each $i \in \mathcal{S}$ **do** \\ Find the critical bid
  Run Algorithm 2 on $\mathcal{N} \setminus \{i\}$ to get the winner set $\mathcal{S}'$ with k' being the worker's index in $\mathcal{S}'$ with the largest bid.
  $b_c = \min\{b_c, b_{k'+1}\}$.
end for
For each $i \in \mathcal{S}$, $p_i = \frac{b_c w_i}{\sum_{i\in\mathcal{N}\setminus\mathcal{S}} w_i}$.
return $\mathbf{p}$.

It is easy to show that problem (14) is reducible to a reverse binary knapsack problem, which is NP-hard. Based on Lemma 1, Theorem 1 follows.

B. Mechanism Design

From Theorem 1, problem (13) is computationally hard when the cardinality of $\mathcal{N}$ is large. To tackle this challenge, we propose a computationally efficient mechanism (see Algorithms 2 and 3), namely *differentially private data auction (DPDA)*, which is truthful and individually rational and can find the set of winners close to the optimal allocation $\mathbf{x}^*$ for problem (13), as discussed in Section III-C.

In Algorithm 2, the idea is to first find the solution C of the fractional relaxation of problem (14), i.e.,

$$\begin{aligned} \text{minimize } & \sum_{i\in\mathcal{N}} b_i\epsilon_i(\mathbf{x}) \\ \text{subject to } & \sum_{i\in\mathcal{N}} w_i x_i \geq W, \\ & 0 \leq x_i \leq 1, \quad \forall i \in \mathcal{N}, \end{aligned} \tag{15}$$

which is chosen as the target cost. Based on the target cost C, the set of winners can be determined by choosing the smallest set of workers with the total cost greater than or equal to C. Because problem (15) is less constrained than problem (14) and thus C is a lower bound of the solution to problem (14). To find this smallest set of workers, we explore the solution structure of problem (15). Based on the relationship between problem (15) and problem (14), we discover the property of *monotonicity* (see the proof of Theorem 4 in the supplementary material), based on which the set of winners can be found by gradually adding the workers into the winner set until the total cost is greater than or equal to the target cost (see the main loop (line 6-10) in Algorithm 2). Essentially, we want to find the smallest k such that $\sum_{i\leq k} b_i w_i/(\sum_{i\geq k+1} w_i) \geq C$, i.e., $k = \min\{j : \sum_{i\leq j} b_i w_i/(\sum_{i\geq j+1} w_i) \geq C,\ \forall j \in \mathcal{N}\}$, and all the workers with $i \leq k$ are in the winner set. Note that due to the externalities, this monotonicity property is *hidden* in problem (14), which makes our problem more technically challenging than the existing auction works on mobile crowdsensing.

In Algorithm 3, we leverage the critical value approach in Auction theory [26]. The idea is to determine the critical bid b_c such that a worker will not be selected if her bid is larger than or equal to b_c. Specifically, we first remove worker i from the worker set $\mathcal{N}$ and find out the smallest bid by which the worker would lose the auction (line 5 in Algorithm 3). Note that the bids are ordered in the increasing order. The critical bid is determined based on the supremum of all these bids (line 6 in Algorithm 3). Using this critical bid, we determine the payment for each winner based on their weights (line 8 in Algorithm 3). From the analysis of DPDA in Section III-C, we can see that the solution given by Algorithms 2 and 3 is feasible and close to the optimal solution to problem (13).

For the complexity of Algorithm 2, we need to solve C for problem (15), which is a linear fractional program. To efficiently solve C, we can transform problem (15) into a linear program based on the following lemma.

Lemma 2: Problem (15) is equivalent to the following linear program:

$$\begin{aligned} \textit{minimize}\ & \sum_{i\in\mathcal{N}} b_i w_i y_i \\ \textit{subject to}\ & \sum_{i\in\mathcal{N}} w_i y_i \geq W z, \\ & 0 \leq y_i \leq z, \quad \forall i \in \mathcal{N}, \\ & \sum_{i\in\mathcal{N}} w_i z - \sum_{i\in\mathcal{N}} w_i y_i = 1. \end{aligned} \tag{16}$$

Proof: To show the equivalence, we will show that any feasible point in problem (15) is also feasible in problem (16) with the same objective value and vice versa. We note that if $\mathbf{x}$ is feasible in problem (15), then $y_i = \frac{x_i}{\sum_{i\in\mathcal{N}} w_i(1-x_i)},\ \forall i \in \mathcal{N}$ and $z = \frac{1}{\sum_{i\in\mathcal{N}} w_i(1-x_i)}$ are feasible in problem (16), yielding the same objective value $\sum_{i\in\mathcal{N}} b_i w_i y_i = \sum_{i\in\mathcal{N}} b_i \epsilon_i(\mathbf{x})$. It follows that the optimal value of problem (15) is greater than or equal to the optimal value of problem (16). Conversely, note that $z > 0$ in problem (16). If y_i and z are feasible in problem (16), then $x_i = y_i/z$ is feasible in problem (15) with the same objective value $\sum_{i\in\mathcal{N}} b_i \epsilon_i(\mathbf{x}) = \sum_{i\in\mathcal{N}} b_i w_i y_i$. Therefore, the optimal value of problem (15) is less than or equal to the optimal value of problem (16). Therefore, problem (15) is equivalent to problem (16). □

Based on Lemma 2, we can solve C by solving a linear program (16). Note that the computational complexity of Algorithm 2 consists of two parts: solving a linear program (16) (line 3) and finding the set of winners (line 6-10). To solve (16) efficiently, we can use many solvers for linear programs, e.g., CPLEX [33], which can solve the linear program (16) in polynomial time [34]. To find the set of winners, it takes at most $O(N)$ time in the worst case. Therefore, Algorithm 2 can determine the winner set for problem (13) in polynomial time. For Algorithm 3, it needs to run Algorithm 2 for each winner, and the worst case is to run N times, which means that it is also solvable in polynomial time.

C. Analysis of DPDA

In this section, we will prove that DPDA is truthful, individually rational, and α-approximation with respect to the optimal cost.

First, we analyze the truthfulness of DPDA.

Theorem 2: DPDA is truthful.

Proof: To show DPDA is truthful, it is sufficient to show that users cannot improve their utilities by deviating their bids from their true valuations. Note that in DPDA, the winner is determined by the ranking of her bid in the set $\mathcal{N}$ and the higher the ranking, the lower the chance of being selected. Moreover, the critical bid determined by Algorithm 3 does not depend on the value of winners' bids. In what follows, we discuss the cases with an untruthful bid $\tilde{b}_i$ of worker i.

- Overbidding $\tilde{b}_i > v_i$. In this case, the ranking of worker i may move backward. If she could win the auction by truthfully bidding v_i and she remains in the winner set by overbidding, then her utility will remain the same because the critical bid b_c determined by Algorithm 3 will remain the same; if she loses the auction by overbidding, her utility will be zero. If she loses the auction by truthfully bidding, then she will still lose by overbidding. In either case, worker i cannot improve her utility.
- Underbidding $\tilde{b}_i < v_i$. In this case, the ranking of worker i may move forward in the group. If she could win the auction by truthfully bidding v_i, then her utility cannot be improved since she must still remain in the winner set and the critical bid remains the same. If she loses the auction by truthfully bidding but underbidding helps her become a winner, her utility would be $u_i = \frac{(b_c - v_i)w_i}{\sum_{i\in\mathcal{N}\setminus\mathcal{S}} w_i}$. Since she is not originally in the winner set, it means that $v_i \geq b_c$, which leads to her utility $u_i \leq 0$.

Therefore, DPDA is truthful. □

Next, we analyze the individual rationality of DPDA.

Theorem 3: DPDA is individually rational.

Proof: For each worker in the winner set, we have

$$p_i = \frac{b_c w_i}{\sum_{i\in\mathcal{N}\setminus\mathcal{S}} w_i} \geq \frac{b_i w_i}{\sum_{i\in\mathcal{N}\setminus\mathcal{S}} w_i} = c_i,$$

since $b_c \geq b_i,\ \forall i \in \mathcal{S}$. For all workers who lose the auction, $p_i - c_i = 0$. Therefore, we have $p_i - c_i \geq 0$ for all the workers, i.e., DPDA is individually rational. □

Then, we analyze the approximation ratio of DPDA. The idea is to first characterize the optimal solution to problem (15), which, however, is still challenging, due to the externalities. To tackle this challenge, we explore the structure of problem (15) and discover the *hidden monotonicity* property after transforming problem (15) into an equivalent problem. Based on this finding, we show that DPDA satisfies the accuracy requirement of problem (13) and derive the approximation ratio of DPDA by using the relationship between the outputs of DPDA and the optimal solution to problem (13). The results are summarized in the following theorem.

Theorem 4: DPDA satisfies the accuracy requirement (i.e., $\delta(\mathbf{x}) \leq \Delta$) and is α-approximation with respect to the optimal cost, where $\alpha = \frac{(b_k+C)w_k}{C\sum_{i\geq k} w_i - \sum_{i\leq k-1} b_i w_i} \geq 1$.

We next refine the approximation ratio α under the following "small bidders" assumption.

Assumption 1: In a "small bidders" scenario where the workers' bids are generally much smaller than the target payment C computed by solving (16), it is satisfied that $C > \beta b_{max}$ with $b_{max} = \max_i b_i$, and $\beta > N$ being a large positive constant.

Remarks: The definition of the "small bidder" in Assumption 1 is used in the literature (e.g., [35]). Such an assumption is applicable to practical crowdsensing applications, such as consumer surveys, product reviews, and voting events, all of which involve a large number of workers, with the individual cost of each worker much smaller compared to the total budget of the platform.

Corollary 1: Under Assumption 1, DPDA is α'-approximation with respect to the optimal cost, where $\alpha' = \frac{1}{1-k/(\beta+1)} \geq 1$.

The proof of Theorem 4 and Corollary 1 are provided in the supplementary material. This result showcases a concrete scenario that the more significant the "small bidders" effect, the larger the value of β, and therefore the better will the performance of our DPDA algorithm be. The numerical results in Section V (see Fig. 5) further validate the Corollary 1.

IV. Incentive Mechanism: The Case With Privacy-Proactive Workers

In the model above, workers are allowed to inject noise over their sensing data locally to avoid revealing private information to the untrustworthy platform. However, since the noise level is specified by the platform, workers lose control to some extent on determining the exact privacy protection level of her data. In this section, we consider the scenario where privacy-proactive workers possess intrinsic requirements on the data privacy levels assigned by the platform. We first present the problem formulation, and then present an auction mechanism developed based on DPDA, followed by the performance analysis.

A. Problem Formulation

Along the same line as in Section III, we aim to devise a mechanism that minimizes platform's total payment subject to the accuracy requirement for the aggregated result. The incentive mechanism outcome should also satisfy workers' privacy level requirements in addition to other properties including *truthfulness, individual rationality* and *computational efficiency*. We reformulate the optimization problem (13) as follows:

$$\begin{aligned} \text{minimize } & \sum_{i\in\mathcal{N}} p_i \\ \text{subject to } & p_i \geq b_i \epsilon_i(\mathbf{x}), \quad \forall i \in \mathcal{N}, \ \text{(Individual rationality)} \\ & \delta(\mathbf{x}) \leq \Delta, \ \text{(Accuracy requirement)} \\ & \epsilon_i(\mathbf{x}) \leq g_i, \ \forall i \in \mathcal{N}, \ \text{(Privacy level requirement)} \\ & x_i \in \{0,1\}, \quad \forall i \in \mathcal{N}. \end{aligned} \tag{17}$$

In our study, we made the following mild assumptions to ensure the problem (17) is feasible.

Assumption 2: Given the bid g_i and weight w_i of each worker $i \in \mathcal{N}$, the platform determines the accuracy requirement Δ such that the following condition is satisfied,

$$\Delta \geq 3(w_i/g_i)^2, \quad i \in \mathcal{N}. \tag{18}$$

This assumption assumes that the platform's accuracy requirement should not be too small relative to the ratio $\frac{w_i}{g_i}$ of a worker. The rationale behind is that in the privacy-proactive case, we may not be able to achieve as high accuracy of aggregation results as in the privacy-passive case, since we can not recruit workers with high skill levels (who can contribute more in terms of result accuracy), but with too strict privacy level requirements. We have the following lemma.

Lemma 3: The payment minimization problem (17) has a feasible solution under Assumption 2.

Proof: Given the accuracy requirement introduced in (17), we have $\sum_{i\in\mathcal{N}} w_i x_i \geq \sum_{i\in\mathcal{N}} w_i - (\Delta/3)^{1/2}$. Under the privacy constraint introduced in (17), we have $\sum_{i\in\mathcal{N}} w_i x_i \leq \sum_{i\in\mathcal{N}} w_i - \frac{w_i x_i}{g_i}$. When $\frac{g_i}{w_i} \geq \sqrt{3/\Delta}$ holds, it is easy to see that we have,

$$\sum_{i\in\mathcal{N}} w_i - (\Delta/3)^{1/2} \leq \sum_{i\in\mathcal{N}} w_i - \frac{w_i}{g_i} \leq \sum_{i\in\mathcal{N}} w_i - \frac{w_i x_i}{g_i}$$

which concludes the proof. □

B. Mechanism Design

The incentive mechanism developed in Section 3 cannot be directly applied in the privacy-proactive scenario due to the additional constraints of workers' privacy levels in (17). Moreover, as both the unit privacy cost and privacy level requirement are assumed to be private information known only by the worker, each worker needs to submit a two-dimensional bid, of which the value could deviate from the true value due to workers' strategic behaviors. Wet present an auction-based incentive mechanism, namely *enhanced differentially private data auction (EDPDA)*, that addresses the new challenges. Similar to the analysis conducted for the privacy-passive scenario, we first show that the reformulated problem (17) can be reduced to a reverse binary knapsack problem, and then relax the integer variable condition to obtain a solvable linear program based on the following Lemma 4.

Lemma 4: The fractional relaxation of Problem (17) is reducible to the following linear program:

$$\begin{aligned} \textit{minimize} \ & \sum_{i\in\mathcal{N}} b_i w_i y_i \\ \textit{subject to} \ & \sum_{i\in\mathcal{N}} w_i y_i \geq W z, \\ & \sum_{j\in\mathcal{N}} w_j y_j + \frac{w_i}{g_i} y_i \leq \sum_{j\in\mathcal{N}} w_j z, \quad \forall i \in \mathcal{N}, \\ & 0 \leq y_i \leq z, \quad \forall i \in \mathcal{N}, \\ & \sum_{i\in\mathcal{N}} w_i z - \sum_{i\in\mathcal{N}} w_i y_i = 1, \end{aligned} \tag{19}$$

where W is as defined as in (14).

The proof of Lemma 4 is provided in the supplementary material.

The set of winners are determined in a greedy manner as described in Algorithm 4. Specifically, we first filter out a set of k workers whose privacy level requirements are guaranteed to be satisfied (line 4-5). Then we use the solution to problem (19) as a budget target and follow the same procedures that has been used in Algorithm 2 to filter out winners without exhausting the budget target. The payment for each winner $i \in \mathcal{S}$ is computed via Algorithm 3 introduced in Section III.

Algorithm 4 Enhanced Differentially Private Data Auction: Winner Determination

1: **Input:** worker set $\mathcal{N}$, weight of each worker $w_i, \forall i \in \mathcal{N}$, bid of each worker $\{b_i, g_i\}, \forall i \in \mathcal{N}$.
2: **Output:** winner set $\mathcal{S}$.
3: Sort the workers in the increasing order of $\frac{w_i}{g_i}$ $i \in \mathcal{N}$.
4: Find the largest integer k such that $\frac{w_k}{g_k} \leq \sum_{i=k+1}^{N} w_i$, and define the set $\mathcal{S}' = \{1, \ldots, k\}$.
5: Resort the workers in $\mathcal{S}'$ in the increasing order of $b_i w_i,\ i \in \mathcal{S}'$.
6: Find the target cost C by solving problem (17).
7: Let $l = 1$, $x_1 = 1$ and $x_i = 0,\ \forall i = 2, \ldots, N$.
8: Set $\mathcal{S} = \{1\}$ and compute $C' = b_1 \epsilon_1(\mathbf{x})$.
9: **while** $C' < C$ and $l \leq k$ **do** // Find the set of winners
10: $l = l + 1$.
11: Set $x_l = 1$ and $\mathcal{S} = \mathcal{S} \cup \{l\}$.
12: $C' = \sum_{i=1}^{l} b_i \epsilon_i(\mathbf{x})$.
13: **end while**
14: **return** $\mathcal{S}$.

C. Analysis of EDPDA

We next show that EDPDA is truthful, individually rational, and meets the data privacy level requirements of all participated workers.

First, we analyze the truthfulness of EDPDA.

Theorem 5: EDPDA is truthful.

Proof: We show the truthfulness of EDPDA by discussing the untruthful bidding behaviors regarding to the unit privacy cost v_i and the privacy constraint E_i separately. For the unit privacy cost, the proof provided for Theorem 2 is sufficient to show that bid deviation from workers' true cost would not bring utility gain. We here focus on the discussion on the untruthfully bidding of workers' privacy constraint.

- Overbidding $\tilde{g}_i > E_i$. In this case, the ranking of worker i after the execution of line 4 in Algorithm 4 may move backward. If worker i is not within $\mathcal{S}'$ via bidding the true value of E_i, she will still be filtered out after moving backward. If she is within $\mathcal{S}'$ and remains in the set after moving backward, she will be ranked for another time regarding her unit privacy bid v_i, which may affect her final utility. Therefore, in either case, overbidding of E_i would not bring benefits to the worker.
- Underbidding $\tilde{g}_i < E_i$. In this case, the ranking of worker i after the execution of line 4 in Algorithm 4 may move forward in the group. If she is already within $\mathcal{S}'$, moving forward would not affect her utility. If she is not within $\mathcal{S}'$ originally, moving forward may help her win the auction, while her privacy constraint would be violated, which is in contrast to her intend.

In summary, we conclude that EDPDA is truthful. □

The individual rationality of EDPDA can be proved by the same procedure as for Theorem 3 given that the payment for the winners are determined via Algorithm 3. Next, we show that EDPDA guarantees that workers' privacy constraints are satisfied.

Theorem 6: EDPDA guarantees that participated workers' privacy constraints are satisfied.

Proof: In Algorithm 4 (line 3-5), we first filter out a set of workers $\mathcal{S}'$ whose privacy constraints are guaranteed to be satisfied, based on which we further determine the winner set $\mathcal{S} \subseteq \mathcal{S}'$. Notice that the resorting of workers within $\mathcal{S}'$ (line 6) and the following procedure would not violate the privacy constraints. According to (11), for each winner $i \in \mathcal{S}$, we have

$$\epsilon_i = \frac{w_i}{\sum_{j\in\mathcal{N}\setminus\mathcal{S}} w_j} \leq \frac{w_i}{\sum_{j\in\mathcal{N}\setminus\mathcal{S}'} w_j} \leq g_i.$$

Since the truthfulness has been proved, we have $\epsilon_i \leq E_i,\ \forall i \in \mathcal{S}$, i.e., each winner's privacy constraint is satisfied under EDPDA. □

The derivation of approximation ratio of the incentive mechanism becomes much more challenging after we extend the winner selection procedure in order to address workers' intrinsic privacy requirements. We thus leave it to our future work.

V. Performance Evaluation

A. Simulation Setup

In our simulation, we generate workers' bids at random. Specifically, the unit privacy costs are generated uniformly from the interval $[1, 20]$ and the data privacy level requirements are generated uniformly from the interval $[0.01, 0.2]$. The weights of workers are first generated uniformly at random from the interval $[1, 10]$ and then normalized. The number of workers N varies from 100 to 300. The distortion is normalized by some largest distortion $\Delta_{\max}$ such that W is always positive under different distortions. The optimal solutions to the problem (14) and (17) are calculated based on the bisection

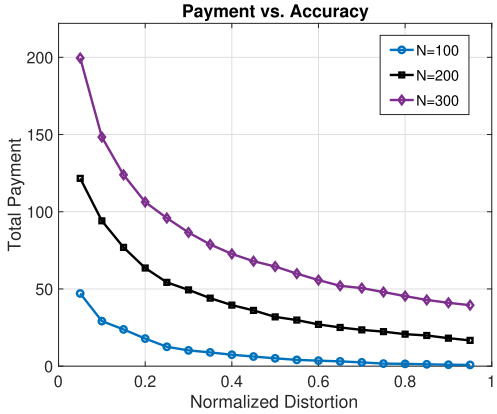


Fig. 2. Payments under different accuracy requirements (privacy-passive case).

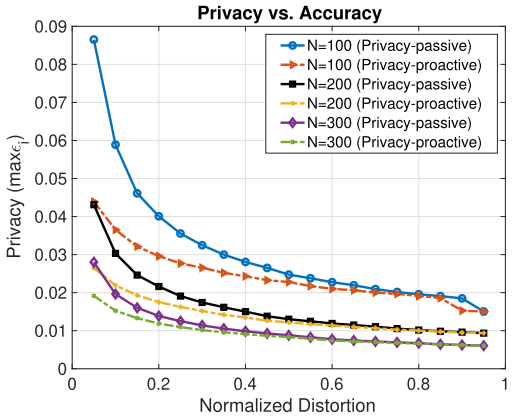


Fig. 3. Relationship between data privacy and the accuracy.

algorithm using the CPLEX optimization solver [33]. To the best of our knowledge, as there are no auction mechanisms for mobile crowdsensing allowing workers to report noisy data while considering the externalities, we examine only the performance of the DPDA algorithm and the EDPDA algorithm we proposed in this paper.

B. Results and Discussions

Payment versus Accuracy. In Fig. 2, we illustrate the payments under different accuracy requirements with different total number of privacy-passive workers. We observe that as the distortion level increases, the total payments decrease, simply because W decreases as Δ increases, i.e., the platform does not need to purchase much privacy from workers. Meanwhile, for the same level of distortion, the total payments increase with the number of workers, because W increases with the number of workers for the same level of distortion based on (14), which requires the platform to select more workers and thereby the total payments increase.

Privacy versus Accuracy. In Fig. 3, we illustrate the relationship between the data privacy and the accuracy. As the privacy of each worker is different, we use the maximum of all the workers' ϵ_i ($\epsilon = \max_{i\in\mathcal{S}} \epsilon_i$) to denote the privacy protection level at the given distortion level. As expected, as the distortion level increases, the data privacy level increases (the smaller ϵ, the higher the privacy protection level), which agrees with our analysis in Section II-D. The results clearly show that privacy-proactive workers in general experience higher privacy level than the privacy-passive workers, which agrees with our expectation since privacy-proactive workers have imposed customized privacy level requirement once enter into the crowdsensing system.

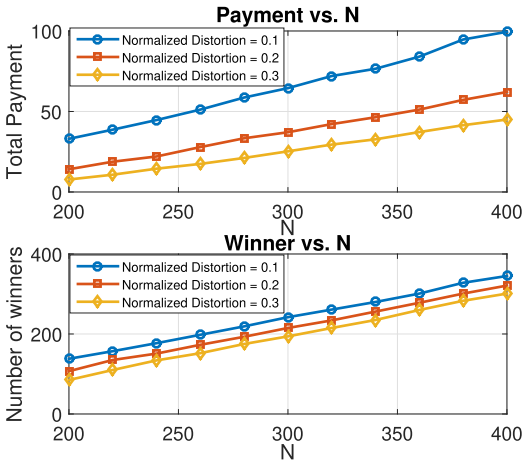
Externalities. Fig. 4 illustrates the effect of externalities. As discussed in Section III-A, the data privacy level of each worker depends on other workers' participations, and when the number of workers changes, it would change workers' privacy levels. As the number of workers increases, the platform needs to hire more workers to maintain the same distortion level. Therefore, we can observe that the increase of total payments and the number of winners as the worker set enlarges. Fig. 4a clearly shows that the higher the distortion level, the lower the total payment and the less the number of winners.

In Fig. 4b, we show the comparison results of the privacy-passive case and privacy-proactive case. We can see that almost the same total payment has to be consumed in the two cases given a fixed normalization distortion with a fixed size of worker set, as W is not influenced by the imposed privacy level requirement g_i of each worker $i \in \mathcal{N}$. Moreover, we observe that the number of winners in the privacy-proactive case is in general less than the number of winners in the privacy-passive case. This is because the additional requirement on privacy level renders a few workers unqualified for being evolved into the private crowdsensing, which leads to the shrink of worker set. Under the effect of externality, the number of winners decreases.

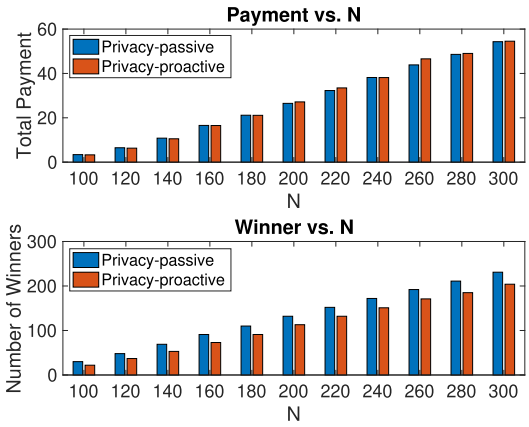
Approximation. In Table I, we illustrate the performance of the proposed DPDA algorithm and EDPDA algorithm respectively by comparing their output total payment with the optimal ones. For each N, we run 100 experiments and in each experiment, we randomly generate the parameters as mentioned in Section V-A. Under different settings, we observe that the total payments generated by these two algorithms are very close to the optimal one and the maximal approximation ratio for each case is around 2.

By comparison, DPDA algorithm outperforms EDPDA in terms of the approximation ratio. This is because that in the privacy-proactive scenario, the platform has to take into account workers' privacy level requirements in addition to the objective of payment minimization. And the winner determination procedure of EDPDA algorithm decouples the two factors by first filtering out workers whose privacy level requirements are satisfied, then selecting out winners that optimize the payment, instead of jointly considering both two factors while choosing the winners.

To further validate the impact of "small bidder" effect, we run the DPDA algorithm under the settings with different values of maximal bid ($b_{max} = 7, 10, 13, 16, 19$) and number of total workers ($N = 200, 300, 400$). As shown in Figure 5, the approximation ratio of the DPDA algorithm is approaching to 1 as b_{max} decreases in all three cases with different value of N, which validates the conclusions in Corollary 1 that the



(a) Privacy-passive case under different accuracy requirements.



(b) Comparison between privacy-passive case and privacy-proactive case (normalized distortion = 0.2).

Fig. 4. Effect of externalities.

TABLE I
APPROXIMATION RATIO OF THE DPDA AND EDPDA (NORMALIZED DISTORTION = 0.2)

(a) DPDA Algorithm

Number of workers N	100	200	300
Average approximation ratio	1.88	1.85	1.85
Minimal approximation ratio	1.45	1.68	1.70
Maximal approximation ratio	2.21	2.23	2.08

(b) EDPDA Algorithm

Number of workers N	100	200	300
Average approximation ratio	1.98	1.89	1.86
Minimal approximation ratio	1.63	1.67	1.70
Maximal approximation ratio	2.75	2.27	2.08

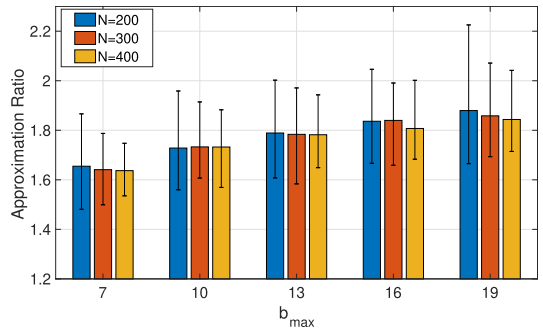


Fig. 5. The impact of the "small bidder" effect on the approximation ratio of the DPDA algorithm.

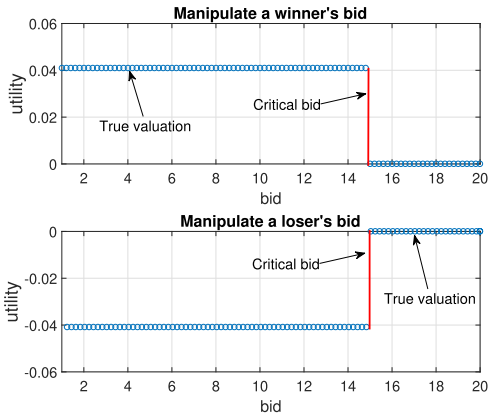


Fig. 6. Truthfulness of the DPDA algorithm.

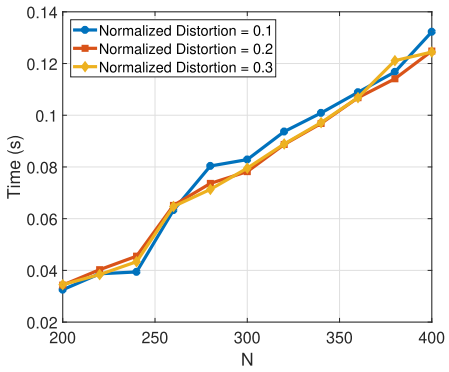


Fig. 7. Computational time of the DPDA algorithm under different settings.

DPDA can achieve a better performance as the "small bidder" effect becomes more significant.

Truthfulness. In Fig. 6, we verify the truthfulness of the proposed DPDA algorithm. We randomly select a winner and a loser in the auction. We fix the bids of the other workers and manipulate the selected worker's bid to evaluate the utility. Fig. 6 illustrates how the utility of the selected worker changes with her bid. As we can see that no matter how the bid changes, a winner or a loser cannot improve her utility and that the best bidding strategy for a worker is to bid truthfully.

Computational Complexity. We next evaluate the computational complexity of the proposed DPDA algorithm. For each N, we examine the average running time of the algorithm

by running 100 experiments, in which the parameters are randomly generated as mentioned in Section V-A. These experiments are run on a PC with a 2.7 GHz Intel Core i7 processor and 16 GB RAM. For the implementation of the DPDA algorithm, we further improve the running efficiency by executing line 4-6 of Algorithm 3 in a parallel manner, so that the payments for the winners can be determined concurrently. In Fig. 7, we can observe that the running time is approximately linear to the network scale, which indicates that the proposed DPDA algorithm can be of high time-efficiency in practice.

VI. Conclusion

We studied privacy-preserving data aggregation for mobile crowdsensing in an auction framework, where the platform plays the role as an auctioneer to recruit workers to complete a sensing task. Under this model, we designed a novel mobile crowdsensing system by leveraging the concept of differential privacy. Specifically, we designed a data aggregation that allows each worker to report a noisy data and can guarantee the use of each worker's data in a differentially private manner. Then, we designed a truthful, individual rational and computationally efficient incentive mechanism that can find a set of workers to approximately minimize the cost of purchasing the private sensing data from workers subject to the accuracy requirement of the aggregated result. We then generalize our results to a privacy-proactive scenario where workers could gain more control of their perceived data privacy protection level by beginning with bidding the lowest acceptable privacy level. We validated the performance of our proposed DPDA algorithm and EDPDA algorithm for the two scenarios through theoretical analysis as well as extensive simulations.

References

[1] L. Yang, M. Zhang, S. He, M. Li, and J. Zhang, "Crowd-empowered privacy-preserving data aggregation for mobile crowdsensing," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jun. 2018, pp. 151–160.

[2] X. Sheng, J. Tang, X. Xiao, and G. Xue, "Sensing as a service: Challenges, solutions and future directions," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3733–3741, Oct. 2013.

[3] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw. (Mobicom)*, 2012, pp. 173–184.

[4] 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. IEEE INFOCOM-IEEE Conf. Comput. Commun.*, Apr. 2014, pp. 1231–1239.

[5] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint," in *Proc. IEEE INFOCOM-IEEE Conf. Comput. Commun.*, Apr. 2014, pp. 1213–1221.

[6] Y. Wen *et al.*, "Quality-driven auction-based incentive mechanism for mobile crowd sensing," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4203–4214, Sep. 2015.

[7] Q. Zhang, Y. Wen, X. Tian, X. Gan, and X. Wang, "Incentivize crowd labeling under budget constraint," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2812–2820.

[8] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2830–2838.

[9] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jun. 2015, pp. 167–176.

[10] H. Jin, L. Su, H. Xiao, and K. Nahrstedt, "INCEPTION: Incentivizing privacy-preserving data aggregation for mobile crowd sensing systems," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, vol. 16, Jul. 2016, pp. 341–350.

[11] M. Zhang, L. Yang, X. Gong, and J. Zhang, "Privacy-preserving crowdsensing: Privacy valuation, network effect, and profit maximization," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[12] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand, "Incentive mechanisms for smartphone collaboration in data acquisition and distributed computing," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1701–1709.

[13] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *Proc. IEEE INFOCOM-IEEE Conf. Comput. Commun.*, Apr. 2014, pp. 745–753.

[14] C. Zhou, Y. Gu, S. He, and Z. Shi, "A robust and efficient algorithm for coprime array adaptive beamforming," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1099–1112, Feb. 2018.

[15] M. H. Cheung, R. Southwell, F. Hou, and J. Huang, "Distributed time-sensitive task selection in mobile crowdsensing," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jun. 2015, pp. 157–166.

[16] S. He, D.-H. Shin, J. Zhang, J. Chen, and P. Lin, "An exchange market approach to mobile crowdsensing: Pricing, task allocation, and Walrasian equilibrium," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 4, pp. 921–934, Apr. 2017.

[17] X. Gong and N. B. Shroff, "Truthful mobile crowdsensing for strategic users with private qualities," in *Proc. 15th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2017, pp. 1–8.

[18] W. Wang, L. Ying, and J. Zhang, "The value of privacy: Strategic data subjects, incentive mechanisms and fundamental limits," in *Proc. ACM SIGMETRICS Int. Conf. Meas. Modeling Comput. Sci.*, 2016, pp. 249–260.

[19] L. Gao, F. Hou, and J. Huang, "Providing long-term participation incentive in participatory sensing," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2803–2811.

[20] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick, "A survey on privacy in mobile participatory sensing applications," *J. Syst. Softw.*, vol. 84, no. 11, pp. 1928–1946, Nov. 2011.

[21] P. Dandekar, N. Fawaz, and S. Ioannidis, "Privacy auctions for recommender systems," *ACM Trans. Econ. Comput.*, vol. 2, no. 3, p. 12, 2014.

[22] A. Ghosh and A. Roth, "Selling privacy at auction," *Games Econ. Behav.*, vol. 91, pp. 334–346, May 2015.

[23] W. Wang, L. Ying, and J. Zhang, "Buying data from privacy-aware individuals: The effect of negative payments," in *Web and Internet Economics*. Berlin, Germany: Springer, 2016, pp. 87–101.

[24] A. Archer and É. Tardos, "Frugal path mechanisms," *ACM Trans. Algorithms*, vol. 3, no. 1, pp. 1–22, Feb. 2007.

[25] J. Hartline and A. Karlin, *Profit Maximization in Mechanism Design*. Cambridge, U.K.: Cambridge Univ. Press, 2007, pp. 331–362.

[26] P. R. Milgrom, *Putting Auction Theory to Work*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[27] C. Dwork, "Differential privacy," in *Automata, Languages and Programming* (Lecture Notes in Computer Science), vol. 4052. Berlin, Germany: Springer, 2006, pp. 1–12.

[28] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand, "Motivating smartphone collaboration in data acquisition and distributed computing," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, pp. 2320–2333, Oct. 2014.

[29] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Jun. 2014, pp. 1187–1198.

[30] C. Meng *et al.*, "Truth discovery on crowd sensing of correlated entities," in *Proc. 13th ACM Conf. Embedded Netw. Sensor Syst.*, Nov. 2015, pp. 169–182.

[31] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations: A Revisit With Applications to Communications, Economics, Engineering, and Finance*. Boston, MA, USA: Birkhäuser, 2012.

[32] X. Wang, W. Cheng, P. Mohapatra, and T. Abdelzaher, "Enabling reputation and trust in privacy-preserving mobile sensing," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2777–2790, Dec. 2014.

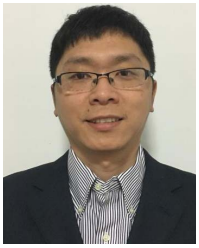
[33] IBM ILOG. *Introducing IBM ILOG CPLEX Optimization Studio V12.5.1*. Accessed: 2020. [Online]. Available: https://www.ibm.com/products/ilog-cplex-optimization-studio

[34] N. Megiddo, "On the complexity of linear programming," in *Proc. 5th World Congr. Adv. Econ. Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1986.

[35] N. Anari, G. Goel, and A. Nikzad, "Budget feasible procurement auctions," *Oper. Res.*, vol. 66, no. 3, pp. 637–652, Jun. 2018.

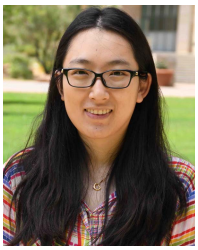
Mengyuan Zhang received the B.S. degree in optical science and engineering from Zhejiang University, Hangzhou, China, in 2011, the M.S. degree from the University of New South Wales, Sydney, NSW, Australia, in 2012, and the Ph.D. degree in control science and engineering from Zhejiang University, in 2020. He is currently working with the Machine Intelligence Technology-Decision Intelligence Laboratory, Alibaba DAMO Academy, Hangzhou.

Lei Yang (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Southeast University, Nanjing, China, in 2005 and 2008, respectively, and the Ph.D. degree from the School of Electrical Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA, in 2012. He was a Post-Doctoral Scholar with Princeton University and an Assistant Research Professor with the School of Electrical Computer and Energy Engineering, Arizona State University. He is currently an Assistant Professor with the Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA. His research interests include big data analytics, AI/ML for cyber-physical systems, edge computing and its applications in the IoT and 5G, data privacy and security in crowdsensing, and optimization and control in mobile social networks. He was a recipient of the Best Paper Award Runner-Up award at IEEE INFOCOM 2014.



Shibo He (Senior Member, IEEE) received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2012. From 2010 to 2011, he was a Visiting Scholar with the University of Waterloo, Waterloo, ON, Canada.

He was an Associate Research Scientist from March 2014 to May 2014, and a Post-Doctoral Scholar from May 2012 to February 2014, with Arizona State University, Tempe, AZ, USA. He is currently a Professor with Zhejiang University. His research interests include wireless sensor networks, crowdsensing, and big data analysis. He coauthored two articles that won the best paper award of IEEE PIRMC 2012 and the IEEE WCNC 2017. He was a recipient of the IEEE Asia-Pacific Outstanding Researcher Award in 2015. He serves on the Editorial Board for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Peer-to-Peer Networking and Application* (Springer), *KSII Transactions on Internet and Information Systems*, and is a Guest Editor for *Computer Communications* (Elsevier), and *International Journal of Distributed Sensor Networks* (Hindawi). He served as the Publicity Chair of the IEEE SECON 2016, the Registration and Finance Chair for ACM MobiHoc 2015, the TPC Co-Chair for IEEE ScalCom in 2014, the TPC Vice Co-Chair for ANT from 2013 to 2014, the Track Co-Chair for the Pervasive Algorithms, Protocols, and Networks of EUSPN in 2013, the Web Co-Chair for IEEE MASS 2013, and the Publicity Co-Chair of IEEE WiSARN in 2010.



Ming Li (Member, IEEE) received the B.E. degree in electrical engineering from Sun Yat-sen University, China, in 2007, the M.E. degree in electrical engineering from the Beijing University of Posts and Communications, China, in 2010, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2014. She is currently an Assistant Professor with the Department of Computer Science and Engineering, The University of Texas at Arlington. Her research interests include mobile computing, the Internet of Things, security, and privacy-preserving computing. Her work won Best Paper awards in GLOBECOM 2015 and DASC 2017, respectively. She received the NSF CAREER Award in 2020. She is a member of the ACM.

Junshan Zhang (Fellow, IEEE) received the Ph.D. degree from the School of ECE, Purdue University, in 2000.

He joined the School of ECEE, Arizona State University, in August 2000, where he has been a Fulton Chair Professor since 2015. His research interests include in the general field of information networks and data science, including communication networks, the Internet of Things (IoT), fog computing, social networks, smart grids. His current research interests include fundamental problems in information networks and data science, including fog computing and its applications in the IoT and 5G, IoT data privacy/security, optimization/control of mobile social networks, cognitive radio networks, stochastic modeling, and control for smart grids.

Prof. Zhang was a recipient of the ONR Young Investigator Award in 2005 and the NSF CAREER Award in 2003. He received the IEEE Wireless Communication Technical Committee Recognition Award in 2016. His articles have won a few awards, including the Kenneth C. Sevcik Outstanding Student Paper Award of ACM SIGMETRICS/IFIP Performance 2016, the Best Paper Runner-up Award of IEEE INFOCOM 2009 and IEEE INFOCOM 2014, and the Best Paper Award at IEEE ICC 2008 and ICC 2017. Building on his research findings, he co-founded Smartiply Inc., a Fog Computing startup company delivering boosted network connectivity and embedded artificial intelligence. He was the TPC co-chair of a number of major conferences in communication networks, including IEEE INFOCOM 2012 and ACM MOBIHOC 2015. He was the General Chair of ACM/IEEE SEC 2017, WiOPT 2016, and IEEE Communication Theory Workshop 2007. He was a Distinguished Lecturer of the IEEE Communications Society. He was an Associate Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Editor of the *Computer Network Journal*, and an Editor *IEEE Wireless Communication Magazine*. He is currently serving as an Editor-at-Large for IEEE/ACM TRANSACTIONS ON NETWORKING and an Editor for *IEEE Network Magazine*.