A Truthful Incentive Mechanism for Mobile Crowdsourcing

by

Boya Song

B.E., Zhejiang University, 2013

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2015

Abstract

In a mobile crowdsourcing system, the platform utilizes ubiquitous smartphones to perform sensing tasks. For a successful mobile crowdsourcing application, the consideration of the heterogeneity of quality of sensing from different users as well as proper incentive mechanism to motivate users to contribute to the system are essential. In this thesis, we introduce quality of sensing into incentive mechanism design. Under a budget constraint, the platform aims to maximize the valuation of the performed tasks, which depends on the quality of sensing of the users. We propose ABSee, an auction-based budget feasible mechanism, which consists of a winning bid selection rule and a payment determination rule. We obtain the approximation ratio of ABSee, which significantly improves the approximation ratio of existing budget feasible mechanisms. ABSee also satisfies the properties of computational efficiency, truthfulness, individual rationality, and budget feasibility. Extensive simulation results show that ABSee provides a higher valuation to the platform when compared with an existing mechanism in the literature.

Preface

I hereby declare that I am the author of this thesis. This thesis is an original, unpublished work under the supervision of Professor Vincent W.S. Wong.

Table of Contents

List of Figures

List of Acronyms

ABSee	Auction-based budget feasible mechanism
GPS	Global Positioning System
CDF	Cumulative Distribution Function
LTE	Long Term Evolution

List of Symbols

$\mathcal{N}$	The set of smartphone users
N	The number of users
Γ	The set of tasks
τ_k	A task in set Γ
M	The number of tasks
$\mathcal{B}_n$	The set of bids user n submits
σ_n	The total number of bids user n submits
$\mathcal{J}_n$	Integer set $\{1, \ldots, \sigma_n\}$
Γ_n^m	A subset of tasks user n can perform
c_n^m	The cost of user n for performing tasks in set Γ_n^m
b_n^m	The reserve price of user n for performing tasks in set Γ_n^m
B_n^m	A bid in $\mathcal{B}_n$, i.e., (Γ_n^m, b_n^m)

$\mathcal{B}$	The set of all bids from all users
$\mathcal{S}$	The set of winning bids
$\mathcal{R}_n(\mathcal{S})$	The set of tasks within the winning bids $\mathcal{S}$ of user n
p_n	The payment to user n
u_n	The utility of user n
q_n	The quality indicator of user n
L	The total number of auctions in which user n is a winner
$\delta_k^{(l)}$	The true value of task τ_k in the lth auction
$\hat{\delta}_k^{(l)}$	The estimated value of task τ_k in the lth auction
$\delta_{k,n}^{(l)}$	The value of task τ_k obtained from user n in the lth auction
$\hat{q}_n^{(l)}$	The quality indicator of user n obtained in the lth auction
$\bar{q}_n^{(l)}$	The quality indicator stored in the historical record after the lth auction
γ	The weight for the most recent quality indicator
$g_k(\mathcal{S})$	The quality of sensing for task τ_k given the winning bids $\mathcal{S}$
$V(\mathcal{S})$	The valuation function the platform obtained from the winning bids $\mathcal{S}$

μ_k	The weight of task τ_k
G	The budget of the platform
$\mathcal{F}$	The winning bid selection rule
$\mathcal{P}$	The payment determination rule
$\boldsymbol{p}$	The payment vector, i.e., $(p_1, \ldots, p_N)$
x_i	The ith winning bid
$\bar{b}^{x_i}$	The reserve price for bid x_i
$\bar{p}^{x_i}$	The payment for bid x_i
$\bar{c}^{x_i}$	The cost for bid x_i
$\bar{\Gamma}^{x_i}$	The subsets of tasks within bid x_i
w	The number of winning bids in $\mathcal{S}$
$V_{x_i}(\mathcal{S}_{i-1})$	The marginal contribution of bid $x_i \in \mathcal{B} \setminus \mathcal{S}_{i-1}$ given a bid subset $\mathcal{S}_{i-1} \subseteq \mathcal{B}$.
θ	The crowd factor of the system
i_k	The kth winning bid when bid x_i is removed from the system
$\mathcal{Q}$	The winning bid set after remove a winning bid

w'	The number of winning bids in $\mathcal{Q}$
$\beta_{i(k)}$	The highest reserve price that the user who has the winning bid x_i can submit to replace bid i_k with x_i in position k
$\rho_{i(k)}$	The highest reserve price that the user who has the winning bid x_i can submit so that the marginal contribution per reserve price of x_i is not less than a selection threshold
α	Approximation ratio
$\widetilde{V}(\mathcal{S})$	The valuation of the platform obtained from the fractional greedy algorithm
$\bar{V}(\mathcal{S}^{(l)})$	The valuation of the platform obtained from the estimated quality indicators in the lth auction

Acknowledgements

I would like to express my deepest appreciation to my supervisor, Prof. Vincent Wong, for his invaluable guidance that helped me to shape my research, and for his persistent encouragement and patience that made me fulfill my master program. Without Prof. Wong's guidance and support, this thesis would have not been possible.

I want to thank my senior Dr. Hamed Shah-Mansouri, who has given me constructive comments and help to my research work. I also would like to thank my colleagues in the Communication Lab: Zehua Wang, Bojiang Ma, Binglai Niu, Zhiyu Dai, Enxin Yao, Yanan Sun, who have provided me with generous support and great company. I sincerely offer my best wishes and regards to all those who have helped me during my graduate study. Finally, I own my heartfelt gratitude to my beloved parents. Without their love, support and understanding, I would have not harvested my invaluable life experiences at UBC.

This research is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Chapter 1

Introduction

In this chapter, we first introduce the background of mobile crowdsourcing and two important issues in mobile crowdsourcing. Then, we present the motivation and contribution of our work. The structure of the thesis is shown at the end of this chapter.

1.1 Mobile Crowdsourcing

Smartphones nowadays are equipped with a variety of sensors (e.g., microphone, camera, global positioning system (GPS)) and have enhanced sensing capabilities. Mobile crowdsourcing exploits the ubiquity of smartphones and utilizes the sensors to monitor the environment [1]. In a mobile crowdsourcing system, the platfrom distributes the sensing tasks to the smartphone users to collect data. Smartphone users perform the tasks and send the results to the platform. Different kind of mobile crowdsourcing applications have been developed. They include monitoring the environment [2, 3, 4], creating wireless network coverage maps [5, 6], and updating the traffic condition [7].

1.2 Incentive Mechanisms and Quality of Sensing

A key factor for a successful mobile crowdsourcing application is the participation of a large number of users. However, performing sensing tasks incurs costs on the smartphone users, such as energy consumption for data sensing, packet transmission charges from the service operator, and manual efforts. In order to motivate smartphone users to participate in mobile crowdsourcing, the platform should make payments to the users to compensate their costs. The applications in [2, 3, 4, 5, 6, 7] are based on voluntary participation of smartphone users, where their participation cannot be guaranteed.

Another important issue in mobile crowdsourcing is that the users provide different quality of sensing. For example, consider a platform which monitors the noise level of a city. Each sensing task corresponds to obtaining the noise level at a sampling location in the city by utilizing the microphone of smartphones. For a specific user, the quality of sensing depends on the type of microphone on his smartphone (e.g., accuracy, calibration) and his manual effort (e.g., taking the smartphone out of his pocket to collect the sample). If the platform intends to obtain a high accuracy of the noise measurement, then it needs to assign tasks and provide payments to the users who can provide a high quality of sensing. However, the quality of sensing of users may not be known by the platform *a priori*. Thus, it is required for the platform to estimate the quality of sensing of users accurately and then select the users accordingly.

1.3 Motivation

Most of the existing works focus on either incentive mechanism design to guarantee the participation level of users (e.g., [8, 9, 10, 11, 12, 13]) or estimation of quality of sensing (e.g., [14, 15, 16, 17]). In this thesis, we consider these two issues jointly. We study a practical scenario where a platform with limited budget aims to maximize the valuation of performed tasks, which depends on the quality of sensing of the users. The platform adopts an incentive mechanism to motivate users to participate. We model the interaction between the platform and the users as an auction and design a winning bid selection rule and a payment determination rule. The platform first announces its tasks to the smartphone users. The users then submit their bids to the platform. Each bid consists of a subset of tasks that the user can perform and the price that the user asks for these tasks. The platform selects the winning bids. Then the users who have at least one winning bid will perform the tasks and receive payment accordingly.

The budget constraint introduces inter-dependence between the winning bid selection rule and payment determination rule. It makes the design of auction mechanism challenging. Although the existing budget feasible mechanisms, e.g., [18, 19], are applicable for any system, they do not take into account the large number of users in a mobile crowdsourcing system. In this case, these mechanisms may not always provide a high valuation to the platform. In our work, we consider a practical scenario where there are many participating users and winning bids in the mobile crowdsourcing system. We capture this property by defining a *crowd factor* θ, where $0 < \theta < 1$. It describes the relative contribution a winning bid can make to the platform. When more bids are submitted and selected as winning bids, the relative contribution

that a winning bid can make to the system decreases. We use the crowd factor to design a novel budget feasible mechanism. The mechanism is applicable not only to mobile crowdsourcing systems but also to other systems where there are a large number of users.

1.4 Contributions

Our main contributions are summarized as follows:

- **Novel model**: We consider quality of sensing and incentive mechanism design jointly by introducing quality of sensing into the valuation function of the platform, which is proven to be a submodular function. Although it is difficult to measure the quality of sensing of a user, it can be estimated accurately in our model. We focus on tasks with continuous values in this thesis, but our model is also suitable for tasks with discrete values, including tasks with binary values.

- **New mechanism**: We propose an Auction-based Budget feaSiblE mEchanism, which is called ABSee. We select the winning bids by employing an iterative approach and using the crowd factor. We rigorously prove that ABSee satisfies the properties of computational efficiency, truthfulness, individual rationality, and budget feasibility.

- **Approximation ratio**: To tackle the computational complexity of the auction design, ABSee adopts a greedy approach. The approximation ratio, which is the ratio between the optimal valuation and the one obtained by the proposed mechanism, is equal to $\frac{2e}{\theta(e-1)}$. Based on the example we provide and simulation results, the approximation ratio can

approach $\frac{2e}{e-1}$ (≈ 3.16) in mobile crowdsourcing systems, which improves the approximation ratio of $\frac{5e}{e-1}$ (≈ 7.91) in the budget feasible mechanism proposed in [19].

- **Extensive simulation**: Simulation results show that a higher valuation can be obtained by adopting ABSee when compared with the budget feasible mechanism in [19] under different scenarios. Moreover, it is shown that the quality of sensing can be estimated accurately.

1.5 Structure of the Thesis

This thesis is organized as follows. In Chapter 2, we introduce the related work. In Chapter 3, we present the budget feasible mechanism, analyze its properties and approximation ratio, and evaluate its performance. Conclusion and future work are given in Chapter 4.

Chapter 2

Related Work

In this chapter, we introduce the related work in incentive mechanism design, quality of sensing, and budget feasible mechanisms.

2.1 Incentive Mechanisms

Different forms of incentive mechanisms have been proposed in the literature. A platform-centric model and a user-centric model are proposed in [8] to maximize the utility of the platform, which is the valuation of performed tasks minus the total payment to the users. An all-pay auction is designed in [11] to maximize the utility, where the users' population is stochastic. In [9], an auction mechanism is proposed to minimize the social cost of smartphone users. An online auction mechanism is designed in [13] for dynamic arrival and departure of smartphone users, with the goal of maximizing the valuation of performed tasks. Auction mechanisms are proposed in [10] for mobile crowdsourcing systems considering the cooperation and competition among the users. The platform is guaranteed to obtain non-negative utility. An incentive mechanism is presented in [12] to provide the long-term participation incentives, which guarantees the users will participate for many times. However, quality of sensing is considered in

none of the mentioned works.

2.2 Quality of Sensing

Existing works in this area mainly focus on tasks with discrete values. An efficient budget allocation algorithm is proposed in [14] to guarantee the accuracy of estimation of tasks with binary values. An online learning approach is adopted in [15] to maximize the quality of sensing to guarantee the robustness of the crowdsourcing system. In the data mining area, the work in [16] uses expectation maximization and maximum likelihood estimation to determine the quality of sensing of tasks with binary values in social sensing applications. Davami *et al.* in [20] compared five trust prediction algorithms and implemented the most accurate one to calculate the quality of sensing of users for a parking space finding application. Beyond that, an optimization problem is formulated in [17] to obtain the estimated value of the tasks and minimize the estimation error by considering tasks with both continuous values and discrete values.

2.3 Incentive Mechanisms and Quality of Sensing

Both incentive mechanism and quality of sensing are considered in a few works. An incentive mechanism is proposed in [21] to guarantee the quality of sensing. However, the platform does not have a budget constraint and only has one task. An auction mechanism is designed in [22] to maximize the valuation of the platform, where the quality of sensing of each user is assumed

to be known by the platform. A sequential Bayesian approach is used in [23] to determine the quality of sensing of each user but the auction mechanism is designed specifically for tasks with binary values with prior information of quality of sensing of the users. A quality-based incentive mechanism is designed in [24]. However, it does not consider the strategic behavior of users, who are interested in maximizing their own utilities and may adopt strategies to manipulate the mechanism. In contrast to the above mentioned works, given the unknown quality of sensing of users, we consider strategic users and propose an incentive mechanism for the platform under a budget constraint.

2.4 Budget Feasible Mechanisms

Budget feasible mechanisms for submodular functions are proposed by Singer [18] and improved by Chen *et al.* [19]. Recently, these mechanisms have been used in several applications. For example, an influence maximization problem is studied by utilizing a coverage model and a budget feasible mechanism in [25]. In our work, we propose a new budget feasible mechanism by utilizing the crowd factor of the system. The proposed mechanism can improve the approximation ratio of other budget feasible mechanisms, summarized in Table 2.1. From the example and simulation experiments to be presented in Sections 3.3 and 3.4, respectively, we show that θ is close to 1 and the approximation ratio approaches 3.16.

Table 2.1: Comparison of budget feasible mechanisms

	AR	**Setting**
[18]	117.7	A mechanism for general submodular functions
[19]	7.91	An improved mechanism for general submodular functions
[25]	31	A mechanism for coverage model (i.e., a special case of submodular functions)
ABSee	$\frac{3.16}{\theta}$	Our proposed mechanism for general submodular functions

AR: Approximation ratio.

Chapter 3

Budget Feasible Mechanism for Mobile Crowdsourcing

In this chapter, we study an auction-based incentive mechanism which considers the heterogeneity of quality of sensing of users. We first present the system model for the auction mechanism and quality of sensing. Then we propose ABSee and provide a walk-through example of ABSee. We prove the mentioned properties of ABSee and analyze the approximation ratio of ABSee. Simulation results show the performance of ABSee outperforms an existing mechanism in the literature and the quality of sensing can be estimated accurately.

3.1 System Model and Problem Formulation

3.1.1 Auction Framework

A mobile crowdsourcing system, as shown in Fig. 3.1, is composed of a platform residing in the cloud computing center and many smartphone users. Users are connected with the cloud via wireless access networks (e.g., WiFi, LTE). The set of smartphone users is denoted by

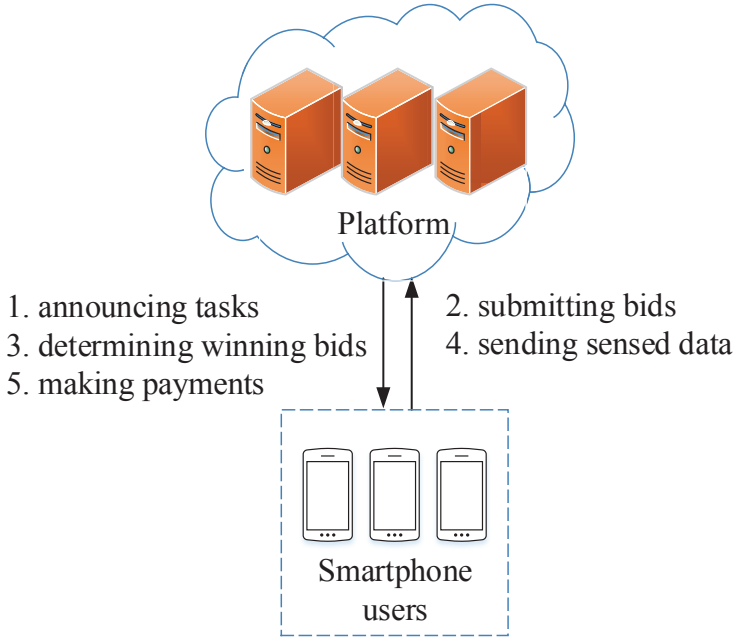

Figure 3.1: A mobile crowdsourcing system consists of a platform and many smartphone users. The auction mechanism is modeled as an interactive process between the platform and the users.

$\mathcal{N} = \{1, \ldots, N\}$, where N is the number of users. The platform announces the sensing tasks to the smartphone users. We use $\Gamma = \{\tau_1, \ldots, \tau_M\}$ to denote the set of tasks and there are M tasks in total.

We model the interactive process between the platform and the users as an auction. Each user $n \in \mathcal{N}$ submits a set of bids $\mathcal{B}_n$. We use $\sigma_n = |\mathcal{B}_n|$ to denote the total number of bids submitted by user n. Each bid is denoted as $B_n^m = (\Gamma_n^m, b_n^m)$, $m \in \mathcal{J}_n = \{1, \ldots, \sigma_n\}$, in which $\Gamma_n^m \subseteq \Gamma$ is a subset of tasks user n can perform, and b_n^m is the corresponding reserve price, which is the price user n asks to be compensated for the cost c_n^m for performing the tasks. Thus, $\mathcal{B}_n = \{B_n^1, \ldots, B_n^{\sigma_n}\} = \{(\Gamma_n^1, b_n^1), \ldots, (\Gamma_n^{\sigma_n}, b_n^{\sigma_n})\}$. We assume the subsets of tasks submitted by user n are pairwise disjoint.

The platform collects all bids $\mathcal{B} = \bigcup_{n \in \mathcal{N}} \mathcal{B}_n$ from the users. Then, it selects a subset $\mathcal{S} \subseteq \mathcal{B}$ as the *winning bids*. Let $\mathcal{R}_n(\mathcal{S}) \subseteq \Gamma$ denote the set of the tasks within the winning bids of user n, i.e.,

$$\mathcal{R}_n(\mathcal{S}) = \left\{ \bigcup_{m \in \mathcal{J}_n} \Gamma_n^m \mid B_n^m = (\Gamma_n^m, b_n^m) \in \mathcal{S} \right\}. \tag{3.1}$$

User n is called a *winner* if $\mathcal{R}_n(\mathcal{S}) \neq \emptyset$. Each winner n performs the tasks within $\mathcal{R}_n(\mathcal{S})$ and sends the sensed data to the platform. The platform then makes payment p_n to winner n for its contribution.

Users are assumed to be strategic. For each bid $B_n^m = (\Gamma_n^m, b_n^m) \in \mathcal{B}_n$, user n decides a reserve price b_n^m for Γ_n^m to maximize its own utility. The utility of user $n \in \mathcal{N}$ is

$$u_n = \begin{cases} p_n - \sum\limits_{m \in \mathcal{J}_n \,:\, \Gamma_n^m \in \mathcal{R}_n(\mathcal{S})} c_n^m, & \text{if } n \text{ is a winner,} \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Note that b_n^m determines whether bid B_n^m of user n can be selected as one of the winning bids or not.

3.1.2 Quality of Sensing

In a mobile crowdsourcing system, the quality of sensing of a user depends on his effort and expertise to perform the tasks and the quality of the sensors of his smartphone. The quality of sensing of user $n \in \mathcal{N}$, i.e., the accuracy of the sensed data, is modeled by a *quality indicator* $q_n > 0$. It is not known *a priori* and needs to be calculated. Notice that user n may have

participated in the auction multiple times if the platform conducts the auction many times. For example, recall the noise map application in Chapter 1. The platform may want to determine the noise level of some locations in the city at different times for several days. It keeps a historical record of the quality indicators of different users. The platform collects all sensed data from the selected winning bids and calculates the estimated value of the tasks. It then measures the quality indicators of the winners and updates the quality indicators in its historical record.

The above model works for any kind of tasks. Consider tasks with continuous values as an example. Assume user n has participated in the auction and has been a winner in L auctions in the past. In the lth auction, where $l \in \{1, \ldots, L\}$, we use $\hat{\delta}_k^{(l)}$ and $\hat{\delta}_{k,n}^{(l)}$ to denote the estimated value of task τ_k performed by the winners and the value of task τ_k obtained from user n, respectively. The platform can estimate $\hat{\delta}_k^{(l)}$ accurately by adopting a truth discovery approach [17]. Let $\hat{q}_n^{(l)}$ denote the quality indicator of user n obtained in the lth auction. Then, $\hat{q}_n^{(l)}$ can be measured as the variance of sensed data provided by user n for tasks within its winning bids:

$$\hat{q}_n^{(l)} = \frac{1}{|\mathcal{R}_n(\mathcal{S}^{(l)})|} \sum_{\tau_k \in \mathcal{R}_n(\mathcal{S}^{(l)})} \left(\hat{\delta}_{k,n}^{(l)} - \hat{\delta}_k^{(l)}\right)^2, \tag{3.3}$$

where $\mathcal{R}_n(\mathcal{S}^{(l)})$ is the set of tasks within the winning bids of user n in the lth auction. Let $\bar{q}_n^{(l)}$ denote the quality indicator stored in the historical record after the lth auction. Then, $\bar{q}_n^{(l)}$ can be obtained as

$$\bar{q}_n^{(l)} = \gamma \hat{q}_n^{(l)} + (1-\gamma)\bar{q}_n^{(l-1)},\ l \in \{1, \ldots, L\}, \tag{3.4}$$

where constant $0 < \gamma < 1$ is the weight for the most recent quality indicator. Notice that $\bar{q}_n^{(0)}$

represents the initial value of quality indicator, which depends on the mobile crowdsourcing application. After the platform has conducted the auction for many rounds, the estimated value approaches the true value. For the sake of simplicity, we abuse the notation and remove the round of auction (i.e., (l)) and use q_n, since we focus on a specific round of auction.

We use $g_k(\mathcal{S})$ to denote the quality of sensing for task τ_k given the winning bids $\mathcal{S}$. It represents the accuracy of the estimated value of task τ_k after aggregating all sensed data from the winning bids. For tasks with continuous values, the accuracy of the estimated value can be defined by the mean square estimation error. We calculate $g_k(\mathcal{S})$ by adopting the maximum likelihood estimation [26]:

$$g_k(\mathcal{S}) = \left(\sum_{n \in \mathcal{N}: \tau_k \in \mathcal{R}_n(\mathcal{S})} \frac{1}{q_n} \right)^{-1}. \tag{3.5}$$

A smaller value of $g_k(\mathcal{S})$ represents a higher accuracy of the estimated value of task τ_k, i.e., better quality of sensing. We illustrate the model for quality of sensing by using tasks with continuous values, but our model is suitable for tasks with discrete values as well by updating (3.3) and (3.5). In statistics, for tasks with discrete values, $\hat{q}_n^{(l)}$ can be estimated by 0-1 loss function or squared loss function and $g_k(\cdot)$ can denote the estimation error rate [17].

3.1.3 Problem Formulation

Each task has a different weight, which can be regarded as the importance of that task to the platform. We use $\mu_k > 0$ to denote the weight of task τ_k. Let $V(\mathcal{S})$ denote the *valuation*

function obtained from the winning bids $\mathcal{S}$. From (3.5), we have

$$\begin{aligned} V(\mathcal{S}) &= \sum_{\tau_k \in \bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S})} \mu_k \log\left(1 + g_k(\mathcal{S})^{-1}\right) \\ &= \sum_{\tau_k \in \bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S})} \mu_k \log\left(1 + \sum_{n\in\mathcal{N}:\tau_k\in\mathcal{R}_n(\mathcal{S})} \frac{1}{q_n}\right). \end{aligned} \tag{3.6}$$

The log term in (3.6) reflects the diminishing marginal returns on the quality of sensing. We first define the submodular function and then prove that $V(\mathcal{S})$ is a non-negative non-decreasing submodular function.

Definition 3.1. *For a finite set $\mathcal{Y}$, the function $f\colon 2^{\mathcal{Y}} \mapsto \mathbb{R}$ is submodular if [27]*

$$f(\mathcal{C} \cup \{y\}) - f(\mathcal{C}) \geq f(\mathcal{D} \cup \{y\}) - f(\mathcal{D}),$$

for any $\mathcal{C} \subseteq \mathcal{D} \subseteq \mathcal{Y}$ and $y \in \mathcal{Y} \setminus \mathcal{D}$. Moreover, a submodular function f is non-decreasing if $f(\mathcal{C}) \leq f(\mathcal{D})$ for any $\mathcal{C} \subseteq \mathcal{D}$.

Lemma 3.1. *The valuation $V(\mathcal{S})$ in (3.6) is a non-negative non-decreasing submodular function.*

Proof. From Definition 3.1, we need to show that

$$V(\mathcal{S} \cup \{B_i^j\}) - V(\mathcal{S}) \geq V(\mathcal{Z} \cup \{B_i^j\}) - V(\mathcal{Z}),$$

for any $\mathcal{S} \subseteq \mathcal{Z} \subseteq \mathcal{B}$ and $B_i^j \in \mathcal{B} \setminus \mathcal{Z}$, where $i \in \mathcal{N}$, $j \in \mathcal{J}_i$, and $\mathcal{B}$ is the set of all submitted

bids. To simplify the expression, let $l_n = \frac{1}{q_n}$, $n \in \mathcal{N}$. According to (3.6), we have

$$V(\mathcal{S}) = \sum_{\tau_k \in \bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S})} \mu_k \log\left(1 + \sum_{n\in\mathcal{N}:\tau_k\in\mathcal{R}_n(\mathcal{S})} l_n\right).$$

Recall $\mathcal{R}_n(\mathcal{S})$ from (3.1) and we denote $\mathcal{R}_n(\mathcal{Z}) = \left\{\bigcup_{m\in\mathcal{J}_n} \Gamma_n^m \mid B_n^m = (\Gamma_n^m, b_n^m) \in \mathcal{Z}\right\}$. Given sets $\mathcal{S}$ and $\mathcal{Z}$, let $s_k = \sum_{n\in\mathcal{N}:\tau_k\in\mathcal{R}_n(\mathcal{S})} l_n$, and $z_k = \sum_{n\in\mathcal{N}:\tau_k\in\mathcal{R}_n(\mathcal{Z})} l_n$. We define

$$\Phi_1 \triangleq \Gamma_i^j \bigcap (\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S})) \bigcap (\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{Z})),$$

$$\Phi_2 \triangleq \Gamma_i^j \bigcap (\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{Z})) \setminus (\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S}),$$

and

$$\Phi_3 \triangleq \Gamma_i^j \setminus \left((\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{Z})) \bigcup (\bigcup_{n\in\mathcal{N}} \mathcal{R}_n(\mathcal{S})) \right).$$

We have $\Phi_1 \bigcup \Phi_2 \bigcup \Phi_3 = \Gamma_i^j$, while $\Phi_1 \bigcap \Phi_2 = \Phi_2 \bigcap \Phi_3 = \Phi_1 \bigcap \Phi_3 = \emptyset$. Then,

$$\begin{aligned} V(\mathcal{S} \cup \{B_i^j\}) - V(\mathcal{S}) = &\sum_{\tau_k\in\Phi_1} \mu_k \log\left(1 + \frac{l_i}{1+s_k}\right) \\ &+ \sum_{\tau_k\in\Phi_2} \mu_k \log(1+l_i) + \sum_{\tau_k\in\Phi_3} \mu_k \log(1+l_i), \end{aligned}$$

$$\begin{aligned} V(\mathcal{Z} \cup \{B_i^j\}) - V(\mathcal{Z}) = &\sum_{\tau_k\in\Phi_1} \mu_k \log\left(1 + \frac{l_i}{1+z_k}\right) \\ &+ \sum_{\tau_k\in\Phi_2} \mu_k \log\left(1 + \frac{l_i}{1+z_k}\right) + \sum_{\tau_k\in\Phi_3} \mu_k \log(1+l_i). \end{aligned}$$

The reason for the difference between the second terms in the above equations is that for any task $\tau_k \in \Phi_2$, it can be performed by bid B_i^j or bids $\mathcal{Z}$ but it cannot be performed by bids $\mathcal{S}$. Since $\mathcal{S} \subseteq \mathcal{Z}$ and $l_n > 0$ for $n \in \mathcal{N}$, we have $0 \leq s_k \leq z_k$. Therefore,

$$V(\mathcal{S} \cup \{B_i^j\}) - V(\mathcal{S}) \geq V(\mathcal{Z} \cup \{B_i^j\}) - V(\mathcal{Z}).$$

It can be observed that $V(\mathcal{S})$ is also non-negative and non-decreasing, which completes the proof. ■

The platform, which has a limited budget G, aims to maximize its valuation $V(\mathcal{S})$ by selecting the set of winning bids $\mathcal{S}$. Formally, it designs a mechanism $\mathcal{M} = (\mathcal{F}, \mathcal{P})$, which consists of a winning bid selection rule $\mathcal{F}$ and a payment determination rule $\mathcal{P}$. Given all bids $\mathcal{B}$ from all users and budget G as input, $\mathcal{F}$ returns a set of winning bids $\mathcal{S}$, and $\mathcal{P}$ returns the payment vector $\boldsymbol{p} = (p_1, \ldots, p_N)$. Our objective is to design a mechanism which satisfies the following properties:

- *Computational Efficiency:* Both the winning bids and payments are determined in polynomial time.
- *Truthfulness:* Each participating user submits its true cost for all of its bids, i.e., $b_n^m = c_n^m$, $m \in \mathcal{J}_n$, $n \in \mathcal{N}$, by which it can maximize its own utility.
- *Individual Rationality:* Each participating user has a non-negative utility, i.e., $u_n \geq 0$, $\forall\, n \in \mathcal{N}$.

- *Budget Feasibility:* The total payment to the users should be within the limited budget G of the platform, i.e., $\sum_{n \in \mathcal{N}} p_n \leq G$.

The importance of computational efficiency, individual rationality, and budget feasibility is obvious while truthfulness is necessary to avoid market manipulation. Since the users are strategic, they will submit bids which maximize their own utilities. However, with truthfulness, the dominant strategy of users is to submit their true costs. These four properties together guarantee the success of the auction mechanism.

3.2 Auction-based Budget Feasible Mechanism

In this section, we first propose a budget feasible mechanism ABSee by designing the winning bid selection rule $\mathcal{F}$ and the payment determination rule $\mathcal{P}$. We then illustrate ABSee in details by an example.

3.2.1 Mechanism Design

Budget feasibility distinguishes our mechanism from many auction mechanisms and makes the mechanism design challenging. With the budget constraint, the winning bid selection rule $\mathcal{F}$ and the payment determination rule $\mathcal{P}$ are inter-dependent. The platform must ensure the total payment for all selected winning bids is within the budget limit, while the payment should be determined carefully to ensure truthfulness. To guarantee that ABSee achieves truthfulness, according to the Myerson's characteristics [28], we have

Proposition 3.1. *An auction mechanism is truthful if and only if:*

- *The winning bid selection rule $\mathcal{F}$ is monotone, i.e., if user n wins the auction by bidding b_n^m for Γ_n^m, it can also win the auction by bidding $\hat{b}_n^m \leq b_n^m$ for Γ_n^m.*
- *User n is paid the threshold payment for winning bid $B_n^m = (\Gamma_n^m, b_n^m)$, $m \in \mathcal{J}_n$, which refers to the highest reserve price the user can submit to win this bid.*

Since $V(\mathcal{S})$ is a submodular function, we can adopt a greedy approach to design the monotone winning bid selection rule $\mathcal{F}$. Given a subset of bids $\mathcal{C} \subseteq \mathcal{B}$, the marginal contribution of bid $B_n^m = (\Gamma_n^m, b_n^m) \in \mathcal{B} \setminus \mathcal{C}$ is defined as

$$V_{B_n^m}(\mathcal{C}) = V(\mathcal{C} \cup \{B_n^m\}) - V(\mathcal{C}),\ m \in \mathcal{J}_n, n \in \mathcal{N}. \tag{3.7}$$

In $\mathcal{F}$, we sort the bids based on their marginal contributions per reserve price. The ith bid in the sorted list, denoted by x_i, has the largest marginal contribution per reserve price in subset $\mathcal{B} \setminus \mathcal{S}_{i-1}$, i.e.,

$$x_i = \underset{B_n^m \in \mathcal{B} \setminus \mathcal{S}_{i-1}}{\arg\max} \frac{V_{B_n^m}(\mathcal{S}_{i-1})}{b_n^m}, \tag{3.8}$$

where $\mathcal{S}_{i-1} = \{x_1, x_2, \ldots, x_{i-1}\}$ and $\mathcal{S}_0 \triangleq \emptyset$. To simplify the expression, for bid $x_i \in \mathcal{B}$, let $\bar{b}^{x_i}$, $\bar{c}^{x_i}$, and $\bar{p}^{x_i}$ denote the reserve price, the cost, and the payment, respectively. Considering the submodularity of $V(\mathcal{S})$, the sorting implies:

$$\frac{V_{x_1}(\mathcal{S}_0)}{\bar{b}^{x_1}} \geq \frac{V_{x_2}(\mathcal{S}_1)}{\bar{b}^{x_2}} \geq \cdots \geq \frac{V_{x_{|\mathcal{B}|}}(\mathcal{S}_{|\mathcal{B}|-1})}{\bar{b}^{x_{|\mathcal{B}|}}}. \tag{3.9}$$

The platform selects the winning bids from set $\mathcal{B}$ within limited budget G. To be selected as a winning bid, the marginal contribution per reserve price of bid x_i must not be less than $\frac{V(\mathcal{S}_i)}{\theta G}$, where

$$\theta = 1 - \frac{V^{\max}}{V(\mathcal{S})}, \tag{3.10}$$

and

$$V^{\max} = \max_{B_n^m \in \mathcal{B}} V(\{B_n^m\}). \tag{3.11}$$

We use θG as a proportion of budget and later show that (3.10) is used to guarantee budget feasibility. For a fixed value of θ, let w be the largest index such that $\frac{V_{x_w}(\mathcal{S}_{w-1})}{\bar{b}^{x_w}} \geq \frac{V(\mathcal{S}_w)}{\theta G}$, i.e.,

$$\bar{b}^{x_w} \leq \theta G \frac{V_{x_w}(\mathcal{S}_{w-1})}{V(\mathcal{S}_w)}. \tag{3.12}$$

Then, $\mathcal{S} = \{x_1, x_2, \ldots, x_w\}$ is the set of winning bids. However, notice that we cannot obtain θ directly since the winning bids set $\mathcal{S}$ is not known *a priori*. We adopt an iterative approach to determine θ and $\mathcal{S}$ together. The value of θ in iteration t is denoted by $\theta^{(t)}$. In each iteration t, the winning bids set $\mathcal{S}$ is obtained by utilizing $\theta^{(t)}$ and then $\mathcal{S}$ is used to calculate $\theta^{(t+1)}$. When $\theta^{(t)}$ increases, there will be more winning bids added to set $\mathcal{S}$ and thus $V(\mathcal{S})$ increases. In this case, both $\theta^{(t)}$ and $V(\mathcal{S})$ increase monotonically. Since $\mathcal{S} \subseteq \mathcal{B}$ is finite, the iteration converges. When θ decreases, $V(\mathcal{S})$ will be always decreasing and similar results can be obtained. Thus, the convergence of the iteration is guaranteed.

According to the Myerson's characteristics, the goal of the payment determination rule $\mathcal{P}$ is to pay each user the threshold payment for its winning bid. For each winning bid $x_i \in \mathcal{S}$,

similar to the winning bid selection rule $\mathcal{F}$, we sort the bids in set $\mathcal{B}' \triangleq \mathcal{B} \setminus \{x_i\}$, based on their marginal contributions per reserve price. We use $\mathcal{Q}_k$ to denote the first k bids in this sorting, and i_k to denote the kth bid, i.e., $i_k = \arg\max_{B_n^m \in \mathcal{B}' \setminus \mathcal{Q}_{k-1}} \frac{V_{B_n^m}(\mathcal{Q}_{k-1})}{b_n^m}$. Then, $V_{i_k}(\mathcal{Q}_{k-1})$ is the marginal contribution of the kth bid in this sorting. Similarly, we can obtain

$$\frac{V_{i_1}(\mathcal{Q}_0)}{\bar{b}^{i_1}} \geq \frac{V_{i_2}(\mathcal{Q}_1)}{\bar{b}^{i_2}} \geq \cdots \geq \frac{V_{i_{|\mathcal{B}|-1}}(\mathcal{Q}_{|\mathcal{B}|-2})}{\bar{b}^{i_{|\mathcal{B}|-1}}}. \tag{3.13}$$

We use w' to denote the largest index such that

$$\bar{b}^{i_{w'}} \leq \theta G \frac{V_{i_{w'}}(\mathcal{Q}_{w'-1})}{V(\mathcal{Q}_{w'})}. \tag{3.14}$$

Let $\beta_{i(k)}$ denote the highest reserve price that the user who has the winning bid x_i can submit to replace bid i_k with x_i in the kth position, and $\rho_{i(k)}$ denote the highest reserve price that this user can submit so that the marginal contribution per reserve price of bid x_i is not less than $\frac{V(\mathcal{Q}_{k-1} \cup \{x_i\})}{\theta G}$, i.e.,

$$\beta_{i(k)} = \frac{V_{x_i}(\mathcal{Q}_{k-1})\bar{b}^{i_k}}{V_{i_k}(\mathcal{Q}_{k-1})},$$

$$\rho_{i(k)} = \theta G \frac{V_{x_i}(\mathcal{Q}_{k-1})}{V(\mathcal{Q}_{k-1} \cup \{x_i\})}.$$

Since x_i can be placed in any position k from 1 to $w' + 1$ to be selected as a winning bid, the payment $\bar{p}^{x_i}$ for x_i is

$$\bar{p}^{x_i} = \max_{k \in \{1,\ldots,w'+1\}} \{\min(\beta_{i(k)}, \rho_{i(k)})\}. \tag{3.15}$$

The user who has submitted bid x_i receives the payment $\bar{p}^{x_i}$.

Algorithm 3.1: ABSee Budget Feasible Mechanism

1 Input $V(\cdot)$, G, $\mathcal{B}$
`/* Winning bid selection rule` $\mathcal{F}$ `*/`
2 $\theta^{(1)} \leftarrow \frac{1}{2}$, $t \leftarrow 0$, $V^{\max} \leftarrow \max_{B_n^m \in \mathcal{B}} V(\{B_n^m\})$
3 **do**
4 $\quad t \leftarrow t+1$, $i \leftarrow 1$
5 $\quad \mathcal{S} \leftarrow \emptyset$, $x_i \leftarrow \arg\max_{B_n^m \in \mathcal{B}} \frac{V(B_n^m(\emptyset))}{b_n^m}$
6 $\quad$ **while** $\bar{b}^{x_i} \leq \theta^{(t)} G \frac{V_{x_i}(\mathcal{S})}{V(\mathcal{S} \cup \{x_i\})}$ **do**
7 $\quad\quad \mathcal{S} \leftarrow \mathcal{S} \cup \{x_i\}$
8 $\quad\quad i \leftarrow i+1$
9 $\quad\quad x_i \leftarrow \arg\max_{B_n^m \in \mathcal{B} \setminus \mathcal{S}} \frac{V_{B_n^m}(\mathcal{S})}{b_n^m}$
10 $\quad$ **end**
11 $\quad$ **if** $\mathcal{S} = \emptyset$ **then break**;
12 $\quad \theta^{(t+1)} \leftarrow 1 - \frac{V^{\max}}{V(\mathcal{S})}$
13 **while** $\theta^{(t+1)} \neq \theta^{(t)}$;
14 $\theta \leftarrow \theta^{(t)}$
`/* Payment determination rule` $\mathcal{P}$ `*/`
15 **for** $n \in \mathcal{N}$ **do**
16 $\quad p_n \leftarrow 0$
17 **end**
18 **for** $x_i \in \mathcal{S}$ **do**
19 $\quad \mathcal{B}' \leftarrow \mathcal{B} \setminus \{x_i\}$, $\mathcal{Q} \leftarrow \emptyset$, $k \leftarrow 0$
20 $\quad$ **do**
21 $\quad\quad k \leftarrow k+1$
22 $\quad\quad i_k \leftarrow \arg\max_{B_n^m \in \mathcal{B}' \setminus \mathcal{Q}} \frac{V_{B_n^m}(\mathcal{Q})}{b_n^m}$
23 $\quad\quad \beta_{i(k)} \leftarrow \frac{V_{x_i}(\mathcal{Q}) \bar{b}^{i_k}}{V_{i_k}(\mathcal{Q})}$, $\rho_{i(k)} \leftarrow \theta G \frac{V_{x_i}(\mathcal{Q})}{V(\mathcal{Q} \cup \{x_i\})}$
24 $\quad\quad \bar{p}^{x_i} \leftarrow \max\{\bar{p}^{x_i}, \min(\beta_{i(k)}, \rho_{i(k)})\}$
25 $\quad\quad \mathcal{Q} \leftarrow \mathcal{Q} \cup \{i_k\}$
26 $\quad$ **while** $\bar{b}^{i_k} \leq \theta G \frac{V_{i_k}(\mathcal{Q} \setminus \{i_k\})}{V(\mathcal{Q})}$;
27 $\quad$ **for** $n \in \mathcal{N}$ **do**
28 $\quad\quad$ **if** $x_i \in \mathcal{B}_n$ **then** $p_n \leftarrow p_n + \bar{p}^{x_i}$;
29 $\quad$ **end**
30 **end**
31 **return** $(\mathcal{S}, \boldsymbol{p})$

Our proposed ABSee budget feasible mechanism is shown in Algorithm 3.1. Steps 2 to 14 show the winning bid selection rule $\mathcal{F}$. We set $\theta^{(1)} = \frac{1}{2}$ for initialization. An iterative approach

is adopted to calculate θ. In each iteration (i.e., Steps 4 to 12), the winning bids are selected in a greedy manner according to their marginal contribution per reserve price. The value of θ is obtained from Step 14. Steps 15 to 30 show the payment determination rule $\mathcal{P}$. The payment to each user is initialized to 0. Then, for each winning bid, the threshold payment is given to the corresponding user in Steps 18 to 30 according to (3.15). Note that Step 24 is executed $w' + 1$ times for each winning bid x_i. Finally, we obtain the set of winning bids $\mathcal{S}$ and the payment vector $\boldsymbol{p}$.

3.2.2 Walk-Through Example

We use the example in Fig. 3.2 to describe how ABSee works. In this figure, squares represent the users and circles represent the tasks. The numbers above the squares denote the quality indicator q_n of each user. The number below the circles denote the weight μ_k of each task. Each user $n \in \{1, 2, 4, 5\}$ has two bids, denoted as B_n^1, B_n^2, while user 3 only has one bid, i.e., B_3^1. Assume the platform has a budget $G = 50$.

We first calculate $V^{\max}$: $V^{\max} = V(\{B_4^2\}) = 20.5$.

<u>Winning bid selection:</u>

1. According to Step 2 in Algorithm 3.1, we initialize $\theta^{(1)} = \frac{1}{2}$ and calculate $V^{\max} = V(\{B_4^2\}) = 20.5$. Then, according to Steps 3 to 13, winning bids are selected and θ is updated in each iteration.

2. $t = 1$. $\theta^{(1)} = \frac{1}{2}$.

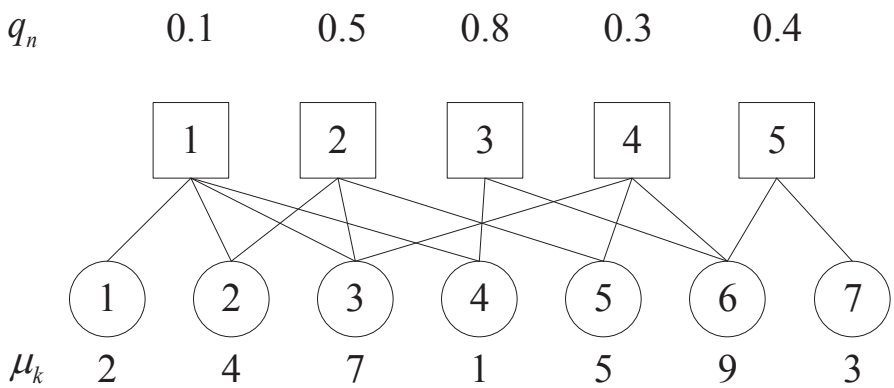

Figure 3.2: Illustration for ABSee. $\mathcal{N} = \{1, 2, 3, 4, 5\}$, $\Gamma = \{\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7\}$, $q_1 = 0.1$, $q_2 = 0.5$, $q_3 = 0.8$, $q_4 = 0.3$, $q_5 = 0.4$, $\mu_1 = 2$, $\mu_2 = 4$, $\mu_3 = 7$, $\mu_4 = 1$, $\mu_5 = 5$, $\mu_6 = 9$, $\mu_7 = 3$, $\mathcal{B}_1 = \{(\{\tau_1, \tau_2\}, 6), (\{\tau_3, \tau_4\}, 6)\}$, $\mathcal{B}_2 = \{(\{\tau_2\}, 5), (\{\tau_3, \tau_5\}, 10)\}$, $\mathcal{B}_3 = \{(\{\tau_4, \tau_6\}, 2)\}$, $\mathcal{B}_4 = \{(\{\tau_3\}, 4), (\{\tau_5, \tau_6\}, 8)\}$, $\mathcal{B}_5 = \{(\{\tau_6\}, 2), (\{\tau_7\}, 2)\}$.

- $\mathcal{S} = \emptyset$: $\frac{V_{B_1^1}(\emptyset)}{b_1^1} = \frac{V(B_1^1)}{b_1^1} = \frac{14.4}{6} = 2.4$, $\frac{V_{B_1^2}(\emptyset)}{b_1^2} = 3.2$, $\frac{V_{B_2^1}(\emptyset)}{b_2^1} = 0.9$, $\frac{V_{B_2^2}(\emptyset)}{b_2^2} = 1.3$, $\frac{V_{B_3^1}(\emptyset)}{b_3^1} = 4.1$, $\frac{V_{B_4^1}(\emptyset)}{b_4^1} = 2.6$, $\frac{V_{B_4^2}(\emptyset)}{b_4^2} = 2.6$, $\frac{V_{B_5^1}(\emptyset)}{b_5^1} = 5.6$, $\frac{V_{B_5^2}(\emptyset)}{b_5^2} = 1.9$.

  $x_1 = B_5^1$, $5.6 = \frac{V_{B_5^1}(\emptyset)}{b_5^1} \geq \frac{V(\emptyset \cup \{B_5^1\})}{\theta^{(1)} G} = 0.4$.

- $\mathcal{S} = \{B_5^1\}$: $\frac{V_{B_1^1}(\{B_5^1\})}{b_1^1} = 2.4$, $\frac{V_{B_1^2}(\{B_5^1\})}{b_1^2} = 3.2$, $\frac{V_{B_2^1}(\{B_5^1\})}{b_2^1} = 0.9$, $\frac{V_{B_2^2}(\{B_5^1\})}{b_2^2} = 1.3$, $\frac{V_{B_3^1}(\{B_5^1\})}{b_3^1} = \frac{\log 2.25 + 9 \log(1 + \frac{1.25}{1+2.5})}{2} = 1.8$, $\frac{V_{B_4^1}(\{B_5^1\})}{b_4^1} = 2.6$,
  $\frac{V_{B_4^2}(\{B_5^1\})}{b_4^2} = \frac{5 \log 4.3 + 9 \log(1 + \frac{3.3}{1+2.5})}{8} = 1.7$, $\frac{V_{B_5^2}(\{B_5^1\})}{b_5^2} = 1.9$.

  $x_2 = B_1^2$, $3.2 = \frac{V_{B_1^2}(\{B_5^1\})}{b_1^2} \geq \frac{V(\{B_5^1, B_1^2\})}{\theta^{(1)} G} = \frac{30.4}{25} = 1.2$.

- $\mathcal{S} = \{B_5^1, B_1^2\}$: $\frac{V_{B_1^1}(\{B_5^1, B_1^2\})}{b_1^1} = 2.4$, $\frac{V_{B_2^1}(\{B_5^1, B_1^2\})}{b_2^1} = 0.9$, $\frac{V_{B_2^2}(\{B_5^1, B_1^2\})}{b_2^2} = 0.7$, $\frac{V_{B_3^1}(\{B_5^1, B_1^2\})}{b_3^1} = 1.4$, $\frac{V_{B_4^1}(\{B_5^1, B_1^2\})}{b_4^1} = 0.5$, $\frac{V_{B_4^2}(\{B_5^1, B_1^2\})}{b_4^2} = 1.7$, $\frac{V_{B_5^2}(\{B_5^1, B_1^2\})}{b_5^2} = 1.9$.

  $x_3 = B_1^1$, $2.4 = \frac{V_{B_1^1}(\{B_5^1, B_1^2\})}{b_1^1} \geq \frac{V(\{B_5^1, B_1^2, B_1^1\})}{\theta^{(1)} G} = \frac{44.8}{25} = 1.8$.

- $\mathcal{S} = \{B_5^1, B_1^2, B_1^1\}$: $\frac{V_{B_2^1}(\{B_5^1, B_1^2, B_1^1\})}{b_2^1} = 0.1$, $\frac{V_{B_2^2}(\{B_5^1, B_1^2, B_1^1\})}{b_2^2} = 0.7$, $\frac{V_{B_3^1}(\{B_5^1, B_1^2, B_1^1\})}{b_3^1} = 1.4$, $\frac{V_{B_4^1}(\{B_5^1, B_1^2, B_1^1\})}{b_4^1} = 0.5$, $\frac{V_{B_4^2}(\{B_5^1, B_1^2, B_1^1\})}{b_4^2} = 1.7$, $\frac{V_{B_5^2}(\{B_5^1, B_1^2, B_1^1\})}{b_5^2} = 1.9$.

$x_4 = B_5^2$, $1.88 = \frac{V_{B_5^2}(\{B_5^1, B_1^2\, B_1^1\})}{b_5^2} < \frac{V(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{\theta^{(1)}G} = \frac{48.6}{25} = 1.94$.

Thus, $\mathcal{S} = \{B_5^1, B_1^2, B_1^1\}$. $\theta^{(2)} = 1 - \frac{V^{\max}}{V(\mathcal{S})} = 0.54$. $\theta^{(2)} \neq \theta^{(1)}$.

3. $t = 2$. $\theta^{(2)} = 0.54$.

- $\mathcal{S} = \emptyset$: $x_1 = B_5^1$, $5.6 = \frac{V_{B_5^1}(\emptyset)}{b_5^1} \geq \frac{V(\emptyset \cup \{B_5^1\})}{\theta^{(2)}G} = 0.4$.
- $\mathcal{S} = \{B_5^1\}$: $x_2 = B_1^2$, $3.2 = \frac{V_{B_1^2}(\{B_5^1\})}{b_1^2} \geq \frac{V(\{B_5^1, B_1^2\})}{\theta^{(2)}G} = \frac{30.4}{27} = 1.1$.
- $\mathcal{S} = \{B_5^1, B_1^2\}$: $x_3 = B_1^1$, $2.4 = \frac{V_{B_1^1}(\{B_5^1, B_1^2\})}{b_1^1} \geq \frac{V(\{B_5^1, B_1^2, B_1^1\})}{\theta^{(2)}G} = \frac{44.8}{27} = 1.7$.
- $\mathcal{S} = \{B_5^1, B_1^2, B_1^1\}$: $x_4 = B_5^2$, $1.88 = \frac{V_{B_5^2}(\{B_5^1, B_1^2, B_1^1\})}{b_5^2} \geq \frac{V(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{\theta^{(2)}G} = \frac{48.6}{27} = 1.79$.
- $\mathcal{S} = \{B_5^1, B_1^2, B_1^1, B_5^2\}$: $\frac{V_{B_2^1}(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{b_2^1} = 0.1$, $\frac{V_{B_2^2}(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{b_2^2} = 0.7$, $\frac{V_{B_3^1}(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{b_3^1} = 1.4$, $\frac{V_{B_4^1}(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{b_4^1} = 0.5$, $\frac{V_{B_4^2}(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{b_4^2} = 1.7$.

  $x_5 = B_4^2$, $1.7 = \frac{V_{B_5^2}(\{B_5^1, B_1^2, B_1^1\})}{b_5^2} < \frac{V(\{B_5^1, B_1^2, B_1^1, B_5^2, B_4^2\})}{\theta^{(2)}G} = \frac{62.0}{27} = 2.3$.

Thus, we obtain $\mathcal{S} = \{B_5^1, B_1^2, B_1^1, B_5^2\}$. $\theta^{(3)} = 1 - \frac{V^{\max}}{V(\mathcal{S})} = 0.58$. $\theta^{(3)} \neq \theta^{(2)}$.

4. $t = 3$. $\theta^{(3)} = 0.58$.

- $\mathcal{S} = \emptyset$: $x_1 = B_5^1$, $5.6 = \frac{V_{B_5^1}(\emptyset)}{b_5^1} \geq \frac{V(\emptyset \cup \{B_5^1\})}{\theta^{(3)}G} = 0.4$.
- $\mathcal{S} = \{B_5^1\}$: $x_2 = B_1^2$, $3.2 = \frac{V_{B_1^2}(\{B_5^1\})}{b_1^2} \geq \frac{V(\{B_5^1, B_1^2\})}{\theta^{(3)}G} = \frac{30.4}{29} = 1.0$.
- $\mathcal{S} = \{B_5^1, B_1^2\}$: $x_3 = B_1^1$, $2.4 = \frac{V_{B_1^1}(\{B_5^1, B_1^2\})}{b_1^1} \geq \frac{V(\{B_5^1, B_1^2, B_1^1\})}{\theta^{(3)}G} = \frac{44.8}{29} = 1.5$.
- $\mathcal{S} = \{B_5^1, B_1^2, B_1^1\}$: $x_4 = B_5^2$, $1.88 = \frac{V_{B_5^2}(\{B_5^1, B_1^2, B_1^1\})}{b_5^2} \geq \frac{V(\{B_5^1, B_1^2, B_1^1, B_5^2\})}{\theta^{(3)}G} = \frac{48.6}{29} = 1.68$.

- $\mathcal{S} = \{B_5^1, B_1^2, B_1^1, B_5^2\}$: $x_5 = B_4^2$, $1.7 = \frac{V_{B_5^2}(\{B_5^1, B_1^2, B_1^1\})}{b_5^2} < \frac{V(\{B_5^1, B_1^2, B_1^1, B_5^2, B_4^2\})}{\theta^{(3)}G} = \frac{62.0}{29} = 2.1$.

Thus, we still have $\mathcal{S} = \{B_5^1, B_1^2, B_1^1, B_5^2\}$. $\theta^{(4)} = 1 - \frac{V^{\max}}{V(\mathcal{S})} = 0.58$. $\theta^{(4)} = \theta^{(3)}$.

5. According to Step 14 in Algorithm 3.1, we obtain $\theta = 0.58$.

<u>Payment determination:</u>

1. $x_1 = B_5^1$: Winning bids and the next bid after the winning bids are $B_3^1, B_1^2, B_1^1, B_4^2$.

   $\beta_{1(1)} = \frac{V_{x_1}(\mathcal{Q}_0)\bar{b}^1{}_1}{V_{1_1}(\mathcal{Q}_0)} = \frac{9\log(1+\frac{1}{0.4})\times 2}{10\log(1+1/0.8)} = 2.78$, $\rho_{1(1)} = \frac{\theta G V_{x_1}(\mathcal{Q}_0)}{V(\mathcal{Q}_0 \cup \{x_1\})} = \theta G = 28.88$, $\beta_{1(2)} = 2.18$, $\rho_{1(2)} = 13.09$, $\beta_{1(3)} = 2.80$, $\rho_{1(3)} = 5.83$, $\beta_{1(4)} = 3.47$, $\rho_{1(4)} = 0.14$. Thus, $\bar{p}^{x_1} = 2.80$.

2. $x_2 = B_1^2$: Winning bids and the next bid after the winning bids are $B_5^1, B_4^1, B_1^1, B_5^2, B_3^1, B_4^2$.

   $\beta_{2(1)} = 3.40$, $\rho_{2(1)} = 28.88$, $\beta_{2(2)} = 7.48$, $\rho_{2(2)} = 18.19$, $\beta_{2(3)} = 4.49$, $\rho_{2(3)} = 9.63$, $\beta_{2(4)} = 5.73$, $\rho_{2(4)} = 6.66$, $\beta_{2(5)} = 6.05$, $\rho_{2(5)} = 6.17$, $\beta_{2(6)} = 6.65$, $\rho_{2(6)} = 5.45$. Thus, $\bar{p}^{x_2} = 7.48$.

3. $x_3 = B_1^1$: Winning bids and the next bid after the winning bids are $B_5^1, B_1^2, B_5^2, B_4^2, B_2^1$.

   $\beta_{3(1)} = 2.55$, $\rho_{3(1)} = 28.88$, $\beta_{3(2)} = 4.50$, $\rho_{3(2)} = 16.19$, $\beta_{3(3)} = 7.66$, $\rho_{3(3)} = 9.27$, $\beta_{3(4)} = 8.62$, $\rho_{3(4)} = 8.55$, $\beta_{3(5)} = 16.37$, $\rho_{3(5)} = 6.71$. Thus, $\bar{p}^{x_3} = 8.55$.

4. $x_4 = B_5^2$: Winning bids and the next bid after the winning bids are $B_5^1, B_1^2, B_1^1, B_4^2$.

   $\beta_{4(1)} = 0.67$, $\rho_{4(1)} = 28.88$, $\beta_{4(2)} = 1.18$, $\rho_{4(2)} = 7.22$, $\beta_{4(3)} = 1.57$, $\rho_{4(3)} = 3.17$, $\beta_{4(4)} = 2.25$, $\rho_{4(4)} = 2.23$. Thus, $\bar{p}^{x_4} = 2.23$.

Then, $p_1 = \bar{p}^{x_2} + \bar{p}^{x_3} = 7.48 + 8.55 = 16.03$, $p_5 = \bar{p}^{x_1} + \bar{p}^{x_4} = 2.80 + 2.23 = 5.03$, $p_2 = p_3 = p_4 = 0$.

3.3 Mechanism Analysis

In this section, we first prove that ABSee satisfies all of the properties introduced in Section 3.1. Then, we calculate its approximation ratio.

3.3.1 Properties of Proposed Mechanism

Theorem 3.1. *ABSee is computationally efficient.*

Proof. Since calculating $V(\mathcal{S})$ takes $O(|\mathcal{B}|\, M)$ time, the bid with the maximum marginal contribution per reserve price takes $O(|\mathcal{B}|^2\, M)$ time to find. Both the winning bid selection rule $\mathcal{F}$ and the payment determination rule $\mathcal{P}$ have nested loops. The while loop in $\mathcal{F}$ takes $O(|\mathcal{B}|^3\, M)$ time because the maximum number of the winning bids is $|\mathcal{B}|$. Then, $\theta^{(t)}$ is updated if more winning bids are selected so that the do-while loop also runs at most $|\mathcal{B}|$ times. Thus, the running time of $\mathcal{F}$ is $O(|\mathcal{B}|^4\, M)$. Similar to $\mathcal{F}$, the running time of $\mathcal{P}$ is also $O(|\mathcal{B}|^4\, M)$. ■

Note that the running time shown above is conservative. In practice, the number of winning bids is much less than $|\mathcal{B}|$.

Theorem 3.2. *ABSee is truthful.*

Proof. It is required to show that ABSee satisfies the Myerson's characteristics. When the auction is conducted multiple times, we obtain quality indicator q_n, $n \in \mathcal{N}$. In the greedy

approach in $\mathcal{F}$, since a lower reserve price can only put the bid in the same or a prior position, the monotonicity is guaranteed. Notice that the monotonicity is not influenced by the value of θ. The iteration in the winning bid selection rule $\mathcal{F}$ cannot change the monotonicity of the greedy approach. Thus, we only need to show that the users receive the threshold payments for their winning bids. The proof follows the approach presented in [18].

Consider winning bid x_i. From (3.15), let $r \leq w' + 1$ be the index such that $\bar{p}^{x_i} = \max_{k\in\{1,\ldots,w'+1\}}\{\min(\beta_{i(k)}, \rho_{i(k)})\} = \min(\beta_{i(r)}, \rho_{i(r)})$. Recall w' from (3.14). We show that the user who has submitted bid x_i wins the auction by bidding $\bar{b}^{x_i} \leq \bar{p}^{x_i}$ and loses the auction if $\bar{b}^{x_i} > \bar{p}^{x_i}$.

If $\bar{b}^{x_i} \leq \bar{p}^{x_i}$, we have $\bar{b}^{x_i} \leq \beta_{i(r)}$. Bid x_i can be placed in the first $w' + 1$ positions in the sorted list given by (3.13). We also have $\bar{b}^{x_i} \leq \rho_{i(r)}$. Thus, it will be chosen as a winning bid.

If $\bar{b}^{x_i} > \bar{p}^{x_i}$, we have two cases.

Case 1: $\beta_{i(r)} \leq \rho_{i(r)}$, i.e., $\bar{p}^{x_i} = \beta_{i(r)}$. In this case, x_i will be placed after the rth position in the sorted list. If $\beta_{i(r)} = \max_{k\in\{1,\ldots,w'+1\}}\beta_{i(k)}$, x_i cannot be a winning bid. Otherwise, if $\beta_{i(r)} < \beta_{i(k)}$ for some $k \in \{1, \ldots, w' + 1\}$, we have

$$\rho_{i(k)} < \beta_{i(r)} = \bar{p}^{x_i} < \bar{b}^{x_i}.$$

Thus, x_i cannot be placed at the kth position to be selected as a winning bid.

Case 2: $\beta_{i(r)} > \rho_{i(r)}$, i.e., $\bar{p}^{x_i} = \rho_{i(r)}$. If $\rho_{i(r)} = \max_{k\in\{1,\ldots,w'+1\}}\rho_{i(k)}$, x_i cannot be a winning

bid. Otherwise, if $\rho_{i(r)} < \rho_{i(k)}$ for some $k \in \{1, \ldots, w' + 1\}$, we have

$$\beta_{i(k)} < \rho_{i(r)} = \bar{p}^{x_i} < \bar{b}^{x_i}.$$

Thus, x_i cannot be a winning bid.

Since the winning bid selection rule $\mathcal{F}$ is monotone and each user receives the threshold payment for its winning bids, we conclude that ABSee is truthful. ■

Theorem 3.3. *ABSee is individually rational.*

Proof. If a user is not a winner, its utility is zero as shown in (3.2). If user $n \in \mathcal{N}$ is a winner, from (3.2), its utility can also be written as:

$$u_n = p_n - \sum_{m \in \mathcal{J}_n : \Gamma_n^m \in \mathcal{R}_n(\mathcal{S})} c_n^m = \sum_{x_i \in \mathcal{B}_n \cap \mathcal{S}} (\bar{p}^{x_i} - \bar{c}^{x_i}).$$

Notice that x_i is a winning bid. Thus, user n will have non-negative utility if $\bar{p}^{x_i} - \bar{c}^{x_i} \geq 0$ for any of its winning bid x_i. Since ABSee is truthful, we have $\bar{b}^{x_i} = \bar{c}^{x_i}$ for a winning bid x_i. If there exists k from 1 to $w' + 1$ such that $\bar{b}^{x_i} \leq \min(\beta_{i(k)}, \rho_{i(k)})$, we have $\bar{b}^{x_i} \leq \bar{p}^{x_i}$. We obtain $\bar{p}^{x_i} - \bar{c}^{x_i} \geq 0$ for all x_i, and thus $u_n \geq 0$.

For winning bid x_i, the payment determination rule $\mathcal{P}$ implies that $\mathcal{Q}_k = \mathcal{S}_k$, for $k < i$, thus $V_{x_i}(\mathcal{S}_{i-1}) = V_{x_i}(\mathcal{Q}_{i-1})$. Recall that bid x_i is sorted in the ith position in (3.9) among bids $\mathcal{B}$ and bid i_k is sorted in the kth position in (3.13) among bids $\mathcal{B} \setminus \{x_i\}$. Consider the case of

$k = i$, we have

$$\bar{b}^{x_i} \leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{S}_{i-1} \cup \{x_i\})} = \theta G \frac{V_{x_i}(\mathcal{Q}_{k-1})}{V(\mathcal{Q}_{k-1} \cup \{x_i\})} = \rho_{i(k)}. \tag{3.16}$$

From (3.8), we have $\frac{V_{x_i}(\mathcal{S}_{i-1})}{\bar{b}^{x_i}} \geq \frac{V_{i_k}(\mathcal{S}_{i-1})}{\bar{b}^{i_k}}$. We can obtain that

$$\bar{b}^{x_i} \leq \frac{V_{x_i}(\mathcal{S}_{i-1})\bar{b}^{i_k}}{V_{i_k}(\mathcal{S}_{i-1})} = \frac{V_{x_i}(\mathcal{Q}_{k-1})\bar{b}^{i_k}}{V_{i_k}(\mathcal{Q}_{k-1})} = \beta_{i(k)}. \tag{3.17}$$

From (3.16) and (3.17), we have $\bar{b}^{x_i} \leq \min\{\beta_{i(k)}, \rho_{i(k)}\}$ when $k = i$. This results in $\bar{b}^{x_i} \leq \bar{p}^{x_i}$ and completes the proof. ■

To prove the budget feasibility of ABSee, we use the following lemma.

Lemma 3.2. *The payment $\bar{p}^{x_i}$ for each winning bid x_i satisfies $\bar{p}^{x_i} \leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{Q}_{w'})}$.*

Proof. For a winning bid x_i, from the submodularity of the valuation function, we have

$$V_{x_i}(\mathcal{S}_{i-1}) \geq V_{x_i}(\mathcal{Q}_{k-1}),\ \forall\, k \geq i. \tag{3.18}$$

Let r be the index for which $\bar{p}^{x_i} = \min\{\beta_{i(r)}, \rho_{i(r)}\}$. We now prove that $r \geq i$. By contradiction, assume $r < i$, we have $x_r = \arg\max_{B_n^m \in \mathcal{B} \setminus \mathcal{S}_{r-1}} \frac{V_{B_n^m}(\mathcal{S}_{r-1})}{b_n^m}$ from (3.8). Thus,

$$\frac{V_{x_r}(\mathcal{S}_{r-1})}{\bar{b}^{x_r}} > \frac{V_{x_i}(\mathcal{S}_{r-1})}{\bar{b}^{x_i}}.$$

We assume the strict inequality holds to simplify the proof. The same results can be obtained

if the equality is considered. From the definition of $\beta_{i(r)}$, we have

$$\frac{V_{i_r}(\mathcal{Q}_{r-1})}{\bar{b}^{i_r}} = \frac{V_{x_i}(\mathcal{Q}_{r-1})}{\beta_{i(r)}}.$$

Since $\mathcal{Q}_k = \mathcal{S}_k$ and $x_k = i_k$ for $k < i$, we have $V_{x_i}(\mathcal{S}_{r-1}) = V_{x_i}(\mathcal{Q}_{r-1})$, $V_{x_r}(\mathcal{S}_{r-1}) = V_{i_r}(\mathcal{Q}_{r-1})$, and $\bar{b}^{x_r} = \bar{b}^{i_r}$. According to these equalities, we conclude that $\bar{p}^{x_i} \leq \beta_{i(r)} < \bar{b}^{x_i}$, which is contradictory with Theorem 3.3. Hence, we have $r \geq i$.

If $r \leq w'$,

$$\frac{V_{i_r}(\mathcal{Q}_{r-1})}{\bar{b}^{i_r}} \geq \frac{V_{i_{w'}}(\mathcal{Q}_{w'-1})}{\bar{b}^{i_{w'}}}.$$

From (3.18), we have

$$\begin{aligned}
\bar{p}^{x_i} &\leq \beta_{i(r)} \\
&= \frac{V_{x_i}(\mathcal{Q}_{r-1})\bar{b}^{i_r}}{V_{i_r}(\mathcal{Q}_{r-1})} \\
&\leq \frac{V_{x_i}(\mathcal{Q}_{r-1})\bar{b}^{i_{w'}}}{V_{i_{w'}}(\mathcal{Q}_{w'-1})} \\
&\leq \theta G \frac{V_{x_i}(\mathcal{Q}_{r-1})}{V(\mathcal{Q}_{w'})} \\
&\leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{Q}_{w'})}.
\end{aligned} \tag{3.19}$$

If $r = w' + 1$, we have

$$\begin{aligned}\bar{p}^{x_i} &\leq \rho_{i(r)} \\ &\leq \theta G \frac{V_{x_i}(\mathcal{Q}_{w'})}{V(\mathcal{Q}_{w'} \cup \{x_i\})} \\ &\leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{Q}_{w'})}. \end{aligned} \tag{3.20}$$

From (3.19) and (3.20), we conclude that $\bar{p}^{x_i} \leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{Q}_{w'})}$, which completes the proof. ■

By utilizing Lemma 3.2, we now prove the budget feasibility.

Theorem 3.4. *ABSee is budget feasible.*

Proof. To prove $\sum_{n \in \mathcal{N}} p_n \leq G$, it is equivalent to prove that $\sum_{x_i \in \mathcal{S}} \bar{p}^{x_i} \leq G$. According to the payment determination rule $\mathcal{P}$, we have

$$V(\mathcal{S}) - V(\mathcal{Q}_{W'}) \leq \max_{B_n^m \in \mathcal{B}} V(\{B_n^m\}).$$

We obtain

$$\begin{aligned} 1 - \frac{V(\mathcal{Q}_{W'})}{V(\mathcal{S})} &= \frac{V(\mathcal{S}) - V(\mathcal{Q}_{W'})}{V(\mathcal{S})} \\ &\leq \frac{\max_{B_n^m \in \mathcal{B}} V(\{B_n^m\})}{V(\mathcal{S})}. \end{aligned}$$

Then,

$$\begin{aligned}\frac{V(\mathcal{Q}_{W'})}{V(\mathcal{S})} &\geq 1 - \frac{\max_{B_n^m \in \mathcal{B}} V(\{B_n^m\})}{V(\mathcal{S})} \\ &= \theta.\end{aligned}$$

Thus, $V(\mathcal{Q}_{w'}) \geq \theta V(\mathcal{S})$. From Lemma 3.2, we know that for each winning bid x_i, the payment $\bar{p}^{x_i} \leq \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{Q}_{w'})}$. Then, we have

$$\begin{aligned}\bar{p}^{x_i} &\leq \frac{V_{x_i}(\mathcal{S}_{i-1})\theta G}{V(\mathcal{Q}_{W'})} \\ &\leq \frac{V_{x_i}(\mathcal{S}_{i-1})\theta G}{\theta V(\mathcal{S})} \\ &= \frac{V_{x_i}(\mathcal{S}_{i-1})G}{V(\mathcal{S})}.\end{aligned}$$

According to (3.7) and $\mathcal{S} = \{x_1, x_2, \ldots, x_w\}$, we have $\sum_{x_i \in \mathcal{S}} V_{x_i}(\mathcal{S}_{i-1}) = V_{x_1}(\emptyset) + V_{x_2}(\mathcal{S}_1) + \cdots + V_{x_w}(\mathcal{S}_{w-1}) = V(\mathcal{S})$. We conclude that $\sum_{x_i \in \mathcal{S}} \bar{p}^{x_i} \leq \sum_{x_i \in \mathcal{S}} V_{x_i}(\mathcal{S}_{i-1}) \frac{G}{V(\mathcal{S})} = G$. ■

We have $\mathcal{Q}_{w'}$ as the winning bids selected from bids $\mathcal{B} \setminus \{x_i\}$. From $\theta \leq \frac{V(\mathcal{Q}_{w'})}{V(\mathcal{S})}$ in the above proof, we observe that θ describes the lower bound of the relative valuation when removing a winning bid from the system. Given budget G and the number of tasks M, when more users participate in the system and submit more bids, a winning bid has a smaller relative contribution to the system. We formally define θ as follows.

Definition 3.2. *We define* $\theta \triangleq 1 - \frac{V^{max}}{V(\mathcal{S})}$, *where* $V^{max} = \max_{B_n^m \in \mathcal{B}} V(\{B_n^m\})$ *from (3.11).* θ *is called the* ***crowd factor****.*

From Algorithm 1, we have $0 < \theta < 1$. For a successful mobile crowdsourcing application, there are usually a large number of participating users and the platform selects many winning bids. Thus, $V(\mathcal{S})$ is much larger than $V^{\max}$ (i.e., $V(\mathcal{S}) \gg V^{\max}$) and θ approaches 1.

Consider the following scenario as an example. Each user has the same quality indicator and submits the same number of bids, i.e., $q_n = q$, $\sigma_n = \sigma$, $n \in \mathcal{N}$. We use $\bar{\Gamma}^{x_i}$ to denote the subset of tasks within bid $x_i \in \mathcal{B}$. The subsets of tasks within all bids are pairwise disjoint, i.e., $\bar{\Gamma}^{x_i} \cap \bar{\Gamma}^{x_i'} = \emptyset$, $x_i, x_i' \in \mathcal{B}$. The number of tasks within each bid is also the same, i.e., $\left|\bar{\Gamma}^{x_i}\right| = \Upsilon$, $x_i \in \mathcal{B}$. All tasks have the same weight, i.e., $\mu_k = \mu$, $\tau_k \in \Gamma$. The reserve price $\bar{b}^{x_i}$ is equal to b. Then, we have $V_{x_i}(\mathcal{S}_{i-1}) = V(\{x_i\}) = \sum_{\tau_k \in \bar{\Gamma}^{x_i}} \mu_k \log(1 + \frac{1}{q}) = \Upsilon \mu \log(1 + \frac{1}{q})$, which is denoted by V. Thus, we obtain $V(\mathcal{S}) = \sum_{x_i \in \mathcal{S}} V_{x_i}(\mathcal{S}_{i-1}) = |\mathcal{S}|\, V$, and $V^{\max} = V$. From the inequalities (3.9) and (3.12), the largest index of a winning bid (i.e., w) satisfies $\frac{V}{b} \geq \frac{wV}{\theta G}$. Thus, we obtain $w = \left\lfloor \frac{\theta G}{b} \right\rfloor$. When $\frac{G}{b}$ is large, the number of winning bids $|\mathcal{S}| = w \approx \frac{\theta G}{b}$. From (3.10), we have $\theta = 1 - \frac{V^{\max}}{V(\mathcal{S})} = 1 - \frac{V}{wV} = 1 - \frac{1}{w} \approx 1 - \frac{b}{\theta G}$. We solve the equation $\theta = 1 - \frac{b}{\theta G}$ to obtain the value of θ. Hence, when $\frac{4b}{G} < 1$, we have

$$\theta = \frac{1}{2}\left(1 + \sqrt{1 - \frac{4b}{G}}\right).$$

For $w \approx \frac{\theta G}{b}$, the number of submitted bids should be at least $\frac{\theta G}{b}$. Since $N\sigma = |\mathcal{B}|$, the number of users should be at least $\frac{\theta G}{\sigma b}$. Thus, when $\frac{4b}{G} \ll 1$ and $N > \frac{G}{\sigma b}$, i.e., $N \gg \frac{4}{\sigma}$, the crowd factor θ approaches 1. Since the conditions $\frac{4b}{G} \ll 1$ and $N \gg \frac{4}{\sigma}$ can easily be satisfied in a practical mobile crowdsourcing system, θ is close to 1 with high probability. In Section 3.4, we will show that θ is close to 1 under different scenarios in a mobile crowdsourcing system.

3.3.2 Approximation Ratio Analysis

A mechanism is called α-approximate if it determines a subset $\mathcal{S} \subseteq \mathcal{B}$ such that $opt(\mathcal{B}) \leq \alpha V(\mathcal{S})$, where $opt(\mathcal{B})$ is the optimal value of the following optimization problem:

$$\begin{aligned} \underset{\mathcal{S}\subseteq\mathcal{B}}{\text{maximize}} \quad & V(\mathcal{S}) \\ \text{subject to} \quad & \sum_{x_i\in\mathcal{S}} \bar{c}^{x_i} \leq G. \end{aligned} \tag{3.21}$$

Problem (3.21) is a budgeted submodular function maximization problem [29]. Similar to the Knapsack problem, problem (3.21) is also an NP-hard problem and the optimal solution cannot be obtained in polynomial time.

To determine the approximation ratio of our proposed mechanism, we first introduce a fractional greedy algorithm [19] to solve problem (3.21). Similar to our proposed winning bid selection rule $\mathcal{F}$, we select bids with the highest marginal contribution per reserve price until we cannot add more bids due to budget limit. We assume that the contribution of a bid can be fractional. Let H be the largest index such that $\sum_{i=1}^{H} \bar{c}^{x_i} \leq G$. We define $C^{x_{H+1}} \triangleq G - \sum_{i=1}^{H} \bar{c}^{x_i}$ and $V'_{x_{H+1}}(\mathcal{S}_H) \triangleq V_{x_{H+1}}(\mathcal{S}_H)\frac{C^{x_{H+1}}}{\bar{c}^{x_{H+1}}}$. The valuation obtained by adopting the fractional greedy algorithm can be defined as:

$$\widetilde{V}(\mathcal{S}) \triangleq \sum_{i=1}^{H} V_{x_i}(\mathcal{S}_{i-1}) + V'_{x_{H+1}}(\mathcal{S}_H). \tag{3.22}$$

We have the following lemma from [19]:

Lemma 3.3. *The fractional greedy solution has an approximation ratio of* $\frac{e}{e-1}$ *for problem*

(3.21). That is,

$$opt(\mathcal{B}) \leq \left(\frac{e}{e-1}\right) \widetilde{V}(\mathcal{S}),$$

where $opt(\mathcal{B})$ is the optimal value given bids set $\mathcal{B}$.

By utilizing Lemma 3.3, we have the following theorem:

Theorem 3.5. *ABSee achieves an approximation ratio of $\frac{2e}{\theta(e-1)}$.*

Proof. Recall that the winning bids $\mathcal{S} = \{x_1, x_2, \ldots, x_w\}$. For any $i \in \{w+1, \ldots, H\}$, we have

$$\frac{\bar{c}^{x_i}}{V_{x_i}(\mathcal{S}_{i-1})} \geq \frac{\bar{c}^{x_{w+1}}}{V_{x_{w+1}}(\mathcal{S}_{w+1})} > \frac{\theta G}{V(\mathcal{S}_{w+1})}.$$

Notice that x_i is no longer a winning bid when $i > w$. We obtain

$$\bar{c}^{x_i} > \theta G \frac{V_{x_i}(\mathcal{S}_{i-1})}{V(\mathcal{S}_{w+1})}$$

and

$$C^{x_{H+1}} > \theta G \frac{V'_{x_{H+1}}(\mathcal{S}_H)}{V(\mathcal{S}_{w+1})}.$$

Then, we have

$$\theta G \frac{\sum_{i=w+1}^{H} V_{x_i}(\mathcal{S}_{i-1}) + V'_{x_{H+1}}(\mathcal{S}_H)}{\sum_{i=1}^{w+1} V_{x_i}(\mathcal{S}_{i-1})} < \sum_{i=w+1}^{H} \bar{c}^{x_i} + C^{x_{H+1}} \leq G.$$

Thus, we obtain

$$\sum_{i=w+1}^{H} V_{x_i}(\mathcal{S}_{i-1}) + V'_{x_{H+1}}(\mathcal{S}_H) < \frac{\sum_{i=1}^{w+1} V_{x_i}(\mathcal{S}_{i-1})}{\theta}. \tag{3.23}$$

From (3.22) and (3.23), we have

$$\begin{aligned}
\widetilde{V}(\mathcal{S}) &= \sum_{i=1}^{H} V_{x_i}(\mathcal{S}_{i-1}) + V'_{x_{H+1}}(\mathcal{S}_H) \\
&= \sum_{i=1}^{w} V_{x_i}(\mathcal{S}_{i-1}) + \sum_{i=w+1}^{H} V_{x_i}(\mathcal{S}_{i-1}) + V'_{x_{H+1}}(\mathcal{S}_H) \\
&< \sum_{i=1}^{w} V_{x_i}(\mathcal{S}_{i-1}) + \frac{\sum_{i=1}^{w} V_{x_i}(\mathcal{S}_{i-1}) + V_{x_{w+1}}(\mathcal{S}_w)}{\theta} \\
&\leq (1+\frac{1}{\theta}) \sum_{i=1}^{w} V_{x_i}(\mathcal{S}_{i-1}) + \frac{V^{\max}}{\theta}.
\end{aligned}$$

Recall that $V(\mathcal{S}) = \sum_{i=1}^{w} V_{x_i}(\mathcal{S}_{i-1})$ and $\theta = 1 - \frac{V^{\max}}{V(\mathcal{S})}$. We have

$$\begin{aligned}
\widetilde{V}(\mathcal{S}) &< \frac{V(\mathcal{S}) + V(\mathcal{S}) - V^{\max}}{V(\mathcal{S}) - V^{\max}} V(\mathcal{S}) + \frac{V(\mathcal{S})}{V(\mathcal{S}) - V^{\max}} V^{\max} \\
&= \frac{2V(\mathcal{S})}{V(\mathcal{S}) - V^{\max}} V(\mathcal{S}) \\
&= \frac{2}{\theta} V(\mathcal{S}).
\end{aligned}$$

From Lemma 3.3, we have $opt(\mathcal{B}) \leq \frac{e}{e-1}\widetilde{V}(\mathcal{S})$. Then, we can bound the optimal value as

$$opt(\mathcal{B}) \leq \frac{e}{e-1}\widetilde{V}(\mathcal{S}) < \frac{2e}{\theta(e-1)} V(\mathcal{S}).$$

■

Thus, ABSee achieves an approximation ratio of $\frac{2e}{\theta(e-1)}$. When θ is close to 1, the approximation ratio approaches $\frac{2e}{e-1}$.

3.4 Performance Evaluation

In this section, we first evaluate the performance of the approach for estimating the quality of sensing of each user. We then compare the performance of ABSee with another budget feasible mechanism GREEDY-SM [19], which also satisfies all the desirable properties.

We assume that tasks and users are randomly distributed within a 1 km $\times$ 1 km region. A user can perform a task if the distance between the user and the task is less than 50 m. We assume each user submits 3 bids (i.e., $\sigma_n = 3$). The cost of user n for Γ_n^m (i.e., c_n^m) is $\eta_n|\Gamma_n^m|$, in which η_n is uniformly distributed over [1, 5]. The weight of task τ_k (i.e., μ_k) in (3.6) is uniformly distributed over [1, 10]. All results are obtained by averaging over 100 instances.

3.4.1 Quality of Sensing

The platform calculates and keeps a record of the quality indicators of the participating users. To simulate this process, consider the noise map application. We assume each task corresponds to the noise level of a specific position. The real noise level of task τ_k in the lth auction, denoted by $\delta_k^{(l)}$, is uniformly distributed over [0, 5]. The sensing data $\hat{\delta}_{k,n}^{(l)}$ provided by user n is generated from a Gaussian distribution $N(\delta_k^{(l)}, q_n)$, where the quality indicator q_n of each user is uniformly distributed over [0, 1]. After collecting sensing data for task τ_k, the platform calculates the estimated value $\hat{\delta}_k^{(l)}$ and updates the estimated quality indicator $\bar{q}_n^{(l)}$. At the beginning, the platform assumes that all users have the same estimated quality indicator as 0.5. The weight γ in (3.4) is selected as 0.5.

We use $V(\mathcal{S})$ to denote the valuation that the platform can obtain given the quality indicator

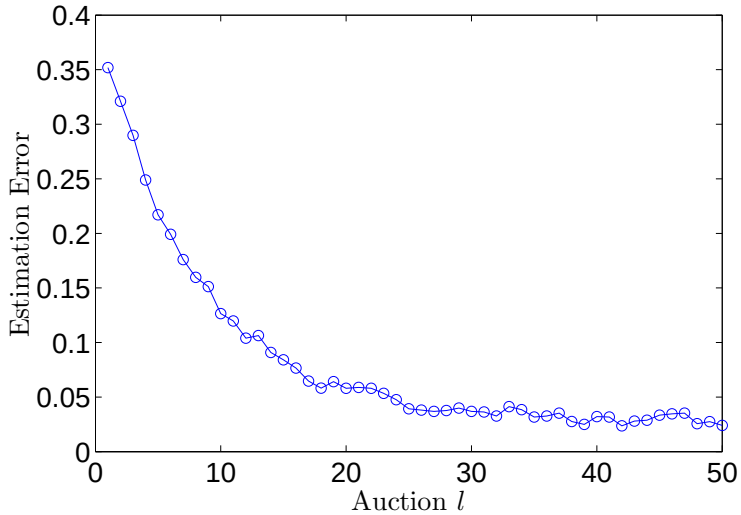

Figure 3.3: Estimation error versus the number of auctions l ($M = 500, N = 1000, G = 5000$). The smaller value of the estimation error results in the higher accuracy.

q_n of the users is known by the platform. Similarly, the valuation that the platform obtains according to its historical record of the estimated quality indicator $\bar{q}_n^{(l)}$ is denoted by $\bar{V}(\mathcal{S}^{(l)})$. Then, the estimation error of the quality of sensing can be calculated as $\frac{\left|V(\mathcal{S})-\bar{V}(\mathcal{S}^{(l)})\right|}{V(\mathcal{S})}$. Fig. 3.3 shows the estimation error of the quality of sensing. Results show that when the platform conducts more auctions, it obtains more accurate estimation of the quality indicators and selects the users more properly to perform the tasks.

3.4.2 ABSee

Figs. 3.4, 3.5, and 3.6 show the valuation function $V(\mathcal{S})$ obtained from both mechanisms. We can see that ABSee significantly improves the valuation obtained from GREEDY-SM. In Fig. 3.4, the value of $V(\mathcal{S})$ increases when there are more tasks. According to (3.6), an increase in the number of tasks can provide a higher valuation to the platform. The platform can also

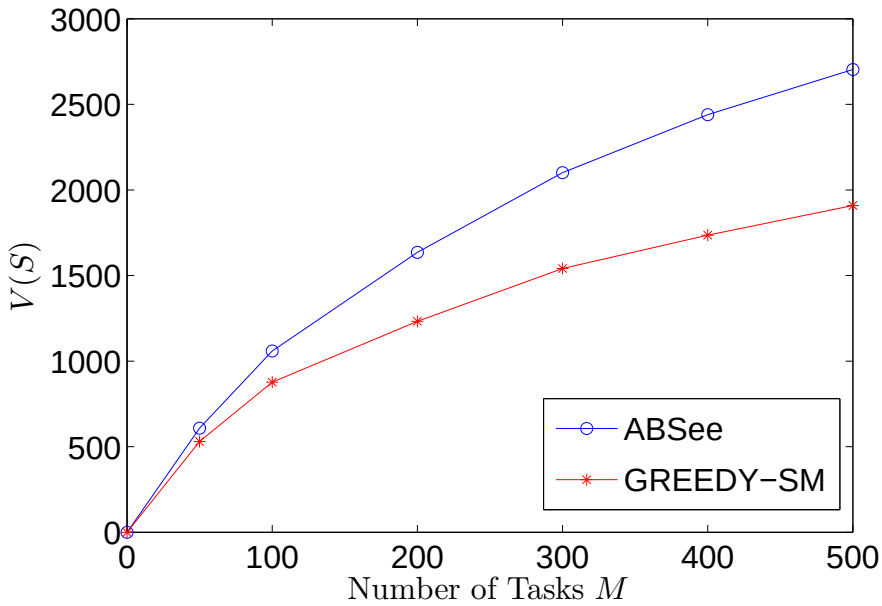

Figure 3.4: Valuation function $V(\mathcal{S})$ versus the number of tasks M ($N = 500$, $G = 500$).

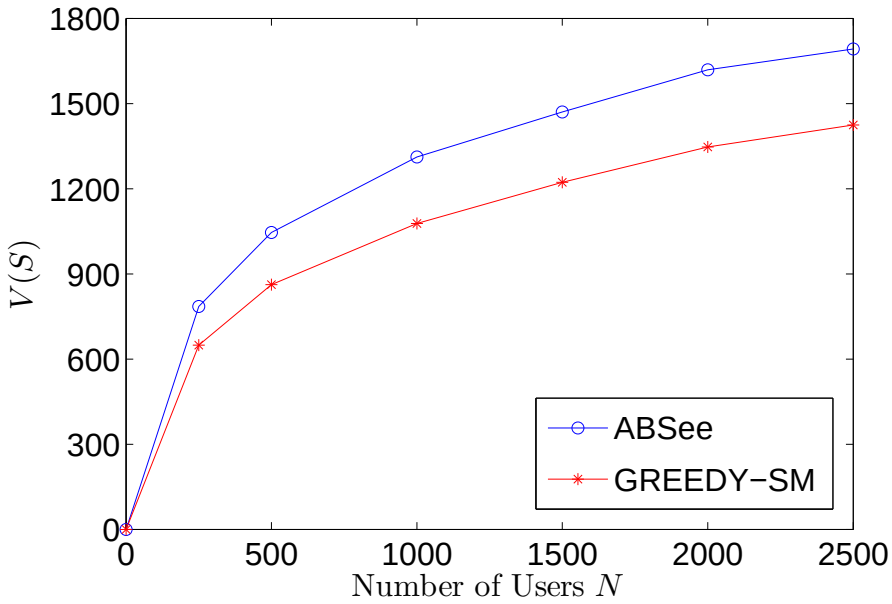

Figure 3.5: Valuation function $V(\mathcal{S})$ versus the number of users N ($M = 100$, $G = 500$).

obtain a higher valuation when the number of users increases as shown in Fig. 3.5. In this case, the number of users who can perform the tasks becomes larger. Therefore, those bids

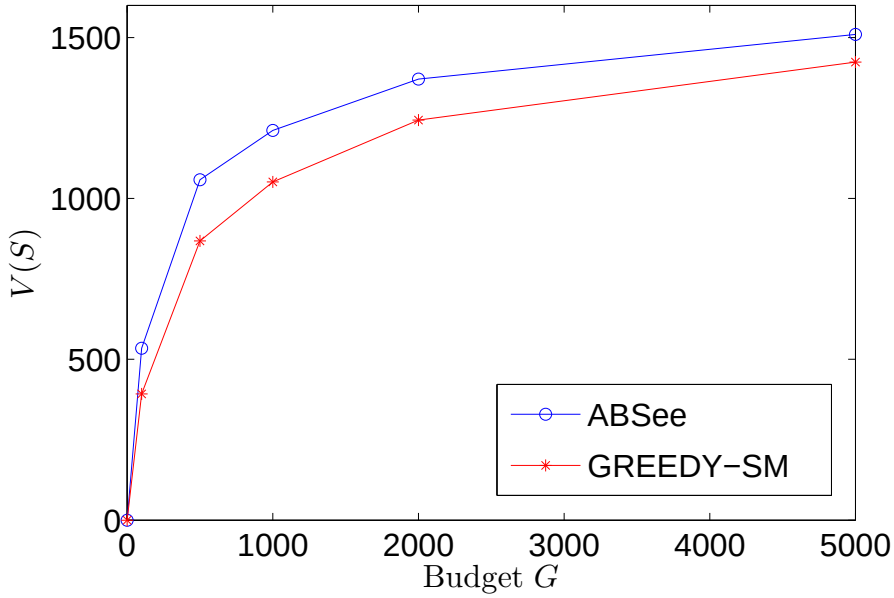

Figure 3.6: Valuation function $V(\mathcal{S})$ versus budget G ($M = 100$, $N = 500$).

with lower reserve price but higher quality of sensing can be chosen. From Fig. 3.6, the more budget the platform has, the higher valuation it can obtain. This is because it can select more winning bids to have the tasks performed.

Fig. 3.7 shows the crowd factor θ versus the number of users. When there are more users or more budget, the number of winning bids increases and thus the crowd factor becomes larger, which captures a practical mobile crowdsourcing system. Fig. 3.8 further shows the cumulative distribution function (CDF) of θ. Among 100 experiments, we can see that θ is close to 1 with high probability.

In Fig. 3.9, we verify the approximation ratio of ABSee. Since problem (3.21) is an NP-hard problem, it is time consuming to obtain the optimal value. Thus, we cannot compare the optimal value and the value obtained from ABSee directly. We circumvent issue by comparing

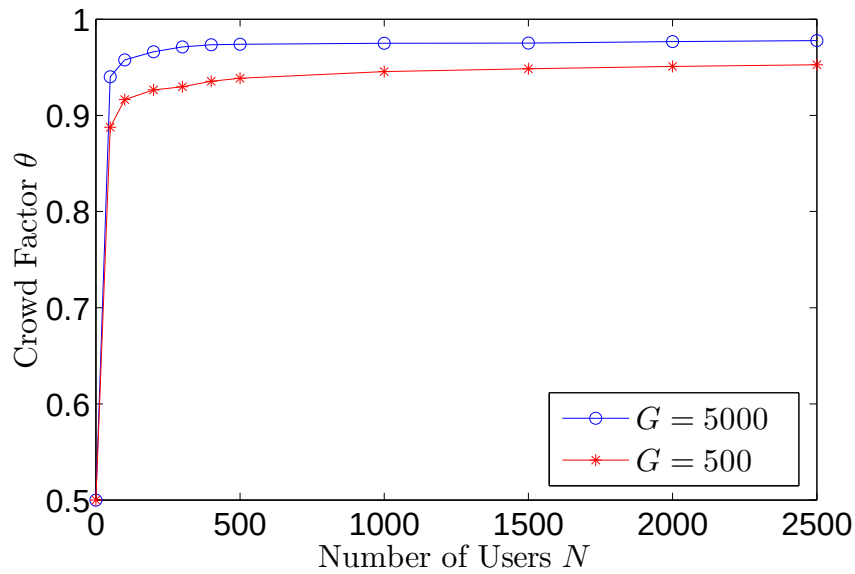

Figure 3.7: Crowd factor θ versus the number of users N ($M = 500$).

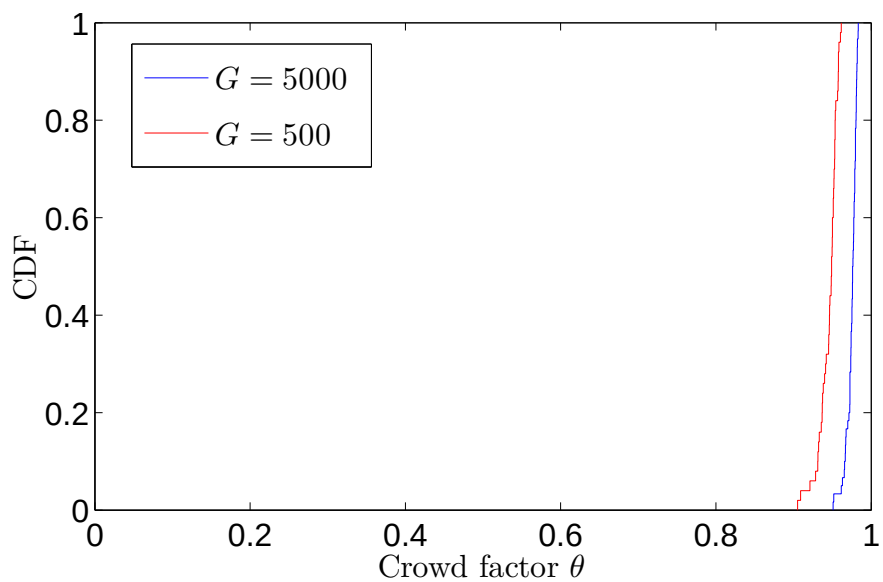

Figure 3.8: CDF of θ ($M = 500, N = 1000$).

$\widetilde{V}(\mathcal{S})$ and $V(\mathcal{S})$. Recall that $\widetilde{V}(\mathcal{S})$ is the valuation obtained from a fractional greedy algorithm. We know $opt(\mathcal{B}) \leq \frac{e}{e-1}\widetilde{V}(\mathcal{S}) < \frac{2e}{\theta(e-1)}V(\mathcal{S})$ from Section 3.3. Thus, we can validate the

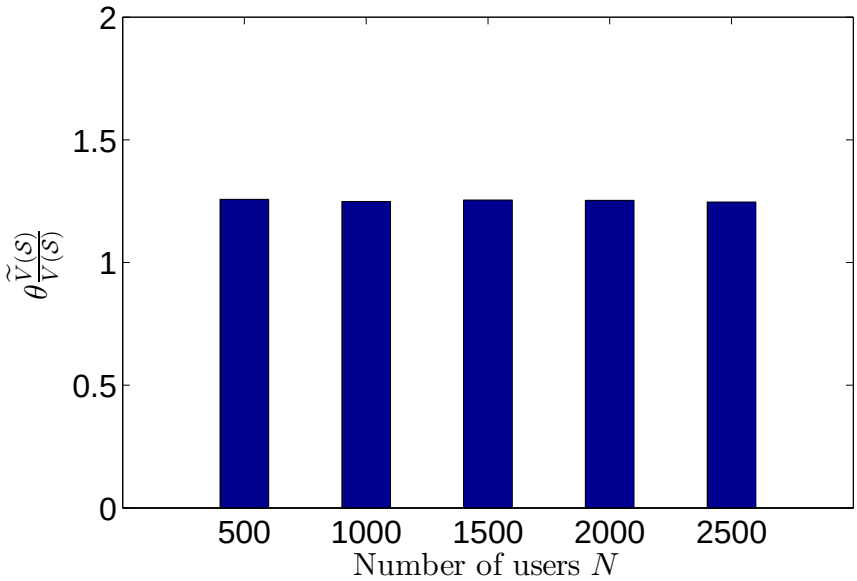

Figure 3.9: $\theta\frac{\widetilde{V}(\mathcal{S})}{V(\mathcal{S})}$ versus the number of users N ($M = 100$, $G = 100$). We have $\theta\frac{\widetilde{V}(\mathcal{S})}{V(\mathcal{S})} \leq 2$, which validates the approximation ratio.

approximation ratio by showing $\theta\frac{\widetilde{V}(\mathcal{S})}{V(\mathcal{S})} < 2$. Fig. 3.9 confirms that $\theta\frac{\widetilde{V}(\mathcal{S})}{V(\mathcal{S})}$ is always smaller than 2. Note that θ is close to 1 when the number of users and budget are large. Thus, the approximation ratio $\frac{opt(\mathcal{B})}{V(\mathcal{S})}$ is always less than $\frac{2e}{\theta(e-1)}$, which approaches to $\frac{2e}{e-1}$.

In Fig. 3.10, we compare the running time between ABSee and GREEDY-SM. ABSee is slightly slower than GREEDY-SM since it needs to calculate the crowd factor θ iteratively. However, it significantly improves the valuation of the platform.

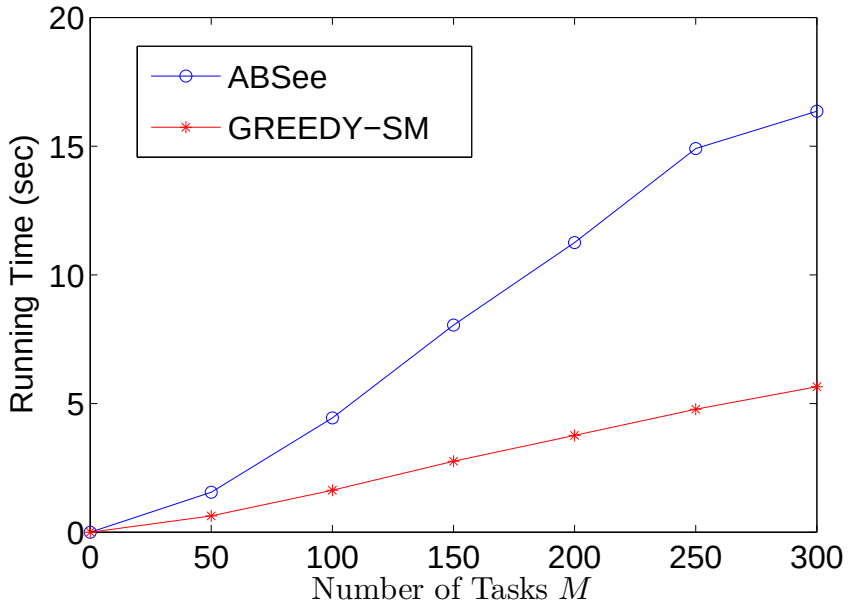

Figure 3.10: Running time versus the number of tasks M ($N = 500$, $G = 100$).

Chapter 4

Conclusion and Future Work

In this chapter, we conclude the thesis by summarizing the research work. We also suggest some possible extensions for future wok.

4.1 Conclusion

In this thesis, we considered quality of sensing of the smartphone users in a mobile crowdsourcing system. The platform aims to maximize the valuation of the performed tasks with a limited budget. The platform estimates the quality of sensing of the users and keeps a historical record. We designed an auction-based budget feasible mechanism called ABSee, which selects the winning bids and determines the payment to the users. In addition to budget feasibility, we proved that ABSee satisfies computational efficiency, truthfulness, and individual rationality. We also proved that ABSee can achieve an approximation ratio of $\frac{2e}{e-1}$ in the mobile crowdsourcing system. Simulation results showed that the quality of sensing of the users can be estimated accurately and the platform can obtain a higher valuation when it implements ABSee in comparison with GREEDY-SM [19].

4.2 Future Work

In terms of future work, possible extensions are as follows:

1. In this thesis, we consider the platform collects sensing data from many smartphone users with strategic behaviors. In this case, there is only one mobile crowdsourcing service consumer, i.e., the platform, and multiple mobile crowdsourcing service providers, i.e., smartphone users. In practice, there may exist many service consumers, all of which aim to maximize their valuations. Thus, developing a double auction mechanism would be a choice to solve this problem.

2. In this thesis, we propose a novel budget feasible mechanism for the platform with a budget constraint. We can further explore the mechanism design when the platform aims to minimize the total payment to the users, i.e., considering frugality in auctions.

3. In this thesis, smartphone users are assumed to submit some random subsets of tasks within the bids to simplify the model so that we can focus on mechanism design and improve the valuation of the platform. However, from the perspective of the users, they can develop some strategies and find specific combinations of tasks to maximize their utilities. This would be another interesting topic to explore.

Bibliography

[1] R. Ganti, F. Ye, and H. Lei, “Mobile crowdsensing: Current state and future challenges,” *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, Nov. 2011.

[2] SmartCitizen, “Smartcitizen,” 2015, online; accessed 10-July-2015. [Online]. Available: https://smartcitizen.me/.

[3] buUuk Pte Ltd, “Weatherlah,” 2015, online; accessed 10-July-2015. [Online]. Available: http://www.weatherlah.com/.

[4] Technische Universität Darmstadt, “Noisemap,” 2015, online; accessed 10-July-2015. [Online]. Available: https://play.google.com/store/apps/details?id=de.tudarmstadt.tk.noisemap.

[5] OpenSignal, “Opensignal,” 2015, online; accessed 10-July-2015. [Online]. Available: http://opensignal.com/.

[6] Sensorly, “Sensorly,” 2015, online; accessed 10-July-2015. [Online]. Available: http://www.sensorly.com.

[7] Waze Mobile, “Waze,” 2015, online; accessed 10-July-2015. [Online]. Available: https://www.waze.com.

[8] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. of ACM MobiCom*, Istanbul, Turkey, Aug. 2012.

[9] Z. Feng, Y. Zhu, Q. Zhang, L. Ni, and A. Vasilakos, "TRAC: Truthful auction for location-aware collaborative sensing in mobile crowdsourcing," in *Proc. of IEEE INFOCOM*, Toronto, Canada, Apr. 2014.

[10] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. of IEEE INFOCOM*, Hong Kong, Apr. 2015.

[11] T. Luo, H.-P. Tan, and L. Xia, "Profit-maximizing incentive for participatory sensing," in *Proc. of IEEE INFOCOM*, Toronto, Canada, Apr. 2014.

[12] L. Gao, F. Hou, and J. Huang, "Providing long-term participation incentive in participatory sensing," in *Proc. of IEEE INFOCOM*, Hong Kong, Apr. 2015.

[13] D. Zhao, X.-Y. Li, and H. Ma, "How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint," in *Proc. of IEEE INFOCOM*, Toronto, Canada, Apr. 2014.

[14] L. Tran-Thanh, M. Venanzi, A. Rogers, and N. R. Jennings, "Efficient budget allocation with accuracy guarantees for crowdsourcing classification tasks," in *Proc. of ACM Int'l Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*, Saint Paul, MN, May 2013.

[15] K. Han, C. Zhang, and J. Luo, "BLISS: Budget limited robust crowdsensing through online learning," in *Proc. of IEEE SECON*, Singapore, Jun. 2014.

[16] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proc. of ACM IPSN*, Beijing, China, Apr. 2012.

[17] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proc. of ACM SIGMOD*, Snowbird, UT, Jun. 2014.

[18] Y. Singer, "Budget feasible mechanisms," in *Proc. of IEEE Symposium on Foundations of Computer Science (FOCS)*, Las Vegas, NV, Oct. 2010.

[19] N. Chen, N. Gravin, and P. Lu, "On the approximability of budget feasible mechanisms," in *Proc. of ACM-SIAM Symposium on Discrete Algorithms (SODA)*, San Francisco, CA, Jan. 2011.

[20] E. Davami and G. Sukthankar, "Improving the performance of mobile phone crowdsourcing applications," in *Proc. of ACM Int'l Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*, Istanbul, Turkey, May 2015.

[21] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *Proc. of IEEE INFOCOM*, Turin, Italy, Apr. 2013.

[22] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *Proc. of ACM MobiHoc*, Hangzhou, China, Jun. 2015.

[23] Q. Zhang, Y. Wen, X. Tian, X. Gan, and X. Wang, "Incentivize crowd labeling under budget constraint," in *Proc. of IEEE INFOCOM*, Hong Kong, Apr. 2015.

[24] D. Peng, F. Wu, and G. Chen, "Pay as how well you do: A quality based incentive mechanism for crowdsensing," in *Proc. of ACM MobiHoc*, Hangzhou, China, Jun. 2015.

[25] Y. Singer, "How to win friends and influence people, truthfully: Influence maximization mechanisms for social networks," in *Proc. of ACM WSDM*, Seattle, WA, Feb. 2012.

[26] I. Koutsopoulos and M. Halkidi, "Lifetime maximization in wireless sensor networks with an estimation mission," in *Proc. of IEEE GLOBECOM*, Honolulu, HI, Nov. 2009.

[27] D. M. Topkis, *Supermodularity and Complementarity*. Princeton University Press, 1998.

[28] R. B. Myerson, "Optimal auction design," *Mathematics of Operations Research*, vol. 6, no. 1, pp. 58–73, Feb. 1981.

[29] A. Krause and C. Guestrin, "A note on the budgeted maximization of submodular functions," *CMU Technical Report*, pp. CMU–CALD–05–103, 2005.