Privacy-Preserving Crowdsourced Spectrum Sensing

Xiaocong Jin, *Member, IEEE*, and Yanchao Zhang, *Senior Member, IEEE*

***Abstract*— Dynamic spectrum access is promising for mitigating worldwide wireless spectrum shortage. Crowdsourced spectrum sensing (CSS) refers to recruiting ubiquitous mobile users to perform real-time spectrum sensing at specified locations and has great potential in mitigating the drawbacks of current spectrum database operations. Without strong incentives and location privacy protection in place, however, mobile users will be reluctant to act as mobile crowdsourcing workers for spectrum-sensing tasks. In this paper, we first formulate participant selection in CSS systems as a reverse auction problem, in which each participant's true cost for spectrum sensing is closely tied to his current location. Then, we demonstrate how the location privacy of CSS participants can be easily breached under the framework. Finally, we present PriCSS, a novel framework for a CSS service provider to select CSS participants in a differentially privacy-preserving manner. In this framework, we propose PriCSS$^-$ and PriCSS$^+$, two different schemes under distinct design objectives and assumptions. PriCSS$^-$ is an approximately truthful scheme that achieves differential location privacy and an approximate minimum payment, while PriCSS$^+$ is a truthful scheme that achieves differential location privacy and an approximate minimum social cost. The detailed theoretical analysis and simulation studies are performed to demonstrate the efficacy of both schemes.**

***Index Terms*— Crowdsourced spectrum sensing, differential privacy, location privacy, mechanism design.**

I. Introduction

DYNAMIC spectrum access (DSA) is an emerging paradigm for mitigating worldwide wireless spectrum shortage. A DSA system consists of licensed primary users and unlicensed secondary users. A secondary user can use a licensed channel currently not used by its primary user. With DSA in place, secondary users have more channels to use, and primary users can profit by sharing their under-utilized licensed spectrums.

Avoiding harmful interference with primary users is the first principal in DSA systems. FCC advocates a solution based on spectrum databases, each currently administrated by private entities such as Google and Microsoft. Each spectrum database administrator accepts registrations from primary users and leverages a well-known propagation model to predict the coverage boundary of each primary user. Each secondary user needs to inquire the spectrum database about the channel occupancy at a chosen location before transmitting there. Current spectrum databases have well-known drawbacks [2]. First, the signal propagation models in use are not accurate, leading to either severe under-utilization of the spectrum or interference with primary users. Second, current spectrum databases cannot provide the quality information of channels, which can significantly vary in space and time. Last, the locations of primary and secondary users cannot be validated, so a spectrum database administrator may return wrong spectrum occupancy information to secondary users.

Crowdsourced spectrum sensing (CSS) is very promising for mitigating the drawbacks of current spectrum databases. In the CSS approach, a spectrum database administrator recruits distributed mobile users to sense a given channel around a specified location and decides the channel occupancy by aggregating sensing results. With CSS in place, the spectrum database administrator can avoid the prohibitive cost of deploying and maintaining a dedicated large-scale sensor network for spectrum sensing. The feasibility of CSS is backed up by a few trends. First, the number of mobile devices are expected to hit 10 billion in 2019 [3], which implies sufficient geographic coverage especially in populated metropolitan areas where DSA systems are expected to play significant roles. Second, future mobile devices are very likely to be capable of spectrum sensing given the expected pervasiveness of DSA-based wireless systems. For example, practical spectrum sensing is demonstrated in [4] through a cheap off-the-shelf device connected via USB to commodity smartphones. Last, mobile devices are increasingly powerful in self-localization, communication, and computation, which has fostered the explosive popularity of mobile crowdsourcing applications [5].

A typical CSS system works as follows. The spectrum database administrator publishes spectrum-sensing tasks either periodically or randomly. Each spectrum-sensing task involves one or multiple channels, a pre-determined set of geographic locations, and the specified sensing time. The sensing results from designated locations can be aggregated to jointly determine the channel occupancy at the specified time. Each mobile user in the CSS system can independently decide his capability of performing the sensing tasks. Given possibly many CSS participants, the spectrum database administrator can select some for each sensing task.

There are many challenges for pushing the promising CSS system above into practice. For example, strong incentives must be provided to stimulate self-interested mobile users for spectrum sensing. Designing incentive mechanisms for CSS systems is a non-trivial task. On the one hand, different

Manuscript received May 6, 2017; revised December 1, 2017 and February 28, 2018; accepted March 18, 2018; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor X.-Y. Li. Date of publication April 20, 2018; date of current version June 14, 2018. This work was supported by the U.S. National Science Foundation under Grant CNS-1619251, Grant CNS-1514381, Grant CNS-1421999, and Grant CNS-1320906. A preliminary version of the content of this paper was presented in INFOCOM 2016 [1]. *(Corresponding author: Yanchao Zhang.)*

X. Jin was with Arizona State University, Tempe, AZ 85287 USA. He is now with Google LLC, Mountain View, CA 94043 USA (e-mail: xcjin@asu.edu).

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

Digital Object Identifier 10.1109/TNET.2018.2823272

users may want different rewards for the same sensing task. For instance, a user far away from the allocated location may require more to compensate for his longer driving time and higher fuel consumption; a user may also lie about his travel distance to a specific sensing location to gain more. On the other hand, the spectrum database administrator wants to minimize the overall participants' cost (i.e., social cost) for every sensing task as long as the sensing quality is sufficient. Another significant challenge lies in the location privacy of mobile users. Since spectrum-sensing tasks involve rich spatiotemporal information, the whereabouts of CSS participants can be easily exposed, which would severely discourage mobile users wary of their location privacy.

This paper presents PriCSS, a novel framework for a spectrum database administrator to select spectrum-sensing participants in a differentially privacy-preserving manner. Our specific contributions are as follows. First, we formulate participant selection in CSS systems as a reverse auction problem where each participant's true cost for spectrum sensing is closely tied to his current location. Second, we demonstrate a location-privacy attack under the previous formulation to motivate the need for protecting location privacy in CSS systems. Third, we propose PriCSS$^-$, an approximately truthful mechanism based on [6] that can achieve an approximate minimum payment and also differential location privacy. Fourth, we propose PriCSS$^+$, a truthful mechanism that can achieve an approximate minimum social cost and differential location privacy. Last, we thoroughly evaluate both schemes through theoretical and simulation studies.

Our evaluations confirm that PriCSS$^-$ and PriCSS$^+$ can achieve the following common objectives.

- **Differential location privacy**. Both PriCSS$^-$ and PriCSS$^+$ can prevent any internal or external attacker with arbitrary knowledge from inferring the locations of mobile participants.

PriCSS$^-$ has the following additional nice properties.

- **Approximate payment minimization**. The payment here refers to the sum of payments the spectrum database administrator makes to all spectrum-sensing participants to fulfill all the sensing tasks.
- **Approximate truthfulness**. PriCSS$^-$ participants have little incentive to lie about their spectrum-sensing cost.

PriCSS$^+$ achieves the following additional objectives at the cost of higher overhead in comparison with PriCSS$^-$.

- **Approximate social cost minimization**. The social cost is the sum of CSS participants' real costs for completing all spectrum-sensing tasks [7].
- **Truthfulness**. Each PriCSS$^+$ participant has no incentive to lie about his spectrum-sensing cost.

The rest of this paper is organized as follows. Section II briefs the related work. Section III introduces the system and adversary models. Section IV formulates CSS participant selection without considering location privacy. Section V discusses the potential location privacy breach in CSS participant selection. Section VI details the design of PriCSS$^-$ and PriCSS$^+$. Section VII gives theoretical analysis. Section VIII presents performance evaluations. Section IX concludes this paper and points out potential future work.

II. Related Work

A. Privacy and Security in DSA Systems

There are some elegant schemes on location privacy in CSS systems [8]–[11]. The majority of these schemes focus on preventing the spectrum database administrator from inferring physical sensing locations from received sensing reports. Jin *et al.* [8] proposed DPSense to enable privacy-preserving assignment of spectrum-sensing tasks based on differentially private location trajectories. DPSense aims to prevent an honest-but-curious spectrum-sensing service provider from inferring the exact locations for CSS participants. In contrast, the spectrum database administrator is fully trusted in this paper, and our goal is to prevent malicious CSS participants from inferring others' locations according to the selection results. Hence, DPSense cannot be applied in our context.

Some schemes aim to provide location-proof verification or privacy protection for centralized dynamic spectrum access [12]–[15]. PriCSS is based on a crowdsourcing model and has totally different goals from these schemes.

Some other schemes aim to detect false sensing reports [16]–[21] or spectrum misuse [22]–[25]. PriCSS focuses on the pre-sensing phase and is orthogonal to these nice efforts.

B. Privacy in Spatial Crowdsourcing Systems

Some schemes aim to provide privacy protection in the context of general spatial crowdsourcing systems. To *et al.* [26] introduced a framework for protecting the location privacy of workers participating in spatial crowdsourcing tasks. In our context, the sensing locations are pre-determined and publicly known. PriCSS seeks to hide the current locations of sensing participants when competing for spectrum-sensing tasks, so we address a very different problem. Jin *et al.* [6] proposed an approximately truthful scheme for general crowdsourcing task assignment based on a reverse auction model. Motivated and built upon this nice work [6], PriCSS$^-$ ensures approximate truthfulness, an approximate minimum payment, and also differential location privacy in CSS systems.

C. Incentive Mechanism Design and Differential Privacy

Numerous efforts [7], [27]–[29] have been made on incentive mechanism design for crowdsourcing worker selection. Our work differs from this line of work by specifically addressing spectrum sensing and also location privacy.

Differential privacy [30]–[33] has been recently introduced into DSA research. The work in [34] and [35] target differentially private spectrum auctions. In contrast, our work aims at CSS systems, where a reverse auction framework is formulated and participants with the lowest bids are preferred. In addition, we further identify a location-privacy attack where the adversary could identify the victim's possible locations and propose two solutions to address it accordingly.

Our work is also related to differentially private combinatorial optimization [36], in which an elegant solution for the weighted set cover problem is given. The actual set of elements to be covered is private information [36], while all

spectrum-sensing tasks need to be covered are publicly known to everyone in our scenario. Therefore, the solution in [36] is not directly applicable in our specific context.

III. System and Adversary Models

A. System Model

PriCSS is run by a spectrum database administrator whose functionalities, however, go far beyond those of current spectrum database administrators. Specifically, the PriCSS administrator accepts registrations from primary users and answers the spectrum-occupancy queries from secondary users. In addition, the PriCSS administrator can manage the spectrum of itself or other licensed users by issuing spatiotemporal spectrum permits [24], [25] which allow secondary users to use specific channels at designated locations and time.

The PriCSS administrator relies on mobile crowdsourcing to obtain fine-grained information for its managed spectrum. Crowdsourcing spectrum-sensing tasks eliminates the need for the PriCSS administrator to deploy and manage a large-scale sensor network dedicated to spectrum sensing. More specifically, to determine the realtime quality and occupancy of a specific channel in a certain area, the PriCSS administrator recruits mobile users there, referred to as PriCSS participants, to perform spectrum sensing at a set of designated locations. The PriCSS administrator can then make a decision by fusing sensing reports. This sensing method is known as cooperative spectrum sensing and has been widely studied. The sensing locations usually should be far apart from each other to ensure high spatial diversity and thus high sensing quality. For the purpose of this paper, we hereby assume that the PriCSS administrator has pre-determined the sensing locations of each sensing task according to existing methods such as [37].

Each PriCSS participant is a mobile user who owns an advanced mobile device capable of spectrum sensing. He registers with the PriCSS administrator under his real identity to receive rewards for performing spectrum sensing. Each PriCSS participant also has a unique pseudonym or identifier which is visible to other participants in the system.

B. Adversary Model

We assume that the PriCSS administrator is fully trusted in preserving the real identity and bids of PriCSS participants. The adversary can be internal or external to PriCSS. An internal attacker corresponds to a PriCSS participant. We assume that internal attackers are honest-but-curious (HBC) in the sense that they faithfully fulfill promised sensing tasks but have interests in finding out the locations of other PriCSS participants. We also assume that PriCSS participants may lie about their spectrum-sensing costs to claim more rewards, but they are rational and only lie if they can benefit from lies. Such HBC and rational assumptions are commonly adopted in the literature to model the attackers not performing denial-of-service attacks. In contrast, an external attacker does not participate in PriCSS but tries to infer the locations of PriCSS participants from public information.

We assume that the adversary has arbitrary background knowledge for attempting to breach the location privacy. For example, both internal and external attackers know the details of the system operations, and they may also collude. We intend to offer differential location privacy to PriCSS participants under this common adversary model in the security literature.

As mentioned in Section II, there can be many other security and privacy issues in CSS systems. We resort to the rich literature for effective defenses, e.g., detecting fake sensing results [16]–[21] and spectrum misuse [22]–[25].

IV. Participant Selection Without Privacy Protection

We first formulate participant selection in PriCSS as a reverse-auction problem without considering location privacy. For this purpose, we assume that there are totally n PriCSS participants in a large geographic region such as the Los Angeles metropolitan area. Each participant has a unique integer index in $\mathcal{N} = \{1, \ldots, n\}$, which corresponds to his system pseudonym in practice.

We assume that the PriCSS administrator issues K sensing tasks. Each task $k \in [1, K]$ contains one or more channels to sense, a time window in which the sensing should be done, and $\mu_k \geq 1$ sensing locations which are determined by the PriCSS administrator according to existing results such as [37]. Finally, we denote the j-th subtask of task k by $t_{k,j}$, all the μ_k subtasks of task k by $T_k = \{t_{k,j} | j \in [1, \mu_k]\}$, and all the $\sum_{k=1}^{K} \mu_k$ subtasks by $\mathcal{T} = \{t_{k,j} | k \in [1, K], j \in [1, \mu_k]\}$. Each participant can bid for multiple sensing tasks per his schedule and itinerary. Since all subtasks of the same sensing task need to be performed in the same (and generally short) time window, we require that each participant perform at most one subtask for each sensing task.

The cost for spectrum sensing is modeled as follows. The PriCSS administrator publishes a constant factor η to compensate each PriCSS participant for his resource (power, communication, and computation) consumption and human effort incurred for each sensing subtask. Another constant θ is also published as the travel compensation per unit distance for gas consumption, driving time, etc. For simplicity, we use Euclidean distance to approximate travel distance between two points. Assume that a participant chooses to perform m subtasks in a round trip of total Euclidean distance d. His true sensing cost is defined as $v = m\eta + \theta d$. For example, if a participant is currently at position l_1 and wants to perform two subtasks a and b which are located at l_a and l_b, respectively. Then d equals $\mathrm{Euclidean}(l_1, l_a) + \mathrm{Euclidean}(l_a, l_b) + \mathrm{Euclidean}(l_b, l_1)$. Therefore, his true sensing cost for the two subtasks is simply $2\eta + \theta d$. Each participant knows this cost model for computing his sensing cost, and the PriCSS administrator can modify the model based on user feedbacks. Our cost model can use true travel distance instead as long as each participant has a reliable way to estimate it, e.g., via navigation software. In this latter case, the location-inference attack in Section V corresponds to location privacy breach in the worst-case scenario.

The PriCSS administrator aims to select n_k unique participants for each spectrum-sensing task $k \in [1, K]$. Since

PriCSS participants compete to perform spectrum-sensing tasks in return for rewards, it is reasonable to model participant selection in PriCSS under a reverse combinatorial auction framework [38]. In this framework, the PriCSS administrator serves as an auctioneer to auction the sensing tasks, and each participant $i \in [1, n]$ acts as a bidder for the sensing tasks.

We outline the auction procedure as follows. The PriCSS administrator broadcasts the subtask list $\mathcal{T}$ and expects each interested participant i to reply with one bid $b_i = (L_i, c_i)$, where $L_i \subset \mathcal{T}$, and c_i is his claimed cost to perform the sensing subtasks L_i. We assume that c_i is limited in the range of $[c_{\min}, c_{\max}]$, where $c_{\min}$ and $c_{\max}$ are reasonable minimum and maximum possible sensing costs, respectively. Each participant follows two rules to place his bid. First, he can bid for no more than one subtask for each sensing task. Second, he can bid for multiple sensing tasks. The first rule is necessary to prevent strategic manipulation of the bids. For example, participants A and B both bid for the same subtasks $t_{1,1}$ and $t_{1,2}$. If bidding truthfully, A will be allocated with $t_{1,1}$, and B will be allocated $t_{1,2}$. However, A might find out that if he is assigned with $t_{1,2}$, he can gain more rewards. Thus, A could purposely lie about the cost of $t_{1,1}$ to give away the sensing opportunity of $t_{1,1}$ to B. Since B has already been assigned with one subtask for this specific sensing task, B is excluded for consideration of task assignment of $t_{1,2}$. In this way, A purposely lies about one sensing cost to win the other sensing subtask and gains more. Such attacks can be effectively thwarted by the first rule. The second rule is to allow participants to perform multiple spectrum-sensing tasks during a round trip so that the total cost for performing bundled spectrum-sensing tasks can be reduced.

Given the bid set $\mathcal{B} = \{b_i | i \in [1, n]\}$, the administrator determines the auction outcome, denoted by $\overrightarrow{x}(\mathcal{B}) = \{x_1, x_2, \ldots, x_n\}$, where x_i is an indicator for participant i:

$$x_i = \begin{cases} 1, & i \text{ wins the subtask bundle } L_i, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Correspondingly, the administrator selects a winner set $\mathcal{W}$ such that all subtasks in $\mathcal{T}$ can be fulfilled.

Each participant i also holds a true valuation about his performing cost for the claimed subtask set L_i, which is calculated with the aforementioned cost model and denoted by v_i. If his bid b_i is accepted, his utility is defined as $u_i = p_i x_i - v_i$, where p_i is the payment received from the administrator. We normalize u_i to 0 if participant i is not a winner. All participants know the allocation algorithm and payment scheme in advance, and each participant wants to choose his strategy to maximize his own utility. So the claimed cost c_i might not equal v_i for each participant.

The participants could have different performing-cost valuations due to diverse sensing and travel costs involved for the same spectrum-sensing task bundle. Since each participant decides his own bundle to bid for, we aim to design a truthful mechanism so that participants have no incentive in lying about the claimed cost.

Problem Formulation: We formulate participation selection in PriCSS without considering location privacy as follows.

$$\begin{aligned} \textbf{minimize } & \sum_{i \in \mathcal{W}} c_i \\ \textbf{subject to } & |(\bigcup_{i \in \mathcal{W}} L_i) \bigcap T_k| = \mu_k, \quad \forall k \in [1, K], \\ & |L_i \bigcap T_k| \leq 1, \quad \forall k \in [1, K], \ \forall i \in \mathcal{W}. \\ & |L_i| \leq \gamma, \quad \forall i \in \mathcal{W}. \end{aligned} \tag{2}$$

The first condition above indicates that participants in the winner set can fulfill all the K spectrum-sensing tasks. The second one requires that each participant bid at most one subtask for each spectrum-sensing task. The third one is to limit the number of sensing tasks a participant can perform in a single round. γ is a constant and specified by the administrator. The overall objective is to select a subset of participants with the lowest sum of sensing costs under the given constraints. An alternative objective can be finding the lowest sum of total payments by replacing c_i with p_i.

The above problem formulation can be essentially treated as a minimum weighted set cover problem [39], which is knowingly NP-hard. So our basic problem here is also NP-hard, which can be solved by an iterative approximation algorithm as follows. We define the average contributory cost of a participant as his original claimed cost over the subtasks that he bids for and are not yet allocated to other participants. In each iteration, the PriCSS administrator selects a new participant who has the minimum contributory cost among the remaining participants. The algorithm terminates when all the constraints are satisfied. We say that one participant *outbid*s another if the former is chosen earlier than the latter.

V. Your Location Is No Secret

In this section, we exemplify some attacks to infer PriCSS participants' locations when they are selected under the previous reverse auction framework. The location of a participant here refers to his *base location* (e.g., home or workplace) where he may stay for a long time each day, and the base location serves as the reference point for the participant to derive his cost for any interested spectrum-sensing tasks. We assume that each participant starts from his base location and returns there after performing spectrum-sensing tasks.

We also assume that the PriCSS administrator publicizes each spectrum-sensing auction result to ensure the public that its participant selection is unbiased. The publicized information only includes the system identifier of each participant winning one or multiple sensing subtasks. The real identity, claimed cost, and received payment of each winning participant are still kept confidential. Making the auction result public can also help the winners achieve greater self-esteem and public recognition, for which there are numerous examples in practice. For instance, an Amazon user can get his product reviews seen and voted by others, and those contributing highly voted reviews can get free products to test and keep.

The key insight for location-inference attacks is that a participant's claimed sensing cost is tied to his round-trip

Euclidean distance according to the aforementioned public cost model $v = m\eta + \theta d$, which corresponds to performing m subtasks in a round trip of total Euclidean distance d. Even if the claimed cost of each participant is hidden, the attackers can still infer the locations of some participants from the auction results and the changes in auction participation. We give some attack examples in what follows to highlight the need for preserving location privacy. We consider two rounds of auctions, which involve identical channels and sensing locations but different sensing times. This is practical because the PriCSS administrator may want to know the occupancy and quality of each channel in each service area according to a periodic, on-demand, or random schedule.

Case 1 (Single Task): We first consider a simple case in which each participant can bid for a single sensing task. Since each participant can perform no more than one sensing subtask for any sensing task, the bid of each participant is hence for a single subtask.

For example, consider three participants $\{A, B, C\}$ bidding for the same subtask. According to the aforementioned cost model, their true sensing costs are $v_A = \eta + \theta d_A$, $v_B = \eta + \theta d_B$, and $v_C = \eta + \theta d_C$, respectively, where d_A, d_B, and d_C denote their respective round-trip Euclidean distance to the subtask location. Assume that the base locations of A, B, and C do not change. Nor do d_A, d_B, and d_C. In addition, we temporarily assume that the claimed cost of each participant equals his true sensing cost, which can be technically guaranteed later. So we have $c_A = v_A$, $c_B = v_B$, and $c_C = v_C$. Assuming that $d_A > d_B > d_C$, we have $c_A > c_B > c_C$. According to our formulation in Eq. (2), participant C will be selected as the winner in the first round. In the second round (say, next day), assuming that C no longer competes for this subtask for some reason such as work schedule change, so only A and B bid. Then B wins in the second round. The PriCSS administrator publishes the participant selection result in each round.

An external attacker can infer from the public information that $c_A > c_B > c_C$ and hence $d_A > d_B > d_C$, which are something a sensitive user does not want to disclose.

Internal attackers can infer much more information. For example, assume that B is an attacker. Since B knows his own distance d_B and $d_C < d_B$, he can infer that participant C must be inside the *suspicion region*, which is the circle centered at the subtask location with radius d_B. If C additionally participates in other sensing subtasks whose locations are also public, B can draw other suspicion regions for C and infer that C is in the intersection area of the suspicion regions with overwhelming probability. B can also speed up his inference and improve the inference accuracy by colluding with other participants in the PriCSS system.

Case 2 (Multiple Tasks): We also give a more complicated example corresponding to the more general case in Eq. (2), in which each participant can bid for multiple subtasks with a single claimed cost. As shown in Fig. 1, our example involves four sensing tasks $T_1 \sim T_4$, each involving a single subtask. So we can use $T_1 \sim T_4$ to denote the four subtasks as well. The number associated with each dotted line in Fig. 1 represents the Euclidean distance between the two

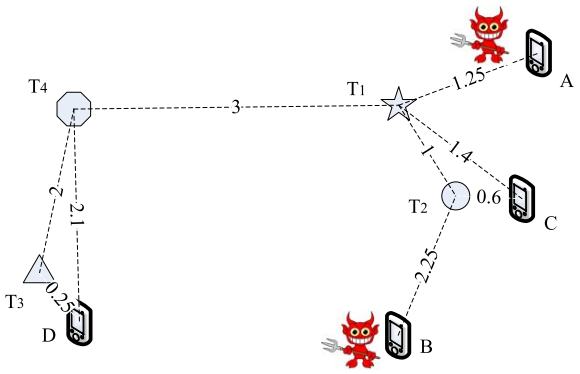


Fig. 1. An example of location-inference attacks.

end locations. Let η be 0.5 and θ be 1 for the aforementioned cost model $v = m\eta + \theta d$, where m denotes the number of chosen subtasks, and d denotes the round-trip Euclidean distance. The bids submitted by $A \sim D$ are as follows: $b_A = \{\{T_1\}, 3\}, b_B = \{\{T_2\}, 5\}, b_C = \{\{T_1, T_2\}, 4\}, b_D = \{\{T_3, T_4\}, 5.35\}$. According to our formulation in Eq. (2), the winner set is $\mathcal{W} = \{C, D\}$. In the second round, assuming that C leaves the area or simply skips the auction, the winner set is $\mathcal{W}' = \{D, A, B\}$. Assume that the PriCSS administrator publishes a re-ordered winner set in each round to conceal each winner's selection order. For example, $\{D, C\}$ and $\{B, D, A\}$ are published as the two rounds' results.

There can be many attack strategies for the above scenario. Due to space limitations, we only discuss one case here, in which A and B collude to infer C's location. The attack involves two steps. First, the attackers need to infer the sensing task bundle that C bids for. Second, the attackers need to estimate the claimed cost of C. The first step can be achieved by studying the difference between the two winner sets, $\mathcal{W}$ and $\mathcal{W}'$. From the attackers' point of view, D's bid must have covered only T_3 and T_4. Otherwise, the winner set would have been changed. It follows that C's bid must have covered at least T_1 and T_2. The remaining question is whether C's bid also covers either or both T_3 and T_4.

There are two possible cases now. In the first case, we assume that D outbids C in the first auction and thus gets T_3 and T_4, so C can only contribute to tasks T_1 and T_2. Since C outbids both A and B, his average contributory cost should be smaller than the smallest of A and B's average contributory cost, which corresponds to $c_C/2 < c_A = 3$ or $c_C < 6$. From Fig. 1, the minimum round-trip cost for C to perform T_2, T_1 and T_4 sequentially must be larger than 6 and is incurred when C first visits the T_2 location, then the T_1 location and the T_4 location, and finally C's location. The additional cost is higher if T_3 is involved. So C's bid covers T_1 and T_2 only. Setting $m = 2$, $\eta = 0.5$, and $\theta = 1$ in the cost model $c_C = m\eta + \theta d_C$, the attackers have $c_C = 1 + d_C$ and thus $d_C < 5$. Since the distance between T_1 and T_2 is 1, the sum of Euclidean distance from C to T_1 and T_2 is smaller than 4. So the attackers can infer that C must be inside the ellipse with T_1 and T_2 locations as two foci and the major-axis length equal to 4. C's location can be further narrowed down if additional information is available.

Case 3 (Other Information): In addition to the two exemplary attacks on location privacy, the participants very close to some subtask locations are likely to have lower claimed costs and higher chances to always win the sensing tasks at those locations, as the aforementioned approximate solution to our formulation in Eq. (2) is a deterministic process. Therefore, if a participant appears much more frequently than other participants in repeated auctions for the same sensing subtasks, the attackers can infer that the participant must be very close to one of the subtask locations. This kind of location privacy breach should also be prevented.

Case 4 (Attacks for Different Formulations): The location privacy breach can also apply to other formulations besides our formulation in Section IV. For example, it is also viable for the administrator to adopt a uniform pricing strategy, i.e., all the winners selected are equally paid with the same payment [6], [34], [40]. In this case, the attack can be even more straightforward. Still consider three participants $\{A, B, C\}$ bidding for three subtasks $\{T_1, T_2, T_3\}$. The bids submitted by $A \sim C$ are as follows: $b_A = \{\{T_1\}, 5\}, b_B = \{\{T_1, T_2\}, 10\}, b_C = \{\{T_2, T_3\}, 9\}$. Based on a heuristic approach we will cover in Section VI-B.1, the administrator will pay 9 for both A and C. However, when C's bid value changes from 9 to 11, the administrator accordingly has to pay 11 for C and one of A and B. In other words, the change of a single bidder's bid value could directly impact the final payment value, thus leading to the breach of his location privacy.

VI. Participant Selection With Differential Location Privacy

Till now we have formulated participant selection in PriCSS as an NP-hard problem and described an approximate solution. We have also demonstrated a few attacks under the basic formulation and solution, which can severely endanger the location privacy of PriCSS participants. In this section, we incorporate differential location privacy into the previous formulation and propose an advanced formulation for participant selection in the PriCSS system to simultaneously achieve approximate social cost minimization, truthfulness, and differential location privacy. In what follows, we first outline some background knowledge to facilitate the presentation and understanding of our scheme. Then we present our advanced formulation with differential location privacy.

A. Background

Definition 1: An auction is truthful if and only if any bidder's (expected) utility of bidding its true valuation v_i is at least its (expected) utility of bidding any other value c_i [41],

$$u_i(v_i, c_{-i}) \geq u_i(c_i, c_{-i}). \tag{3}$$

In the above equation, u_i is the utility of bidder i and c_{-i} is the cost vector for all bidders except i.

Definition 2: Let s_i denote the strategy when player i behaves truthfully. A mechanism is said to be ξ-truthful if for every player i, for any strategy $s_i' \neq s_i$ and any other players' strategy profile s_{-i} [34],

$$\mathbb{E}[u_i(s_i, s_{-i})] \geq \mathbb{E}[u_i(s_i', s_{-i})] - \xi, \tag{4}$$

where $\xi > 0$ is a small constant.

Definition 3: A mechanism satisfies the voluntary participation condition if agents who bid truthfully never incur a net loss, i.e., $\text{profit}_i(v_i, (c_{-i}, v_i)) \geq 0$ for all agents i, true value v_i, and other agents' bids c_{-i} [42].

Clearly, the voluntary participation condition is a desired property of our scheme design.

Theorem 1: A decreasing output function admits a truthful payment scheme satisfying voluntary participation if and only if $\int_0^\infty x_i(c_{-i}, u)du \leq \infty$ for all i, c_{-i}. In this case, we can take the payments to be [42]

$$p_i(c_{-i}, c_i) = c_i x_i(c_{-i}, c_i) + \int_{c_i}^{\infty} x_i(c_{-i}, u)du. \tag{5}$$

Differential privacy is a powerful tool to provide statistical guarantee on the privacy leakage induced by publishing outputs based on sensitive input data sets. The basic idea is that for two almost identical input data sets, the output of the mechanism are nearly identical. The formal definition of differential privacy is as follows [30].

Definition 4: A randomized function $\mathcal{M}$ gives ϵ-differential privacy if for all data sets D_1 and D_2 differing on at most one element, and all $\mathcal{S} \subseteq Range(\mathcal{M})$,

$$\Pr[\mathcal{M}(D_1) \in \mathcal{S}] \leq \exp(\epsilon) \times \Pr[\mathcal{M}(D_2) \in \mathcal{S}]. \tag{6}$$

Approximate differential privacy relaxes on the strict requirement and allows a small additive term in the bound [43].

Definition 5: A randomized function $\mathcal{M}$ gives δ-approximate ϵ-differential privacy if for all data sets D_1 and D_2 differing on at most one element, and all $\mathcal{S} \subseteq Range(\mathcal{M})$,

$$\Pr[\mathcal{M}(D_1) \in \mathcal{S}] \leq \exp(\epsilon) \times \Pr[\mathcal{M}(D_2) \in \mathcal{S}] + \delta. \tag{7}$$

The parameter δ ensures that although not all events can satisfy the strong guarantee as specified by Eq. (6), the alternation is only for very low probability cases. Hence, it is desired that ϵ and δ be as close to 0 as possible.

The exponential mechanism is a powerful tool to facilitate mechanism design via differential privacy [31]. The query function defined as $q(A, r)$ maps a pair of the input data set A and candidate outcome r to a real valued "score," with the understanding that the higher score is, the better performance the mechanism can achieve. Specifically, it is defined as

$$\Pr[\varepsilon_q^\epsilon(A) = r] \propto \exp(\epsilon q(A, r)). \tag{8}$$

The exponential mechanism gives $2\epsilon\Delta$ differential privacy, where Δ is the largest change in q by a single change of the input in A.

The following theorem suggests that the probability of a highly suboptimal output is exponentially low [36].

Theorem 2: The exponential mechanism, when used to select an output $r \in R$, gives $2\epsilon\Delta$-differential privacy, letting

R_{OPT} be the subset of R achieving $q(A,r) = \max_r q(A,r)$, ensures that

$$\Pr[q(A,\varepsilon_q^\epsilon(A)) < \max_r q(A,r) - \frac{\ln(|R|/|R_{\mathsf{OPT}}|)}{\epsilon} - \frac{t}{\epsilon}] \leq \exp(-t). \quad (9)$$

B. Differentially Private Participant Selection

Due to the NP-hardness of the basic problem, we first adapt a basic approximately truthful scheme [6] to achieve the minimum overall payment, low computation complexity, and differential privacy in our framework. Then we approach it from a different angle and propose a truthful scheme that achieves the minimum overall social cost, low computation complexity, and differential location privacy.

1) PriCSS$^-$ Design: This scheme requires a defined set of all possible prices $\mathcal{P}$. The prices in $\mathcal{P}$ should be compatible with the cost model we defined earlier. For our model, $\mathcal{P}$ is defined as $\{\rho | \rho \bmod 10 = 0,\ \rho \in [\rho_{\min}, \rho_{\max}]\}$, where $\rho_{\min}$ ($\rho_{\max}$) is the minimum (maximum) price that can guarantee the fulfillment of all sensing tasks. Hence, all the values in the price set $\mathcal{P}$ can guarantee 100% task fulfillment and are multiples of 10. Accordingly, we round up the cost for each participant to the nearest multiples of 10 and make the rounding number to be each participant's bidding value. To avoid the usage of different letters to denote the bidding value before and after rounding, we simply abuse the notation and denote c_i as the bidding value for participant i. In this solution, we adopt a uniform payment structure such that the administrator pays every winner equally. We consider each price ρ in $\mathcal{P}$ as a possible payment price. To keep the overall payment low, the administrator's objective is to find the winner set of minimum cardinality that satisfies Eq. (10). Let the winner set be $\mathcal{W}_\rho$ for a certain price $\rho \in \mathcal{P}$. Then the overall objective here for the administrator is as follows, which is a little different from what is defined in Eq. (2):

$$\begin{aligned} &\textbf{minimize } \rho|\mathcal{W}_\rho| \\ &\textbf{subject to } |(\bigcup_{i\in\mathcal{W}_\rho} L_i)\bigcap T_k| = \mu_k, \quad \forall k \in [1,K], \\ &\qquad |L_i \bigcap T_k| \leq 1, \quad \forall k \in [1,K],\ \forall i \in \mathcal{W}_\rho, \\ &\qquad |L_i| \leq \gamma, \quad \forall i \in \mathcal{W}_\rho, \\ &\qquad \rho \in \mathcal{P}. \end{aligned} \quad (10)$$

A straightforward solution to the above formulation is as follows. We determine the minimum winner set $\mathcal{W}_\rho$ for each candidate price ρ in $\mathcal{P}$. This sub-problem is clearly a minimum set cover problem too and thus NP-hard. We adopt a heuristic approach to solve this problem. We first initialize the winner set as an empty set. Then we choose a participant who can contribute most to the remaining task set for each round of the selection. If there are multiple participants who can equally contribute, we randomly select one. The selection of winners ends when all the tasks are covered. Since the minimum winner set $\mathcal{W}_\rho$ is available for every ρ, we can easily pick the best ρ with the minimum $\rho|\mathcal{W}_\rho|$.

Algorithm 1 Participant Selection in PriCSS$^-$

Input: Universe set $\mathcal{T}$ of sensing tasks, price set $\mathcal{P}$, ϵ.

Output: Winner set $\mathcal{W}$, final payment ρ^*.

1: **for all** ρ in $\mathcal{P}$ **do**
2: $\mathcal{W}_\rho \leftarrow \varnothing$, $T_\mathcal{W} \leftarrow \varnothing$, $\mathcal{B} = \bigcup_{i\in\mathcal{N}, b_i \leq \rho}\{b_i\}$;
3: **while** $|\mathcal{T} - T_\mathcal{W}| > 0$ **do**
4: $\mathcal{I} = \text{argmax}_{i:b_i\in\mathcal{B}}|(\mathcal{T} - T_\mathcal{W})\bigcap L_i|$
5: Select $i_{\max} \in \mathcal{I}$ with probability $1/|\mathcal{I}|$
6: $\mathcal{B} \leftarrow \mathcal{B} - b_{i_{\max}}$;
7: $\mathcal{W}_\rho \leftarrow \mathcal{W}_\rho \bigcup i_{\max}$;
8: $T_\mathcal{W} \leftarrow T_\mathcal{W} \bigcup L_{i_{\max}}$;
9: **end while**
10: **end for**
11: randomly pick a price ρ^* according to the distribution $\Pr[\rho = x] = \frac{\exp(-\frac{\epsilon x|\mathcal{W}x|}{2c_{\max}\sum_{k\in[1,K]} n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon x|\mathcal{W}y|}{2c_{\max}\sum_{k\in[1,K]} n_k})}$, $\forall x \in \mathcal{P}$
12: $\mathcal{W} = \mathcal{W}_{\rho^*}$
13: **return** $\mathcal{W}$, ρ^*

The above heuristic solution is simple enough but cannot protect the location privacy of participants. We thus proceed to incorporate the exponential mechanism to randomize the price selection result in the above heuristic approach. We first define the quality score as follows:

$$q(\rho) = -\frac{\rho|\mathcal{W}_\rho|}{2c_{\max}\sum_{k\in[1,K]} n_k}. \quad (11)$$

Then the administrator picks a payment price generating a higher total payment with lower probability and vice versa according to the exponential mechanism. In this way, we can simultaneously achieve the approximate minimization of the total payment and differential location privacy. The detailed algorithm is shown in **Algorithm 1**.

Algorithm 1 works as follows. From Line 2 to Line 10, we loop around all the possible ρ values in $\mathcal{P}$ to determine a winner set of approximately minimum cardinality that can fulfill all the sensing tasks. Specifically, in Line 2, we initialize the bids and fulfilled task set. In Line 4, $(\mathcal{T} - T_\mathcal{W})\bigcap L_i$ is used to calculate the pure contribution of each candidate with consideration to all completed sensing tasks. In Line 5, we randomly choose one winner among all the candidates with the most contributions with equal probability. The loop terminates once all the tasks are covered by winners. After determining the winner sets for all possible prices, the administrator finally utilizes the exponential mechanism to pick one final price ρ^* from $\mathcal{P}$ according to the probability defined in Line 11. The algorithm in the end outputs the final payment value ρ^* and the corresponding winner set $\mathcal{W}$.

There can be some variations in how the scheme works in practice. In Line 4, we select the winner with equal possibility among a set of candidates with equal contributions to the tasks. In practice, the selection can be based on different criteria. For example, reputation systems are typically built into crowdsourcing systems [44], in which each user has

a reputation score indicating the quality of his past contributions. If a certain user contributes high quality data, he is likely to gain a good reputation. With such a reputation system, our winner selection can choose the candidate with the highest reputation among the set of candidates contributing equally. In this way, the quality of the sensing results can be consistently high. Other selection criteria can be incorporated similarly.

As we prove in Section VII-A, PriCSS$^-$ achieves only approximate truthfulness and also requires a fixed candidate price set $\mathcal{P}$. So we further propose PriCSS$^+$ to achieve truthfulness at the cost of higher computation overhead.

2) PriCSS$^+$ Design: The objective of the PriCSS$^+$ administrator is still to select a set of participants for bundled spectrum-sensing tasks, and we refer to Section IV for the notation. Different from PriCSS$^-$, PriCSS$^+$ aims to achieve exact truthfulness and to minimize the overall social cost instead of the total payment. We first define a ranking metric to characterize the administrator's preference for participants, which applies to each participant $i \in [1, n]$:

$$r(c_i) = \frac{c_i}{|(\mathcal{T} - T_{\mathcal{W}}) \bigcap L_i|}, \tag{12}$$

where the set $T_{\mathcal{W}}$ denotes the set of subtasks included in the current winning bids, i.e., $T_{\mathcal{W}} = \bigcup_{i \in \mathcal{W}} L_i$.

The rationale of this definition is as follows. The administrator always tends to select the participant with the lowest claimed cost per subtask that has not yet been included in $T_{\mathcal{W}}$. In each iteration, each participant's ranking preference is calculated. Then for any remaining participant i who has not been included in the winner list, we adopt the following quality score for the exponential mechanism,

$$q(c_i, x_i) = -r(c_i). \tag{13}$$

The "$-$" sign is placed to fit the exponential mechanism in our reverse auction model. It is clear that the smaller $r(c_i)$, the higher the quality score of participant i. This effect is preferred during the winner selection.

The details of the proposed allocation scheme are shown in **Algorithm 2**. According to the exponential mechanism, the probability of participant i being selected as a winner is

$$\Pr(x_i = 1) \propto \exp(-\epsilon' r(c_i)) \ , \tag{14}$$

where ϵ' is specified as $\frac{\epsilon}{\Delta \cdot e \ln(e/\delta)}$. Δ is the maximum input difference for c_i, which equals $c_{\max} - c_{\min}$. ϵ and δ are parameters to balance the privacy leakage and efficiency (in terms of social cost minimization in our scenario). Line 11 in **Algorithm 2** can thus be derived by considering all the unselected participants. It essentially normalizes the overall participants' selection probability. Based on the selection probability for each remaining participant, participant i is selected as the winner in this iteration. We then remove his bid b_i from $\mathcal{B}$ and include i in the winner set $\mathcal{W}$.

We resort to Theorem 1 for the truthful payment design. Each winner i is paid by the administrator

Algorithm 2 Participant Selection in PriCSS$^+$

Input: Universe set $\mathcal{T}$ of sensing tasks, set $\mathcal{B} = \bigcup_{i \in \mathcal{N}} \{b_i\}$ of all submitted bids.
Output: Winner set $\mathcal{W}$, social cost c.

1: Initialization: $\epsilon' \leftarrow \frac{\epsilon}{\Delta \cdot e \ln(e/\delta)}$, $\mathcal{W} \leftarrow \varnothing$, $c \leftarrow 0$, $T_{\mathcal{W}} \leftarrow \varnothing$;
2: **while** $|\mathcal{T} - T_{\mathcal{W}}| > 0$ **do**
3: **for all** b_i in $\mathcal{B}$ **do**
4: **if** $L_i \subseteq T_{\mathcal{W}}$ **then**
5: $\mathcal{B} \leftarrow \mathcal{B} - \{b_i\}$;
6: **else**
7: $r(c_i) = \frac{c_i}{|(\mathcal{T} - T_{\mathcal{W}}) \bigcap L_i|}$;
8: **end if**
9: **end for**
10: **for all** b_i in $\mathcal{B}$ **do**
11: $\Pr[\mathcal{W} \leftarrow \mathcal{W} \bigcup \{i\}] = \frac{\exp(-\epsilon' \cdot r(c_i))}{\sum_{b_j \in \mathcal{B}} \exp(-\epsilon' \cdot r(c_j))}$;
12: **end for**
13: Select b_i according to the computed probability distribution.
14: **if** b_i is selected **then**
15: $\mathcal{B} \leftarrow \mathcal{B} - \{b_i\}$;
16: $\mathcal{W} \leftarrow \mathcal{W} \bigcup \{i\}$;
17: $c = c + c_i$;
18: $T_{\mathcal{W}} \leftarrow T_{\mathcal{W}} \bigcup L_i$;
19: **end if**
20: **end while**
21: **return** $\mathcal{W}$, c

with the amount

$$p_i(c_{-i}, c_i) = c_i x_i(c_{-i}, c_i) + \int_{c_i}^{c_{max}} x_i(c_{-i}, u) du \ , \tag{15}$$

where $x_i(c_{-i}, c_i)$ represents the probability that participant i is selected to perform the sensing task bundle L_i when i's claimed cost is c_i and others' claimed cost vector is c_{-i}.

In contrast to PriCSS$^-$, PriCSS$^+$ involves a much more complicated process for payment calculations. When the candidate price range is large, it could be computationally expensive to obtain the payment values for each selected candidate. Such higher computation overhead is the inevitable price for getting exact truthfulness and the minimum total social cost. The PriCSS administrator can select which scheme to use according to its performance goals, computational capacity, the availability of participants, and the decision-time constraint.

VII. Performance Analysis

In this section, we analyze the theoretical performance of both PriCSS$^-$ and PriCSS$^+$.

A. Properties of PriCSS$^-$

1) Differential Location Privacy: Since in our framework, bidding values are directly associated with CSS participant locations, we refer to the differential bidding-value privacy as differential location privacy for convenience.

Theorem 3: PriCSS$^-$ achieves ϵ-differential location privacy.

Proof: In two consecutive auction rounds, assume that there are two bidding sets $\mathcal{B}$ and $\mathcal{B}'$ that differ by one single element.

$$\begin{aligned}
&\frac{\Pr(\mathcal{M}(\mathcal{B})=x)}{\Pr(\mathcal{M}(\mathcal{B}')=x)}\\
&=\frac{\exp(-\frac{\epsilon x|\mathcal{W}_x|}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\exp(-\frac{\epsilon x|\mathcal{W}'_x|}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&\quad\cdot\frac{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y|\mathcal{W}'_y|}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y|\mathcal{W}_y|}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&=\exp(-\frac{\epsilon x(|\mathcal{W}_x|-|\mathcal{W}'_x|)}{2c_{\max}\sum_{k\in[1,K]}n_k})\\
&\quad\cdot\frac{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y|\mathcal{W}'_y|}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y|\mathcal{W}_y|}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&\le\exp(\frac{\epsilon}{2})\cdot\frac{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y(|\mathcal{W}_y|-\sum_{k\in[1,K]}n_k)}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon y|\mathcal{W}_y|}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&\le\exp(\frac{\epsilon}{2})\cdot\exp(\frac{\epsilon}{2})\\
&=\exp(\epsilon). \qquad (16)
\end{aligned}$$

■

The inequalities above hold because by changing a single bidding value, the upper bound of the winner set size change is $\sum_{k\in[1,K]}n_k$ and for any y in $\mathcal{P}$, $y\le c_{\max}$. So we conclude that PriCSS$^-$ achieves ϵ-differential location privacy.

2) Approximate Overall Payment Minimization: Recall that for each price in $\mathcal{P}$, we formulate the problem as a minimum set cover problem. According to [39], the size of the winner set returned by PriCSS$^-$ and the size of the smallest winner set satisfy $|\mathcal{W}_\rho|\le H(\sum_{k\in[1,K]}n_k)|\mathcal{W}_{\mathrm{OPT}}|$ for any $\rho\in\mathcal{P}$, where $H(\sum_{k\in[1,K]}n_k)=1+\frac{1}{2}+\cdots+\frac{1}{\sum_{k\in[1,K]}n_k}$. We utilize this property to prove the following theorem.

Theorem 4: Let $O(x)$ denote the total payment for a price x selected by PriCSS$^-$. The expectation of $O(x)$ is guaranteed to be within a limited bound determined by the optimal (minimum) payment O_{OPT} such that $\mathbb{E}_{x\in\mathcal{P}}[O(x)]\le H(\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}+\ln(e+\frac{\epsilon|\mathcal{P}|H(\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}}{2c_{\min}})\frac{6c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}$.

Proof: Let $O_{\min}$ and $O_{\max}$ be the minimum and maximum total payment generated by our scheme. We also define the following sets: $\mathcal{B}_t=\{x|O(x)>O_{\min}+t\}$, $\overline{\mathcal{B}}_t=\{x|O(x)\le O_{\min}+t\}$ and $\mathcal{B}_{2t}=\{x|O(x)>O_{\min}+2t\}$ for some constant $t>0$. Then we have

$$\begin{aligned}
\Pr(x\in\mathcal{B}_{2t})&\le\frac{\Pr(x\in\mathcal{B}_{2t})}{\Pr(x\in\overline{\mathcal{B}}_t)}\\
&=\frac{\sum_{x\in\mathcal{B}_{2t}}\frac{\exp(-\frac{\epsilon O(x)}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon O(y)}{2c_{\max}\sum_{k\in[1,K]}n_k})}}{\sum_{x\in\overline{\mathcal{B}}_t}\frac{\exp(-\frac{\epsilon O(x)}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{y\in\mathcal{P}}\exp(-\frac{\epsilon O(y)}{2c_{\max}\sum_{k\in[1,K]}n_k})}}\\
&=\frac{\sum_{x\in\mathcal{B}_{2t}}\exp(-\frac{\epsilon O(x)}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\sum_{x\in\overline{\mathcal{B}}_t}\exp(-\frac{\epsilon O(x)}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&\le\frac{|\mathcal{B}_{2t}|}{|\overline{\mathcal{B}}_t|}\times\frac{\exp(-\frac{\epsilon(O_{\min}+2t)}{2c_{\max}\sum_{k\in[1,K]}n_k})}{\exp(-\frac{\epsilon(O_{\min}+t)}{2c_{\max}\sum_{k\in[1,K]}n_k})}\\
&=\frac{|\mathcal{B}_{2t}|}{|\overline{\mathcal{B}}_t|}\exp(-\frac{\epsilon t}{2c_{\max}\sum_{k\in[1,K]}n_k}). \qquad (17)
\end{aligned}$$

So the expectation of $O(x)$ is

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\in\mathcal{P}}[O(x)]\\
&=\sum_{x\in\overline{\mathcal{B}}_{2t}}O(x)\Pr(\rho=x)+\sum_{x\in\mathcal{B}_{2t}}O(x)\Pr(\rho=x)\\
&\le O_{\min}+2t+O_{\max}\frac{|\mathcal{B}_{2t}|}{|\mathcal{B}_t|}\exp(-\frac{\epsilon t}{2c_{\max}\sum_{k\in[1,K]}n_k})\\
&\le O_{\min}+2t+O_{\max}|\mathcal{P}|\exp(-\frac{\epsilon t}{2c_{\max}\sum_{k\in[1,K]}n_k}). \qquad (18)
\end{aligned}$$

For any $t\ge\ln(\frac{O_{\max}|\mathcal{P}|}{t})\frac{2c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}$, we can further simplify the above equation as $\mathbb{E}_{x\in\mathcal{P}}[O(x)]\le O_{\min}+3t$. One of the qualifying values for t is $\ln(e+\frac{\epsilon|\mathcal{P}|O_{\max}}{2c_{\max}\sum_{k\in[1,K]}n_k})\frac{2c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}$. We will prove the validness of t after the proof of the theorem. Hence, based on the chosen value of t,

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\in\mathcal{P}}[O(x)]\le O_{\min}&+\ln(e+\frac{\epsilon|\mathcal{P}|O_{\max}}{2c_{\max}\sum_{k\in[1,K]}n_k})\\
&\times\frac{6c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}. \qquad (19)
\end{aligned}$$

Assume that the optimal (minimum) total payment O_{OPT} is achieved when $\rho=\rho_0$, i.e., $O_{\mathrm{OPT}}=\rho_0|\mathcal{W}_{\mathrm{OPT}}|$. Then

$$\begin{aligned}
O_{\min}\le\rho_0|\mathcal{W}_{\rho_0}|&\le H(\sum_{k\in[1,K]}n_k)\rho_0|\mathcal{W}_{\mathrm{OPT}}|\\
&=H(\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}. \qquad (20)
\end{aligned}$$

Therefore, $O_{\max}\le\frac{c_{\max}}{c_{\min}}\sum_{k\in[1,K]}n_kO_{\min}\le\frac{c_{\max}}{c_{\min}}\sum_{k\in[1,K]}n_kH(\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}$, we can draw the conclusion

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\in\mathcal{P}}[O(x)]\le H(&\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}\\
&+\ln(e+\frac{\epsilon|\mathcal{P}|H(\sum_{k\in[1,K]}n_k)O_{\mathrm{OPT}}}{2c_{\min}})\\
&\times\frac{6\sum_{k\in[1,K]}n_kc_{\max}}{\epsilon}. \qquad (21)
\end{aligned}$$

Finally, we show the validness proof of the t value we selected.

$$\begin{aligned}
&\ln(\frac{O_{\max}|\mathcal{P}|}{t})\frac{2c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}\\
&\le\ln(\frac{O_{\max}|\mathcal{P}|\epsilon}{2c_{\max}\sum_{k\in[1,K]}n_k})\frac{2c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}\\
&\le\ln(e+\frac{\epsilon|\mathcal{P}|O_{\max}}{2c_{\max}\sum_{k\in[1,K]}n_k})\frac{2c_{\max}\sum_{k\in[1,K]}n_k}{\epsilon}=t. \qquad (22)
\end{aligned}$$

■

It is worth noting that the total payment guarantee is dependent on the price set $\mathcal{P}$ according to O_{OPT}'s definition. In other words, if $\rho_{\min}$ can be lowered, O_{OPT} might as well be further lowered. On the other hand, a lower $\rho_{\min}$ might lead to the failure of task fulfillment, so it is critical to select a reasonable and working $\rho_{\min}$ that guarantees 100% task fulfillment and generates a reasonably small O_{OPT}. How to determine the value of $\rho_{\min}$ is left as future work. Generally, we can assume that the administrator can rely on past transaction data as a guideline to configure this value, and it can change the value accordingly based on other factors such as the change of population density.

3) Approximate Truthfulness:

Theorem 5: PriCSS$^-$ is $\epsilon(c_{\max}-c_{\min})$-truthful.

Proof: We use $\mathcal{B}$ and $\mathcal{B}'$ to denote two bid profiles that differ in only one worker's bid. As proved in Theorem 2, we have $\Pr(\mathcal{M}(\mathcal{B})=x)\le\exp(\epsilon)\Pr(\mathcal{M}(\mathcal{B}')=x)$, $\forall x\in\mathcal{P}$. Additionally, when ϵ is small enough, $\exp(-\epsilon)\ge 1-\epsilon$. Hence, the expectation of any worker i's utility is

$$\begin{aligned}\mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B})}[u_i(x)] &= \sum_{x\in\mathcal{P}} u_i(x)\Pr(\mathcal{M}(\mathcal{B})=x)\\ &\ge \sum_{x\in\mathcal{P}} u_i(x)\exp(-\epsilon)\Pr(\mathcal{M}(\mathcal{B}')=x)\\ &= \exp(-\epsilon)\mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B}')}[u_i(x)]\\ &\ge (1-\epsilon)\mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B}')}[u_i(x)]\\ &= \mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B}')}[u_i(x)]-\epsilon\mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B}')}[u_i(x)]\\ &\ge \mathop{\mathbb{E}}_{x\sim\mathcal{M}(\mathcal{B}')}[u_i(x)]-\epsilon(c_{\max}-c_{\min}). \quad (23)\end{aligned}$$

Therefore, PriCSS$^-$ achieves $\epsilon(c_{\max}-c_{\min})$-truthfulness.

So by deviating the bid from his true value, the bidder's max utility gain is no more than $\epsilon(c_{\max}-c_{\min})$, where ϵ is typically small (e.g., 0.1). ■

B. Properties of PriCSS$^+$

1) Differential Location Privacy:

Theorem 6: For any $\delta\le 1/2$, PriCSS$^+$ preserves $((e-1)\epsilon'\Delta\ln(e\delta^{-1}),\delta)$-differential location privacy.

Proof: To facilitate the proof, we first define Q_i as the subtask set that participant i can still contribute to, i.e., $Q_i=(\mathcal{T}-T_{\mathcal{W}})\bigcap L_i$. In two consecutive auction rounds, assume that there are two bidding vectors $\{c_1,c_2,\ldots,c_l,\ldots,c_n\}$ and $\{c_1',c_2',\ldots,c_l',\ldots,c_n'\}$ that differ by only one single element at the lth index. $c_i=c_i'$ for all $i\in[1,n]$ except $i=l$. Differential privacy suggests that with these two bidding vectors as input, the probability that the outputs of the mechanism, i.e., the winner sets $\mathcal{W}$ and $\mathcal{W}'$, are approximately the same. The rationale of our proof is to obtain an exponential upper-bound for $\Pr[\mathcal{W}=\{w_1,w_2,\ldots,w_p\}]/\Pr[\mathcal{W}'=\{w_1,w_2,\ldots,w_p\}]$, where $\mathcal{W}$ and $\mathcal{W}'$ are the two ordered winner lists, i.e., w_i is always selected as a winner before w_j for any $j>i$. We give our formal proof below:

$$\begin{aligned}&\frac{\Pr[\mathcal{W}=\{w_1,w_2,\ldots,w_p\}]}{\Pr[\mathcal{W}'=\{w_1,w_2,\ldots,w_p\}]}\\ &=\prod_{i=1}^{p}\frac{\exp(-\epsilon'\cdot c_i/|Q_i|)/\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j/|Q_j|)}{\exp(-\epsilon'\cdot c_i'/|Q_i|)/\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j'/|Q_j|)}\\ &=\prod_{i=1}^{p}\frac{\exp(-\epsilon'\cdot c_i/|Q_i|)}{\exp(-\epsilon'\cdot c_i'/|Q_i|)}\cdot\prod_{i=1}^{p}\frac{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j'/|Q_j|)}{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j/|Q_j|)}\\ &=\exp(\epsilon'\frac{c_l'-c_l}{|Q_l|})\prod_{i=1}^{p}\frac{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j'/|Q_j|)}{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j/|Q_j|)}, \quad (24)\end{aligned}$$

where $\pi_1=\emptyset$ and $\pi_i=\{w_1,w_2,\ldots,w_{i-1}\}(i>1)$. If $c_l<c_l'$, the second term is smaller than 1. Then

$$\frac{\Pr[\mathcal{W}=\{w_1,w_2,\ldots,w_p\}]}{\Pr[\mathcal{W}'=\{w_1,w_2,\ldots,w_p\}]}<\exp(\epsilon'\Delta), \quad (25)$$

where Δ is the maximum difference of the bid values for the same set of task bundles.

If $c_l>c_l'$, the first term is smaller than 1. We denote $\alpha_j=c_j-c_j'$, then

$$\begin{aligned}&\frac{\Pr[\mathcal{W}=\{w_1,w_2,\ldots,w_p\}]}{\Pr[\mathcal{W}'=\{w_1,w_2,\ldots,w_p\}]}\\ &<\prod_{i=1}^{p}\frac{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j'/|Q_j|)}{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j/|Q_j|)}.\\ &=\prod_{i=1}^{p}\frac{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot c_j'/|Q_j|)}{\sum_{j\in\mathcal{N}\setminus\pi_i}\exp(-\epsilon'\cdot\alpha_j/|Q_j|)\exp(-\epsilon'\cdot c_j'/|Q_j|)}\\ &=\prod_{i=1}^{p}\mathbb{E}_{j\in\mathcal{N}\setminus\pi_i}[\exp(\epsilon'\cdot\alpha_j/|Q_j|)]. \quad (26)\end{aligned}$$

Note that for all $\eta\le 1$, $e^{\eta}\le 1+(e-1)\eta$. Therefore, for all $\epsilon'\le 1/\Delta$,

$$\begin{aligned}&\frac{\Pr[\mathcal{W}=\{w_1,w_2,\ldots,w_p\}]}{\Pr[\mathcal{W}'=\{w_1,w_2,\ldots,w_p\}]}\\ &\le\prod_{i=1}^{p}\mathbb{E}_{j\in\mathcal{N}\setminus\pi_i}(1+(e-1)\cdot\epsilon'\cdot\alpha_j)\\ &\le\exp((e-1)\epsilon'\sum_{i=1}^{p}\mathbb{E}_{j\in\mathcal{N}\setminus\pi_i}\alpha_j). \quad (27)\end{aligned}$$

So if $\sum_{i=1}^{p}\mathbb{E}_{j\in\mathcal{N}\setminus\pi_i}\alpha_j$ is upper-bounded, the theorem is established. Based on the proofs in [36], we have $\Pr(\sum_{i=1}^{p}\mathbb{E}_{j\in\mathcal{N}\setminus\pi_i}\alpha_j>\Delta\ln(e\delta^{-1}))\le\delta$. ■

2) Approximate Social Cost Minimization:

Theorem 7: With probability of at least $1-1/n^{\mathcal{O}(1)}$, PriCSS$^+$ can assign spectrum-sensing tasks to a set of winners with a social cost of at most $\gamma\mathsf{OPT}+\mathcal{O}(\ln(n))$, where OPT denotes the optimal (minimum) social cost, and n is the number of participants.

Proof: Let $\mathcal{W}_{\mathsf{OPT}}$ denote the set of winners in the auction with the minimum social cost. We denote an arbitrary set of winners as $\mathcal{W}$ and number the winners according to the order of being selected, i.e., $\mathcal{W}=\{w_1,w_2,\ldots,w_l\}$.

For each $i\in\mathcal{W}$, we define a set $\mathcal{W}_i$, with the following constraints ($\forall j\in\mathcal{W}_i$):

1) $j\in\mathcal{W}_{\mathsf{OPT}}$;
2) $Q_j\bigcap Q_i\neq\varnothing$;
3) $|Q_j-(Q_j\bigcap Q_i)|=0$;
4) $Q_j\neq\varnothing$ before i is selected as one winner.

The above constraints suggest that in this arbitrary selection $\mathcal{W}$, the reason that a participant j is not listed is that there is a participant i with a conflicting task set with that of participant j, and i wins. Note that in Eq. (9), the q function

corresponds to the inverse and unified cost in our scenario. Therefore, by taking $t = \mathcal{O}(\ln(n))$, we have

$$-\frac{c_i}{|Q_i|} \geq -\frac{c_j}{|Q_j|} - \mathcal{O}(\ln n) \tag{28}$$

with a probability of at least $1 - 1/n^{\mathcal{O}(1)}$.

Since $|Q_j|$ is upper bounded by a constant γ where $\gamma < n$ when n is large, we have

$$c_j \geq \frac{c_i}{|Q_i|} \cdot |Q_j| - \mathcal{O}(\ln n) \tag{29}$$

with a probability of at least $1 - 1/n^{\mathcal{O}(1)}$.

Summing all j ($j \in \mathcal{W}_i$) together, we have

$$\begin{aligned} \sum_{j \in \mathcal{W}_i} c_j &\geq (\frac{c_i}{|Q_i|} - \mathcal{O}(\ln n)) \cdot \sum_{j \in \mathcal{W}_i} c|Q_j| \\ &\geq \frac{c_i}{\gamma} - \mathcal{O}(\ln n). \end{aligned} \tag{30}$$

with a probability of at least $1 - 1/n^{\mathcal{O}(1)}$. The last step holds because $\sum_{j \in \mathcal{W}_i} |Q_j| \geq 1$.

Summing all $i \in \mathcal{W}$, we have

$$\begin{aligned} \sum_{j \in \mathcal{W}_{\mathsf{OPT}}} c_j &= \sum_{i \in \mathcal{W}} (\sum_{j \in \mathcal{W}_i} c_j + \sum_{j \in \mathcal{W}_{\mathsf{OPT}} \bigcap \mathcal{W}_i} c_j) \\ &\geq \sum_{i \in \mathcal{W}} \frac{c_i}{\gamma} - \mathcal{O}(\ln n). \end{aligned} \tag{31}$$

This concludes the proof. ■

3) Truthfulness: We finally prove that PriCSS$^+$ is truthful. Based on Theorem 1, we need to show that the selection of PriCSS$^+$ is monotone decreasing with an appropriate payment scheme.

Lemma 8: In PriCSS$^+$, for each participant i, the probability that i is assigned with the interested spectrum-sensing task bundle is monotone decreasing with his claimed cost c_i.

Proof: Due to the randomized property of our scheme, we simply prove that the probability that i is assigned with the interested spectrum-sensing task bundle is decreasing when his claimed cost c_i increases in each round of winner selection.

$$\begin{aligned} &\Pr(\mathcal{W} \leftarrow \mathcal{W} \bigcup \{i\}) \\ &= \frac{\exp(-\epsilon' \cdot r(c_i))}{\sum_{b_j \in \mathcal{B}} \exp(-\epsilon' \cdot r(c_j))} \\ &= \frac{\exp(-\epsilon' \cdot r(c_i))}{\sum_{b_j \in \mathcal{B} \setminus \{c_i\}} \exp(-\epsilon' \cdot r(c_j)) + \exp(-\epsilon' \cdot r(c_i))} \\ &= 1 - \frac{\sum_{c_j \in \mathcal{B} \setminus \{c_i\}} \exp(-\epsilon' \cdot r(c_j))}{\sum_{b_j \in \mathcal{B} \setminus \{c_i\}} \exp(-\epsilon' \cdot r(c_j)) + \exp(-\epsilon' \cdot r(c_i))} \end{aligned} \tag{32}$$

In the above equation, if we increase c_i, $r(c_i)$ also increases. Then the exponential term of c_i decreases, causing the overall equation value to decrease. This indicates that if we increase c_i, the probability that $\mathcal{W}$ includes i in every round decreases if i has not been included in previous rounds. ■

We thus have the following theorem established.

Theorem 9: PriCSS$^+$ is truthful.

VIII. Performance Evaluation

In this section, we conduct simulations to evaluate PriCSS$^-$ and PriCSS$^+$. Since they have different design objectives, we evaluate them individually. Specifically, PriCSS$^-$ is evaluated with regard to the total payment and location-privacy leakage, and PriCSS$^+$ is evaluated with regard to the total social cost and location-privacy leakage.

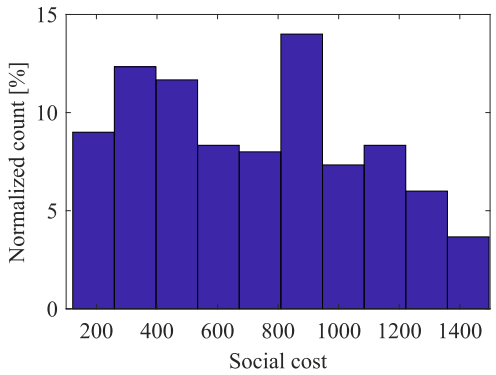


Fig. 2. Social cost distribution for a randomly generated topology with 300 participants.

Our simulation setting is as follows. We simulate a square urban area of 1km by 1km. The system administrator issues sensing tasks in response to the queries of secondary users, each with a transmission radius of 300m. The base locations of participants are uniformly distributed, and we vary the number of participants from 300 to 1000 in simulations. In our simulation, the preferred sensing locations are chosen beforehand according to the specific diversity requirement as discussed in Section III. To minimize the overall sensing cost, we want the subtask locations to be as far from each other as possible. We specify a minimum separation distance of 100m for the subtasks in each sensing task. The number of sensing locations (or subtasks) for each sensing task is fixed as 5 in our simulations. We vary the number of sensing tasks K in one round of auction from 3 to 9. Each sensing task is characterized by the locations of the corresponding secondary users, which are uniformly distributed within the region. We also set the modeling parameters $\eta = 100$ reward units and $\theta = 1$ unit per meter. The parameter γ is specified as 3. In addition, we set the bidding cost range $[c_{\min}, c_{\max}]$ to [100,1500]. For PriCSS$^-$, we set the price set as $\mathcal{P} = \{\rho | \rho \bmod 10 = 0,\ \rho \in [900, 1500]\}$. Note that other configurations of η and θ lead to similar performance. We omit other cases here due to limited space. The privacy parameter ϵ is chosen as 0.1 or 2 unless otherwise stated, and δ is set to 0.25. The simulations are done in MATLAB, and each result represents the average of 200 runs. Fig. 2 shows the social cost distribution for a randomly generated topology with 300 participants. The social cost for each participant is associated with the task bundle he is interested in. We can clearly see that the cost is not uniformly distributed across the range, which is different from the simulations conducted in [6]. In addition, the total normalized count does not add up to 1 for the range we show here due to the upper bound of cost (1500) we have imposed in our system. In other words, a small portion of participants are filtered out due to their high social cost.

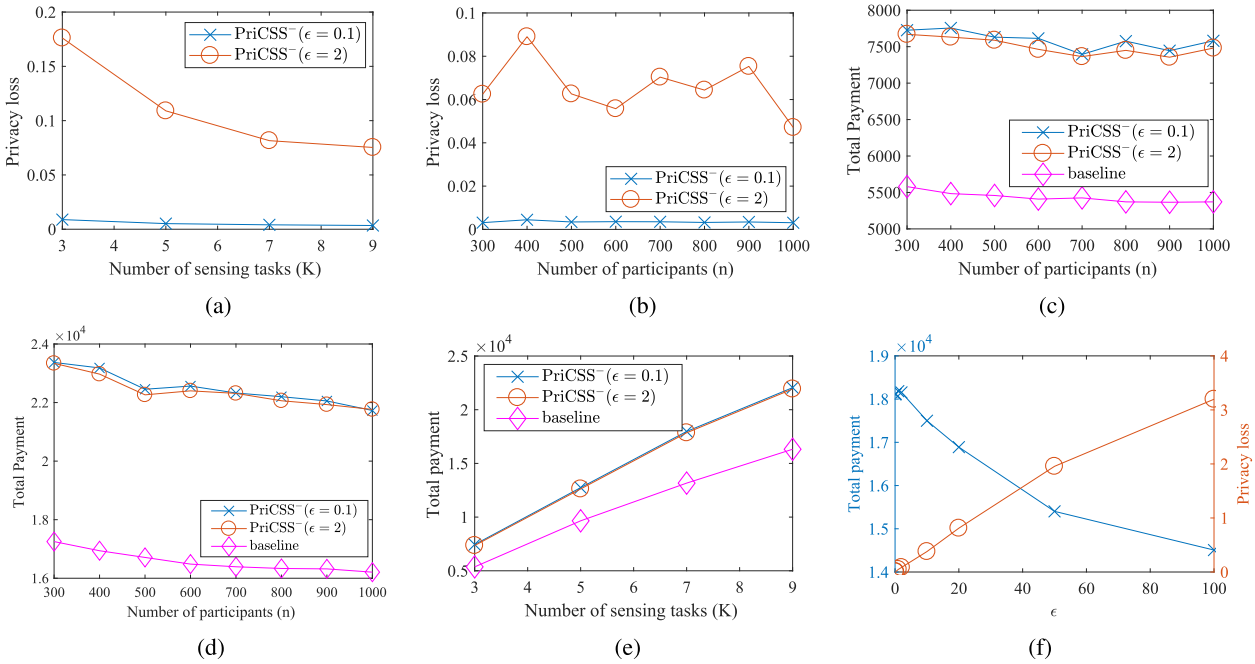


Fig. 3. Performance evaluation for PriCSS$^-$. (a) Privacy loss with 900 participants. (b) Privacy loss with 9 sensing tasks. (c) Total payment with 3 sensing tasks. (d) Total payment with 9 sensing tasks. (e) Total payment with 900 participants. (f) Trade-off between privacy loss and total payment.

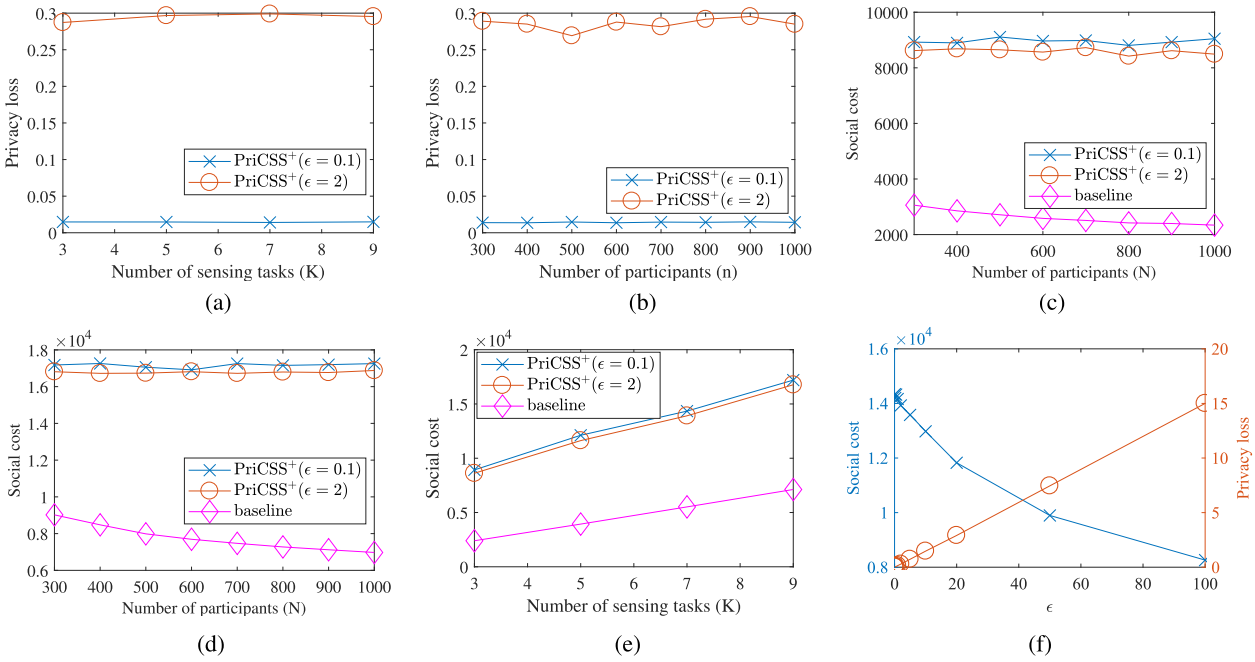


Fig. 4. Performance evaluation for PriCSS$^+$. (a) Privacy loss with 900 participants. (b) Privacy loss with 9 sensing tasks. (c) Total payment with 3 sensing tasks. (d) Total payment with 9 sensing tasks. (e) Total payment with 900 participants. (f) Trade-off between privacy loss and total cost.

We use two performance metrics. The first is the location-privacy loss, defined according to **Definition** 4:

$$\epsilon = \max_{\mathcal{S}} \ln \frac{\Pr[\mathcal{M}(D_1) \in \mathcal{S}]}{\Pr[\mathcal{M}(D_2) \in \mathcal{S}]}, \tag{33}$$

where D_1 and D_2 correspond to two cost vectors for all the participants that differ by one element. Intuitively, the smaller ϵ, the less impact the change of a single cost on the auction results, the better individual sensing-cost privacy is protected, and the more location privacy each participant enjoys. The second metric is the total payment for PriCSS$^-$ or the social (or true sensing) cost of the winners for PriCSS$^+$, which is desired to be as low as possible. For the purpose of comparison, we also show the total payment (or social cost) induced using the approximation algorithm without privacy considerations introduced in Section VI-B.1 for PriCSS$^-$ (or Section IV for PriCSS$^+$) and use the label "baseline" in the figures.

A. Evaluations of PriCSS$^-$

We first evaluate the location-privacy loss in PriCSS$^-$. As we know from Section VII-A, PriCSS$^-$ preserves ϵ-differential location privacy. Fig. 3a shows the privacy loss with 900 participants by varying the number of sensing tasks. Clearly, when ϵ is smaller, the privacy loss is also smaller. Also, the achieved privacy loss for different ϵ values is obviously well under the guarantee. Moreover, the more sensing tasks, the better privacy protection PriCSS$^-$ can deliver. We conjecture that a larger number of sensing tasks lead to a larger range of social cost value and hence better privacy protection under PriCSS$^-$. Fig. 3b shows the privacy loss with nine sensing tasks by varying the number of participants. Although we do not observe a clear trend in this figure, we can see that the achieved privacy loss is still well under the theoretical guarantee.

We then evaluate the total payment in PriCSS$^-$. Fig. 3c and Fig. 3d show the total payment with three and nine sensing tasks, respectively. In general, the total payment decreases with the increase of participants because the more participants we have, the better choice in terms of minimizing the total payment the system can generally make. Also, the total payment declines with the increase of ϵ. The baseline algorithm achieves the lowest overall payment because it has no privacy protection in place. Fig. 3e shows the total payment with 900 participants in PriCSS$^-$. Clearly, the total payment increases almost linearly with K, the number of sensing tasks.

Lastly, we show in Fig. 3f the trade-off between the privacy loss and the total payment in PriCSS$^-$. We see that with the increase of ϵ, the privacy loss increases, while the total payment decreases.

B. Evaluations of PriCSS$^+$

We first evaluate the location-privacy loss in PriCSS$^+$. As proved in Section VII, PriCSS$^+$ preserves $((e-1)\epsilon'\Delta\ln(e\delta^{-1}), \delta)$-differential location privacy, where ϵ' is specified as $\frac{\epsilon}{\Delta\cdot e\ln(e/\delta)}$. This is equivalent to achieving $(\frac{e-1}{e}\epsilon, \delta)$ differential privacy. Fig. 4a and Fig. 4b shows the achievable privacy loss in PriCSS$^+$, which is obviously much lower than the theoretical result. Specifically, when $\epsilon = 0.1$, we can observe almost a constant privacy loss of 0.01, which is far lower than the theoretical value $\frac{e-1}{10e} \approx 0.06$. Similar conclusions can be drawn with $\epsilon = 2$. This indicates that when there is any change of a single cost value for any participant, there is rarely any chance that the auction result can change. Hence, we can safely conclude that the attackers can no longer infer the participants' locations by performing the attacks in Section V or by adopting other attack strategies.

We show the social cost incurred using PriCSS$^+$ and the baseline algorithm for three and nine sensing tasks in Fig. 4c and Fig. 4d, respectively. For the baseline algorithm, we observe that as the number of participants increases, the social cost tends to decrease due to increased competition among participants. The trend of decrease, however, cannot be found with PriCSS$^+$ for both $\epsilon = 0.1$ and $\epsilon = 2$ cases. We conjecture that with PriCSS$^+$ in place, the advantage of cost-efficient participants who claim lower sensing costs in the hope of winning is weakened by the increased number of participants. In other words, their ranking metrics play less significant roles when the number of participants increases. Still, we see that the social cost when $\epsilon = 0.1$ is slightly worse than that when $\epsilon = 2$. This is the expected trade-off between privacy and utility: the larger ϵ, the heavier weight on the ranking metric, and the lower the social cost. In Fig. 4e, we also show the social cost for different numbers of sensing tasks when there are 900 participants. As expected, when the number of sensing tasks increases, the social cost also increases.

Finally, we show the trade-off between the privacy loss and the total social cost in Fig. 4f. Similar to PriCSS$^-$, we can observe the same trend with the interactions among the three values: privacy loss, social cost and ϵ. The social cost can drop dramatically if ϵ increases.

IX. Conclusion

In this paper, we presented PriCSS, a novel framework for a spectrum database administrator to select spectrum-sensing participants in a differentially privacy-preserving manner. In this framework, we proposed PriCSS$^-$ and PriCSS$^+$, two different schemes under distinct design objectives and assumptions. PriCSS$^-$ is an approximately truthful scheme that can achieve both differential location privacy and an approximate minimum payment, while PriCSS$^+$ is a truthful scheme that can achieve both differential location privacy and an approximate minimum social cost. Detailed theoretical analysis and simulation studies demonstrated the efficacy of both schemes.

There are several possible directions that we would like to explore in our future work. First, since people are becoming more and more concerned with location privacy, it is preferred that we design a mechanism that protects location privacy from the administrator as well. Note that although reference [8] provides one solution, it might need accurate probabilistic modeling in practice. Also, how to minimize the negative performance impact introduced by privacy protection mechanisms is challenging. Second, it is also more practical if we can incorporate a reputation system. If such a reputation mechanism exists, how to incorporate reputations into participant selection can be an interesting problem to solve. Lastly, since on-time sensing is critical for spectrum sensing, how to incorporate the timing factor into sensing task allocation deserves thorough investigations.

Acknowledgment

This work was mainly done when X. Jin was a Ph.D. student at Arizona State University.

References

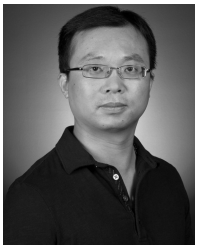
[1] X. Jin and Y. Zhang, "Privacy-preserving crowdsourced spectrum sensing," in *Proc. INFOCOM*, Apr. 2016, pp. 1–6.

[2] T. Zhang, N. Leng, and S. Banerjee, "A vehicle-based measurement framework for enhancing whitespace spectrum databases," in *Proc. MobiCom*, Sep. 2014, pp. 17–28.

[3] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper*, Cisco, San Jose, CA, USA, Mar. 2017.

[4] A. Nika, Z. Zhang, X. Zhou, B. Zhao, and H. Zheng, "Towards commoditized real-time spectrum monitoring," in *Proc. HotWireless*, Sep. 2014, pp. 25–30.
[5] J. Sun, R. Zhang, X. Jin, and Y. Zhang, "SecureFind: Secure and privacy-preserving object finding via mobile crowdsourcing," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1716–1728, Mar. 2016.
[6] H. Jin, L. Su, B. Ding, K. Nahrstedt, and N. Borisov, "Enabling privacy-preserving incentives for mobile crowd sensing systems," in *Proc. ICDCS*, Jun. 2016, pp. 344–353.
[7] 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*, Apr. 2014, pp. 1231–1239.
[8] X. Jin, R. Zhang, Y. Chen, T. Li, and Y. Zhang, "DPSense: Differentially private crowdsourced spectrum sensing," in *Proc. CCS*, Oct. 2016, pp. 296–307.
[9] S. Li *et al.*, "Location privacy preservation in collaborative spectrum sensing," in *Proc. INFOCOM*, Mar. 2012, pp. 729–737.
[10] Z. Gao, H. Zhu, S. Li, S. Du, and X. Li, "Security and privacy of collaborative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 106–112, Dec. 2012.
[11] W. Wang and Q. Zhang, "Privacy-preserving collaborative spectrum sensing with multiple service providers," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1011–1019, Feb. 2015.
[12] Y. Li, L. Zhou, H. Zhu, and L. Sun, "Privacy-preserving location proof for securing large-scale database-driven cognitive radio networks," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 563–571, Aug. 2016.
[13] Z. Gao, H. Zhu, Y. Liu, M. Li, and Z. Cao, "Location privacy in database-driven cognitive radio networks: Attacks and countermeasures," in *Proc. INFOCOM*, Apr. 2013, pp. 2751–2759.
[14] M. Clark and K. Psounis, "Can the privacy of primary networks in shared spectrum be protected?" in *Proc. INFOCOM*, Apr. 2016, pp. 1–6.
[15] Y. Dou *et al.*, "P^2-SAS: Privacy-preserving centralized dynamic spectrum access system," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 173–187, Jan. 2017.
[16] R. Zhang, J. Zhang, Y. Zhang, and C. Zhang, "Secure crowdsourcing-based cooperative pectrum sensing," in *Proc. INFOCOM*, Apr. 2013, pp. 2526–2534.
[17] X. O. Wang, W. Cheng, P. Mohapatra, and T. Abdelzaher, "ARTSense: Anonymous reputation and trust in participatory sensing," in *Proc. INFOCOM*, Apr. 2013, pp. 2517–2525.
[18] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. INFOCOM*, Apr. 2008, pp. 1876–1884.
[19] A. Min, K. Shin, and X. Hu, "Attack-tolerant distributed sensing for dynamic spectrum access networks," in *Proc. ICNP*, Oct. 2009, pp. 294–303.
[20] H. Li and Z. Han, "Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3554–3565, Nov. 2010.
[21] S. Li, H. Zhu, Z. Gao, X. Guan, and K. Xing, "YouSense: Mitigating entropy selfishness in distributed collaborative spectrum sensing," in *Proc. INFOCOM*, Apr. 2013, pp. 2535–2543.
[22] Y. Liu, P. Ning, and H. Dai, "Authenticating primary users' signals in cognitive radio networks via integrated cryptographic and wireless link signatures," in *Proc. SP*, May 2010, pp. 286–301.
[23] V. Kumar, J. Park, and K. Bian, "Blind transmitter authentication for spectrum security and enforcement," in *Proc. CCS*, Nov. 2014, pp. 787–798.
[24] X. Jin, J. Sun, R. Zhang, Y. Zhang, and C. Zhang, "SpecGuard: Spectrum misuse detection in dynamic spectrum access systems," in *Proc. INFOCOM*, Apr. 2015, pp. 172–180.
[25] X. Jin, J. Sun, R. Zhang, and Y. Zhang, "SafeDSA: Safeguard dynamic spectrum access against fake secondary users," in *Proc. CCS*, Oct. 2015, pp. 304–315.
[26] H. To, G. Ghinita, and C. Shahabi, "A framework for protecting worker location privacy in spatial crowdsourcing," in *Proc. VLDB*, Jun. 2014, pp. 919–930.
[27] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. INFOCOM*, Apr. 2015, pp. 2830–2838.
[28] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. MobiCom*, Aug. 2012, pp. 173–184.
[29] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint," in *Proc. INFOCOM*, Apr. 2014, pp. 1213–1221.
[30] C. Dwork, "Differential privacy," in *Proc. ICALP*, Jul. 2006, pp. 1–12.
[31] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proc. FOCS*, Oct. 2007, pp. 94–103.
[32] Z. Huang and S. Kannan, "The exponential mechanism for social welfare: Private, truthful, and nearly optimal," in *Proc. FOCS*, Oct. 2012, pp. 140–149.
[33] J. Sun, R. Zhang, J. Zhang, and Y. Zhang, "PriStream: Privacy-preserving distributed stream monitoring of thresholded percentile statistics," in *Proc. INFOCOM*, Apr. 2016, pp. 1–9.
[34] R. Zhu, Z. Li, F. Wu, K. Shin, and G. Chen, "Differentially private spectrum auction with approximate revenue maximization," in *Proc. MobiHoc*, Aug. 2014, pp. 185–194.
[35] R. Zhu and K. G. Shin, "Differentially private and strategy-proof spectrum auction with approximate revenue maximization," in *Proc. INFOCOM*, Apr. 2015, pp. 918–926.
[36] A. Gupta, K. Ligett, F. McSherry, A. Roth, and K. Talwar, "Differentially private combinatorial optimization," in *Proc. SODA*, Nov. 2009, pp. 1106–1125.
[37] Y. Selen, H. Tullberg, and H. Kronander, "Sensor selection for cooperative spectrum sensing," in *Proc. DySPAN*, Oct. 2008, pp. 1–11.
[38] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, *Algorithmic game theory*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[39] V. Vazirani, *Approximation Algorithms*. Berlin, Germany: Springer-Verlag, 2001.
[40] N. Immorlica, A. Karlin, M. Mahdian, and K. Talwar, "Balloon popping with applications to ascending auctions," in *Proc. FOCS*, Oct. 2007, pp. 104–112.
[41] J. Jia, Q. Zhang, Q. Zhang, and M. Liu, "Revenue generation for truthful spectrum auction in dynamic spectrum access," in *Proc. MobiHoc*, May 2009, pp. 1–12.
[42] A. Archer and É. Tardos, "Truthful mechanisms for one-parameter agents," in *Proc. FOCS*, Oct. 2001, pp. 482–491.
[43] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Proc. EUROCRYPT*, May 2006, pp. 486–503.
[44] Y. Zhang and M. van der Schaar, "Reputation-based incentive protocols in crowdsourcing applications," in *Proc. INFOCOM*, Mar. 2012, pp. 2140–2148.

Xiaocong Jin (M'17) received the B.E. degree in information engineering from Shanghai Jiao Tong University, China, in 2009, the M.S. degree in information, production, and systems engineering from Waseda University, Japan, in 2010, the M.S. degree in signal and information processing from Shanghai Jiao Tong University in 2012, and the Ph.D. degree in electrical engineering from Arizona State University, USA. He is currently with Google LLC as a Software Engineer. His primary research interests are network and distributed system security and privacy, wireless networking, and mobile computing.



Yanchao Zhang (SM'11) received the B.E. degree in computer science and technology from the Nanjing University of Posts and Telecommunications in 1999, the M.E. degree in computer science and technology from the Beijing University of Posts and Telecommunications in 2002, and the Ph.D. degree in electrical and computer engineering from the University of Florida in 2006. He is currently an Associate Professor with the School of Electrical, Computer and Energy Engineering, Arizona State University. His primary research interests are network and distributed system security, wireless networking, and mobile computing. He received the NSF CAREER Award in 2009. He was the TPC Co-Chair of the Communication and Information System Security Symposium and the IEEE GLOBECOM 2010. He is an Editor of the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the IEEE WIRELESS COMMUNICATIONS.