On Sybil-proof Optimal Auction Design for Crowdsensing

Swastik Brahma and Eunice Michael
Department of Computer Science
Tennessee State University
Nashville, TN 37209
Emails: {sbrahma, emichae2}@tnstate.edu

Abstract—The concept of crowdsensing enables a sensing task to be performed by outsourcing the task to crowd participants who are carrying devices with built-in sensors, such as sensors embedded in smartphones. This paper presents the design of a Sybil-proof optimal auction mechanism for incentive based crowdsensing that can not only ensure truthful revelation of crowd participants’ participation costs (which are sent as bids), but also disincentivize them from adopting Sybil behavior where participants may opt to participate in the auction mechanism using multiple fake identities. Simulation results are also provided to gain insights into the developed auction-based crowdsensing mechanism.

I. INTRODUCTION

Many of today’s sensing applications allow a number of users carrying devices with built-in sensors, such as sensors built in smartphones, automobiles and smart homes to contribute sensing data towards a sensing task. For instance, today’s smart phones are embedded with various sensors, such as camera, microphone, accelerometer, GPS, which can be used in an information acquisition process. An advantage of such architectures is that they do not need a dedicated sensing infrastructure for different inference tasks, thereby providing cost effectiveness. Another advantage of such architectures is that they allow ubiquitous coverage.

Systems and applications that rely on utilizing an infrastructure where sensing measurements of participating users are used are poised to revolutionize many sectors of our life. Some example application domains include environmental monitoring [14], green computing [9], target localization and tracking [7], [12], [17], [20], [23], healthcare [16] (such as predicting and tracking disease patterns/outbreaks), and tracking traffic patterns [10], [25]. For instance, the OpenSense project [14] involves the design of a sensing infrastructure for real-time air quality monitoring using heterogeneous sensors owned by the general public. GreenGPS [9] uses data from sensors installed in automobiles to map fuel consumption on city streets and construct fuel efficient routes between arbitrary end-points. Various systems to estimate object locations and to track them using smartphone sensors have also been proposed. For instance, [17], [20] utilize built-in sensors in smartphones such as camera, digital compass and GPS, to estimate a target location as well as monitor the velocity of moving objects. [7], [12], [23] use proximity sensors built-in smartphones to track objects (such as lost/stolen devices) installed with electronic tags (such as Bluetooth or RFID tags).

Many of the existing sensing applications and systems (for example, [1], [7], [17], [20], [23], [25]), however, assume voluntary participation of users (crowd participants). While participating in a sensing task, crowd participants consume their own resources such as energy and processing power, which can result in an insufficient number of participants unless suitable incentives are provided. To address such a concern, market-based mechanisms have been explored by past work [2]–[4], [13], [18]. In [18], the authors explored the possibility of using economic concepts for sensor management without explicitly formulating the problem. The authors in [4] used the concept of the Walrasian equilibrium [19] to model market based sensor management. In [13], the authors proposed a market based dynamic bit allocation scheme for target tracking in energy constrained wireless sensor networks (WSNs) using quantized data. However, the mechanisms proposed in [4], [13] are not truthful and are, therefore, prone to market manipulations. To address this concern, some past work (for example, [2], [3]) has considered the problem of designing auction based mechanisms for crowdsensing that can ensure truthful revelation of participation costs (which are sent as bids in the auction mechanism) of crowd participants. However, the aforementioned literature has not considered the design of incentive-based crowdsensing mechanisms when crowd participants may exhibit Sybil behavior by assuming multiple fake identities to participate in a crowdsensing task.

The goal of this paper is to present an optimal auction based mechanism for crowdsensing that can not only prevent crowd participants from gaining an undue advantage by bidding falsified participation costs, but also prevent them from making an undue profit by sending (potentially falsified) bids using multiple fake identities. We refer to such crowd participants who can potentially send multiple bids (which can be falsified) using fake identities as Sybil crowd participants, and to an auction mechanism that can disincentivize such behavior as a Sybil-proof auction. The concept of a Sybil attack, in which one physical entity can present itself using multiple identities, was originally described by [6] in the context of peer-to-peer networks. Such attacks have been studied in the context of communication networks, such as to understand network resource usage under such attacks [15], [24], spectrum allocation...
in wireless networks \[22\], and resource allocation in cloud platforms \[21\]. The impact of Sybil behavior in social welfare maximizing auctions (such as VCG auctions \[11\]) has been discussed in \[5\]. However, to the best of our knowledge, the problem of designing a revenue maximizing optimal auction based mechanism for providing incentives for crowdsensing when crowd participants may opt to adopt Sybil behavior has not been addressed by any prior work, which is the focus of this paper.

The rest of the paper is organized as follows. Section II describes our system model and formulates the auction design problem. Section III analyzes the problem and presents the optimal auction-based crowdsensing mechanism that can incentivize crowd participants to honestly report their participation costs without exhibiting Sybil behavior. Section IV provides simulation results to gain insights into the developed mechanism. Finally, Section V concludes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a crowdsensing platform (CP) that wants to perform a sensing task by outsourcing it, with the CP deriving a benefit \( v_{CP} \) from having the task performed\(^1\). Consider also that the crowdsensing platform allows creation of \( N \) accounts (identities), say numbered \( \{1, \ldots, N\} \), ideally by \( N \) different crowd participants to participate in the sensing task. Suppose that every crowd participant \( i \) has a private value estimate \( v_i \) (that reflects its participation cost) for performing the sensing task. The CP (who acts as the buyer) is assumed to be unaware of the true valuations (participation costs) of the crowd participants (who acts as the sellers) so that the crowd participants have to announce their valuations to the CP in the form of bids, with \( \psi = [v_1, \ldots, v_N] \) being the vector of announced value estimates. This gives the crowd participants an opportunity to lie about their valuations hoping for an extra benefit. We assume that the CP’s uncertainty about the value estimate of bidder \( i \) can be described by a continuous probability distribution \( f_i: [a_i, b_i] \rightarrow \mathbb{R}_+ \) over a finite interval \( [a_i, b_i] \), where \( a_i \) is the lowest possible valuation of \( i \) and \( b_i \) is the highest possible valuation of \( i \) with \( -\infty \leq a_i \leq b_i \leq \infty \). \( F_i: [a_i, b_i] \rightarrow [0, 1] \) denotes the cumulative distribution function, where \( F_i(v_i) = \int_{a_i}^{v_i} f_i(t)dt \).

The optimal auction based crowdsensing mechanism can then be described by two functions- \( a) \) \( q(\psi) = [q_1(\psi), \ldots, q_N(\psi)] \), where \( q_i(\psi) \) is the probability of selecting crowd participant \( i \) to perform the sensing task, and \( b) \) \( p(\psi) = [p_1(\psi), \ldots, p_N(\psi)] \), where \( p_i(\psi) \) is the payment made to crowd participant \( i \). The utility of the CP then becomes

\[
U_{CP}(q, p) = \mathbb{E} \left[ v_{CP} \sum_{i=1}^{N} q_i(\psi) - \sum_{i=1}^{N} p_i(\psi) \right]
\]

where, \( v_{CP}(\cdot) \) is the benefit the CP derives from having the sensing task performed, with the expectation in (1) taken over all possible combinations of the crowd participants’ valuations.

The utility of crowd participant \( i \) for a given true participation cost \( v_i \) from the auction mechanism becomes

\[
U_i(p_i, q_i, v_i) = \mathbb{E}[p_i(\psi) - v_i q_i(\psi)]
\]

Consider also that a crowd participant \( i \) can choose to act as a Sybil by sending a bid \( w_i \) (which need not be equal to the true valuation \( v_i \)) \( k \) times using \( k \) fake identities to receive the utility \( k\tilde{U}_i(p_i, q_i, w_i) \).

A. The Optimization Problem

Based on the above definitions, the optimal auction design problem becomes determining the functions \( p \) and \( q \) so as to maximize the utility of the CP (1), subject to certain constraints. Specifically, the optimization problem can be expressed as follows.

\[
\begin{align*}
\max_{p, q} & \quad U_{CP}(p, q) \\
\text{s.t.} & \quad U_i(p_i, q_i, v_i) \geq 0 \quad (3a) \\
& \quad U_i(p_i, q_i, v_i) \geq k\tilde{U}_i(p_i, q_i, w_i) \quad (3b) \\
& \quad \sum_{i=1}^{N} q_i \leq 1 \quad (3c)
\end{align*}
\]

The constraints above can be explained as follows.

- **Individual-Rationality (IR) constraint** (3a), which rationalizes participation by ensuring every crowd participant’s utility to be non-negative.
- **Incentive-Compatibility (IC) constraint** (3b), which ensures that the utility of every crowd participant from announcing its true valuation \( v_i \) without exhibiting Sybil behavior is greater than or equal to the utility the crowd participant receives from announcing a valuation \( w_i \) (which need not be equal to \( v_i \)) using \( k \) fake identities.
- **Selection constraint** (3c), which ensures that the sensing task is outsourced to at most one crowd participant.

In the next section, we analyze the aforementioned optimization problem and present our proposed auction-based crowdsensing mechanism.

III. SYBIL-PROOF OPTIMAL AUCTION-BASED CROWDSENSING

In this section, we analyze the optimization problem described in Section II-A and present the optimal auction-based crowdsensing mechanism that can achieve the desired properties described earlier. We define

\[
Q_i^k(q_i, v_i) = k \cdot \mathbb{E}[q_i(v_i, \psi_{-i})]
\]

for every crowd participant \( i \) who can act as a Sybil by sending a valuation \( v_i \) \( k \) times using \( k \) fake identities. In (4), \( (v_i, \psi_{-i}) \) denotes \( N \) valuations, with \( v_i \) being the valuation of the \( i^{th} \) crowd participant who can act as a Sybil and \( \psi_{-i} \) denoting all the remaining valuations. Thus, \( Q_i^k(q_i, v_i) \) is the conditional

---

\(^1\)For instance, \( v_{CP} \) can reflect the valuation for finding a lost/stolen object in \[7\], \[12\], \[23\].
probability that \( i \) is assigned the sensing task by the CP given that it sends valuation \( v_i \) \( k \) times using \( k \) fake identities.

Our first result is a simplified characterization of the IC constraint (3b) presented in Section II-A.

**LEMMA 1:** The IC constraint holds only if the following two conditions holds.
1. if \( v_i \leq w_i \), then \( Q_S(q_i, w_i) \leq Q_S(q_i, v_i) \) \hspace{1cm} (5a)
2. \( U_i(p_i, q_i, v_i) = U_i(p_i, q_i, b_i) + \int_{v_i}^{b_i} Q_S(q_i, w_i) dw_i \) \hspace{1cm} (5b)

**Proof:** Suppose \( v_i \leq w_i \). Also, suppose that, while \( v_i \) is the true valuation, crowd participant \( i \) sends the falsified valuation \( w_i \) \( k \) times using \( k \) fake identities, which makes the utility of \( i \) to be

\[
k E[p_i(w_i, v_{\ldots}) - q_i(w_i, v_{\ldots})v_i]
= k E[p_i(w_i, v_{\ldots}) - q_i(w_i, v_{\ldots})v_i] + k(w_i - v_i)E[q_i(w_i, v_{\ldots})]
= kU_i(p_i, q_i, w_i) + (w_i - v_i)Q_S(q_i, w_i)
\]

The IC constraint (3b) states that the expected utility of a crowd participant by reporting its true valuation once without adopting fake identities is greater than or equal to the utility obtained by reporting a falsified valuation using multiple fake identities. Thus, we must have,

\[
U_i(p_i, q_i, v_i) \geq kU_i(p_i, q_i, w_i) + (w_i - v_i)Q_S(q_i, w_i)
\]

From (7), we get,

\[
(w_i - v_i)Q_S(q_i, w_i) \leq U_i(p_i, q_i, v_i) - U_i(p_i, q_i, v_i)
\]

Similarly, considering \( w_i \) to be the true participation cost of crowd participant \( i \) while \( i \) sends the falsified valuation \( v_i \) to the CP \( k \) times using \( k \) fake identities, we get,

\[
U_i(p_i, q_i, v_i) \geq kU_i(p_i, q_i, v_i) - (w_i - v_i)Q_S(q_i, v_i)
\]

From (9), we get,

\[
(w_i - v_i)Q_S(q_i, v_i) \geq U_i(p_i, q_i, v_i) - U_i(p_i, q_i, v_i)
\]

Using (8) and (10), we get,

\[
(w_i - v_i)Q_S(q_i, w_i) \leq U_i(p_i, q_i, v_i) - U_i(p_i, q_i, v_i) \leq (w_i - v_i)Q_S(q_i, w_i)
\]

From (11), we can derive (5a). Moreover, defining \( \delta = w_i - v_i \), we can write the inequalities in (11) for any \( \delta \rightarrow 0 \) as,

\[
Q_S(q_i, w_i)\delta \leq U_i(p_i, q_i, w_i - \delta) - U_i(p_i, q_i, w_i) \leq \delta Q_S(q_i, w_i - \delta)
\]

Therefore, \( Q_S(q_i, w_i) \) is a decreasing function of \( w_i \), and thus Riemann integrable, based on which we get,

\[
\int_{v_i}^{b_i} Q_S(q_i, w_i) dw_i = U_i(p_i, q_i, v_i) - U_i(p_i, q_i, b_i)
\]

which proves (5b). This concludes the proof of the lemma.

Next, based on Lemma 1, the optimization problem presented in Section II-A can be simplified as follows.

**THEOREM 1:** In the optimal auction-based crowdsensing mechanism, \( q \) should maximize

\[
\sum_{i=1}^{N} \int_{T} \left[ v_{CP} - v_i - k \frac{F_i(v_i)}{(f_i(v_i))^k} \right] q_i(v) f(v) dv
\]

subject to constraint (3c), where \( T \) denotes the set of all possible combinations of crowd participants’ valuations i.e., \( T = [a_1, b_1] \times \cdots \times [a_N, b_N] \), and the payment to crowd participant \( i \) should be given by,

\[
p_i(v) = v_i q_i(v) + k \int_{v_i}^{b_i} q_i(w_i, v_{\ldots}) dw_i
\]

**Proof:** We can rewrite the CP’s utility (1) as,

\[
U_{CP}(p, q) = \sum_{i=1}^{N} \int_{T} \left[ v_i q_i(v) - p_i(v) \right] f(v) dv
\]

Now, we have,

\[
\int_{T} \left[ v_i q_i(v) - p_i(v) \right] f(v) dv
= - \int_{a_i}^{b_i} U_i(p_i, q_i, v_i) f_i(v_i) dv_i
\]

Substituting (17) into (16), we get,

\[
U_{CP}(p, q) = \sum_{i=1}^{N} \int_{T} \left[ v_i q_i(v) - k \frac{F_i(v_i)}{(f_i(v_i))^k} \right] q_i(v) f(v) dv
\]

In (18), \( p \) only appears in the last term of the objective function of the CP. Also, from the IR constraint (3a) we know that for every crowd participant \( i \), \( U_i(p_i, q_i, b_i) \geq 0 \). Thus, the best possible value of the last term in (18) can be obtained, which is zero since the CP seeks to maximize its objective function, as well as the IR constraint can be satisfied by having \( U_i(p_i, q_i, b_i) = 0 \), which implies, using (5b),

\[
U_i(p_i, q_i, v_i) - k \int_{v_i}^{b_i} Q_S(q_i, w_i) dw_i = 0
\]

Using (2) and (19), we get (15). This proves the theorem.
Therefore, based on Theorem 1, upon receiving a set of participation costs (valuations) from crowd participants, the CP can find the crowd participant to whom the sensing task should be outsourced and the corresponding payment to be made so as to solve the optimization problem described in Section II-A in the following manner.

- Based on the set of valuations, \( v = [v_1, \cdots, v_N] \), received from the crowd participants, the CP computes the following quantity for all \( i \in [1, N] \):

\[
\eta_i(v_i) = v_{CP} - v_i - k \frac{v_i}{f_i(v_i)}
\]  

(20)

If \( \max_{i \in [1, N]} \eta_i(v_i) < 0 \), the CP does not outsource the sensing task to any crowd participant; otherwise, the CP selects the crowd participant with the highest \( \eta_i(v_i) \) for performing the task. In other words, if \( \eta_i(v_i) = \max_{i \in [1, N]} \eta_i(v_i) \), then the solution to the selection probability \( q \) is \( q_i(v) = 1 \) and \( q_j(v) = 0 \), \( \forall j \in [1, \cdots, i-1, i+1, \cdots, N] \). Ties can be broken arbitrarily without affecting the utility of the CP.

- Based on the payment formula (15), crowd participants who are not selected for performing the task, do not receive any payments. This follows from the fact that if a crowd participant \( i \) is not selected based on the participation cost \( v_i \), then we have \( q_i(w_i, v_{-i}) = 0 \), \( \forall w_i \in [v_i, b_i] \). The payment of the crowd participant who is assigned the task can be found using (15).

In the next section, we provide simulation results to gain insights into the developed auction mechanism.

IV. SIMULATION RESULTS

In this section, we study the dynamics of our developed auction-based crowdsensing mechanism. In Figure 1, we show how the utility of the crowdsensing platform varies with varying number of crowd participants. In the figure, we consider the valuations of the crowd participants to be uniformly distributed over the range \([5, 15]\) and the number of fake identities every participant can opt to use to be \( k = 2 \). As can be seen from the figure, the utility of the CP increases as the number of crowd participants increases. This is due to the fact that as the number of crowd participants increases, the chances of the CP finding a crowd participant who requires less payment to perform the task increase. In other words, competition among the crowd participants increases as the number of crowd participants increases, thereby lowering the payment needed to have the task performed in a crowdsourced manner, resulting in the utility of the CP to increase. The figure also shows the utility of the CP as the benefit the CP receives from having the sensing task performed, \( v_{CP} \), varies. As is intuitive, the utility of the CP increases as \( v_{CP} \) increases which can be seen from the figure.

In Figure 2, we show how the utility of the CP varies as the number of fake identities (\( k \)) that every crowd participant can use varies. The figure considers \( N = 10 \), the valuations of the crowd participants to be uniformly distributed over the range \([5, 15]\), and \( v_{CP} = 10 \). As can be seen from the figure, the utility of the CP decreases as \( k \) increases. This is because, as the number of fake identities that every crowd participant can use increases, the payment required to disincentive such behavior follows an increasing trend based on the payment mechanism (15), thereby resulting in the utility of the CP to decrease.

V. CONCLUSIONS

This paper considered the problem of incentive-based crowdsensing where a crowdsensing platform (CP) provides incentives to crowd participants who are carrying devices with built-in sensors (such as smartphones) to perform a sensing task in a crowdsourced manner. Specifically, the paper designed a Sybil-proof optimal auction-based crowdsensing mechanism that can not only ensure truthful revelation of participation costs of crowd participants, but also prevent them from exhibiting Sybil behavior by ensuring that participants are unable to make an undue profit by sending multiple (potentially falsified) bids to the CP. Simulation results provide insights into the auction-based crowdsensing mechanism designed in the paper.
REFERENCES


