Sybil-Proof Online Incentive Mechanisms for Crowdsensing

Jian Lin, Ming Li, Dejun Yang, Guoliang Xue

***Abstract*—Crowdsensing leverages the rapid growth of sensor-embedded smartphones and human mobility for pervasive information collection. To incentivize smartphone users to participate in crowdsensing, many auction-based incentive mechanisms have been proposed for both offline and online scenarios. It has been demonstrated that the Sybil attack may undermine these mechanisms. In a Sybil attack, a user illegitimately pretends multiple identities to gain benefits. Sybil-proof incentive mechanisms have been proposed for the offline scenario. However, the problem of designing Sybil-proof online incentive mechanisms for crowdsensing is still open. Compared to the offline scenario, the online scenario provides users one more dimension of flexibility, i.e., active time, to conduct Sybil attacks, which makes this problem more challenging. In this paper, we design Sybil-proof online incentive mechanisms to deter the Sybil attack for crowdsensing. Depending on users' flexibility on performing their tasks, we investigate both single-minded and multi-minded cases and propose *SOS* and *SOM*, respectively. *SOS* achieves computational efficiency, individual rationality, truthfulness, and Sybil-proofness. *SOM* achieves individual rationality, truthfulness, and Sybil-proofness. Through extensive simulations, we evaluate the performance of *SOS* and *SOM*.**

I. INTRODUCTION

In recent years, the rapid proliferation of smartphones with rich embedded sensors has attracted great attentions from both academy and industry [7]. The effectiveness of crowdsensing to collect data enabled numerous crowdsensing applications in a wide variety of domains, such as transportation [28], marketing [6], environmental monitoring [22], cellular coverage map [19], and etc.

The number of participating users is a critical factor for the success of a crowdsensing application. Most of the early systems [18, 22] assume that the smartphone users contribute to the platform voluntarily. In practice, however, smartphone users taking part in crowdsensing cause extra cost while performing the sensing tasks, e.g., sensing time, battery expenditure, transmission expense and potential privacy threats from the exposure of their locations. Therefore, it is necessary to design incentive mechanisms to stimulate users to participate in crowdsensing.

J. Lin, M. Li are with Colorado School of Mines, Golden, CO 80401 USA. D. Yang (corresponding author) is with Colorado School of Mines, Golden, CO 80401 USA and Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023 China (e-mail: {jilin, mili, djyang}@mines.edu). G. Xue is with Arizona State University, Tempe, AZ 85287 USA (e-mail: xue@asu.edu). This research was supported in part by NSF grants 1421685, 1444059, 1717197, 1717315, and Jiangsu Key Laboratory of Big Data Security & Intelligent Processing, NJUPT. The information reported here does not reflect the position or the policy of the federal government.

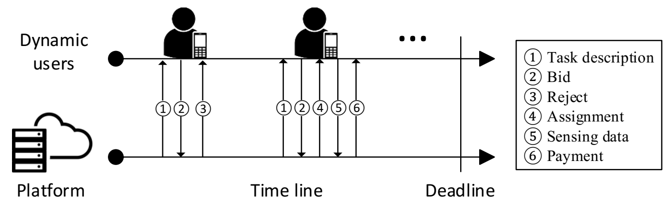


Fig. 1. Online crowdsensing system

At present, a number of auction-based incentive mechanisms have been proposed for crowdsensing. These incentive mechanisms model the interaction between the crowdsensing platform and the smartphone users as a reverse auction. The buyer is the platform, and the sellers are smartphone users who bid to perform sensing tasks. In these mechanisms, the platform selects users according to their submitted bids. Most of existing incentive mechanisms (e.g., [10, 15, 31, 32, 36]) focus on *offline* scenario in which smartphone users are required to submit their bids at the beginning of the auction, and the platform selects a subset of users according to some criteria for different objectives, e.g, maximize social welfare and platform utility. Some works (e.g., [4, 5, 8, 9, 27, 37, 39, 40, 40]) consider a more practical yet dynamic *online* scenario in which smartphone users participate in the system in a random order, as shown in Fig. 1. Once a user arrives, the platform has to make irrevocable decisions on whether to select it and how much it should be paid without knowing future information. However, none of above mechanisms take into consideration the Sybil attack [2], also known as false-name attack [34]. In recent years, the potential threat from Sybil attack has been investigated in various areas, such as cloud resource allocation [26], social networks [23, 24], and crowdsourced mobile apps [25]. These works focus on designing detection methods to eliminate Sybil attackers. The impact of Sybil attack in auctions has been analyzed in [33, 34].

In both offline and online crowdsensing systems, a user may submit multiple bids under different fictitious identities in the hope to increase its utility. This attack is easy to conduct (e.g., creating multiple accounts) but difficult to detect. The vulnerability to Sybil attack may make a mechanism fail to achieve its desired properties, and the fairness of the system will be jeopardized, since the increase of an attacker's utility may decrease other users' utility.

Lin *et al.* [16] are the first to investigate the Sybil attack in crowdsensing. They proposed two Sybil-proof incentive mechanisms in the offline scenario. However, their mechanisms cannot be directly applied to the online scenario, *since*

978-1-5386-4128-6/18/$31.00

the online scenario provides users one more dimension of flexibility to conduct Sybil attacks, i.e., active time window (to be elaborated latter). Existing online incentive mechanisms for crowdsensing [4, 5, 9, 27, 37–39] are all vulnerable to Sybil attack. Among them, the VCG-based incentive mechanism [5] is not Sybil-proof, since the VCG auction has been proved not Sybil-proof in [34]. The mechanism proposed in [4] is not Sybil-proof, since a user can increase its critical value and thus increase its payment by submitting multiple bids. For mechanisms in [9, 30, 37, 38], a user can increase its utility by changing from a loser to a winner via Sybil attack. We will use examples to demonstrate the vulnerability of existing online mechanisms to Sybil attack in Section IV. *Therefore, the problem of designing Sybil-proof online incentive mechanisms for crowdsensing remains open.*

In this paper, we focus on designing Sybil-proof online incentive mechanisms for crowdsensing. A mechanism is Sybil-proof if, participating in crowdsensing using a single identity is a dominant strategy of each user. The main contributions of this paper are as follows:

- To the best of our knowledge, we are the first to investigate Sybil attack in online incentive mechanisms for crowdsensing.
- We analyze existing online incentive mechanisms and demonstrate that they are all vulnerable to Sybil attack.
- Depending on users' flexibility on performing their tasks, we consider both the single-minded and multi-minded cases. We design *SOS* and *SOM* for these two cases, respectively. In order to design *SOS*, we provide a sufficient condition for an online mechanism to be Sybil-proof. We prove that *SOS* achieves computational efficiency, individual rationality, truthfulness, and Sybil-proofness, and that *SOM* achieves individual rationality, truthfulness, and Sybil-proofness.

The remainder of this paper is organized as follows. In Section II, we review the related work. In Section III, we introduce the system model and the objectives. In Section IV, we analyze the vulnerability of existing online mechanisms to Sybil attack. In Section V and Section VI, we present two online mechanisms for single-minded case and multi-minded case and prove their desired properties, respectively. Performance evaluations are presented in Section VII. We conclude this paper in Section VIII.

II. Related Work

In recent years, a number of auction-based incentive mechanisms have been proposed for crowdsensing. Most of them are offline mechanisms with different objectives e.g., maximizing the utility of the platform under a budget constraint [35], minimizing the social cost [3, 30], and preserving users' privacy [12, 14, 17]. The quality of sensing data has been considered in [11, 13, 20, 29]. Lin *et al.* [16] formalized the Sybil attack in crowdsensing and demonstrated that previous offline mechanisms are all vulnerable to Sybil attack. Two Sybil-proof offline incentive mechanisms are prosed in this paper, i.e., SPIM-S and SPIM-M. However, these two mechanisms cannot be directly applied to the online scenario.

Several works considered the online scenario where smartphone users come to the system in a random order. Among them, some online pricing mechanisms have been analyzed in [8, 40]. In addition, a number of auction-based online incentive mechanisms have been proposed for crowdsensing. Most of them aim to design online mechanisms, which have comparable performance to offline mechanisms. To get the information about upcoming users, two-sage based mechanisms have been proposed in [9, 37]. However, these mechanisms cannot guarantee the consumer sovereignty, since the first batch of users are rejected no matter how they bid. Zhao *et al.* [38, 39] proposed two multi-stage mechanisms, which satisfy consumer sovereignty. Furthermore, Gao *et al.* [5] proposed a VCG-based mechanism to invitivize users to participate in the system for a long-term. Feng *et al.* [4] considered a system with dynamic users and dynamic tasks. The users' privacy has been considered in [27]. However, none of existing online mechanisms take into consideration the Sybil attack.

Recently, the impact of Sybil attack has been widely analyzed in areas including virtual machine instance allocation [26], social networks [23] and crowdsourced mobile apps [25]. As pioneers, Yokoo *et al.* [34] analyzed the effects of Sybil attack on combinatorial auctions. They proved that VCG auction is not Sybil-proof in this paper. In addition, the price-oriented rationing-free protocols, which characterize the Sybil-proof protocols for combinatorial auction have been proposed in [33].

The problem of designing Sybil-proof online incentive mechanisms for crowdsensing is still open. All existing online incentive mechanisms are vulnerable to Sybil attack as explained in Section I, and we will show their vulnerabilities to Sybil attack in Section IV.

III. Model and Problem Formulation

A. System Model

In this paper, we consider a crowdsensing system consisting of a platform and a crowd of smartphone users $\mathcal{U} = \{1, 2, \ldots, n\}$, where n is unknown. The platform first publicizes a set $\mathcal{T} = \{\tau_1, \tau_2, \ldots, \tau_m\}$ of m sensing tasks, aiming at finding some users to complete these tasks before a specified deadline T, which is divided into slots of equal size. Each task $\tau_i \in \mathcal{T}$ has a value v_i to the platform. We use *bundle* to refer to any subset of $\mathcal{T}$. There is a function $V(\mathcal{B})$ to calculate the value of bundle $\mathcal{B}$ to the platform, i.e., $V(\mathcal{B}) = \sum_{\tau_i \in \mathcal{B}} v_i, \mathcal{B} \subseteq \mathcal{T}$. Each user i has a *active time window* within which it promises to complete the tasks if it is assigned, and a task set $\tilde{\Gamma}_i \subseteq \mathcal{T}$, which i can complete within its active time window. Let $\tilde{a}_i \in \{1, \ldots, T\}$ denote the begin of active time window and $\tilde{d}_i \in \{1, \ldots, T\}, \tilde{d}_i \geq \tilde{a}_i$ denote the end of active time window. Note that the platform has to make decision to each user i by $\tilde{d}_i$. As with [16], we assume that each user i has a cost function $\tilde{c}_i(\mathcal{B})$, which determines the cost for i to perform all tasks in bundle $\mathcal{B}$. The cost function $\tilde{c}_i(\cdot)$ satisfies the following properties:

- $\tilde{c}_i(\emptyset) = 0$;
- $\tilde{c}_i(\{\tau_j\}) = \infty, \forall \tau_j \in \mathcal{T} \setminus \tilde{\Gamma}_i$;
- $\tilde{c}_i(\mathcal{B}') \leq \tilde{c}_i(\mathcal{B}''), \forall \mathcal{B}' \subseteq \mathcal{B}'' \subseteq \mathcal{T}$;
- $\tilde{c}_i(\mathcal{B}) \leq \tilde{c}_i(\mathcal{B}') + \tilde{c}_i(\mathcal{B}''), \forall \mathcal{B}', \mathcal{B}'' \subseteq \mathcal{T}$ and $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$.

The first two properties depict user's capability. The third property implies that a user may incur more cost by performing more tasks. The last property means that a user's cost of performing a set of tasks is not greater than that of performing these tasks separately. These four properties together closely characterize a user's cost when participating in crowdsensing.

Depending on users' flexibility on performing their task sets, we consider two cases in this paper: *single-minded* (SM) and *multi-minded* (MM). For the SM case, each user $i \in \mathcal{U}$ is willing to perform only $\tilde{\Gamma}_i$ and behaves in a "win all or nothing" manner. For the MM case, each user $i \in \mathcal{U}$ is willing to perform any subset of $\tilde{\Gamma}_i$ and behaves in a flexible manner.

We use the sealed-bid reverse auction to model the interaction between the platform and the smartphone users. In our model, the buyer is the platform buying sensing services, and the sellers are smartphone users bidding for performing tasks. User i is a winner if it is assigned tasks, and a loser otherwise. Let $\beta_i = (a_i, d_i, \Gamma_i, b_i)$ denote the *bid* of user i. Similar to most online crowdsensing systems [4, 39], we assume that a user cannot announce an earlier arrival or a later departure, i.e., $\tilde{a}_i \leq a_i \leq d_i \leq \tilde{d}_i$. A user is active at time slot t if $a_i \leq t \leq d_i$. Note that, b_i is a value in the SM case, while b_i is a cost function in the MM case. We call user i's bid β_i is true if $\beta_i = (\tilde{a}_i, \tilde{d}_i, \tilde{\Gamma}_i, \tilde{c}(\tilde{\Gamma}_i))$ in the SM case; $\beta_i = (\tilde{a}_i, \tilde{d}_i, \tilde{\Gamma}_i, \tilde{c}(\cdot))$ in the MM case. At each time slot t, each newly arriving user i, i.e., $a_i = t$ submits its bid to the platform, which is not necessarily to be true. Given the bids of all active users at any time slot t, the platform will assign each active user i a bundle $\mathcal{A}_i^t \subseteq \tilde{\Gamma}_i$ to complete. Note that $\mathcal{A}_i^t = \emptyset$ means user i is not assigned any task to perform at time t. In addition, the platform calculates the payment p_i^t to user i for time slot t. Let $\mathcal{A}^t$ and p^t denote the assignment profile and the payment profile of all active users at time slot t, respectively. Besides, let $\mathcal{A}_i = \bigcup_{t \in [a_i, d_i]} \mathcal{A}_i^t$ and $p_i = \sum_{t=a_i}^{d_i} p_i^t$ denote the overall assignment and overall payment to user i, respectively. At last, let $\mathcal{A} = (\mathcal{A}^1, \ldots \mathcal{A}^T)$ denote the overall assignment profile, and $p = (p^1, \ldots p^T)$ denote the overall payment profile. The platform pays users once it receives the results of assigned tasks. Note that $p_i = 0$, if $\mathcal{A}_i = \emptyset$ or user i fails to perform assigned tasks. The utility of i in SM case is

$$\tilde{u}_i = \begin{cases} p_i - \tilde{c}_i(\mathcal{A}_i), & \text{if } \mathcal{A}_i = \tilde{\Gamma}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The utility of i in MM case is

$$\tilde{u}_i = \begin{cases} p_i - \tilde{c}_i(\mathcal{A}_i), & \text{if } \mathcal{A}_i \subseteq \tilde{\Gamma}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The utility of the platform is

$$u_0 = V(\bigcup_{i \in \mathcal{U}} \mathcal{A}_i) - \sum_{i \in \mathcal{U}} p_i. \tag{3}$$

B. Attack Model

In this paper, we assume that all users are selfish but rational. Hence it is possible that user i submits a false bid to maximizes its utility. Specifically, a false bid may have a false active time window, i.e., $a_i \neq \tilde{a}_i$ or $d_i \neq \tilde{d}_i$. In addition, b_i may not be true, i.e., $b_i \neq \tilde{c}_i(\tilde{\Gamma}_i)$ in the SM case; $b_i \neq \tilde{c}_i(\cdot)$ in the MM case. Furthermore, user i could also misreport its task set i.e., $\Gamma_i \neq \tilde{\Gamma}_i$.

We also assume that an attacker could conduct Sybil attack at any time slot in its active time by submitting multiple bids under fictitious identities. As a simple case, attacker i could submit two bids under two identities i' and i'', respectively. Note that i could submit these two bids simultaneously or at different time slots in its active time window. This case is sufficient to represent the general Sybil attack. We extend the definition of attacker's utility in the following two cases.

Single-Minded Case: Each attacker i is only willing to perform $\tilde{\Gamma}_i$. Attacker i submits $\beta_{i'} = (a_{i'}, d_{i'}, \Gamma_{i'}, b_{i'})$ and $\beta_{i''} = (a_{i''}, d_{i''}, \Gamma_{i''}, b_{i''})$ within $[\tilde{a}_i, \tilde{d}_i]$ using identities i' and i'', respectively, where $\Gamma_{i'} \cup \Gamma_{i''} = \tilde{\Gamma}_i$. Attacker i's utility is

$$u_i = \begin{cases} p_{i'} + p_{i''} - \tilde{c}_i(\tilde{\Gamma}_i), & \text{if } \mathcal{A}_{i'} \cup \mathcal{A}_{i''} = \tilde{\Gamma}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Multi-Minded Case: Each attacker is willing to perform any subset of its task set. Attacker i submits $\beta_{i'} = (a_{i'}, d_{i'}, \Gamma_{i'}, b_{i'})$ and $\beta_{i''} = (a_{i''}, d_{i''}, \Gamma_{i''}, b_{i''})$ within $[\tilde{a}_i, \tilde{d}_i]$ using identities i' and i'', respectively, where $\Gamma_{i'} \subseteq \tilde{\Gamma}_i$ and $\Gamma_{i''} \subseteq \tilde{\Gamma}_i$. Attacker i's utility is

$$u_i = \begin{cases} p_{i'} + p_{i''} - \tilde{c}_i(\mathcal{A}_{i'} \cup \mathcal{A}_{i''}), & \text{if } \mathcal{A}_{i'} \cup \mathcal{A}_{i''} \subseteq \tilde{\Gamma}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

It is obvious that attacker i has an incentive to conduct Sybil attack if $u_i > \tilde{u}_i$ in either case. We will use examples to show the vulnerability of existing online incentive mechanisms to Sybil attack in Section IV. Note that any user can be an attacher, we use attacker and user interchangeably in the rest of this paper.

C. Desired Properties and Objective

In this paper, we consider the following properties:

- **Computational Efficiency:** A mechanism is computaionally efficient if it terminates in polynomial time.
- **Individual Rationality:** A mechanism is individually rational if each user has a non-negative utility when bidding its true bid.
- **Truthfulness:** A mechanism is truthful if any user's utility is maximized when bidding its true bid including both true active time window and true cost.
- **Sybil-Proofness:** A mechanism is Sybil-proof if any user's utility is maximized when bidding its bid using a single identity.

The objective of this paper is to design Sybil-proof online incentive mechanisms satisfying above properties. The main notations are summarized in Table I.

TABLE I
NOTATIONS

Symbol	Description
$\mathcal{T}$, m, τ_j	set of tasks, number of tasks, and one task
$\mathcal{U}$, n, i	set of users, number of users, and one user
$\mathcal{U}^t$	active users at time slot t
$V(\mathcal{B})$	value of bundle $\mathcal{B}$ to the platform
v_i^t	marginal value of user i at time slot t
$\tilde{u}_i$, u_i	i's utility and i's utility via Sybil attack
u_0	platform utility
T, t, $t[i]$	deadline, each time slot, and time slot i wins
$\tilde{a}_i$, a_i	i's true and submitted start of its active time window
$\tilde{d}_i$, d_i	i's true and submitted end of its active time window
$\tilde{\Gamma}_i$, Γ_i, Γ_i^t	i's true and submitted task set, and tasks i can perform at time t
$\tilde{c}(\cdot)$, $c(\cdot)$	i's true and submitted cost function
$\mathcal{A}_i^t$, $\mathcal{A}^t$, $\mathcal{A}$	i's and all active users' assignment at time slot t, and overall assignment profile
p_i^t, p^t, p	i's and all active users' payment at time slot t, and overall payment profile
$\mathbb{B}_i^t(\mathcal{B})$	set of subsets of Γ_i^t that intersect with $\mathcal{B}$

IV. VULNERABILITY OF EXISTING ONLINE MECHANISMS TO SYBIL ATTACK

In this section, we use examples to show existing online incentive mechanisms are vulnerable to Sybil attack.

A. Mechanism Classification

We classify existing online incentive mechanisms into three categories according to their vulnerabilities to Sybil attack. The first category is the VCG-based mechanism [5]. The second category is the critical value-based mechanism in [4] where each winner is paid its critical value. The third category consists of threshold-based mechanisms [9, 27, 37–39]. The mechanism in the first category is not Sybil-proof since VCG auction is proved not Sybil-proof in [34]. Next, we analyze the vulnerabilities to Sybil attack for the last two categories.

B. Vulnerabilities of Critical Value-based Mechanisms

The mechanism in [4] executes a reverse auction round by round. In one round of auction, the mechanism sorts all active users in a non-decreasing order by their bids. The first m_t users will be selected as winners, where m_t is the number of tasks at time slot t. At last, the payment to each winner i is set to its critical value. Let t_i denote the time slot i wins. Winner i's critical value is the largest of all the m_t-th users' bids at each time slot $t \in [t_i, d_i]$. It is obvious that at any time slot $t \in [t_i, d_i]$ in which the m_t-th user does not exist, i can submit a higher bid using a fictitious identity to take the m_t-th place. Therefore, i can increase its critical value and thus increase its payment via Sybil attack.

C. Vulnerabilities of Threshold-based Mechanisms

Within this category, we further divide the mechanisms into two groups. The first group comprises mechanisms in [9, 37]. These two mechanisms are two-stage mechanisms in which the first arrived $\lfloor \frac{n}{e} \rfloor$ users are rejected and their bids are used as the sample for the next stage, where n is the number of participating users and e is the base of the natural logarithm. In the first stage, the largest v_i/b_i value will be used as a threshold for the user selection in the next stage, where v_i is user i's marginal value and b_i is its bid. In the second stage, the first user i whose v_i/b_i value is no less than the threshold will be selected as a winner. We use the example in TABLE II. to show that these mechanisms are not Sybil-proof. In this example, the value of τ_1 is 1 and the value of τ_2 is 3. The mechanism will reject the first user (User1) since $n = 5$ and $\lfloor \frac{5}{e} \rfloor = 1$. Next, assume user 1 conducts Sybil attack by submitting two bids $\beta_{i'} = (1, 1, \{\tau_1\}, 2)$ and $\beta_{i''} = (2, 2, \{\tau_2\}, 2)$ under fictitious identities $1'$ and $1''$, respectively. In this case, the first two users (User $1'$ and User 2) are rejected since $n = 6$ and $\lfloor \frac{6}{e} \rfloor = 2$, and the threshold is 0.5. User i'' is the third arrived user whose $v_{i''}/b_{i''} = 1.5 > 0.5$, and thus it is selected as a winner. Because these mechanisms satisfy individual rationality, user i'' has a non-negative utility. Therefore, these mechanisms are not Sybil-proof, since a user can increase its utility by changing from a loser to a winner via Sybil attack.

The second group comprises of an improved two-stage mechanism [27] and a multi-stage mechanism called OMG [38, 39]. Given the deadline T and budget B, these two mechanisms will set the cutoff time of the first stage by $\frac{T}{2^{\lfloor \ln T \rfloor}}$ and $\frac{T}{2^{\lfloor \log_2 T \rfloor}}$, respectively, and allocate the first stage a stage-budget $\frac{B}{2^{\lfloor \ln T \rfloor}}$ and $\frac{B}{2^{\lfloor \log_2 T \rfloor}}$, respectively. We take the OMG as an example and show that it is still not Sybil-proof using the example in TABLE II. In the first stage, OMG selects users iteratively according to users' marginal density, v_i/b_i, where v_i is user i's marginal value, and b_i is its bid. In each iteration, the user with the largest marginal value will be selected, and if its marginal density is not less than a preset density threshold ρ^* and its bid does not exceed the stage-budget, it will be selected as a winner. In this example, density threshold $\rho^* = 0.5$, $T = 8$ and $B = 16$, and thus the first stage ends at time slot 1 and the stage-budget in the first stage is 2. The value of τ_1 is 1 and the value of τ_2 is 2. In this case, user 1 will not be selected since its bid is greater than the stage-budget. Next, assume user 1 conducts Sybil attack by submitting two bids $\beta_{i'} = (1, 2, \{\tau_1\}, 2)$ and $\beta_{i''} = (2, 2, \{\tau_2\}, 2)$ under fictitious identities $1'$ and $1''$, respectively. In this case, user $1'$ will win with non-negative utility since OMG is individually rational. Therefore, OMG is not Sybil-proof, since a user can increase its utility by changing from a loser to a winner via Sybil attack.

TABLE II
EXAMPLE SHOWING VULNERABILITIES TO SYBIL ATTACK

User	Bid (a_i, d_i, Γ_i, b_i)
1	(1,2,$\{\tau_1, \tau_2\}$, 4)
2	(2,2,$\{\tau_1\}$, 2)
3	(3,4,$\{\tau_1\}$, 3)
4	(3,3,$\{\tau_2\}$, 4)
5	(3,4,$\{\tau_1, \tau_2\}$, 4)

V. *SOS*: SYBIL-PROOF ONLINE INCENTIVE MECHANISM FOR SINGLE-MINDED CASE

In this section, we design and analyze *SOS*, a Sybil-proof online incentive mechanism for SM case.

A. Design Rationale

In SM case, a user could maximize its utility by submitting multiple subsets of its task set using multiple identities with different active time window in the hope that all the identities will be selected as winners. In order to design Sybil-proof mechanisms, we provide a sufficient condition for an online mechanism to be Sybil-proof in the following theorem.

Theorem 1: An online mechanism is *Sybil-proof* if it satisfies the following two conditions: If any user i pretends two identities i' and i'', and both i' and i'' are selected as winners with assignment $\mathcal{A}_{i'}$ and $\mathcal{A}_{i''}$, respectively within $[\tilde{a}_i, \tilde{d}_i]$, then

1) i should be selected as a winner with assignment $\mathcal{A}_i = \mathcal{A}_{i'} \cup \mathcal{A}_{i''}$ within $[\tilde{a}_i, \tilde{d}_i]$ while using only one identity;
2) $p_i \geq p_{i'} + p_{i''}$.

Proof: Assume user i pretends two identities i' and i'', and both i' and i'' are winners within its true active time window with assignment $\mathcal{A}_{i'}$ and $\mathcal{A}_{i''}$, respectively. According to 1), i would have been a winner within its true active time window with assignment $\mathcal{A}_i = \mathcal{A}_{i'} \cup \mathcal{A}_{i''}$. According to (1), user i's utility is $\tilde{u}_i = p_i - \tilde{c}_i(\mathcal{A}_i)$. According to (4), user i's utility via Sybil attack is $u_i = p_{i'} + p_{i''} - \tilde{c}_i(\mathcal{A}_i)$. Because of 2), we have $p_i \geq p_{i'} + p_{i''}$, and thus $\tilde{u}_i \geq u_i$. Therefore, user i cannot increase its utility via Sybil attack. ■

The following theorem will be used to guarantee the truthfulness of *SOS*.

Theorem 2: [4] An online mechanism is truthful iff:

- The winner selection rule is monotone: If user i wins the auction by bidding $\beta_i = (a_i, d_i, \Gamma_i, b_i)$, it also wins by bidding $\beta'_i = (a'_i, d'_i, \Gamma'_i, b'_i)$, where $a'_i \leq a_i, d'_i \geq d_i, \Gamma_i \subseteq \Gamma'_i, b'_i \leq b_i$;
- Each winner is paid the critical value, which is the smallest value such that user i would lose the auction if it bids higher than this value.

In order to guarantee Sybil-proofness and truthfulness, *SOS* should satisfy both Theorem 1 and Theorem 2.

B. Design of Mechanism

In this section, we describe the details of *SOS*, which is comprised of two subroutines: winner selection with assignment and payment determination.

The winner selection with assignment is illustrated in Algorithm 1. It selects winners iteratively at each time slot until all tasks are assigned or deadline T is reached. Let $\mathcal{R}^t$ denote the set of currently unassigned tasks and $v_i^t = V(\mathcal{R}^t \cap \Gamma_i)$ denote the marginal value of user i to the platform at time slot t. At each time slot t, Algorithm 1 selects the user with the largest criterion value, $v_i^t - b_i$, from all active users in $\mathcal{U}^t$. If its criterion value is non-negative, this user will be put into winner set $\mathcal{W}$ and assigned the tasks it submits. Otherwise, it will not be assigned tasks. All active users' task assignments constitute the assignment profile $\mathcal{A}^t$ of time slot t. The assignment profile of every time slot constitutes the overall assignment profile $\mathcal{A}$. The outcome of Algorithm 1 are $\mathcal{A}$ and $\mathcal{W}$.

The payment determination is illustrated in Algorithm 2. The input are user i's ID, the time slot $t[i]$ in which i is assigned tasks, and the set $\mathcal{R}^{t[i]}$ of unassigned tasks at the beginning of time slot $t[i]$. For each time slot in $[t[i], d_i]$, Algorithm 2 calculates the highest price i can bid in order to be a winner. At last, the payment to user i is set to the highest price among these prices. Note that

$$p_i = \arg \max_{t \in [t[i], d_i]} \{v_i^t - c^t\}, \tag{6}$$

where $c^t = \max\{0, v_{i_j}^t - b_{i_j}\}$, and i_j is the user with the largest criterion value at time slot t when i is not in $\mathcal{U}^t$.

The main algorithm of *SOS* is illustrated in Algorithm 3. Note that, there is at most one winner at each time slot according to Algorithm 1. Therefore, *SOS* iterates all time slots and calculates the payment for the winner at each time slot using Algorithm 2.

Algorithm 1: *SOS*-WSA($\mathcal{T}$, T)

$\mathcal{W} \leftarrow \emptyset$, $\mathcal{A}^t \leftarrow \emptyset, \forall t \in [1, T]$, $t \leftarrow 1$, $\mathcal{R}^t \leftarrow \mathcal{T}$;
while $\mathcal{R}^t \neq \emptyset$ **and** $t \leq T$ **do**
  $\mathcal{U}^t \leftarrow$ the set of active users at time slot t;
  $\mathcal{A}_j^t \leftarrow \emptyset, \forall j \in \mathcal{U}^t$;
  $i \leftarrow \arg\max_{j \in \mathcal{U}^t}(v_j^t - b_j)$;
  if $b_i \leq v_i^t$ **then**
    $\mathcal{W} \leftarrow \mathcal{W} \cup \{i\}$, $\mathcal{A}_i^t \leftarrow \Gamma_i$, $\mathcal{R}^{t+1} \leftarrow \mathcal{R}^t \setminus \Gamma_i$;
  end
  $t \leftarrow t + 1$;
end
$\mathcal{A} \leftarrow (\mathcal{A}^1, \ldots \mathcal{A}^T)$;
return $(\mathcal{A}, \mathcal{W})$.

Algorithm 2: *SOS*-PD(i, $t[i]$, $\mathcal{R}^{t[i]}$)

$t \leftarrow t[i], p_i^{t[i]} \leftarrow 0$;
while $t \leq d_i$ **and** $\mathcal{R}^t \neq \emptyset$ **do**
  $\mathcal{U}^t \leftarrow$ the set of active users at time slot t;
  $\mathcal{U}^t \leftarrow \mathcal{U}^t \setminus \{i\}$;
  $i_j \leftarrow \arg\max_{j \in \mathcal{U}^t}(v_j^t - b_j)$;
  if $b_{i_j} \leq v_{i_j}^t$ **then**
    $p_i^{t[i]} \leftarrow \max\{p_i^{t[i]}, v_i^t - (v_{i_j}^t - b_{i_j})\}$;
    $\mathcal{R}^{t+1} \leftarrow \mathcal{R}^t \setminus \Gamma_i$;
  else
    $p_i^{t[i]} \leftarrow \max\{p_i^{t[i]}, v_i^t\}$;
  end
  $t \leftarrow t + 1$;
end
return $p_i^{t[i]}$.

Algorithm 3: *SOS*($\mathcal{T}$, T)

$t \leftarrow 1$, $\mathcal{R}^t \leftarrow \mathcal{T}$;
$(\mathcal{A}, \mathcal{W}) \leftarrow$ *SOS*-WSA($\mathcal{T}$, T);
while $\mathcal{R}^t \neq \emptyset$ **and** $t \leq T$ **do**
  $i \leftarrow j \in \mathcal{W}$ s.t. $\mathcal{A}_j^t \neq \emptyset$;
  $p_i^t \leftarrow$ *SOS*-PD(i, t, $\mathcal{R}^t$);
  $\mathcal{R}^{t+1} \leftarrow \mathcal{R}^t \setminus \mathcal{A}_i^t$;
  $t \leftarrow t + 1$;
end
return $(\mathcal{A}, p)$.

C. Analysis of SOS

In this section, we prove the properties of *SOS* in the following theorem.

Theorem 3: *SOS* is computationally efficient, individually rational, truthful and Sybil-proof in SM case.

We prove this theorem with the following lemmas.

Lemma 1: SOS is computationally efficient.

Due to space limit, we omit the proof for this lemma.

Lemma 2: SOS is individually rational.

Proof: For any winner i, assume it is selected at time slot $t[i]$ with its true bid, i.e., $b_i = \tilde{c}(\tilde{\Gamma}_i)$. If there exists a winner j at time slot $t[i]$ when i is not in $\mathcal{U}^{t[i]}$, we have $v_i^{t[i]} - b_i \geq v_j^{t[i]} - b_j \geq 0$ since i was the winner at time slot $t[i]$. Then according to Line 7 in Algorithm 2, we have $p_i^{t[i]} \geq v_i^{t[i]} - (v_j^{t[i]} - b_j) \geq b_i$. If there is no winner at time slot $t[i]$ when i is not in $\mathcal{U}^{t[i]}$, we have $p_i \geq v_i^{t[i]} \geq b_i$ according to Line 10. Therefore, $u_i = p_i - \tilde{c}(\tilde{\Gamma}_i) = p_i - b_i \geq 0$, and *SOS* is individually rational. ■

Lemma 3: SOS is truthful.

Proof: We first prove that user i cannot increase its utility by submitting a false task set. We then prove that user i cannot increase its utility by submitting a false active time window or a false cost. If user i submits a false task set $\Gamma_i \subset \tilde{\Gamma}_i$, the utility of i is 0 according to (1). On the contrary, if $\Gamma_i \backslash \tilde{\Gamma}_i \neq \emptyset$, user i will not be paid since it cannot finish all the tasks in Γ_i. Thus, there is no incentive for i to submit a false task set.

To prove that user i cannot increase its utility by submitting a false active time window or a false cost, it suffices to prove that the selection rule of *SOS* is monotone and the payment to each winner is its critical value according to Theorem 2. Obviously, the criterion value of a user will increase with the decrease of user's cost. Meanwhile, due to the submodularity of user's marginal value, the criterion value of a user at each time slot will not decrease if it bids a wider active time window. Therefore, the selection rule of *SOS* is monotone. Next, we prove that the payment p_i to winner i is its critical value. Assume i was selected at time slot $t[i]$ with b_i, and thus $p_i = p_i^{t[i]} \geq b_i$. If user i bids $\tilde{b}_i > p_i$, it is obvious that i still loses at any time slot $t \in [a_i, t[i])$ since its criterion value is less than that when i bids b_i but loses in $[a_i, t[i])$. At any time slot $t \in [t[i], d_i]$, i still loses since there always exists a user i_j such that $v_i^t - \tilde{b}_i < v_{i_j}^t - b_{i_j}$ according to (6). If user i bids $\tilde{b}_i < p_i$, i wins at least within $[t[i], d_i]$, according to (6), if not earlier. Therefore, p_i is the critical value for user i. ■

Lemma 4: SOS is Sybil-proof.

Proof: We prove *SOS* is Sybil-proof by proving it satisfies the sufficient conditions in Theorem 1. Assume user i submits $(a_{i'}, d_{i'}, \Gamma_{i'}, b_{i'})$ and $(a_{i''}, d_{i''}, \Gamma_{i''}, b_{i''})$ using two fictitious identities i' and i'', respectively, where $\Gamma_{i'} \cup \Gamma_{i''} = \tilde{\Gamma}_i$.

We first prove that *SOS* satisfies the first condition in Theorem 1. We assume that both i' and i'' are selected as winner at time $t[i']$ and $t[i'']$ (w.l.o.g. $t[i'] < t[i'']$), respectively. It implies that $\Gamma_{i'} \subset \tilde{\Gamma}_i$ and $\Gamma_{i'} \subset \tilde{\Gamma}_i$, since one will make the other lose otherwise. In addition, we have $v_{i'}^{t[i']} \geq b_{i'}$ and $v_{i''}^{t[i'']} \geq b_{i''}$, since both i' and i'' are winners. If user i use a single identity and submits its true bid $(\tilde{a}_i, \tilde{d}_i, \tilde{\Gamma}_i, b_i)$. At time slot $t[i']$, we have $v_i^{t[i']} - b_i \geq v_{i'}^{t[i']} + v_{i''}^{t[i'']} - b_i \geq v_{i'}^{t[i']} + v_{i''}^{t[i'']} - b_{i'} - b_{i''}$. The first inequation lies in the fact that $\Gamma_{i'} \cup \Gamma_{i''} = \tilde{\Gamma}_i$. The second inequation is based on the fourth property of the cost function. Therefore, we have $v_i^{t[i']} - b_i \geq v_{i'}^{t[i']} - b_{i'}$ since $v_{i''}^{t[i'']} - b_{i''} \geq 0$. This implies that i wins at $t[i']$ at the latest while using a single identity. Therefore, the first condition in Theorem 1 is satisfied.

We next prove that *SOS* satisfies the second condition in Theorem 1. We know that user i wins at $t[i] \leq t[i']$. Let t_c, t'_c and t''_c denote the time that determine the payment of i, i' and i'' according to (6), respectively. We know that $t_c \in [t[i], d_i]$, $t'_c \in [t[i'], d_{i'}]$, and $t''_c \in [t[i''], d_{i''}]$ according to Algorithm 2. We then prove by cases. In Case 1, $t_c \in [t[i'], d_i]$ and $t'_c \leq t''_c$. According to (6), we have $p_i \geq v_i^{t'_c} - c^{t'_c} \geq v_{i'}^{t'_c} + v_{i''}^{t''_c} - c^{t'_c} \geq p_{i'} + p_{i''}$. The first inequation results from the fact $t'_c \in [t[i], d_i]$. The second inequation is based on the fact $\Gamma_{i'} \cup \Gamma_{i''} = \tilde{\Gamma}_i$. The third inequation is based on the fact $p_{i'} = v_{i'}^{t'_c} - c^{t'_c}$ and $p_{i''} \leq v_{i''}^{t''_c}$ according to (6). In Case 2, $t_c \in [t[i'], d_i]$ and $t'_c > t''_c$. Similar to Case 1, we can prove that $p_i \geq v_i^{t''_c} - c^{t''_c} \geq v_{i'}^{t'_c} + v_{i''}^{t''_c} - c^{t''_c} \geq p_{i'} + p_{i''}$. In Case 3, $t_c \in [t[i], t[i'])$. Because $t_c < t[i']$, we have $p_i \geq \max\{v_i^{t'_c} - c^{t'_c}, v_i^{t''_c} - c^{t''_c}\} \geq p_{i'} + p_{i''}$ based on the proofs of Case 1 and Case 2. Therefore, we have $p_i \geq p_{i'} + p_{i''}$. Hence, the second condition in Theorem 1 is satisfied.

Therefore, *SOS* is Sybil-proof according to Theorem 1. We can use a similar proof for the case where a user pretends more than two identities. ■

VI. *SOM*: Sybil-Proof Online Incentive Mechanism for Multi-Minded Case

In this section, we design and analyze *SOM*, a Sybil-proof online incentive mechanism for MM case.

A. Design Rationale

In MM case, a user is willing to perform any subset of its task set and tries to maximize its utility by submitting multiple bids under fictitious identities. To guarantee that each user submits its true cost function, *SOM* gives the payment to each user, which is independent of its own cost function. The time-truthfulness is based on the monotonic task assignment rule and the submodularity of users' marginal value. To achieve Sybil-proofness, we extend the characterization of Sybil-proof mechanisms in [33] to the online scenario.

B. Design of Mechanism

The main algorithm of *SOM* is shown in Algorithm 4. It selects users iteratively at each time slot until all tasks are assigned or deadline T is reached. Given the sensing tasks $\mathcal{T}$ and deadline T, *SOM* outputs the overall assignment profile $\mathcal{A}$ and the overall payment profile p.

Let $\Gamma_i^t \subseteq \tilde{\Gamma}_i$ denote a set of unassigned tasks that i can perform at time slot t. Let $\mathbb{B}_i^t(\mathcal{B}) = \{\mathcal{B}' | \mathcal{B}' \subseteq \Gamma_i^t, \mathcal{B}' \cap \mathcal{B} \neq \emptyset\}$. At each time slot t, *SOM* first calculates the payment $p_{i,\mathcal{B}}^t$ to each active user i for any bundle $\mathcal{B} \subseteq \Gamma_i^t$. Note that the payment to user i for any bundle $\mathcal{B} \subseteq \Gamma_i^t$ is independent of its cost function $c_i(\cdot)$, i.e.,

$$p_{i,\mathcal{B}}^t = V(\mathcal{B}) - \max\{0, \max_{j \neq i, \mathcal{B}' \in \mathbb{B}_j^t(\mathcal{B})} (V(\mathcal{B}') - c_j(\mathcal{B}'))\}. \quad (7)$$

Algorithm 4: *SOM($\mathcal{T}$, T)*

```
F ← ∅, A^t ← ∅, ∀t ∈ [1, T], t ← 1;
while F ≠ T and t ≤ T do
    U^t ← the set of active users at time slot t;
    Γ_i^t ← Γ_i \ F, ∀i ∈ U^t;
    foreach i ∈ U^t do
        Calculate the payment to i for any bundle B ⊆ Γ_i^t, p_{i,B}^t ←
        V(B) − max{0, max_{j≠i, B'∈𝔹_j^t(B)} (V(B') − c_j(B'))};
    end
    foreach i ∈ U^t do
        A_i^t ← arg max_{B⊆Γ_i^t} (p_{i,B}^t − c_i(B));
        p_i^t ← p_{i,A_i^t}^t;
        F ← F ∪ A_i^t;
    end
    t ← t + 1;
end
A = A^1, . . . A^T);
p = (p^1, . . . , p^T);
return A and p.
```

At last, *SOM* will assign each active user i a set of tasks $\mathcal{A}_i^t$, which is a bundle $\mathcal{B} \subseteq \Gamma_i^t$ maximizing its utility based on the calculated payment, i.e.,

$$\mathcal{A}_i^t = \arg\max_{\mathcal{B}\subseteq\Gamma_i^t}(p_{i,\mathcal{B}}^t - c_i(\mathcal{B})). \tag{8}$$

The payment p_i^t to each user i at time slot t for assignment $\mathcal{A}_i^t$ is $p_{i,\mathcal{A}_i^t}^t$. Note that $p_i^t = p_{i,\mathcal{A}_i^t}^t = 0$, if $\mathcal{A}_i^t = \emptyset$.

C. Analysis of SOM

In this section, we prove the properties of *SOM* in the following theorem.

Theorem 4: *SOM* is individually rational, truthful and Sybil-proof in MM case.

We prove this theorem with the following lemmas.

Lemma 5: SOM is individually rational.

Proof: The utility of any active user i at any time slot t is 0 when the assignment $\mathcal{A}_i^t = \emptyset$ according to (2). According to (8), at any time slot t, *SOM* assigns any active user i a bundle $\mathcal{A}_i^t$ maximizing its utility. It implies that $\mathcal{A}_i^t \neq \emptyset$ only if $u_i = p_{i,\mathcal{A}_i^t}^t - \tilde{c}_i(\mathcal{A}_i^t) > 0$, and thus the utility of any user i is non-negative. Therefore, *SOM* is individually rational. ■

Lemma 6: SOM is truthful.

Proof: We first prove that user i cannot increase its utility by submitting a false task set. Then, we prove that user i cannot increase its utility by submitting a false active time window or a false cost function. Assume user i submits a false bid $\beta_i = (a_i, d_i, \Gamma_i, c_i(\cdot))$. By (7), at any time slot t, the payment to i for any bundle $\mathcal{B} \subseteq \Gamma_i^t$ is calculated independently of i's own cost function. If $\Gamma_i \subset \tilde{\Gamma}_i$, the payment to i for any subset of Γ_i is the same as that when i submits $\tilde{\Gamma}_i$. In addition, at each time slot t, *SOM* assigns i a bundle maximizing its utility by (8). Therefore, user i cannot increase its utility by submitting $\Gamma_i \subset \tilde{\Gamma}_i$. On the contrary, if $\Gamma_i \setminus \tilde{\Gamma}_i \neq \emptyset$, it will not be assigned tasks in $\Gamma_i \setminus \tilde{\Gamma}_i \neq \emptyset$, since its utility is negative according to the second property of the cost function. Therefore, user i has no incentive to submit a false task set.

Next, we prove that user i has no incentive to submit a false active time window, i.e., $a_i > \tilde{a}_i$ or $d_i < \tilde{d}_i$. By (7), *SOM* only considers Γ_i^t for each active user i at any time slot t. Besides, the size of Γ_i^t is non-increasing with time. Thus, a narrow time window will not increase a user's chance to be a winner. Therefore, a user has no incentive to submit a false time window.

At last, a false cost function $c_i(\cdot) \neq \tilde{c}_i(\cdot)$ can only affect the result of (8) according to *SOM*. Let $\mathcal{A}_i^t$ and $\tilde{\mathcal{A}}_i^t$ denote the assignments to i when i submits $c_i(\cdot)$ and the true cost function $\tilde{c}_i(\cdot)$, respectively. It is obvious that the utility of i will not change if $\mathcal{A}_i^t = \tilde{\mathcal{A}}_i^t$. If $\mathcal{A}_i^t \neq \tilde{\mathcal{A}}_i^t$, we have $p_{i,\tilde{\mathcal{A}}_i^t} - \tilde{c}_i(\tilde{\mathcal{A}}_i^t) \geq p_{i,\mathcal{A}_i^t} - \tilde{c}_i(\mathcal{A}_i^t)$ because of both $\mathcal{A}_i^t$ and $\tilde{\mathcal{A}}_i^t$ are the subset of Γ_i^t, and $\tilde{\mathcal{A}}_i^t$ is the bundle maximizing i's utility. Thus, user i cannot increase its utility by submitting $c_i(\cdot)$.

Therefore, *SOM* is truthful. ■

Lemma 7: SOM is Sybil-proof.

Proof: We assume that user i pretends two identities i' and i'' who are assigned $\mathcal{A}_{i'}$ and $\mathcal{A}_{i''}$, respectively. Let $u_i(\mathcal{A}_i)$ denote the utility of user i when assigned $\mathcal{A}_i$. For any time slot $t \in [a_i, d_i]$, we have $u_i(\mathcal{A}_i^t) \geq u_i(\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t)$ since $\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t \subseteq \Gamma_i^t$ and *SOM* assigns i the bundle that maximizes its utility. Next, we prove that $u_i(\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t) \geq p_{i'}^t + p_{i''}^t - c_i(\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t)$. Let $m_{i'}$ denote $\max_{j\neq i',\mathcal{B}\in\mathbb{B}_j^t(\mathcal{A}_{i'}^t)}(V(\mathcal{B}) - c_j(\mathcal{B}))$, and $m_{i''}$ denote $\max_{j\neq i'',\mathcal{B}\in\mathbb{B}_j^t(\mathcal{A}_{i''}^t)}(V(\mathcal{B}) - c_j(\mathcal{B}))$. By (7), the payments to i' and i'' at any time slot t are

$$p_{i'}^t = p_{i',\mathcal{A}_{i'}^t}^t = V(\mathcal{A}_{i'}^t) - \max\{0, m_{i'}\},$$
$$p_{i''}^t = p_{i'',\mathcal{A}_{i''}^t}^t = V(\mathcal{A}_{i''}^t) - \max\{0, m_{i''}\}.$$

Let $\tilde{\mathcal{A}}_i^t = \mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t$, the payment to i when assigned $\tilde{\mathcal{A}}_i^t$ is

$$p_{i,\tilde{\mathcal{A}}_i^t}^t = V(\tilde{\mathcal{A}}_i^t) - \max\left\{0, \max_{j\neq i,\mathcal{B}\in\mathbb{B}_j^t(\tilde{\mathcal{A}}_i^t)}(V(\mathcal{B}) - c_j(\mathcal{B}))\right\}.$$

In addition, we know that

$$\bigcup_{\forall j\in\mathcal{S}^t, j\neq i}\mathbb{B}_j^t(\tilde{\mathcal{A}}_i^t) = \bigcup_{\forall j\in\mathcal{S}^t, j\neq i'}\mathbb{B}_j^t(\mathcal{A}_{i'}^t) \cup \bigcup_{\forall j\in\mathcal{S}^t, j\neq i''}\mathbb{B}_j^t(\mathcal{A}_{i''}^t).$$

Let m_i denote $\max_{j\neq i,\mathcal{B}\in\mathbb{B}_j^t(\tilde{\mathcal{A}}_i^t)}(V(\mathcal{B}) - c_j(\mathcal{B}))$. Thus we have $m_i = \max\{m_{i'}, m_{i''}\}$, and thus $m_i \leq m_{i'} + m_{i''}$. In addition, we can prove that $\mathcal{A}_{i'}^t \cap \mathcal{A}_{i''}^t = \emptyset$ by contradiction. Assume $\mathcal{A}_{i'}^t \cap \mathcal{A}_{i''}^t \neq \emptyset$, the payments to i' is

$$p_{i',\mathcal{A}_{i'}^t}^t = V(\mathcal{A}_{i'}^t) - \max\{0, m_{i'}\} \leq V(\mathcal{A}_{i'}^t) - (V(\mathcal{A}_{i''}^t) - c_i(\mathcal{A}_{i''}^t))$$

The inequality results from the fact $V(\mathcal{A}_{i''}^t) - c_i(\mathcal{A}_{i''}^t) \leq m_{i'}$, since $\mathcal{A}_{i''}^t \in \bigcup_{\forall j\in\mathcal{S}^t, j\neq i'}\mathbb{B}_j^t(\mathcal{A}_{i'}^t)$. Similarly, we have $p_{i'',\mathcal{A}_{i''}^t}^t \leq V(\mathcal{A}_{i''}^t) - (V(\mathcal{A}_{i'}^t) - c_i(\mathcal{A}_{i'}^t))$. Therefore, the summation of the utilities of i' and i'' at any time slot t is

$$\begin{aligned}
&p_{i',\mathcal{A}_{i'}^t}^t - c_i(\mathcal{A}_{i'}^t) + p_{i'',\mathcal{A}_{i''}^t}^t - c_i(\mathcal{A}_{i''}^t) \\
&\leq V(\mathcal{A}_{i'}^t) - (V(\mathcal{A}_{i''}^t) - c_i(\mathcal{A}_{i''}^t)) - \tilde{c}_i(\mathcal{A}_{i'}^t) \\
&\quad + V(\mathcal{A}_{i''}^t) - (V(\mathcal{A}_{i'}^t) - c_i(\mathcal{A}_{i'}^t)) - c_i(\mathcal{A}_{i''}^t) = 0.
\end{aligned}$$

It implies that $\mathcal{A}_{i'}^t = \mathcal{A}_{i''}^t = \emptyset$ since *SOM* is individually rational, which contradicts the assumption. Therefore, $\mathcal{A}_{i'}^t \cap$

$\mathcal{A}_{i''}^t = \emptyset$. We also have $V(\tilde{\mathcal{A}}_i^t) = V(\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t) = V(\mathcal{A}_{i'}^t) + V(\mathcal{A}_{i''}^t)$, since $\tilde{\mathcal{A}}_i^t = \mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t$. By (7), the payment to i is

$$\begin{aligned} p_{i,\tilde{\mathcal{A}}_i^t}^t &= V(\tilde{\mathcal{A}}_i^t) - \max\{0, m_i\} \\ &\geq V(\tilde{\mathcal{A}}_i^t) - (m_{i'} + m_{i''}) \\ &= V(\mathcal{A}_{i'}^t \cup \mathcal{A}_{i''}^t) - (m_{i'} + m_{i''}) \\ &= V(\mathcal{A}_{i'}^t) - m_{i'} + V(\mathcal{A}_{i''}^t) - m_{i''} \\ &\geq p_{i',\mathcal{A}_{i'}^t}^t + p_{i'',\mathcal{A}_{i''}^t}^t. \end{aligned}$$

In addition, we have $c_i(\tilde{\mathcal{A}}_i^t) \leq c_i(\mathcal{A}_{i'}^t) + c_i(\mathcal{A}_{i''}^t)$ because of the fourth property of the cost function. By (2) and (5), the utility of i when using two identities is not greater than that obtained by using a single identity at any time slot t. User i's utility via Sybil attack is not greater than that obtained by using a single identity.

Therefore, *SOM* is Sybil-proof. ■

Remark: We can use a proof similar to that in Lemma 7 to prove that $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ for any two users i and j. In addition, *SOM* does not satisfy computational efficiency, since at each time slot t it calculates the payments to each active user i for every subset of Γ_i^t, and the time complexity is exponential to the largest $|\Gamma_i^t|$ for all $i \in \mathcal{S}^t$. In reality, however, the number of tasks each user can perform is very small because of various constraints, e.g., travel budget [21], and thus the execution time of *SOM* is still practical.

VII. Performance Evaluation

In this section, we compare the performances of *SOS* and *SOM* with three benchmarks. The first benchmark is an online mechanism adapted from [4] for SM case, denoted by *Greedy*. Note that, this mechanism is not Sybil-proof. The second benchmark is *SPIM-S* [16], which is a Sybil-proof mechanism for SM case. The third benchmark is *SPIM-M* [16], which is a Sybil-proof mechanism for MM case. The performance metrics include *total payment*, *platform utility* and *Sybil-proofness*.

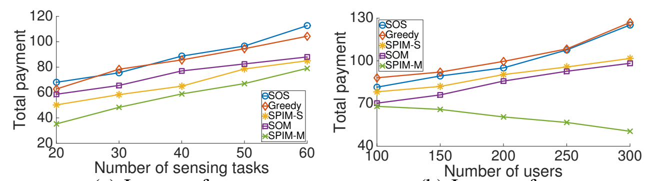
A. Evaluation Setup

For a fair comparison with *SPIM-S* and *SPIM-M*, we use the same dataset, which is a real-world dataset consisting of the traces of taxi drivers [1]. As in [16], we consider a crowdsensing system in which the tasks are measuring the Wi-Fi signal strength at specific locations. In this system, tasks are represented by GPS locations of the taxi drivers in the dataset, and users are all the taxi drivers.

In our evaluation, we randomly select locations on taxi drivers' traces as the sensing tasks. The value of each task is uniformly distributed over $[1, 5]$, and users' cost for each task is uniformly distributed over $[1, 5]$. We set the deadline (T) to 60 (min), each user i's $\tilde{a}_i$ is uniformly distributed over $[0, 60]$, and the active time window ($\tilde{d}_i - \tilde{a}_i$) of each user is uniformly distributed over $[0, 5]$. To evaluate the impact of the number of sensing tasks (m) on the performance metrics, we fix the number of users (n) at 200 and vary m from 20 to 60 with a step of 10. To evaluate the impact of the number of users on the performance metrics, we fix m at 150 and vary n from 100 to 300 with a step of 50. All the results are averaged over 1000 independent runs.

B. Evaluation of Total Payment

The impacts of m and n on the total payment to users are shown in Fig. 2. In Fig. 2 (a), we see that the total payment of both offline and online mechanisms increase with the increase of m. This is because the platform may recruit more users when m increase, and thus has a higher payment. In addition, we see that the total payment of the online mechanisms (*Greedy*, *SOS*, *SOM*) are higher than that of the offline mechanisms (*SPIM-S*, *SPIM-M*). This is because the online mechanisms may recruit more users when users arrives in different time slots, and thus has a higher payment. In Fig. 2 (b), we observe that the total payment of *SPIM-M* decrease with the increase of n. This is because, with more users, *SPIM-M* may find more low-cost users to perform the tasks. The total payment of the other four mechanisms increase slightly with the increase of n. This is because, with more users, these mechanisms may assign more tasks incurring a higher payment. In addition, we see that the total payment of the online mechanisms are higher than that of the offline mechanisms as explained before. Note that, *SOS* has a similar performance as Greedy, which is not Sybil-proof.

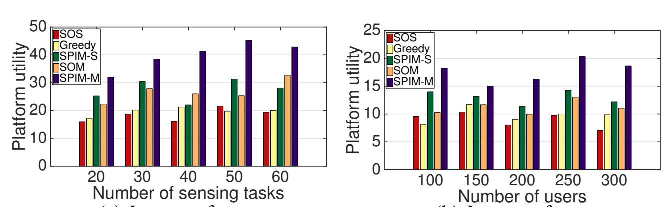


(a) Impact of m (b) Impact of n

Fig. 2. Total payment

C. Evaluation of Platform Utility

The impacts of m and n on the platform utility are shown in Fig. 3. In both Fig. 3 (a) and Fig. 3 (b), we see that the platform utilities achieved by the offline mechanisms (*SPIM-S*, *SPIM-M*) are larger than that achieved by online mechanisms (*Greedy*, *SOS*, *SOM*). This is because offline mechanisms know all users' bids before making decision, while online mechanisms have no information of future users. In addition, we see that *SOM* outperforms *SOS* in term of the platform utility. This is because *SOM* assigns each task to at most one user, and thus avoids paying users to perform duplicated tasks.

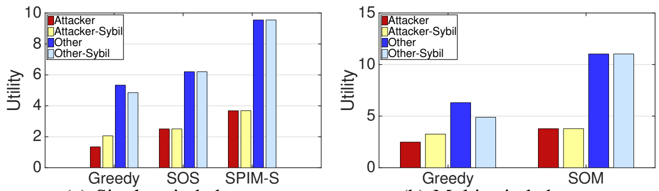


(a) Impact of m (b) Impact of n

Fig. 3. Platform utility

D. Evaluation of Sybil-proofness

Fig. 4 shows the utility of Sybil attacker and the other users without Sybil attack, denoted by *Attacker* and *Other*, and that via Sybil attack, denoted by *Attacker-Sybil* and *Other-Sybil*. In both Fig. 4 (a) and Fig. 4 (b), we see that the attacker increase its utility while decreasing the utility of others' in *Greedy*. This is because *Greedy* is vulnerable to Sybil attack. However, the attacker cannot increase its utility in *SPIM-S*, *SOS* and *SOM*, since they are Sybil-proof.



(a) Single-minded case (b) Multi-minded case

Fig. 4. Sybil-proofness

VIII. Conclusion

In this paper, we demonstrated that existing online incentive mechanisms for crowdsensing are all vulnerable to Sybil attack. We proposed two Sybil-proof online incentive mechanisms *SOS* and *SOM* for single-minded case and multi-minded case, respectively. Specifically, *SOS* achieves computational efficiency, individual rationality, truthfulness and Sybil-proveness. *SOM* achieves individual rationality, truthfulness and Sybil-proveness. We rigorously proved the desired properties of the mechanisms and evaluated their performances through extensive simulations.

References

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

[2] J. R. Douceur, "The sybil attack," in *IPTPS*, 2002, pp. 251–260.

[3] 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 *INFOCOM*, 2014, pp. 1231–1239.

[4] Z. Feng, Y. Zhu, Q. Zhang, H. Zhu, J. Yu, J. Cao, and L. M. Ni, "Towards truthful mechanisms for mobile crowdsourcing with dynamic smartphones," in *ICDCS*, 2014, pp. 11–20.

[5] L. Gao, F. Hou, and J. Huang, "Providing long-term participation incentive in participatory sensing," in *INFOCOM*, 2015, pp. 2803–2811.

[6] "Gigwalk," http://www.gigwalk.com/.

[7] B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Computing Surveys*, vol. 48, no. 1, p. 7, 2015.

[8] K. Han, Y. He, H. Tan, S. Tang, H. Huang, and J. Luo, "Online pricing for mobile crowdsourcing with multi-minded users," in *MobiHoc*, 2017, pp. 18:1–18:10.

[9] K. Han, C. Zhang, J. Luo, M. Hu, and B. Veeravalli, "Truthful scheduling mechanisms for powering mobile crowdsensing," *IEEE Trans. Comput.*, vol. 65, no. 1, pp. 294–307, 2016.

[10] L. G. Jaimes, I. Vergara-Laurens, and A. Raij, "A crowd sensing incentive algorithm for data collection for consecutive time slot problems," in *LATINCOM*, 2014, pp. 1–5.

[11] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *MobiHoc*, 2015, pp. 167–176.

[12] H. Jin, L. Su, B. Ding, K. Nahrstedt, and N. Borisov, "Enabling privacy-preserving incentives for mobile crowd sensing systems," in *ICDCS*, 2016, pp. 344–353.

[13] H. Jin, L. Su, and K. Nahrstedt, "Theseus: Incentivizing truth discovery in mobile crowd sensing systems," in *MobiHoc*, 2017, pp. 1:1–1:10.

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

[15] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *INFOCOM*, 2013, pp. 1402–1410.

[16] J. Lin, M. Li, D. Yang, G. Xue, and J. Tang, "Sybil-proof incentive mechanisms for crowdsensing," in *INFOCOM*, 2017, pp. 2088–2096.

[17] J. Lin, D. Yang, M. Li, J. Xu, and G. Xue, "BidGuard: A framework for privacy-preserving crowdsensing incentive mechanisms," in *CNS*, 2016, pp. 439–447.

[18] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda, "Peir, the personal environmental impact report, as a platform for participatory sensing systems research," in *MobiSys*, 2009, pp. 55–68.

[19] "Opensignal," http://opensignal.com/.

[20] D. Peng, F. Wu, and G. Chen, "Pay as how well you do: A quality based incentive mechanism for crowdsensing," in *MobiHoc*, 2015.

[21] L. Pournajaf, L. Xiong, V. Sunderam, and S. Goryczka, "Spatial task assignment for crowd sensing with cloaked locations," in *MDM*, vol. 1, 2014, pp. 73–82.

[22] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, and W. Hu, "Ear-phone: an end-to-end participatory urban noise mapping system," in *IPSN*, 2010, pp. 105–116.

[23] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based sybil defenses," in *SIGCOMM*, 2010, pp. 363–374.

[24] B. Wang, L. Zhang, and N. Z. Gong, "SybilSCAR: Sybil detection in online social networks via local rule based propagation," in *INFOCOM*, 2017, pp. 1099–1107.

[25] G. Wang, B. Wang, T. Wang, A. Nika, H. Zheng, and B. Y. Zhao, "Defending against sybil devices in crowdsourced mapping services," in *MobiSys*, 2016, pp. 179–191.

[26] Q. Wang, B. Ye, B. Tang, S. Guo, and S. Lu, "eBay in the clouds: False-name-proof auctions for cloud resource allocation," in *ICDCS*, 2015, pp. 153–162.

[27] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and G. Wu, "An incentive mechanism with privacy protection in mobile crowdsourcing systems," *Computer Networks*, vol. 102, pp. 157–171, 2016.

[28] "Waze," https://www.waze.com/.

[29] Y. Wen, J. Shi, Q. Zhang, X. Tian, Z. Huang, H. Yu, Y. Cheng, and X. Shen, "Quality-driven auction based incentive mechanism for mobile crowd sensing," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4203–4214, 2015.

[30] J. Xu, J. Xiang, and D. Yang, "Incentive mechanisms for time window dependent tasks in mobile crowdsensing," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6353–6364, 2015.

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

[32] D. Yang, G. Xue, X. Fang, and J. Tang, "Incentive mechanisms for crowdsensing: Crowdsourcing with smartphones," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1732–1744, 2016.

[33] M. Yokoo, "Characterization of strategy/false-name proof combinatorial auction protocols: price-oriented, rationing-free protocol," in *IJCAI*, 2003, pp. 733–742.

[34] M. Yokoo, Y. Sakurai, and S. Matsubara, "The effect of false-name bids in combinatorial auctions: New fraud in internet auctions," *Games and Economic Behavior*, vol. 46, no. 1, pp. 174–188, 2004.

[35] Q. Zhang, Y. Wen, X. Tian, X. Gan, and X. Wang, "Incentivize crowd labeling under budget constraint," in *INFOCOM*, 2015, pp. 2812–2820.

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

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

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

[39] D. Zhao, X.-Y. Li, and H. Ma, "Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 647–661, 2016.

[40] Z. Zheng, Y. Peng, F. Wu, S. Tang, and G. Chen, "An online pricing mechanism for mobile crowdsensing data markets," in *MobiHoc*, 2017, pp. 26:1–26:10.