Incentive Mechanism for Spatial Crowdsourcing With Unknown Social-Aware Workers: A Three-Stage Stackelberg Game Approach

Yin Xu, Mingjun Xiao, *Member, IEEE*, Jie Wu, *Fellow, IEEE*, Sheng Zhang, *Member, IEEE*, and Guoju Gao

Abstract—In this paper, we investigate the incentive problem in Spatial Crowdsourcing (SC), where mobile social-aware workers have unknown qualities and can share their answers to tasks via social networks. The objectives are to recruit high-quality workers and maximize all parties' utilities simultaneously. However, most existing works assume that the qualities of workers are known in advance or cannot take all parties' utilities into account together, especially having not considered the impact of social networks. Thus, we propose an incentive mechanism based on the multi-armed bandit and three-stage Stackelberg game, called TACT. We first design a greedy arm-pulling scheme to recruit workers, which not only can solve the exploration-exploitation dilemma but also takes workers' social relations into account. Based on the recruitment results, we further design the utility functions incorporating with social benefits for workers, and model the payment computation problem as a three-stage Stackelberg game among all participants. Next, we derive the optimal strategy group so that each party can maximize its own utility to form a multi-win situation. Moreover, we theoretically prove the unique existence of Stackelberg equilibrium and the worst regret bound. Finally, we conduct extensive simulations on a real trace to corroborate the performance of TACT.

Index Terms—Incentive mechanism, multi-armed bandit, social network, spatial crowdsourcing, stackelberg game

1 Introduction

With the unprecedented pervasiveness of GPS-enabled smart devices and the development of sharing economy, Spatial Crowdsourcing (SC) has been recognized as a promising paradigm in utilizing the crowd power to address a large-scale of complex tasks [1], [2], [3], [4], [5], [6]. A typical SC system consists of three parties: task requesters, a crowd of workers, and a cloud platform. Requesters can outsource a variety of location-related tasks to workers via the cloud platform, and then workers will physically move to some locations to accomplish the corresponding tasks. Since SC can fulfill a great quantity of spatial tasks that individuals cannot deal with, many commercial SC applications have emerged prosperously, such as Waze [7] for traffic monitoring, Gigwalk [8] for mobile business, and so on.

In this paper, we take wisdom traveling in an open SC system as an example, where a requester (e.g., wisdom traveling service provider) outsources the traffic flow detection task to the crowd via the SC platform, as illustrated in Fig. 1. In such a system, some social network users (called *social-aware workers*) can use their smart devices to collect the traffic data (e.g., road picture, vehicle velocity, congestion index, etc.) on some specified roads. Meanwhile, these workers can share their traffic data with each other through socially-connected networks (e.g., Facebook, Twitter, and WeChat), so as to acquire some social benefits (e.g., replan a better route). When workers upload collected traffic data as their answers to the traffic flow detection task, the platform can aggregate the traffic data and return them to the requester. Due to the openness, various workers can enter or leave the system at any time. These workers might provide the traffic data with different qualities due to their diverse smart devices. Their data qualities are generally *unknown* to the requester and the platform. Here, we focus on the incentive mechanism design for such an open SC system. Due to the unknown social-aware workers, this mechanism design suffers two major challenges:

The first challenge is the recruitment problem of unknown social-aware workers. As a crucial part of the incentive mechanism, the SC platform will recruit some workers with high-quality traffic data so that the requester can obtain valuable answers for the travel recommendation service. However,

- *Yin Xu and Mingjun Xiao are with the School of Computer Science and Technology /Suzhou Institute for Advanced Research / the CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China (USTC), Hefei 230026, China. E-mail: xuyin218@mail.ustc.edu.cn, xiaomj@ustc.edu.cn.*
- *Jie Wu is with the Center for Networked Computing, Temple University, Philadelphia PA 19122 USA. E-mail: jiewu@temple.edu.*
- *Sheng Zhang is with the State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China. E-mail: sheng@nju.edu.cn.*
- *Guoju Gao is with the School of Computer Science and Technology, Soochow University, Suzhou 215006, China. E-mail: gjgao@suda.edu.cn.*

Manuscript received 17 Dec. 2021; revised 2 Feb. 2022; accepted 4 Mar. 2022. Date of publication 8 Mar. 2022; date of current version 30 June 2023.
This work was supported in part by the National Key R&D Program of China under Grant 2018AAA0101204, in part by the National Natural Science Foundation of China (NSFC) under Grants 62172386, 62102275, 61872175, 61872330, 61936015, and 61572457, in part by the NSF under Grants CNS 1824440, CNS 1828363, CNS 1757533, CNS 1618398, CNS 1651947, and CNS 1564128, in part by the Natural Science Foundation of Jiangsu Province in China under Grants BK20191194, BK20210704, BK20131174, and BK2009150, in part by Anhui Initiative in Quantum Information Technologies under Grant AHY150300, and in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 21KJB520025.
(Corresponding author: Mingjun Xiao.)
Digital Object Identifier no. 10.1109/TMC.2022.3157687

Fig. 1. Illustrate of the SC scenario with social-aware workers.

each worker's answer quality is non-public, generally following some unknown distributions. Thus, the platform needs to estimate workers' quality distributions through online learning, i.e., the so-called *exploration* process. Meanwhile, the platform might directly select the workers whose qualities have been explored, a.k.a., the *exploitation* process. In order to maximize the total answer quality, we need to determine an optimal trade-off to balance exploration and exploitation. Moreover, since workers can bring the extra social benefits to the whole SC system by sharing their traffic data with others, we also need to take workers' social relations into account. In order to obtain the larger social benefits, the workers with more socially-connected friends will also be selected preferentially, making the recruitment problem more challenging.

Another challenge is the payment computation method of the incentive mechanism, which involves a complex three-party game issue. In the system, the requester can produce value by receiving the answers to the task from workers via the platform. Since workers and the platform might spend some cost in collecting and aggregating traffic data, the requester needs to pay them some rewards as compensation. When determining the payments, each of them will try to maximize its own utility (i.e., the net profit), forming a three-party game. The requester hopes to decrease the unit price of the received answers, so as to save its expenses. As a countermeasure, workers might affect the quality of recommendations (derived from the aggregated answers) as well as the corresponding income by controlling the frequency of uploading answers. In addition, the workers' social benefits also need to be taken into consideration, which makes the game more challenging.

So far, a wide spectrum of incentive mechanisms have been developed for SC systems [9], [10], [11], [12], [13], [14], [15], but most of them assume that workers' qualities are known in advance. Only a few works have studied the unknown worker recruitment problem with pricing concerns [16], [17], [18], [19]. Nevertheless, they mainly involve two parties and do not take social networks into consideration. In short, none of the previous researches has addressed the two challenges together.

Inspired by the above considerations, we propose TACT, an incenTive mechAnism based on Combinatorial Multi-Armed Bandit (CMAB) and Three-stage Hierarchical Stackelberg (THS) game for SC. More Specifically, we first model the unknown social-aware worker recruitment process as a K-arm CMAB problem and design a greedy arm-pulling scheme by extending the Upper Confidence Bound (UCB) policy, which can solve the exploration-exploitation dilemma (i.e., a trade-off between exploring potential high-quality workers and exploiting online learning results). Next, according to the worker recruitment together with the online learning results, we design the utility functions incorporating with social benefits for workers. Based on these utility functions, we further model the payment computation problem as a THS game to determine the optimal strategy group, in which the requester is the first-tier leader, the platform is the second-tier leader, and recruited social-aware workers are followers. Finally, we analyze the worst regret bound and prove that there exists a Stackelberg Equilibrium (SE) by the game theory, i.e., each participant can obtain its own optimal strategy to enter into a multi-win situation.

To sum up, our multi-fold contributions are as follows.

1) We propose an incentive mechanism (called TACT) based on CMAB and THS, consisting of the unknown social-aware worker recruitment and the optimal strategy group determination. Unlike existing researches, TACT combines the reinforcement learning and Stackelberg game for SC with unknown social-aware workers, which can guarantee efficient recruitment and maximize the utilities of all parties simultaneously.
2) We design a greedy arm-pulling scheme based on CMAB to recruit suitable workers. Not only can it balance the exploration and exploitation processes, but also it takes workers' social relations into account. Moreover, we derive the regret bound.
3) By jointly considering all participants' selfish behaviors and workers' social benefits, we formulate the payment computation process as a THS game and derive the closed-form optimal strategy group for each party. Through the theoretical analysis, there is a unique Stackelberg Equilibrium so that no one has incentives to unilaterally deviate from its optimal decisions.
4) We conduct extensive simulations on real-world traces to demonstrate the significant performance of TACT.

The remainder of the paper is organized as follows. In Section 2, we introduce the system model and the problem formulation. The detailed incentive mechanism design is elaborated in Section 3. In Section 4, we introduce the theoretical analysis in great detail. The simulations and evaluations are illustrated in Section 5. We discuss the related works in Section 6, and conclude the paper in Section 7.

2 Model and Problem

2.1 System Model

We consider an SC system, consisting of a requester (which can be extended to multiple requesters easily), a crowd of unknown social-aware workers, and a cloud platform. We give the definition of each party as follows.

Definition 1 (Requester and Task). *The requester has a long-term SC task (e.g., the traffic flow detection task), which needs to collect data at multiple locations for a long time. The task can be expressed by* $Task \triangleq \langle Loc, T, K, \Delta, \Delta_{min} \rangle$. *Here,* Loc *is a candidate location set where the task allows to be performed. Besides, the long-term task operates in a time-slotted manner, and the whole process is divided into* T *periods. The requester is allowed to decide the number of demanded workers in each period, denoted by* K*, which is a preset constant according to task demands. Also, each period* $t \in \{1, 2, \ldots, T\}$ *has the same duration* Δ *(e.g., one day), and the minimum uploading interval is* Δ_{min} *(e.g., one hour).*

Definition 2 (The Platform). *As a profit-making intermediate agent, the platform would charge some monetary reward from the requester for the statistics service, from which it will extract a proportion of the reward as its own commission to compensate for aggregation cost. Meanwhile, for stimulating workers to try their best to fulfill tasks, the platform will pay remuneration for the recruited workers to make up for their task completion cost.*

Definition 3 (Unknown Social-aware Workers). *There are N workers with unknown quality and they are social network users, denoted by $\mathcal{N}=\{1,2,\ldots,N\}$. A recruited worker $i\in\mathcal{N}$ will travel to some locations selected from Loc and then upload answers (e.g., collected data in traffic flow detection task) to the platform. Moreover, these N workers are available in all T periods.*

In each period t, the number of times that worker i uploads collected data is δ_i^t, where $\delta_i^t\in[0,\Delta/\Delta_{min}]$, called answer frequency*. Let $q_{i,k}^t\in[0,1]$ ($k\in[1,\delta_i^t]$) indicate the observed answer quality of worker i for the k-th answer in the t-th period. The values of $\{q_{i,k}^t|i\in\mathcal{N},t\in[1,T]\}$ follow an independent and identically unknown distribution with an unknown expectation q_i, i.e., $q_i\triangleq\mathbb{E}(q_{i,k}^t)$. Here, q_i denotes the worker i's real quality, which is mainly determined by the worker's device and collection habit. Note that, the higher the answer frequency, the larger the total size of collected data will be. Let $\boldsymbol{\delta}^t=\{\delta_1^t,\delta_2^t,\ldots,\delta_N^t\}$ mean the answer frequency vector of N workers in the t-th period, and we interchangeably use answer and collected data for ease of exposition in the context.*

In the SC system, social-aware workers can benefit from shared information through social relations, which significantly improves the system efficacy and can be defined as:

Definition 4 (Social Network). *A social network is presented by $\Phi=[\phi_{ij}]_{N\times N}$, where $\phi_{ij}>0$ represents the strength of the influence of worker j on worker i. Let $\chi_i=|\{\phi_{ij}|\phi_{ij}\neq 0\}|/(N-1)$ denote the normalized degree centrality of worker i, and we assume that bilateral interactions are symmetric (i.e., $\phi_{ij}=\phi_{ji},\phi_{ii}=0$) due to the reciprocity of social relations. A larger ϕ_{ij} means the stronger social relation and the greater influence.*

Definition 5 (Social Benefit). *There is a significant phenomenon, namely social network effects, which traditionally means that increasing the usage level of a consumer has a positive impact on the usage levels of her socially-connected friends [20], [21]. That is, a worker can obtain a social benefit from shared information. We model the benefit as $G_s(\delta_i^t,\boldsymbol{\delta}_{-i}^t)=\sum_{j\in\mathcal{N}}\phi_{ij}\delta_i^t\delta_j^t$, which is widely adopted to imply the positive externalities and local complementarities [22], [23], [24]. The product form captures that a worker will benefit more as its answers increase or as its social friends increase their answers [20]. Here, $\boldsymbol{\delta}_{-i}^t$ denotes the vector of answer frequency of all other workers except the worker i.*

To stimulate all parties (i.e., the requester, the platform, and workers), we need an incentive mechanism in which each party can take a strategy to balance its income and expenditure. The strategy group is defined as follows.

Definition 6 (Strategy Group). *The requester needs to give the platform a payment, which will be divided in proportion to the answer frequency. Thus, the requester must determine a unit payment in each period (i.e., the price per uploading), denoted by $\beta^t\in\{\beta_{min},\beta_{max}\}$. Likewise, the profit-making platform can also determine a unit recruitment salary for each worker, denoted by $\mathbf{p}^t=\{p_1^t,p_2^t,\ldots,p_N^t\},\forall p_i^t\in\{p_{min},p_{max}\}$. Besides, each worker can elaborately choose the answer frequency δ_i^t as a strategy. In brief, we use a triple $\langle\beta^t,\mathbf{p}^t,\boldsymbol{\delta}^t\rangle$ to indicate the strategy group of three parties in the t-th period. Once the strategy group is determined, the requester would afford the platform the payment (i.e., the unit payment multiplied by the whole answer frequency), and the platform will pay the worker i the salary (i.e., the unit salary multiplied by its answer frequency).*

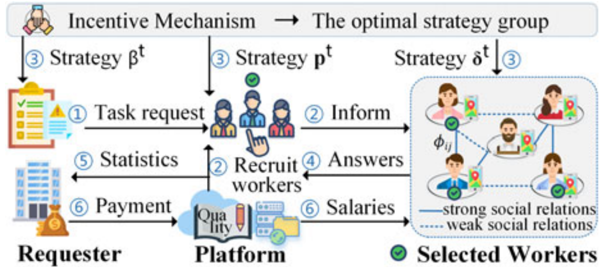


Fig. 2. The workflow in the t-th period.

As illustrated in Fig. 2, the whole workflow includes:

1) The requester publishes a task request to the platform, which will be conducted period by period until T.
2) Based on the selection criteria (Section 3.1), the platform recruits suitable workers with qualities as high as possible, and then informs the task to these workers.
3) In order to maximize their own utilities, the requester, the platform, and the recruited workers will devise the best individual strategy according to the proposed incentive mechanism (Section 3.2), respectively.
4) Each recruited worker carries out the task in some specified locations and returns answers abiding by the determined strategy (i.e., δ_i^t). In addition, these social-aware workers can share their own answers through social networks and gain information from the network.
5) The platform aggregates answers and sends the statistics to the requester. Meanwhile, the learned qualities of recruited workers will be updated for the next recruitment.
6) Based on the strategy group, the requester will offer a payment to the platform, and each recruited worker can obtain the salary from the platform and the social benefit.

Design Goals. In the above system, we concentrate on the incentive mechanism design. First, we need to complete the effective recruitment among unknown social-aware workers. After selecting suitable workers, we need to determine the optimal strategy group for all participants (i.e., the requester, the platform, and the recruited workers). More specifically, two major functionalities of the incentive mechanism need to be achieved as follows.

The unknown social-aware worker recruitment. Faced with the unknown qualities of workers, we need to determine an optimal trade-off to balance exploration and exploitation. In addition, there exist extra social benefits through social

TABLE 1
Description of Major Notations

Variable	Description
i, t, k	the indexes of worker, period, and answer.
N, T	the number of workers and the total period.
$K, \mathcal{W}^t$	the number and the set of recruited workers in t.
Δ, Δ_{min}	the duration of one period and minimum interval.
$\tilde{q}_i^t, \hat{q}_i^t$	the empirical quality and the CUCB-based quality.
Φ, ϕ_{ij}	the social network and the effect strength of j on i.
δ_i^t, χ_i	answer frequency, normalized degree centrality.
$\beta^t, \boldsymbol{p}^t$	the unit payment, unit recruitment salary vector.
U_i^t, U_p^t	the utility of worker i and the platform in t.
U_r^t, ξ	the utility of requester and the system parameter.
n_i^t, φ_i^t	the learned times, the confidence bound in t.
G_s, G_r	the social benefit of worker i, the requester's gain.
C_i, a_i, b_i	the cost function and parameters of the worker i.
C_p, c, d	the cost function and parameters of the platform.
ψ_i^t	indicate whether worker i is recruited in period t.
$\boldsymbol{\delta}^t$	the answer frequency of all workers.

relations. Therefore, the degree centrality of each worker should also be considered. In short, we aim to maximize the total data quality while considering social relations.

The optimal strategy group determination. First, we need to design the utility functions for all participants, during which we need to utilize the online learning results in the worker recruitment. Each participant wishes to maximize its utility by manipulating its own strategy (i.e., β^t or $\boldsymbol{p}^t$ or $\boldsymbol{\delta}^t$): the requester tends to obtain more high-quality answers with a lower payment; the platform also hopes to make a good profit with lower recruitment salary expenses; and each recruited social-aware worker wants to balance the income (including the salary and the social benefit) and its cost by controlling the frequency of uploading answers. Since the utilities of participants affect each other, there exists a complex three-party game. In order to achieve an equilibrium of the game (i.e., a multi-win situation for all participants), we need to derive the optimal strategy group while taking the social network effects into consideration.

Additionally, for ease of reference, we summarize the major notations throughout the paper in Table 1.

2.2 Formulation of Unknown Worker Recruitment

Recruiting unknown workers is actually an online learning and sequential decision-making problem, and the CMAB model is suitable to cope with it. Traditionally, a CMAB model consists of a slot machine with multiple arms [25], [26]. Pulling an arm would earn a reward which follows an unknown distribution. A decision-maker will pull a set of arms (called a super arm) together period by period based on a bandit policy to maximize the cumulative revenue. Therefore, we model our unknown worker recruitment issue as a K-arm CMAB, where each worker is regarded as an arm, its answer quality is seen as the corresponding reward, and recruiting K workers is equivalent to pulling a super arm. In each period, K workers are recruited and their answer qualities would be learned. Recall that, $K \in \{1, 2, \ldots, N\}$ is a preset value decided by the requester according to its task demands. Specially, $K=1$ is an MAB problem.

Based on the above modeling, the worker recruitment problem is transformed into a bandit policy determination problem, which aims to maximize the total task completion quality. A bandit policy Ψ is a sequence of mappings: $\Psi=\{\Psi^1, \ldots, \Psi^t, \ldots, \Psi^T\}$, where $\Psi^t=\{\psi_1^t, \psi_2^t, \ldots, \psi_N^t\}$ and $\psi_i^t \in \{0,1\}$ is an indicator variable to denote whether the worker i is recruited or not in the t-th period. Thus, the recruitment problem can be formulated as follows.

$$Maximize\ \mathbb{E}[Q(\Psi)] = \mathbb{E}\left[\sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{N} \sum\nolimits_{k=1}^{\delta_i^t} q_{i,k}^t \psi_i^t\right] \quad (1)$$

$$Subject\ to: \psi_i^t \in \{0,1\}, \forall i \in \mathcal{N}, \forall t \in [1,T] \quad (2)$$

$$\sum_{i=1}^{N} \psi_i^t = K, \forall t \in [1,T] \quad (3)$$

$$\delta_i^t \in [1, \Delta/\Delta_{min}], \forall i \in \mathcal{N}, \forall t \in [1,T] \quad (4)$$

2.3 Formulation of Optimal Strategy Group

We model the payment computation problem as a Three-stage Hierarchical Stackelberg (THS) game, so as to determine the optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$. In each period, the requester can adjust the unit payment β^t to affect the profit of the platform, and the platform can manipulate the unit salary vector $\boldsymbol{p}^t$ to determine the reward of each recruited worker. Meanwhile, workers can affect the utility of the requester and the social benefits by customizing their own answer frequency δ_i^t, in turn. We first design the utility functions for all parties as follows.

<u>*The utility of each social-aware worker*</u>: Each worker i's utility is the salary from the platform and the social benefit from the social network minus its cost in each period.

$$\begin{aligned} U_i^t(\delta_i^t, p_i^t) &= p_i^t \delta_i^t \psi_i^t + G_s(\delta_i^t, \boldsymbol{\delta}_{-\boldsymbol{i}}^t)\psi_i^t - C_i(\tilde{q}_i^t, \delta_i^t)\psi_i^t \\ &= p_i^t \delta_i^t \psi_i^t + \sum_{j \in \mathcal{N}} \phi_{ij} \delta_i^t \delta_j^t \psi_i^t \psi_j^t - [\psi_i^t \tilde{q}_i^t (a_i(\delta_i^t)^2 + b_i \delta_i^t)]. \end{aligned} \quad (5)$$

The first term is the salary paid by the platform. The second term denotes the social benefit. A social-aware worker can enjoy a benefit from the information shared by the socially-connected workers. The third term displays the answer cost, which is assumed to be a continuous and reversible convex function. In this paper, we adopt a widely-used quadratic function (like in [23], [27], [28]) to model the increasing marginal cost for every additional unit of effort exerted, i.e., $C_i(\tilde{q}_i^t, \delta_i^t) = \tilde{q}_i^t (a_i(\delta_i^t)^2 + b_i \delta_i^t)$. Here, $a_i > 0$ and $b_i \geq 0$ are different among workers owing to the heterogeneity of devices. Since the real quality q_i is unknown, we will adopt a learned empirical quality $\tilde{q}_i^t$, which is the output of the online learning process in the worker recruitment.

<u>*The utility of platform*</u>: The platform's utility is the payment from the requester minus its cost. We mainly consider the aggregation cost and salary cost.

$$U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t) = \beta^t \sum_{i=1}^{N} \delta_i^t \psi_i^t - C_p(\boldsymbol{\delta}^t, \Psi^t) - \sum_{i=1}^{N} p_i^t \delta_i^t \psi_i^t$$

$$= \beta^t \sum_{i=1}^{N} \delta_i^t \psi_i^t - \left[c\left(\sum_{i=1}^{N} \delta_i^t \psi_i^t\right)^2 + d\sum_{i=1}^{N} \delta_i^t \psi_i^t\right] - \sum_{i=1}^{N} p_i^t \delta_i^t \psi_i^t. \quad (6)$$

The first term is the payment paid by the requester. The second term is the cost for aggregating answers. As in [23], [27], [28], we also choose a quadratic function: $C_p(\boldsymbol{\delta}^t, \Psi^t) = c(\sum_{i=1}^N \delta_i^t \psi_i^t)^2 + d\sum_{i=1}^N \delta_i^t \psi_i^t$, where $c > 0, d \geq 0$. The third term denotes the total salary paid to all recruited workers.

The utility of requester: The requester's utility is the difference between the gain and its cost in each period t, i.e.,

$$\begin{aligned} U_r^t(\boldsymbol{\delta}^t, \beta^t) &= G_r(\tilde{q}^t, \boldsymbol{\delta}^t, \Psi^t) - \beta^t \sum_{i=1}^N \delta_i^t \psi_i^t \\ &= \xi \ln\left(1 + \tilde{q}^t \sum_{i=1}^N \delta_i^t \psi_i^t\right) - \beta^t \sum_{i=1}^N \delta_i^t \psi_i^t. \end{aligned} \tag{7}$$

The first term is the gain brought by the statistics of answers, and the second term is the cost used for the payment. Like in [28], [29], we use the typical logarithmic function to model the gain: $G_r(\tilde{q}^t, \boldsymbol{\delta}^t, \Psi^t) = \xi \ln(1 + \tilde{q}^t \sum_{i=1}^N \delta_i^t \psi_i^t)$. Here, $\xi > 0$ is a system parameter and $\tilde{q}^t = \sum_{i=1}^N \tilde{q}_i^t \psi_i^t / \sum_{i=1}^N \psi_i^t$. Note that, $G_r(\cdot)$ satisfies the law of diminishing marginal utility, i.e., G_r rises along with the answer frequency but the increasing rate of G_r decreases with the answer frequency.

Based on the above function designs, we can formulate the optimal strategy group determination as follows.

Formulation of determining the optimal strategy group: In the first stage, the requester plays a role of the first-tier leader. It will determine an optimal unit payment β^{t*}, which ensures that the requester cannot improve its utility by adopting any other strategy β^t when each worker has determined its optimal strategy. Similarly, the platform is the second-tier leader and decides an optimal unit salary vector $\boldsymbol{p}^{t*}$ to obtain maximum profit under the given unit payment β^t. In the third stage, workers are followers and customize an optimal answer frequency vector $\boldsymbol{\delta}^{t*}$ under the given unit salary vector $\boldsymbol{p}^t$. In short, the payment computation process is devoted to finding an optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$, which guarantees that no participant can improve its own utility by unilaterally deviating from its optimal strategy, i.e.,

$$\textbf{Stage I [Requester] :}\ U_r^t(\delta_i^{t*}, \beta^{t*}) \geq U_r^t(\delta_i^{t*}, \beta^t) \tag{8}$$

$$\textbf{Stage II [Platform] :}\ U_p^t(\delta_i^{t*}, \boldsymbol{p}^{t*}, \beta^{t*}) \geq U_p^t(\delta_i^{t*}, \boldsymbol{p}^t, \beta^{t*}) \tag{9}$$

$$\textbf{Stage III [Worker i] :}\ U_i^t(\delta_i^{t*}, p_i^{t*}) \geq U_i^t(\delta_i^t, \boldsymbol{\delta}_{-i}^{t*}, p_i^{t*}) \tag{10}$$

Actually, the three inequalities indicate the equilibrium of the Stackelberg game, which is defined as follows:

Definition 7 (Stackelberg Equilibrium (SE)). *An optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$ for $\forall t \in [1, T]$ constitutes a Stackelberg equilibrium of the THS game* iff *the inequalities Eqs. (8), (9), and (10) are satisfied.*

3 Incentive Mechanism Design

In this section, we propose the incentive mechanism TACT. First, we design a greedy arm-pulling scheme to select workers. Based on the recruitment results, we derive the optimal strategy group by exploiting the backward induction method. In the following subsection, we present the basic idea of the unknown social-aware recruitment and derive the optimal strategy group, followed by a detailed algorithm and a straightforward example.

3.1 Unknown Social-Aware Worker Recruitment

In order to balance the trade-off between exploration (i.e., trying some sub-optimal workers to find the potential optimal workers) and exploitation (i.e., recruiting the current best workers based on the learned results), we design a Combinatorial UCB (CUCB)-based quality index as the selection criteria, taking social networks into consideration. The detailed design is expounded as follows.

When a worker i is recruited in the t-th period (i.e., $\psi_i^t = 1$), i will upload answers abiding by the answer frequency δ_i^t. So, the number of times of i's quality being learned by the platform in the t-th period is actually δ_i^t. Based on this, we first introduce n_i^t for $i \in \mathcal{N}, t \in [1, T]$ to record the number of times that i's quality has been learned. After selected workers upload answers, the value of n_i^t at the end of t is updated using the following equations:

$$n_i^t = \begin{cases} n_i^{t-1} + \delta_i^t, & if\ \psi_i^t = 1 \\ n_i^{t-1}, & if\ \psi_i^t = 0. \end{cases} \tag{11}$$

Then, we introduce $\tilde{q}_i^t$ to record the learned empirical quality for the worker i until the t-th period. After selected workers finish the task, the value of $\tilde{q}_i^t$ is updated:

$$\tilde{q}_i^t = \begin{cases} \frac{\tilde{q}_i^{t-1} n_i^{t-1} + \sum_{k=1}^{\delta_i^t} q_{i,k}^t}{n_i^{t-1} + \delta_i^t}, & if\ \psi_i^t = 1 \\ \tilde{q}_i^{t-1}, & if\ \psi_i^t = 0. \end{cases} \tag{12}$$

Now, we bring in the CUCB-based quality $\hat{q}_i^t$ to recruit workers. $\hat{q}_i^t$ consists of the empirical quality (indicating the learned knowledge from observed answer qualities), a confidence bound (indicating the uncertainty of empiricism), and the centrality (indicating the number of social relations), i.e.,

$$\hat{q}_i^t = \tilde{q}_i^t + \varphi_i^t + \chi_i,\ \varphi_i^t = \sqrt{(K+1)\ln(\textstyle\sum_{j=1}^N n_j^t)/n_i^t}. \tag{13}$$

Here, φ_i^t is the upper confidence bound, which takes a simple heuristic principle called optimism in the face of uncertainty. We can see that φ_i^t decreases rapidly with the increase of n_i^t. This implies that workers who are selected less may get more chances to be recruited in the next period. χ_i is the normalized degree centrality of worker i. A larger χ_i means that the worker i has more socially-connected friends and will have priority to be recruited.

Based on the above CUCB-based qualities, we introduce a greedy arm-pulling scheme to recruit unknown social-aware workers in each period. That is, we always select the top K workers by sorting the CUCB-based qualities $\{\hat{q}_i^t | i \in \mathcal{N}\}$ in a descending order. The selected workers will continue to participate in the following determination of the optimal strategy group.

3.2 Determining the Optimal Strategy Group

Now, we derive the optimal strategy group. The backward induction method is employed to analyze the THS game. First, we analyze Stage III to determine each worker i's optimal answer frequency δ_i^{t*} under a given vector $\boldsymbol{p}^t$. Next, we derive the platform's optimal unit recruitment salary vector $\boldsymbol{p}^{t*}$ under a given β^t in Stage II. Finally, we turn to Stage I to detect the requester's best unit payment β^{t*}. Moreover, we prove that $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$ forms a unique SE, so that no parties would have incentives to unilaterally deviate from their optimal decisions.

First, we begin to analyze the third stage for workers.

Theorem 1. *In Stage III, given any unit recruitment salary p_i^t, the optimal strategy of each social-aware worker i (i.e., the optimal answer frequency) can be determined by*

$$\delta_i^{t*} = \frac{p_i^t - \tilde{q}_i^t b_i}{2\tilde{q}_i^t a_i} + \frac{\sum_{j=1}^K \phi_{ij}\delta_j^t}{2\tilde{q}_i^t a_i}. \tag{14}$$

Proof. We compute the first-order and second-order derivatives of $U_i^t(\delta_i^t, p_i^t)$ with respect to δ_i^t as follows:

$$\frac{\partial U_i^t}{\partial \delta_i^t} = p_i^t + \sum_{j=1}^K \phi_{ij}\delta_j^t - (2\tilde{q}_i^t a_i \delta_i^t + \tilde{q}_i^t b_i). \tag{15}$$

$$\frac{\partial^2 U_i^t}{\partial (\delta_i^t)^2} = -2\tilde{q}_i^t a_i \, < \, 0. \tag{16}$$

Based on Eq. (15) and Eq. (16), the utility of each worker is strictly concave in the feasible region of δ_i^t. We can obtain the optimal strategy by solving $\partial U_i^t / \partial \delta_i^t = 0$ and the optimal answer frequency δ_i^{t*} is derived as Eq. (14).

Theorem 1 shows that the optimal answer frequency of each worker is composed of two terms. The left term $(p_i^t - \tilde{q}_i^t b_i)/2\tilde{q}_i^t a_i$ is independent from other workers' strategies, and the right term $\sum_{j=1}^K \phi_{ij}\delta_j^t / 2\tilde{q}_i^t a_i$ relies on other workers' strategies owing to the underlying social network effects. Here, we need to make the following assumption to ensure that the optimal answer frequency of each worker is bounded in the THS game.

Assumption 1. $\sum_{j=1}^K \phi_{ij}/2\tilde{q}_i^t a_i \, < \, 1, \forall i \in \mathcal{N}$.

Based on the above sufficient assumption, we can guarantee the existence and uniqueness of a Nash Equilibrium (NE) of the third stage in the THS game as follows.

Lemma 1. *Under Assumption 1, Stage III admits the existence and uniqueness of Nash equilibrium.*

Proof. We first prove the existence. Let δ_i^{t*} be the maximum answer frequency in $\boldsymbol{\delta}^{t*}$, i.e., $\delta_i^{t*} \ge \delta_j^{t*}, \forall \delta_j^{t*} \in \boldsymbol{\delta}^{t*}$.

$$\begin{aligned}\delta_i^{t*} &= \frac{p_i^t - \tilde{q}_i^t b_i}{2\tilde{q}_i^t a_i} + \frac{\sum_{j=1}^K \phi_{ij}\delta_j^{t*}}{2\tilde{q}_i^t a_i} \le \frac{p_i^t - \tilde{q}_i^t b_i}{2\tilde{q}_i^t a_i} + \frac{\sum_{j=1}^K \phi_{ij}\delta_i^{t*}}{2\tilde{q}_i^t a_i} \\ &\le \frac{|p_i^t - \tilde{q}_i^t b_i|}{2\tilde{q}_i^t a_i} + \frac{\sum_{j=1}^K |\phi_{ij}|\delta_i^{t*}}{2\tilde{q}_i^t a_i} = \frac{|p_i^t - \tilde{q}_i^t b_i|}{2\tilde{q}_i^t a_i} + \delta_i^{t*}\frac{\sum_{j=1}^K |\phi_{ij}|}{2\tilde{q}_i^t a_i}.\end{aligned} \tag{17}$$

According to Eq. (17) and Assumption 1, we further get $\delta_i^{t*} \le \frac{|p_i^t - \tilde{q}_i^t b_i|}{2\tilde{q}_i^t a_i - \sum_{j=1}^K |\phi_{ij}|}$. Because δ_i^{t*} is the largest value in $\boldsymbol{\delta}^{t*}$, there exists an upper value δ_{upper}^{t*} to guarantee that the strategy space of workers is $[1, \delta_{upper}^{t*}]$, where the recruited workers upload their answers at least once. In addition, $U_i^t(\delta_i^t, p_i^t)$ is continuous and concave, and the second-derivative of U_i^t is less than zero (i.e., Eq. (16)), so there exists a Nash equilibrium in the third stage.

Next, we prove the uniqueness of the NE in the third stage. According to Assumption 1 and Eq. (17), there is

$$-\frac{\partial^2 U_i^t}{\partial (\delta_i^t)^2} = 2\tilde{q}_i^t a_i > \sum\nolimits_{j=1}^K \phi_{ij} = \sum\nolimits_{j=1}^K |\phi_{ij}| = \sum\nolimits_{j=1}^K \left| -\frac{\partial^2 U_i^t}{\partial \delta_i^t \delta_j^t} \right|.$$

Based on the uniqueness theorem [30], the above equation satisfies the dominance-solvability condition. Consequently, the NE of the third stage is unique.

In light of Lemma 1, we can rewrite Eq. (14) of all recruited workers (i.e., $\partial U_i^t / \partial \delta_i^t = 0$) in a matrix form as follows: $\boldsymbol{p}^t + \mathbb{S}\boldsymbol{\delta}^{t*} - \mathbb{A}^t\boldsymbol{\delta}^{t*} - \boldsymbol{y}^t = 0$. Thus, we have

$$\boldsymbol{\delta}^{t*} = (\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t). \tag{18}$$

Here, $\mathbb{A}^t = diag(2\tilde{q}_1^t a_1, 2\tilde{q}_2^t a_2, \ldots, 2\tilde{q}_K^t a_K)$ is a diagonal matrix, $\mathbb{S} = [\phi_{ij}]_{K\times K}$ is a social network, and $\boldsymbol{y}^t = [\tilde{q}_i^t b_i]_{K\times 1} = (\tilde{q}_1^t b_1, \tilde{q}_2^t b_2, \ldots, \tilde{q}_K^t b_K)$. Note that, the vector (e.g., $\boldsymbol{p}^t$) is a column vector by default in this paper.

Then, we analyze the second stage for the platform.

Theorem 2. *In Stage II, given any unit payment β^t, the optimal strategy of the platform (i.e, the optimal unit recruitment salary vector $\boldsymbol{p}^t$) can be determined by*

$$\boldsymbol{p}^{t*} = (2c\mathbb{I}\mathbb{B}^t + 2\mathcal{I})^{-1}(\beta^t\mathbf{1} - d\mathbf{1} + 2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t + \boldsymbol{y}^t), \tag{19}$$

where $\mathbf{1} = [1]_{K\times 1}$ is a K-dimensional vector, $\mathbb{I} = [1]_{K\times K}$ is a matrix, $\mathcal{I}$ is an $K \times K$ identity matrix, and $\mathbb{B}^t = (\mathbb{A}^t - \mathbb{S})^{-1}$.

Proof. The utility function of the platform $U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t)$ can be rewritten in a matrix form as follows:

$$\begin{aligned}U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t) &= \beta^t \sum_{i=1}^K \delta_i^t - c\left(\sum_{i=1}^K \delta_i^t\right)^2 - d\sum_{i=1}^K \delta_i^t - \sum_{i=1}^K p_i^t\delta_i^t \\ &= \beta^t \mathbf{1}^\intercal\boldsymbol{\delta}^t - c(\mathbf{1}^\intercal\boldsymbol{\delta}^t)^\intercal(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - d(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - (\boldsymbol{p}^t)^\intercal\boldsymbol{\delta}^t \\ &= \beta^t \mathbf{1}^\intercal\boldsymbol{\delta}^t - c((\boldsymbol{\delta}^t)^\intercal\mathbf{1})(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - d(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - (\boldsymbol{p}^t)^\intercal\boldsymbol{\delta}^t \\ &= \beta^t \mathbf{1}^\intercal\boldsymbol{\delta}^t - c(\boldsymbol{\delta}^t)^\intercal\mathbb{I}\boldsymbol{\delta}^t - d(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - (\boldsymbol{p}^t)^\intercal\boldsymbol{\delta}^t.\end{aligned} \tag{20}$$

Based on Theorem 1 and Eq. (18), the optimal answer frequency vector $\boldsymbol{\delta}^{t*}$ can be substituted into Eq. (20). Therefore, $U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t)$ can be rewritten further, i.e.,

$$\begin{aligned}U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t) &= \beta^t \mathbf{1}^\intercal\boldsymbol{\delta}^t - c(\boldsymbol{\delta}^t)^\intercal\mathbb{I}\boldsymbol{\delta}^t - d(\mathbf{1}^\intercal\boldsymbol{\delta}^t) - (\boldsymbol{p}^t)^\intercal\boldsymbol{\delta}^t \\ &= \beta^t\mathbf{1}^\intercal(\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t) - d\mathbf{1}^\intercal(\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t) \\ &\quad - (\boldsymbol{p}^t)^\intercal(\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t) \\ &\quad - c((\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t))^\intercal\mathbb{I}(\mathbb{A}^t - \mathbb{S})^{-1}(\boldsymbol{p}^t - \boldsymbol{y}^t).\end{aligned}$$

To simplify the expression of U_p^t, we let $\mathbb{B}^t = (\mathbb{A}^t - \mathbb{S})^{-1}$. Since $\mathbb{A}^t - \mathbb{S}$ is a positive definite matrix and is invertible, there is $(\mathbb{B}^t)^\intercal = ((\mathbb{A}^t - \mathbb{S})^\intercal)^{-1} = (\mathbb{A}^t - \mathbb{S})^{-1} = \mathbb{B}^t$, and we simplify the above equation of $U_p^t(\boldsymbol{\delta}^t, \boldsymbol{p}^t, \beta^t)$ as follows:

$$
\begin{aligned}
&U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)=\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)-d\mathbf{1}^{\mathsf T}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)\\
&\quad-c(\boldsymbol{p}^t-\boldsymbol{y}^t)^{\mathsf T}(\mathbb{B}^t)^{\mathsf T}\mathbb{I}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)-(\boldsymbol{p}^t)^{\mathsf T}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)\\
&=\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{p}^t-\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t-d\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{p}^t+d\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t\\
&\quad-c(\boldsymbol{p}^t-\boldsymbol{y}^t)^{\mathsf T}\mathbb{B}^t\mathbb{I}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)-(\boldsymbol{p}^t)^{\mathsf T}\mathbb{B}^t\boldsymbol{p}^t+(\boldsymbol{p}^t)^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t\\
&=(\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t-d\mathbf{1}^{\mathsf T}\mathbb{B}^t)\boldsymbol{p}^t-c(\boldsymbol{p}^t-\boldsymbol{y}^t)^{\mathsf T}\mathbb{B}^t\mathbb{I}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)\\
&\quad-(\boldsymbol{p}^t)^{\mathsf T}\mathbb{B}^t\boldsymbol{p}^t+(\boldsymbol{p}^t)^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t+d\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t-\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t. \qquad (21)
\end{aligned}
$$

Based on Eq. (21), we can obtain the first-order derivative of $U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)$ with reference to $\boldsymbol{p}^t$ as follows:

$$
\begin{aligned}
\frac{\partial U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)}{\partial \boldsymbol{p}^t}&=(\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t-d\mathbf{1}^{\mathsf T}\mathbb{B}^t)^{\mathsf T}-(\mathbb{B}^t+(\mathbb{B}^t)^{\mathsf T})\boldsymbol{p}^t\\
&\quad-c[\mathbb{B}^t\mathbb{I}\mathbb{B}^t+(\mathbb{B}^t\mathbb{I}\mathbb{B}^t)^{\mathsf T}](\boldsymbol{p}^t-\boldsymbol{y}^t)+\mathbb{B}^t\boldsymbol{y}^t\\
&=\beta^t\mathbb{B}^t\mathbf{1}-d\mathbb{B}^t\mathbf{1}-2c\mathbb{B}^t\mathbb{I}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t)\\
&\quad-2\mathbb{B}^t\boldsymbol{p}^t+\mathbb{B}^t\boldsymbol{y}^t, \qquad (22)
\end{aligned}
$$

where $(\mathbb{B}^t)^{\mathsf T}=\mathbb{B}^t$ and $(\mathbb{B}^t\mathbb{I}\mathbb{B}^t)^{\mathsf T}=(\mathbb{B}^t)^{\mathsf T}\mathbb{I}(\mathbb{B}^t)^{\mathsf T}=\mathbb{B}^t\mathbb{I}\mathbb{B}^t$. Next, we can further obtain the second-order derivative of $U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)$ with regard to $\boldsymbol{p}^t$ based on Eq. (22) as follows:

$$
\begin{aligned}
\partial^2 U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)/\partial(\boldsymbol{p}^t)^2&=-2(c(\mathbb{B}^t\mathbb{I}\mathbb{B}^t)^{\mathsf T}+(\mathbb{B}^t)^{\mathsf T})\\
&=-2(c\mathbb{B}^t\mathbb{I}\mathbb{B}^t+\mathbb{B}^t)\;<\;0. \qquad (23)
\end{aligned}
$$

Since $\mathbb{B}^t$ is a positive definite matrix and $c>0$, there is $c\mathbb{B}^t\mathbb{I}\mathbb{B}^t+\mathbb{B}^t>0$. Therefore, we can get that the second-order derivative $\partial^2 U_p^t/\partial(\boldsymbol{p}^t)^2<0$ always satisfies. That is to say, $U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)$ is a strictly concave function in the feasible region of $\boldsymbol{p}^t$. Therefore, the maximum value (i.e, the optimal unit recruitment salary vector $\boldsymbol{p}^t$) can be computed by solving $\partial U_p^t(\boldsymbol{\delta}^t,\boldsymbol{p}^t,\beta^t)/\partial \boldsymbol{p}^t=0$ as follows:

$$
\begin{aligned}
\boldsymbol{p}^{t*}&=(2c\mathbb{B}^t\mathbb{I}\mathbb{B}^t+2\mathbb{B}^t)^{-1}(\beta^t\mathbb{B}^t\mathbf{1}-d\mathbb{B}^t\mathbf{1}+2c\mathbb{B}^t\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\mathbb{B}^t\boldsymbol{y}^t)\\
&=(2c\mathbb{I}\mathbb{B}^t+2\mathcal{I})^{-1}(\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t). \qquad (24)
\end{aligned}
$$

where $\mathcal{I}$ is an $K\times K$ identity matrix.

Finally, we investigate the first stage for the requester.

Theorem 3. *In Stage I, the optimal strategy of the requester (i.e., the optimal unit payment) can be determined by*

$$
\beta^{t*}=\frac{-(3\tilde{q}^t\nu^t+2)+\sqrt{8\xi(\tilde{q}^t)^2\kappa^t+(\tilde{q}^t\nu^t+2)^2}}{4\tilde{q}^t\kappa^t}, \qquad (25)
$$

where $\kappa^t=\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}$, $\nu^t=\mathbf{1}^{\mathsf T}[\mathbb{C}^t(-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbb{B}^t\boldsymbol{y}^t]$, *and* $\mathbb{C}^t=\mathbb{B}^t(2c\mathbb{I}\mathbb{B}^t+2\mathcal{I})^{-1}$. *Moreover,* $\mathbf{1}=[1]_{K\times 1}$ *is a K-dimensional vector,* $\mathbb{I}=[1]_{K\times K}$ *is a matrix,* $\mathcal{I}$ *is an* $K\times K$ *identity matrix, and* $\mathbb{B}^t=(\mathbb{A}^t-\mathbb{S})^{-1}$.

Proof. According to Theorem 1 and Eq. (18), the optimal answer frequency vector $\boldsymbol{\delta}^{t*}$ can be substituted into Eq. (7). Therefore, $U_r^t(\boldsymbol{\delta}^t,\beta^t)$ can be rewritten as follows:

$$
U_r^t(\boldsymbol{\delta}^t,\beta^t)=\xi\ln(1+\tilde{q}^t\mathbf{1}^{\mathsf T}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t))-\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t(\boldsymbol{p}^t-\boldsymbol{y}^t). \qquad (26)
$$

Then, we substitute Eq. (19) into Eq. (26) and get

$$
\begin{aligned}
U_r^t(\boldsymbol{\delta}^t,\beta^t)&=\xi\ln(1+\tilde{q}^t\mathbf{1}^{\mathsf T}\mathbb{B}^t((2c\mathbb{I}\mathbb{B}^t+2\mathcal{I})^{-1}(\beta^t\mathbf{1}-d\mathbf{1}\\
&\qquad+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\boldsymbol{y}^t))-\beta^t\mathbf{1}^{\mathsf T}\mathbb{B}^t((2c\mathbb{I}\mathbb{B}^t+2\mathcal{I})^{-1}\\
&\qquad\times(\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\boldsymbol{y}^t)\\
&=\xi\ln(1+\tilde{q}^t\mathbf{1}^{\mathsf T}(\mathbb{C}^t\boldsymbol{\varpi}^t-\mathbb{B}^t\boldsymbol{y}^t))-\beta^t\mathbf{1}^{\mathsf T}(\mathbb{C}^t\boldsymbol{\varpi}^t-\mathbb{B}^t\boldsymbol{y}^t). \qquad (27)
\end{aligned}
$$

Here, we define $\mathbb{C}^t\triangleq\mathbb{B}^t(2c\mathbb{I}\mathbb{B}^t+2\mathcal{I})^{-1}$ and $\boldsymbol{\varpi}^t\triangleq\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t$ for notational simplicity. Note that, $\boldsymbol{\varpi}^t$ is a function of β^t essentially. Furthermore, we define $\hat{\boldsymbol{\varpi}}^t\triangleq\mathbb{C}^t\boldsymbol{\varpi}^t-\mathbb{B}^t\boldsymbol{y}^t$, which is a function of $\boldsymbol{\varpi}^t$. Thus, we can rewritten Eq. (27) as $U_r^t(\boldsymbol{\delta}^t,\beta^t)=\xi\ln(1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t)-\beta^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t$. Now, we can compute the first-order derivative of $U_r^t(\boldsymbol{\delta}^t,\beta^t)$ with respect to β^t as follows:

$$
\begin{aligned}
\frac{\partial U_r^t(\boldsymbol{\delta}^t,\beta^t)}{\partial\beta^t}&=\frac{\xi\tilde{q}^t\mathbf{1}^{\mathsf T}}{1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t}\frac{\partial\hat{\boldsymbol{\varpi}}^t}{\partial\boldsymbol{\varpi}^t}\frac{\partial\boldsymbol{\varpi}^t}{\partial\beta^t}\\
&\quad-\beta^t\mathbf{1}^{\mathsf T}\frac{\partial\hat{\boldsymbol{\varpi}}^t}{\partial\boldsymbol{\varpi}^t}\frac{\partial\boldsymbol{\varpi}^t}{\partial\beta^t}-\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t=\frac{\xi\tilde{q}^t\mathbf{1}^{\mathsf T}}{1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t}\cdot\mathbb{C}^t\cdot\mathbf{1}\\
&\quad-\beta^t\mathbf{1}^{\mathsf T}\cdot\mathbb{C}^t\cdot\mathbf{1}-\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t\\
&=\frac{\xi\tilde{q}^t\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}}{1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t}-\beta^t\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}\\
&\quad-\mathbf{1}^{\mathsf T}(\mathbb{C}^t(\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbb{B}^t\boldsymbol{y}^t), \qquad (28)
\end{aligned}
$$

where $\hat{\boldsymbol{\varpi}}^t=\mathbb{C}^t(\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbb{B}^t\boldsymbol{y}^t$. Based on Eq. (28), we continue to obtain the second-order derivative of $U_r^t(\delta_i^t,\beta^t)$ with reference to β^t as follows:

$$
\begin{aligned}
\frac{\partial^2 U_r^t(\boldsymbol{\delta}^t,\beta^t)}{\partial(\beta^t)^2}&=\frac{-\xi\tilde{q}^t\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}}{(1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t)^2}(\tilde{q}^t\mathbf{1}^{\mathsf T}\frac{\partial\hat{\boldsymbol{\varpi}}^t}{\partial\boldsymbol{\varpi}^t}\frac{\partial\boldsymbol{\varpi}^t}{\partial\beta^t})-\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}-\mathbf{1}^{\mathsf T}\frac{\partial\hat{\boldsymbol{\varpi}}^t}{\partial\boldsymbol{\varpi}^t}\\
&\quad\times\frac{\partial\boldsymbol{\varpi}^t}{\partial\beta^t}\\
&=\frac{-\xi(\tilde{q}^t)^2\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}}{(1+\tilde{q}^t\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t)^2}-\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}-\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}\;<\;0.
\end{aligned}
$$

Because $\mathbb{B}^t$ is a positive definite matrix, we can easily get that $\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}\;>\;0$ always holds after some transformations. Therefore, $\frac{\partial^2 U_r^t(\boldsymbol{\delta}^t,\beta^t)}{\partial(\beta^t)^2}\;<\;0$ can be acquired, and the maximum value (i.e., the optimal unit payment β^{t*}) can be obtained by solving $\frac{\partial U_r^t(\boldsymbol{\delta}^t,\beta^t)}{\partial\beta^t}=0$. For ease of presentation, we define $\kappa^t\triangleq\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}$ and $\nu^t\triangleq\mathbf{1}^{\mathsf T}[\mathbb{C}^t(-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbb{B}^t\boldsymbol{y}^t]$. Based on this, there is

$$
\begin{aligned}
\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t&=\mathbf{1}^{\mathsf T}[\mathbb{C}^t(\beta^t\mathbf{1}-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbb{B}^t\boldsymbol{y}^t]\\
&=\beta^t\mathbf{1}^{\mathsf T}\mathbb{C}^t\mathbf{1}+\mathbf{1}^{\mathsf T}\mathbb{C}^t(-d\mathbf{1}+2c\mathbb{I}\mathbb{B}^t\boldsymbol{y}^t+\boldsymbol{y}^t)-\mathbf{1}^{\mathsf T}\mathbb{B}^t\boldsymbol{y}^t=\beta^t\kappa^t+\nu^t.
\end{aligned}
$$

By substituting $\mathbf{1}^{\mathsf T}\hat{\boldsymbol{\varpi}}^t$ into Eq. (28), we can obtain

$$
\frac{\partial U_r^t(\boldsymbol{\delta}^t,\beta^t)}{\partial\beta^t}=0\Leftrightarrow\frac{\xi\tilde{q}^t\kappa^t}{1+\tilde{q}^t(\beta^t\kappa^t+\nu^t)}-\beta^t\kappa^t-(\beta^t\kappa^t+\nu^t)=0
$$

$$
\begin{aligned}
&\Leftrightarrow\frac{2\tilde{q}^t(\kappa^t)^2(\beta^t)^2+3\tilde{q}^t\kappa^t\nu^t\beta^t+2\kappa^t\beta^t+\tilde{q}^t(\nu^t)^2+\nu-\xi\tilde{q}^t\kappa^t}{1+\tilde{q}^t(\beta^t\kappa^t+\nu^t)}=0\\
&\Leftrightarrow 2\tilde{q}^t(\kappa^t)^2(\beta^t)^2+(3\tilde{q}^t\kappa^t\nu^t+2\kappa^t)\beta^t+\tilde{q}^t(\nu^t)^2+\nu^t-\xi\tilde{q}^t\kappa^t=0. \qquad (29)
\end{aligned}
$$

Obviously, Eq. (29) is a quadratic equation of one unknown parameter β^t. Therefore, we can derive the optimal unit payment β^{t*} by solving the roots of Eq. (29). We first compute the discriminant of the quadratic equation (denoted by Dis) as follows:

$$\begin{aligned} Dis &= (3\tilde{q}^t\kappa^t\nu^t+2\kappa^t)^2 - 8\tilde{q}^t(\kappa^t)^2(\tilde{q}^t(\nu^t)^2+\nu^t-\xi\tilde{q}^t\kappa^t) \\ &= (\kappa^t)^2[8\xi(\tilde{q}^t)^2\kappa^t + (\tilde{q}^t\nu^t+2)^2] > 0. \end{aligned} \tag{30}$$

Here, owing to $\kappa^t > 0$ and $\xi > 0$, we have $Dis > 0$. According to the property of quadratic equation, Eq. (29) has two roots by using the standard quadratic-root formula: $\frac{-(3\tilde{q}^t\kappa^t\nu^t+2\kappa^t)}{4\tilde{q}^t(\kappa^t)^2} \pm \frac{\sqrt{(\kappa^t)^2[8\xi(\tilde{q}^t)^2\kappa^t+(\tilde{q}^t\nu^t+2)^2]}}{4\tilde{q}^t(\kappa^t)^2}$. Since we only need the positive value, β^{t*} is derived as Eq. (25).

Algorithm 1. The Incentive Mechanism TACT

Input: $T, N, K, \Delta, \Delta_{min}, \boldsymbol{a}, \boldsymbol{b}, c, d, \xi, \Phi$
Output: $\Psi, \{\beta^{t*}|t\in[1,T]\}, \{\boldsymbol{p}^{t*}|t\in[1,T]\}, \{\boldsymbol{\delta}^{t*}|t\in[1,T]\}$

1: **Initialization**:
2: $\psi_i^t = 0, \beta^t = 0, p_i^t = 0, \delta_i^t = 0, \forall i \in \mathcal{N}, t \in [1,T]$;
3: **Initial exploration phase**:
4: Recruit all social-aware workers when $t = 1$, i.e.,
5: **for** each worker $i \in \mathcal{N}, t = 1$ **do** $\psi_i^1 = 1$, set $q_{i,k}^1 = 1$;
6: Determine the salary for recruited workers: $p_i^{1*} = p_{max}$;
7: Requester: pay $\beta^{1*} = argmin_{\beta^1}(U_p^t \geq 0)$ to the platform;
8: Workers: perform the task; upload and share answers;
9: **for** each worker $i \in \mathcal{N}$ **do** Update $n_i^1, \tilde{q}_i^1, \varphi_i^1, \hat{q}_i^1$;
10: **Learning of quality and the THS game**:
11: **while** each period $t < T$ **do**
12: Sort the workers according to the CUCB-based quality value $\hat{q}_i^t$: $\hat{q}_{w_1}^t \geq \hat{q}_{w_2}^t \geq \ldots \geq \hat{q}_{w_N}^t$; $t \Leftarrow t+1$;
13: Select the top K workers as winners, denoted as $\mathcal{W}^t$;
14: **for** each worker $w_i \in \mathcal{W}^t$ **do** $\psi_{w_i}^t = 1$;
15: Carry out the THS game: determine the optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$ for each party according to Eqs. (18), (19), and (25);
16: Workers: perform the task; upload and share answers;
17: **for** each worker $i \in \mathcal{N}$ **do** Update $n_i^t, \tilde{q}_i^t, \varphi_i^t, \hat{q}_i^t$;
18: **end while**
19: **Return** $\mathcal{W}^t, \Psi, \langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle, \forall t \in [1,T]$.

In brief, we can derive the optimal strategy of the requester (i.e., the optimal unit payment β^{t*}) based on Theorem 3, and then the platform's optimal strategy (i.e., the optimal unit salary vector $\boldsymbol{p}^{t*}$) can be determined by substituting β^{t*} into Eq. (19). Finally, workers' optimal strategies (i.e., the optimal answer frequency vector $\boldsymbol{\delta}^{t*}$) are derived by substituting $\boldsymbol{p}^{t*}$ into Eq. (18). In this way, the optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$ is determined to achieve utility maximization for the SC system.

Remark: Here, the derived optimal strategy of each participant is closely associated with the learned empirical qualities and social relations of recruited workers. Different recruitment results will lead to diverse optimal strategy groups, and great recruitment results will upgrade all participants' utilities. In short, the unknown social-aware worker recruitment is an essential prerequisite for the optimal strategy group determination. On the other hand, like in the recent related studies [27], [28], [31], the Stackelberg game used in this paper is a type of incentive mechanism with complete information, namely, the parameters (e.g., a_i, b_i, c, and d) in the utility functions are public information essentially. In this paper, we assume that the platform is trustworthy, and the requester can obtain these public information via the platform.

3.3 Algorithm Design

We illustrate the proposed incentive mechanism pseudocode in Algorithm 1, which mainly consists of two components: the unknown social-aware worker recruitment and the optimal strategy group determination. At the beginning, we initialize some values as zeros (Lines 1-2). In order to explore all workers' qualities, we execute the initial exploration phase in Lines 3-8. More specifically, without loss of generality, we set the initial qualities of all workers as 1 (i.e., $\psi_i^1 = 1$ and $q_{i,k}^1 = 1, \forall i \in \mathcal{N}$), and then the platform recruits all workers in the first period (Lines 4-5). Next, each worker selects a strategy randomly and the platform pays the maximum salary to workers (Line 6). In Line 7, the requester calculates a minimal unit payment which can ensure that the utility of the platform is non-negative. At the end of the first period, we can evaluate all workers' real qualities to update the related values (i.e., $n_i^1, \tilde{q}_i^1, \varphi_i^1, \hat{q}_i^1$) for the following periods (Line 8).

In Lines 9-17, we start the iteration of exploitation and exploration with the THS game. Through the exploration phase, we can harness the learned quality knowledge to recruit suitable workers and determine the optimal strategy group (i.e., exploitation). More specifically, we first sort the CUCB-based qualities in a descending order and adopt a greedy arm-pulling scheme (Lines 11-14). That is, we greedily select the top K workers as winners to carry out the task. Then, all parties play the THS game to determine their respective optimal strategies according to Eqs. (18), (19), and (25), respectively. In Lines 15-16, TACT continues to learn the quality knowledge to prepare for the next period (i.e., exploration). We conduct the exploitation and exploration alternatively to recruit the more suitable workers.

The above process runs period by period until $t = T$. Finally, the incentive mechanism returns all optimal strategy groups and the sets of winning workers: $\langle \{\beta^{t*}|t \in [1,T]\}, \{\boldsymbol{p}^{t*}|t \in [1,T]\}, \{\boldsymbol{\delta}^{t*}|t \in [1,T]\} \rangle$ and $\{\mathcal{W}^t|t \in [1,T]\}$. In the initial exploration phase, the computational complexity is at most $O(N)$. In the exploitation and exploration phase, there are three main operations of sorting, game, and updating, whose total computational complexity is $O(N\log(N))$. Therefore, the computation complexity of Algorithm 1 is $O(TKN^2\log(N))$.

3.4 A Straightforward Example

For a better understanding, we provide a straightforward example to illustrate the worker recruitment and optimal strategy determination process of the mechanism TACT. We consider the scenario where there are three social-aware workers $\mathcal{N} = \{1,2,3\}$, a requester with the traffic flow detection task, and an SC platform. We assume that the real data quality of each answer in all periods follows the Gaussian distribution, and set $K = 2, \xi = 40, c = 0.1, d = 1$. The information of workers and the actual qualities of each recruited worker in different periods are presented in Fig. 3.

At the beginning, we do not have any information about workers' qualities. Thus, we start the initial exploration phase by recruiting all workers to learn the quality information. According to the Algorithm 1, the platform will give the highest unit salary to all workers, i.e., $p_i^{1*} = 3$, and the

Workers	$\mathbb{E}(q_{i,k}^t)$	Parameters	Network Effects
worker-1	$q_1 = 0.8$	$a_1 = 0.5, b_1 = 0.1$	$\phi_{12} = 0.1, \phi_{13} = 0.2$
worker-2	$q_2 = 0.7$	$a_2 = 0.3, b_2 = 0.1$	$\phi_{21} = 0.1, \phi_{23} = 0.1$
worker-3	$q_3 = 0.6$	$a_3 = 0.2, b_3 = 0.1$	$\phi_{31} = 0.2, \phi_{32} = 0.1$

Other parameters
$K = 2$, $\xi = 40$, $c = 0.1$, $d = 1$

Quality	Period 1			Period 2		Period 3		Period 4		Period 5		Period 6		Period 7		Period 8		Period 9		Period 10	
δ_i^t \ Worker	1	2	3	1	3	2	1	2	1	3	2	3	1	2	3	2	1	1	3	2	1
1	0.58	0.56	0.27	0.69	0.93	0.55	0.64	0.72	0.21	0.74	0.85	0.84	0.38	0.57	0.78	0.54	0.75	0.35	0.33	0.68	0.71
2	0.77	0.44	0.82	0.74	0.25	0.46	0.39	0.67	0.25	0.78	0.35	0.51	0.41	0.89	0.78	0.89	0.68	0.78	0.34	0.49	0.35
3					0.92	0.61		0.75		0.9		0.38			0.44	0.34			0.51	0.67	
4					0.25							0.77							0.72		

Fig. 3. Workers information and real qualities in different periods.

requester needs to pay $\beta^{1*} = 4.6$ to ensure the non-negative utility of the platform. In the end of period 1, the platform can update the empirical quality of each worker: $\tilde{q}_1^1 = (0.58 + 0.77)/2 = 0.675$, $\tilde{q}_2^1 = 0.5$, $\tilde{q}_3^1 = 0.545$. Correspondingly, the CUCB-based quality can be updated as: $\hat{q}_1^1 = 2.981$, $\hat{q}_2^1 = 2.806$, $\hat{q}_3^1 = 2.851$. Due to $\hat{q}_1^1 > \hat{q}_3^1 > \hat{q}_2^1$, the worker-1 and worker-3 are recruited in period 2, i.e., $\mathcal{W}^2 = \{1, 3\}$. After the THS game, the optimal strategy group can be determined as: $\langle \beta^{2*} = 3.2, \boldsymbol{p}^{1*} = \{0.55, 0.54\}, \boldsymbol{\delta}^{t*} = \{1.89, 3.96\} \rangle$. Based on the specific application scene, the answer frequency needs to be rounded to integer, i.e., $\boldsymbol{\delta}^{t*} = \{2, 4\}$. Then, each worker's empirical quality and CUCB-based quality will be updated: $\tilde{q}_1^2 = (2 * 0.675 + 0.69 + 0.74)/4 = 0.695$, $\tilde{q}_2^2 = 0.5$, $\tilde{q}_3^2 = 0.573$ and $\hat{q}_1^2 = 2.727$, $\hat{q}_2^2 = 3.13$, $\hat{q}_3^2 = 2.355$. Owing to $\hat{q}_2^2 > \hat{q}_1^2 > \hat{q}_3^2$, worker-2 and worker-1 will be winners in period 3. Repeating the above steps until $t = T$, the task would be over completely. Here, we only present the first ten periods and the recruitment order is $\{1, 2, 3\}, \{1, 3\}, \{2, 1\}, \{2, 1\}, \{3, 2\}, \{3, 1\}, \{2, 3\}, \{2, 1\}, \{1, 3\}, \{2, 1\}$. The whole process and values of TACT are illustrated in Fig. 4.

4 Theoretical Analysis

In this section, we analyze the regret bound and the existence of the unique Stackelberg equilibrium in TACT.

4.1 Regret Analysis

The regret of online CMAB is the difference between the total quality of TACT and the optimal solution. The optimal total quality can be achieved when we assume that the platform knows all workers' qualities in advance, i.e., $q_i, \forall i \in \mathcal{N}$. Let $\mathcal{W}^t$ and $\mathcal{W}^*$ be the recruited worker set based on Algorithm 1 and the optimal recruited worker set based on the optimal policy, respectively. Thus, we introduce the concept of regret [32] as follows:

$$\mathcal{R} = Q(\mathcal{W}^*) - \mathbb{E}[Q(\mathcal{W}^t)]. \tag{31}$$

Note that, $\mathcal{W}^t$ can be mapped to Ψ^t directly. Let $*$ denote the identifications of the optimal workers under the optimal policy. Clearly, if the platform knows the order: $q_{w_1^*} + \chi_{w_1^*} \geq \dots q_{w_K^*} + \chi_{w_K^*} \dots \geq q_{w_N^*} + \chi_{w_N^*}$, it will always recruit the top K workers as winners in all periods, i.e., $\mathcal{W}^* = \{w_1^*, \dots, w_K^*\}$. For the following derivation, we first define the smallest and largest possible difference of quality values among all non-optimal workers, i.e., $\mathcal{W}^t \neq \mathcal{W}^*$.

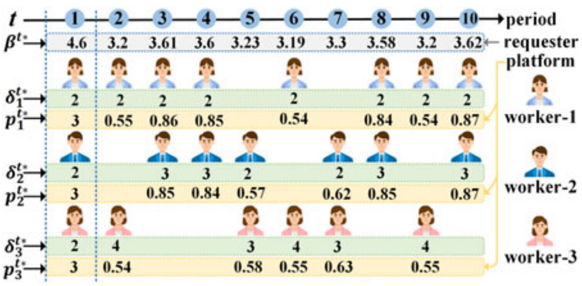


Fig. 4. The whole process of TACT.

$$\nabla_{min} = \sum_{i \in \mathcal{W}^*} (q_i + \chi_i) - max_{\mathcal{W}^t \neq \mathcal{W}^*} \sum_{i \in \mathcal{W}^t} (q_i + \chi_i); \tag{32}$$

$$\nabla_{max}^q = \sum_{i \in \mathcal{W}^*} q_i - min_{\mathcal{W}^t \neq \mathcal{W}^*} \sum_{i \in \mathcal{W}^t} q_i. \tag{33}$$

Next, we introduce γ_i^t as the counter of worker i after the initial exploration period. In each period t $(t \geq 2)$, γ_i^t is updated according to the following rules:

$$\begin{cases} \gamma_i^t \ \ keeps \ \ unchanged, & case\,1 \\ i = argmin_{i' \in \mathcal{W}^t} \gamma_{i'}^{t-1}, \gamma_i^t = \gamma_i^{t-1} + 1, & case\,2. \end{cases} \tag{34}$$

Specifically, each period must occur in one of the two cases: 1) the optimal worker set is selected, i.e., $\mathcal{W}^* = \mathcal{W}^t$; 2) a non-optimal worker set is determined, i.e., $\mathcal{W}^* \neq \mathcal{W}^t$. The γ_i^t will not change in case 1. In case 2, there must exist one worker i' with the minimum counter $\gamma_{i'}^{t-1}$, and we let $\gamma_i^t = \gamma_i^{t-1} + 1$. If multiple workers have the same minimum counter, we arbitrarily choose one. In this way, the sum of the counter γ_i^t is equal to the total number of the non-optimal worker sets, since there always exists a counter to be increased by 1 when $\mathcal{W}^* \neq \mathcal{W}^t$. Then, we use λ_i^t to denote the exact times that worker i has been recruited until the period t, where $\gamma_i^t \leq \lambda_i^t$. Now, we analyze the upper bound of the expected counter $\mathbb{E}[\gamma_i^T]$.

Lemma 2. *At the end of the period T, the expected counter γ_i^T has an upper bound for any worker $i \in \mathcal{N}$, i.e.,*

$$\mathbb{E}[\gamma_i^T] \leq \frac{4(K+1)K^2 \ln(TN\Delta/\Delta_{min})}{\nabla_{min}} + 1 + \frac{K\pi^2}{3\ N^{2(K+1)}}. \tag{35}$$

Proof. According to the update rule of the counter γ_i^T in Eq. (34), we can get the following results:

$$\begin{aligned} \gamma_i^T &= \sum_{t=2}^{T} \Pi\{case\,2\} \leq \vartheta + \sum_{t=2}^{T} \Pi\{case\,2, \gamma_i^t \geq \vartheta\} \\ &\leq \vartheta + \sum_{t=2}^{T} \Pi\left\{ \sum_{i \in \mathcal{W}^t} \hat{q}_i^{t-1} \geq \sum_{i \in \mathcal{W}^*} \hat{q}_i^{t-1}, \gamma_i^t \geq \vartheta \right\} \\ &\leq \vartheta + \sum_{t=2}^{T} \Pi\Bigg\{ \max_{\vartheta \leq \lambda_{w_1}^t \leq \dots \leq \lambda_{w_K}^t \leq t-1} \sum_{j=1}^{K} \hat{q}_{w_j}^{t-1} \\ &\geq \min_{1 \leq \lambda_{w_1}^* \leq \dots \leq \lambda_{w_K}^* \leq t-1} \sum_{j=1}^{K} \hat{q}_{w_j^*}^{t-1} \Bigg\} \\ &\leq \vartheta + \sum_{t=1}^{T} \sum_{\lambda_{w_1}^t = \vartheta}^{t-1} \cdots \sum_{\lambda_{w_K}^t = \vartheta}^{t-1} \sum_{\lambda_{w_1}^* = 1}^{t-1} \cdots \sum_{\lambda_{w_K}^* = 1}^{t-1} \Pi\left\{ \sum_{j=1}^{K} \hat{q}_{w_j}^t \geq \sum_{j=1}^{K} \hat{q}_{w_j^*}^t \right\}. \end{aligned} \tag{36}$$

Here, Π is an indicator function, i.e., $\Pi\{true\} = 1$ and $\Pi\{false\} = 0$. Next, it is necessary to derive the bound of $\Pi\{\cdot\}$ in Eq. (36). That is to say, we need to prove the probability that $\sum_{j=1}^{K} \hat{q}_{w_j}^t \geq \sum_{j=1}^{K} \hat{q}_{w_j^*}^t$ (i.e., $\sum_{j=1}^{K} (\tilde{q}_{w_j}^t + \varphi_{w_j}^t + \chi_{w_j}) \geq \sum_{j=1}^{K} (\tilde{q}_{w_j^*}^t + \varphi_{w_j^*}^t + \chi_{w_j^*})$) holds.

TABLE 2
Evaluation Settings

Parameter name	Values
number of workers, N	[50, 110] (50 in default)
number of periods, T	[1000, 1500] (1000 in default)
number of recruited workers, K	5, **10**, 15, 20, 25, 30
maximal answer frequency, Δ_{min}	6, 8, 12, 15, 18, **20**
system parameter, ξ	[2600, 7000] (3000 in default)
standard deviation of Φ, σ_Φ	[0.02, 0.5] (0.02 in default)

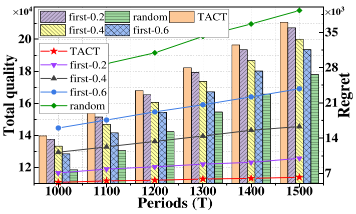


Fig. 5. Total quality and regret vs. Number of periods T ($K = 10, N = 50$).

Based on the proof by contradiction, we can know that at least one of the following equations must hold:

$$\sum_{j=1}^{K}(q_{w_j^*} - \varphi^t_{w_j^*} + \chi_{w_j^*}) \geq \sum_{j=1}^{K}(\tilde{q}^t_{w_j^*} + \chi_{w_j^*}); \tag{37}$$

$$\sum_{j=1}^{K}(q_{w_j} + \varphi^t_{w_j} + \chi_{w_j}) \leq \sum_{j=1}^{K}(\tilde{q}^t_{w_j} + \chi_{w_j}); \tag{38}$$

$$\sum_{j=1}^{K}(q_{w_j} + 2\varphi^t_{w_j} + \chi_{w_j}) > \sum_{j=1}^{K}(q_{w_j^*} + \chi_{w_j^*}). \tag{39}$$

Now, we need to derive the upper bound of the probability when Eqs. (37) and (38) are established. The inequality Eq. (37) implies an underestimate of the optimal recruited worker set, and the upper bound of the probability of Eq. (37) can be deduced as follows:

$$\begin{aligned}
&\mathbf{Pr}\Big\{\sum_{j=1}^{K}(q_{w_j^*} - \varphi^t_{w_j^*} + \chi_{w_j^*}) \geq \sum_{j=1}^{K}(\tilde{q}^t_{w_j^*} + \chi_{w_j^*})\Big\}\\
&\leq \sum_{j=1}^{K}\mathbf{Pr}\{q_{w_j^*} - \varphi^t_{w_j^*} \geq \tilde{q}^t_{w_j^*}\} \leq \sum_{j=1}^{K} e^{-2n^t_{w_j^*}\cdot\varphi^t_{w_j^*}\cdot\varphi^t_{w_j^*}}\\
&\leq \sum_{j=1}^{K} e^{-2n^t_{w_j^*}\left[(K+1)\ln\left(\sum_{j'=1}^{N} n^t_{j'}\right)/n^t_{w_j^*}\right]}\\
&\leq \sum_{j=1}^{K} e^{-2(K+1)\ln(t\cdot N\cdot 1)} \leq K\cdot(tN)^{-2(K+1)}.
\end{aligned} \tag{40}$$

The deduction of Eq. (40) makes use of the Chernoff-Hoeffding bound, which has been widely used in previous works [25], [32]. Each recruited worker must upload answers at least once. The inequality Eq. (38) signifies a drastic overestimate of the sub-optimal recruited worker set. Similarly, we can also get

$$\begin{aligned}
&\mathbf{Pr}\Big\{\sum_{j=1}^{K}(q_{w_j} + \varphi^t_{w_j} + \chi_{w_j}) \leq \sum_{j=1}^{K}(\tilde{q}^t_{w_j} + \chi_{w_j})\Big\}\\
&\leq \sum_{j=1}^{K}\mathbf{Pr}\{q_{w_j} + \varphi^t_{w_j} \leq \tilde{q}^t_{w_j}\} \leq \sum_{j=1}^{K} e^{-2n^t_{w_j}\cdot\varphi^t_{w_j}\cdot\varphi^t_{w_j}}\\
&\leq \sum_{j=1}^{K} e^{-2n^t_{w_j}\left[(K+1)\ln\left(\sum_{j'=1}^{N} n^t_{j'}\right)/n^t_{w_j}\right]} \leq K\cdot(tN)^{-2(K+1)}.
\end{aligned}$$

According to the derived upper bounds (i.e., $K\cdot(tN)^{-2(K+1)}$), we observe that two probabilities will become smaller along with the increase of the period t. Therefore, as long as the number of periods T is large enough, the probability of the underestimate/overestimate is low, and the set $\mathcal{W}^T$ will be close to the optimal set $\mathcal{W}^*$. This indicates that the proposed CMAB-based arm-pulling scheme is convergent.

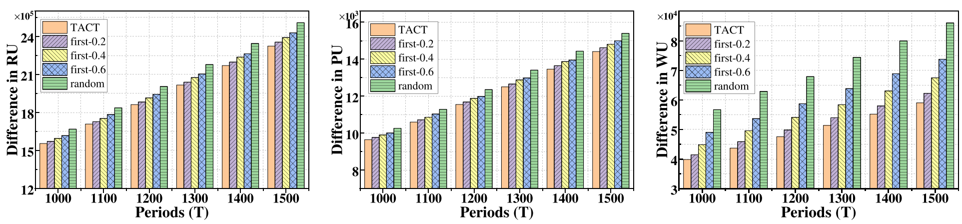
Due to the contradiction, Eq. (39) will be false if Eq. (37) and Eq. (38) are true. Therefore, we need to find a suitable ϑ to make the Eq. (39) impossible.

$$\begin{aligned}
&\sum_{j=1}^{K}(q_{w_j^*} + \chi_{w_j^*}) - \sum_{j=1}^{K}(q_{w_j} + 2\varphi^t_{w_j} + \chi_{w_j})\\
&= \sum_{j=1}^{K}(q_{w_j^*} + \chi_{w_j^*}) - \sum_{j=1}^{K}(q_{w_j} + \chi_{w_j}) - 2\sum_{j=1}^{K}\sqrt{(K+1)\ln(\sum_{j=1}^{N} n^t_j)/n^t_i}\\
&\geq \nabla_{min} - 2\sum_{j=1}^{K}\sqrt{(K+1)\ln(TN\Delta/\Delta_{min})/\vartheta} \geq 0.
\end{aligned} \tag{41}$$

Note that, there exists $n^t_i \geq \lambda^t_i \geq \gamma^t_i \geq \vartheta$. In order to make Eq. (39) invalid, we determine the range of ϑ by solving Eq. (41). Then, we can obtain

$$\vartheta \geq 4(K+1)K^2\ln(TN\Delta/\Delta_{min})/\nabla^2_{min}. \tag{42}$$

Up to now, we can continue to derive the upper bound of expected counter γ^T_i based on Eq. (36).



(a) Requester (b) Platform (c) Workers

Fig. 6. Difference in utility of each party vs. Number of periods T.

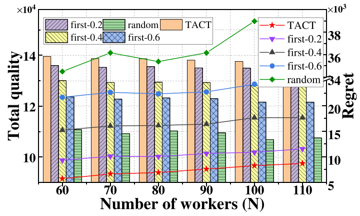


Fig. 7. Total quality and regret vs. Number of workers N ($K = 10, T = 1000$).

$$\mathbb{E}[\gamma_i^T] \leq \lceil \frac{4(K+1)K^2 \ln(TN\Delta/\Delta_{min})}{\nabla_{min}^2} \rceil + \sum_{t=1}^{+\infty} (t-\vartheta)^K (t-1)^K 2K(tN)^{-2(K+1)}$$
$$\leq \frac{4(K+1)K^2 \ln(TN\Delta/\Delta_{min})}{\nabla_{min}^2} + 1 + 2KN^{-2(K+1)} \sum_{t=1}^{+\infty} t^{-2}$$
$$\leq \frac{4(K+1)K^2 \ln(TN\Delta/\Delta_{min})}{\nabla_{min}^2} + 1 + \frac{K\pi^2 \ N^{-2(K+1)}}{3}. \quad (43)$$

The proof of the lemma is now completed.

Based on lemma 2, we further derive the regret bound.

Theorem 4. *At the end of the period T, the upper bound on the regret $\mathcal{R}$ of Algorithm 1 can be tightened to $O(N\Delta_{num}K^3 \ln(TN\Delta_{num}))$, where $\Delta_{num} = \Delta/\Delta_{min}$.*

Proof. According to the lemma 2 and the definition of regret in Eq. (31), we can get that the regret satisfies

$$\mathcal{R} = Q(\mathcal{W}^*) - \mathbb{E}[Q(\mathcal{W}^t)] \leq \left(\sum_{i=1}^{N} \gamma_i^T\right) \nabla_{max}^q (\Delta/\Delta_{min})$$
$$\leq \nabla_{max}^q \frac{\Delta}{\Delta_{min}} N \left(\frac{4(K+1)K^2 \ln(TN\Delta/\Delta_{min})}{\nabla_{min}^2} + 1 + \frac{K\pi^2}{3 \ N^{2(K+1)}} \right)$$
$$= O(N\Delta_{num}K^3 \ln(TN\Delta_{num})), \quad (44)$$

where $\Delta_{num} = \Delta/\Delta_{min}$. The proof is now completed.

4.2 Stackelberg Equilibrium Analysis

Finally, we analyze the existence and uniqueness of Stackelberg equilibrium in the whole THS game.

Theorem 5. *The unique Stackelberg equilibrium exists in the THS game based on the optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$.*

Proof. In the whole THS game, each stage can derive its optimal closed-form solution in any period t: the unit payment strategy β^{t*} of the requester, the unit salary strategy $\boldsymbol{p}^{t*}$ of the platform, and the answer frequency strategies $\boldsymbol{\delta}^{t*}$ of workers. As the role of the first-tie leader in the first stage, the requester can uniquely determine β^{t*} according to Eq. (25) in Theorem 3. It is worth noting that the value of β^{t*} is computed just by some public parameters such as $\boldsymbol{a}$, $\boldsymbol{b}$, c, and d. Recall that, the Stackelberg game is the complete information game. That is, β^{t*} is only related to the constant inputs without knowing other participants' strategies. In light of the proof in Theorem 3, there is a unique maximal value of the equation Eq. (29), so that the requester cannot acquire a larger utility if he/she adopts other strategies (i.e., $\beta^t \neq \beta^{t*}$). Similarly, if β^{t*} has been determined, the optimal strategy $\boldsymbol{p}^{t*}$ of the second-tier leader (i.e., the platform) can be computed uniquely based on Eq. (19); if the strategies of two leaders (i.e., β^{t*} and $\boldsymbol{p}^{t*}$) are fixed, the followers (i.e., workers) can determine their optimal strategies $\boldsymbol{\delta}^{t*}$ based on Eq. (18). In a word, each stage has a unique equilibrium under the optimal strategy group $\langle \beta^{t*}, \boldsymbol{p}^{t*}, \boldsymbol{\delta}^{t*} \rangle$, and no one is willing to apply other strategies which will not make the participant earn the maximum utility. According to Definition 7, we can conclude that the THS game has a unique Stackelberg equilibrium.

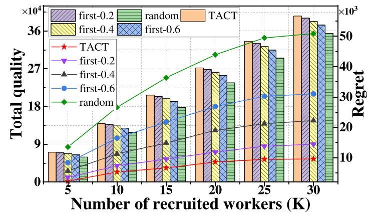


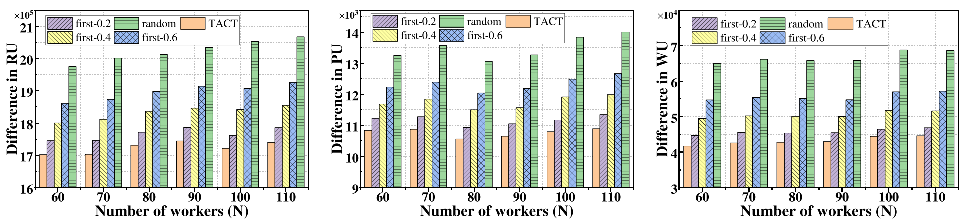
Fig. 9. Total quality and regret vs. Number of recruited workers K ($N = 50, T = 1000$).

5 Performance Evaluation

In this section, we evaluate the performance of the proposed mechanism TACT with extensive simulations.

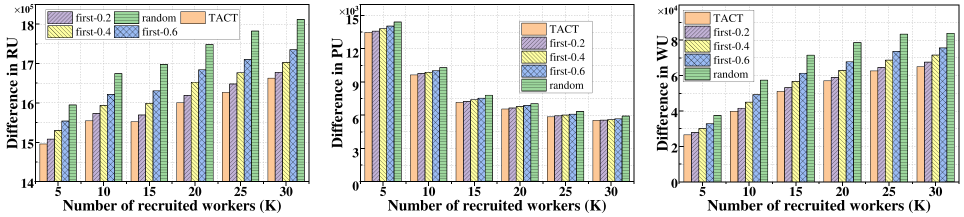
5.1 Evaluation Methodology

Simulation Setup. We adopt a real data trace of Queensu [33]. This dataset is to be used in conjunction with the roma/taxi dataset, and the taxis provide collected data of the areas in Rome where they were located (GPS coordinates of approximately 289 taxicabs). We first choose N taxicabs from the trace as candidate workers, where N is selected from [50, 110]. Then, we select $\Delta_{num} = 20$ since the highest answer frequency in the trace is 17 times one day (Taxi ID 135). We also simulate the social network based on a real data trace from SNAP (Gowalla) [34], which is a location-based social friendship network built by mobile phone users. We randomly select N nodes with edges from the undirected



(a) Requester (b) Platform (c) Workers

Fig. 8. Difference in utility of each party vs. Number of workers N.



(a) Requester (b) Platform (c) Workers

Fig. 10. Difference in utility of each party vs. Number of recruited workers K.

network (node numbers: 196,591). If two workers have no social edge, $\phi_{ij}=0$; otherwise, we set that ϕ_{ij} between any two workers follows a normal distribution with mean = 0 and standard deviation σ_Φ. Here, σ_Φ is produced from 0.02 to 0.5, the total period T changes from [1000, 1500], and the number of recruited workers K is selected from [5, 30]. The default values are $N=50$, $T=1000$, $K=10$. We randomly generate q_i, a_i, b_i from [0.2, 0.98], [2, 4], [0.1, 0.8], respectively. Besides, we adopt truncated Gaussian distribution to simulate the observed answer qualities and set $c=0.1$, $d=0.1$ by default. Table 2 lists some simulation parameters, where default values are in bold fonts.

Compared Algorithms. Since the existing studies do not consider the unknown social-aware workers and the three-party game, they are not applicable to our model. Like in [18], [35], [36], we borrow two basic algorithms: First-ϵ bandit and Random for comparison. Here, "First-ϵ" randomly selects K workers in the first ϵT periods and greedily recruits the top K workers with the highest CUCB-based qualities in the remaining periods. "Random" selects K workers in each period randomly. Additionally, we implement the optimal algorithm (called OPT) as a baseline in which the qualities of all workers are known in advance.

5.2 Evaluation Results

Let RU, PU, and WU denote Requester Utility, Platform Utility, and Workers Utility. "Difference" means the difference in utility between the OPT and each other algorithms. It is unnecessary to observe all workers so that we only choose three workers (i.e., worker-2/6/10) for illustration.

Evaluation of TACT. We first evaluate the impact of the number of periods T and change it from 1000 to 1500 with $K=10$ and $N=50$. Fig. 5 depicts the changing trend of total quality and regret. For the sake of comparison fairness, we adopt the average of observed answer qualities in each period. With the increase of the number of periods, we observe that TACT can achieve a higher total quality and a lower regret than the compared algorithms. Especially, when the number of periods is 1500, the achieved regret of TACT is about 37.7% and 84.2% lower than those of the "First-0.2" and "Random" algorithms on average, respectively. In addition, we evaluate the difference between all implemented algorithms and OPT on each participant's total utility, as presented in Fig. 6. The results indicate that TACT still outperforms compared algorithms in terms of RU, PU, and WU since the gap between TACT and the optimal algorithm is smaller.

Then, we evaluate the performance of TACT by changing the number of workers from 60 to 110, shown in Fig. 7 and Fig. 8. From these figures, we note that a larger N might bring in more low-quality workers and more complex social relations, but TACT can still remain a lower regret and utility gap compared with other algorithms, especially with the random policy. When the number of workers is 110, the regret of the random algorithm is about 344.8% higher than that of TACT on average. Since the random algorithm might always recruit low-quality workers, it has the maximum regret. Additionally, RU, PU, and WU are achieved with a fixed K under the SE, so they are relatively stable when the number of workers increases.

Next, we observe the effect of the number of recruited workers by changing K from 5 to 30. From Fig. 9, we observe that the growth rate of TACT's regret becomes slow. This is because K and N are getting closer. Especially, when the number of recruited workers is 30, the achieved regret of TACT is about 33.7% and 81.2% lower than those of the "First-0.2" and "Random" algorithms on average, respectively. As illustrated in Fig. 10, the gap of PU between OPT and TACT is narrowing, and TACT achieves the lowest utility of the platform compared with other algorithms. Moreover, the difference of WU first rises, this is because the optimal algorithm can recruit more high-quality workers. With the growth of K/N, the learned qualities and social network effects have an attenuate impact and the rate of gap growth has slowed down.

Finally, we investigate the impact of maximal answer frequency Δ_{num} from 6 to 20 with $N=50$ and $T=1000$, and exhibit the performance of different algorithms in terms of total quality, regret, and utility gaps, as depicted in Figs. 11 and 12, respectively. With the growth of Δ_{num}, total quality and the regret will also have an increase, which is consistent with the theoretical analysis results. At the same time, the difference of RU/PU/WU would be larger when the answer frequency is small. This is because many workers cannot apply its optimal strategy and are forced to adopt Δ_{num}. Thus, if we set Δ_{min} large enough, the gaps will be stable.

Evaluation of THS game. We demonstrate the evaluation results of the THS game in any period. As illustrated in Fig. 13(a), we first demonstrate the existence of Stackelberg equilibrium under different system parameters ξ when the

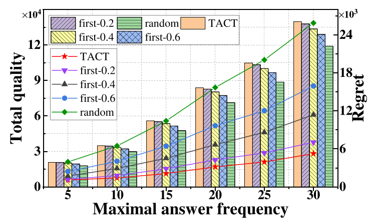
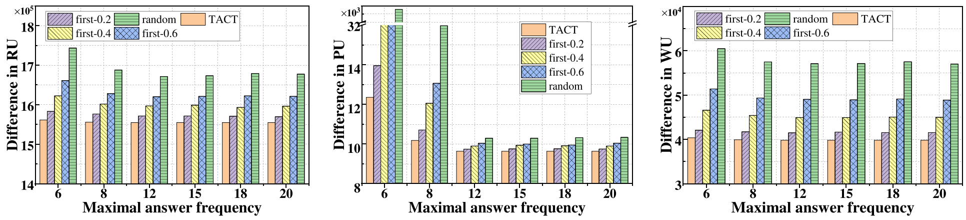


Fig. 11. Total quality and regret vs. Maximal answer frequency Δ_{num} ($N=50, T=1000$).



(a) Requester (b) Platform (c) Workers

Fig. 12. Difference in utility of each party vs. Maximal answer frequency Δ_{num}.

requester changes its strategy in the range [25, 55]. We can see that RU will always find a maximum point, and the larger ξ will harvest the larger RU when we change ξ. Therefore, the requester can find the optimal unit payment β^{t*} to maximize its utility. Besides, both WU and PU increase with the growth of β^t since workers and the platform will get a higher payment from the requester.

Then, we verify the existence of Stackelberg equilibrium for the platform under different system parameters. Because the platform needs to determine the unit salary for each worker, we only change the platform's strategy for worker-1 while other strategies are optimal. As shown in Fig. 14a, PU first climbs up and then declines as p_1^t grows, which signifies that the platform can meet the SE. Due to the strong relations between worker-1 and worker-10, the growth rate of WU-10 is much faster in Fig. 14b.

Next, we display the evaluation results by changing the strategy of worker-10 in the range $[1, 20]$. Clearly, the worker-10 can find the maximum point under different ξ, depicted in Fig. 15a. It demonstrates that all workers can meet an SE, i.e., there is a unique optimal answer frequency for each worker to earn its maximum utility. In Fig. 15b, since WU is influenced by salary, quality, cost parameters (i.e., a_i, b_i), and social network effects according to Eq. (5), worker-2's WU remains unchanged when there is $\phi_{2,10} = 0$.

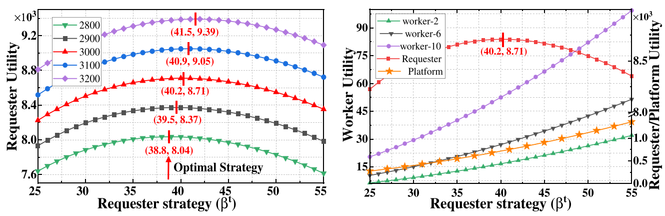
Moreover, we change the standard deviation of the social network σ_Φ from 0.02 to 0.5, and then evaluate the influence of social network effects on all participants' utilities and strategies. Note that, the value of ϕ_{ij} remains invariant if $\phi_{ij} = 0$. With the increase of σ_Φ, the strength of social relations will become stronger among workers and the social benefits of workers will rise correspondingly. Thus, some workers will add the answer frequency in Fig. 16b and their utilities will go up in Fig. 16a. In addition, both PU and RU can get a marked increase, which embodies the positive significance of social networks. Meanwhile, since most workers have higher social benefits, the requester and the platform can appropriately cut down the compensation (i.e., lower the optimal β^t and $\boldsymbol{p}^t$).

Finally, Fig. 17 presents the effect of system parameters and Fig. 18 plots the evaluation results when we change the worker-10's cost parameter b_{10} from 0.1 to 4. When we enlarge the system parameter ξ, the strategies and the utilities of all parties will increase. The reason is similar to the results in Fig. 13a. From Fig. 18b, we can notice that the larger b_i would cause a higher cost of worker i, so that the strategy of worker-10 has an appropriate decline. In order to compensate workers, the requester and the platform will improve the unit payment and the unit salary, respectively. The impact extent of each worker's WU depends on the social network effects among workers.

6 Related Work

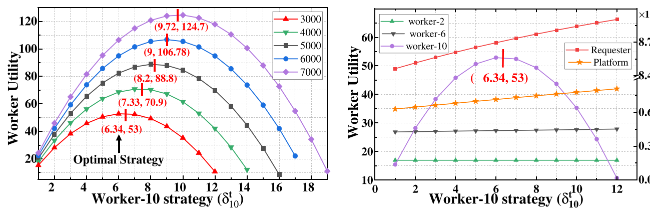
We review the related work from the following aspects:

Incentive Mechanisms: Aware of the importance of incentivizing SC participants, much effort has been devoted to



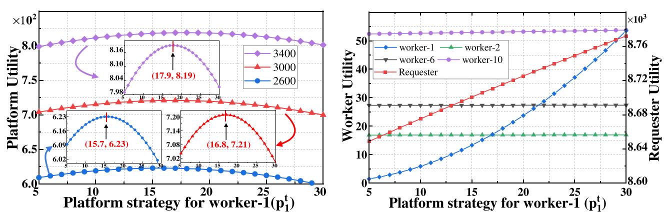
(a) Requester Utility vs. β^t (b) WU, RU, PU vs. β^t

Fig. 13. SE of requester and Impact of β^t.



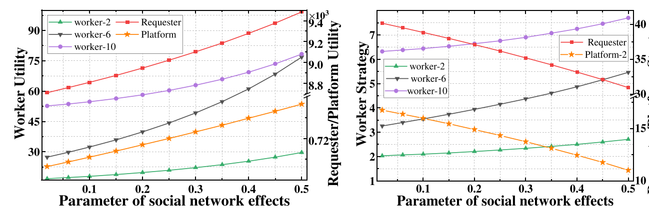
(a) Worker Utility vs. δ_{10}^t (b) WU, RU, PU vs. δ_{10}^t

Fig. 15. SE of worker-10 and Impact of δ_{10}^t.



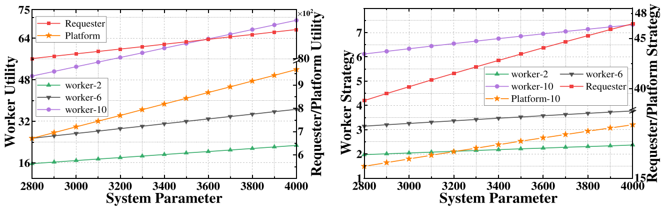
(a) Platform Utility vs. p_1^t (b) WU and RU vs. p_1^t

Fig. 14. SE of platform and Impact of p_1^t.



(a) WU, RU, PU vs. σ_Φ (b) Strategy vs. σ_Φ

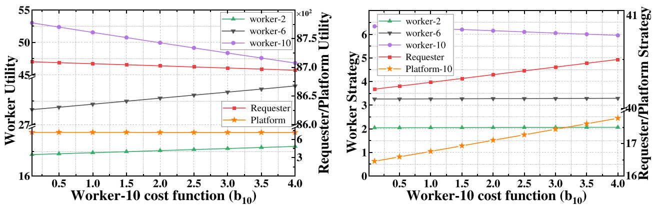
Fig. 16. Impact of social network effects.



(a) WU, RU, PU vs. ξ

(b) Strategy vs. ξ

Fig. 17. Impact of system parameter ξ.



(a) WU, RU, PU vs. b_{10}

(b) Strategy vs. b_{10}

Fig. 18. Impact of worker-10's parameter b_{10}.

designing various incentive mechanisms [9], [10], [11], [12], [13], [14], [15], [23], [27], [28], [37], [38], [39], [40], [41], [42], [43], [44], [45]. Diverse tools have been adopted in the previous works, such as auction mechanism [46], [47], [48], [49], contract theory [10], reputation mechanism [40], deep reinforcement learning [28], [39], and game theory [27], [28]. For example, [11] used two reverse auction models to maximize the social welfare; [37] proposed a pricing strategy on SC tasks to maximize the total revenue of the platform; [28] formulated a multi-leader multi-follower Stackelberg game and designed an incentive mechanism based on deep reinforcement learning to deal with the Markov decision process. However, many of them did not take the social network into account. Owing to the non-ignorable importance of social network effects, it is necessary to design an effective incentive mechanism for socially-aware SC systems.

With the great success of social networks, a few researches on the incentive mechanism with considering social network effects have been studied. For instance, the study in [23] combined the impacts of both the user diversity and social effect into the reward mechanism design, so as to improve the profit of service provides and the social surplus of users; [31] analyzed the behaviors of contributors in the Word-of-Mouth crowdsourcing market; The authors in [27] played a two-stage Stackelberg game among service providers and users, in which the interconnections of service providers and the social influence of users are incorporated in the payoff modeling; Wang *et al.* [38] proposed a dynamic incentive mechanism based on social networks and the SIR epidemic model to recruit sufficient workers, and provides the rewards according to workers' actions. However, most of these researches just concentrate on the interaction of two parties without considering the utilities of requesters.

Only a few studies [29], [50] involve three parties. Nevertheless, [50] still splits the whole process into two double-side problems, and [29] assumes that requesters' payments are fixed. In other words, none of them takes three parties' utilities, the unknown qualities of workers and the impact of social networks into account together, so that they cannot be applied in our system to acquire the multi-win solution.

CMAB Mechanisms: We model the unknown social-aware worker recruitment problem as a novel CMAB problem. The existing algorithms for MAB [32], [35], [36] do not involve the quality learning or the incentive issues. The most related works are [18], [19] in which they study the top K bandit selection problem. [18] combined the auction into the CMAB model to determine the critical payment. The authors in [19] developed a modified Thompson sampling worker selection by comparing workers' extrinsic and intrinsic abilities. Nevertheless, they do not take the number of workers' social relations into consideration and can only guarantee the non-negativity of workers' utilities. More importantly, our exploration phase considers not only the arm's reward but also its number of social relations, so that our selection policy and regret analysis are different from the existing CMAB models. To the best of our knowledge, our paper is the first work considering social benefits, unknown social-aware worker recruitment, and multiple maximization goals at the same time for SC systems.

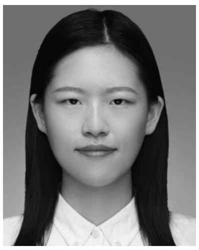
7 Conclusion

In this paper, we study the incentive problem for SC with unknown social-aware workers, and propose the incentive mechanism TACT including the unknown social-aware worker recruitment and the payment computation. We formalize the recruitment problem as a K-arm CMAB model and design a greedy arm-pulling scheme, which can balance the trade-off among exploration, exploitation, and workers' social benefits. After the recruitment process, we model the payment computation problem as a THS game among all participants, in which the social benefits of recruited workers are taken into account. We derive the optimal strategy group for each party so as to constitute a unique SE. In addition, sufficient theoretical analysis and trace-based simulations show its significant performance.

References

[1] Y. Tong, Z. Zhou, Y. Zeng, L. Chen, and C. Shahabi, "Spatial crowdsourcing: A survey," *VLDB J.*, vol. 29, no. 1, pp. 217–250, 2020.
[2] Y. Zhao and Q. Han, "Spatial crowdsourcing: Current state and future directions," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 102–107, Jul. 2016.
[3] X. Zhang, Z. Yang, Y. Liu, and S. Tang, "On reliable task assignment for spatial crowdsourcing," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, no. 1, pp. 174–186, Jan.–Mar. 2019.
[4] C. H. Liu, Z. Dai, H. Yang, and J. Tang, "Multi-task-oriented vehicular crowdsensing: A deep learning approach," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 1123–1132.
[5] K. Xie *et al.*, "Quick and accurate false data detection in mobile crowd sensing," in *Proc. IEEE Conf. Comput. Commun.*, 2019, pp. 2215–2223.
[6] C. Huang, H. Yu, J. Huang, and R. A. Berry, "Strategic information revelation in crowdsourcing systems without verification," in *Proc. IEEE Conf. Comput. Commun.*, 2021, pp. 1–10.
[7] " Waze," 2022. [Online]. Available: https://www.waze.com
[8] " Gigwalk," 2022. [Online]. Available: http://gigwalk.com
[9] Z. Wang *et al.*, "Task-bundling-based incentive for location-dependent mobile crowdsourcing," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 54–59, Feb. 2019.
[10] P. Sun *et al.*, "Towards personalized privacy-preserving incentive for truth discovery in crowdsourced binary-choice question answering," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 1133–1142.

[11] H. Jin, L. Su, D. Chen, H. Guo, K. Nahrstedt, and J. Xu, "Thanos: Incentive mechanism with quality awareness for mobile crowd sensing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 8, pp. 1951–1964, Aug. 2019.
[12] Y. Qu *et al.*, "Posted pricing for chance constrained robust crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 19, no. 1, pp. 188–199, Jan. 2020.
[13] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in Proc. *IEEE Conf. Comput. Commun.*, 2015, pp. 2830–2838.
[14] W. Liu, Y. Yang, E. Wang, and J. Wu, "Dynamic user recruitment with truthful pricing for mobile crowdsensing," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 1113–1122.
[15] H. Wang, S. Guo, J. Cao, and M. Guo, "MeLoDy: A long-term dynamic quality-aware incentive mechanism for crowdsourcing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 4, pp. 901–914, Apr. 2018.
[16] A. Biswas, S. Jain, D. Mandal, and Y. Narahari, "A truthful budget feasible multi-armed bandit mechanism for crowdsourcing time critical tasks." in *Proc. 20th Int. Conf. Auton. Agents Multiagent Syst.*, 2015, pp. 1101–1109.
[17] B. An, M. Xiao, A. Liu, X. Xie, and X. Zhou, "Crowdsensing data trading based on combinatorial multi-armed bandit and stackelberg game," in *Proc. IEEE 37th Int. Conf. Data Eng.*, 2021, pp. 253–264.
[18] G. Gao, H. Huang, M. Xiao, J. Wu, Y. E. Sun, and S. Zhang, "Auction-based combinatorial multi-armed bandit mechanisms with strategic arms," in *Proc. IEEE Conf. Comput. Commun.*, 2021, pp. 1–10.
[19] Y. Wu, F. Li, L. Ma, Y. Xie, T. Li, and Y. Wang, "A context-aware multiarmed bandit incentive mechanism for mobile crowd sensing systems," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7648–7658, Oct. 2019.
[20] X. Gong, L. Duan, X. Chen, and J. Zhang, "When social network effect meets congestion effect in wireless networks: Data usage equilibrium and optimal pricing," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 2, pp. 449–462, Feb. 2017.
[21] O. Candogan, K. Bimpikis, and A. E. Ozdaglar, "Optimal pricing in networks with externalities," *Operations Res.*, vol. 60, no. 4, pp. 883–905, 2012.
[22] I. P. Fainmesser and A. Galeotti, "Pricing network effects," *Rev. Econ. Stud.*, vol. 83, no. 1, pp. 165–198, 2016.
[23] M. H. Cheung, F. Hou, and J. Huang, "Make a difference: Diversity-driven social mobile crowdsensing," in *Proc. IEEE Conf. Comput. Commun.*, 2017, pp. 1–9.
[24] M. Belhaj and F. Deroïan, "The value of network information: Assortative mixing makes the difference," *Games Econ. Behav.*, vol. 126, pp. 428–442, 2021.
[25] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2-3, pp. 235–256, 2002.
[26] F. Li, J. Liu, and B. Ji, "Combinatorial sleeping bandits with fairness constraints," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1799–1813, Jul.–Sep. 2020.
[27] J. Nie, J. Luo, Z. Xiong, D. Niyato, P. Wang, and H. V. Poor, "A multi-leader multi-follower game-based analysis for incentive mechanisms in socially-aware mobile crowdsensing," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1457–1471, Mar. 2021.
[28] Y. Zhan, C. H. Liu, Y. Zhao, J. Zhang, and J. Tang, "Free market of multi-leader multi-follower mobile crowdsensing: An incentive mechanism design by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 19, no. 10, pp. 2316–2329, Oct. 2020.
[29] Y. Li, F. Li, S. Yang, P. Zhou, L. Zhu, and Y. Wang, "Three-stage stackelberg long-term incentive mechanism and monetization for mobile crowdsensing: An online learning approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1385–1398, Apr.–Jun. 2021.
[30] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: J. Econometric Soc.*, vol. 33, no. 3, pp. 520–534, 1965.
[31] R. Wang, F. Zeng, L. Yao, and J. Wu, "Game-theoretic algorithm designs and analysis for interactions among contributors in mobile crowdsourcing with word of mouth," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8271–8286, Sep. 2020.
[32] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial multi-armed bandit: General framework and applications," in *Proc. Int. Conf. Mach. Learn.*, 2013, vol. 28, pp. 151–159.
[33] M. A. Alswailim, H. S. Hassanein, and M. Zulkernine, "CRAWDAD dataset queensu/crowd temperature (v. 2015-11-20): derived from roma/taxi (v. 2014-07-17)," 2015, [Online]. Available: https://doi.org/10.15783/C7CG65.
[34] J. Leskovec and A. Krevl, "SNAP datasets: Gowalla: Stanford large network dataset collection," 2014, [Online]. Available: http://snap.stanford.edu/data/loc-Gowalla.html.
[35] J. Zuo, X. Zhang, and C. Joe-Wong, "Observe before play: Multi-armed bandit with pre-observations," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 7023–7030.
[36] H. Wang, Y. Yang, E. Wang, W. Liu, Y. Xu, and J. Wu, "Combinatorial multi-armed bandit based user recruitment in mobile crowdsensing," in *Proc. IEEE 17th Annu. Int. Conf. Sens., Commun., Netw.*, 2020, pp. 1–9.
[37] Y. Tong, L. Wang, Z. Zhou, L. Chen, B. Du, and J. Ye, "Dynamic pricing in spatial crowdsourcing: A matching-based approach," in *Proc. Int. Conf. Manage. Data*, 2018, pp. 773–788.
[38] Z. Wang, Y. Huang, X. Wang, J. Ren, Q. Wang, and L. Wu, "Socialrecruiter: Dynamic incentive mechanism for mobile crowdsourcing worker recruitment with social networks," *IEEE Trans. Mobile Comput.*, vol. 20, no. 5, pp. 2055–2066, May 2021.
[39] Y. Zhao and C. H. Liu, "Social-aware incentive mechanism for vehicular crowdsensing by deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2314–2325, Apr. 2021.
[40] H. Gao, C. H. Liu, J. Tang, D. Yang, P. Hui, and W. Wang, "Online quality-aware incentive mechanism for mobile crowd sensing with extra bonus," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2589–2603, Nov. 2019.
[41] Y. Chen, X. Wang, B. Li, and Q. Zhang, "An incentive mechanism for crowdsourcing systems with network effects," *ACM Trans. Internet Technol.*, vol. 19, pp. 1–21, 2019.
[42] Q. Liu, L. Zheng, Y. Shen, and L. Chen, "Finish them on the fly: An incentive mechanism for real-time spatial crowdsourcing," in *Proc. Int. Conf. Database Syst. Adv. Appl.*, 2020, vol. 12113, pp. 694–710.
[43] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, 2012, pp. 173–184.
[44] J. Liu, Y. Ji, W. Lv, and K. Xu, "Budget-aware dynamic incentive mechanism in spatial crowdsourcing," *J. Comput. Sci. Technol.*, vol. 32, no. 5, pp. 890–904, 2017.
[45] B. Zhao, X. Liu, W. Chen, W. Liang, X. Zhang, and R. H. Deng, "PRICE: Privacy and reliability-aware real-time incentive system for crowdsensing," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17584–17595, Dec. 2021.
[46] X. Chen, L. Zhang, Y. Pang, B. Lin, and Y. Fang, "Timeliness-aware incentive mechanism for vehicular crowdsourcing in smart cities," *IEEE Trans. Mobile Comput.*, early access, Jan. 19, 2021, doi: 10.1109/TMC.2021.3052963.
[47] F. Restuccia, P. Ferraro, S. Silvestri, S. K. Das, and G. L. Re, "Incentme: Effective mechanism design to stimulate crowdsensing participants with uncertain mobility," *IEEE Trans. Mobile Comput.*, vol. 18, no. 7, pp. 1571–1584, Jul. 2019.
[48] C. Xiang, S. He, K. G. Shin, Y. Qu, and P. Yang, "Incentivizing platform-user interactions for crowdsensing," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8314–8327, May 2021.
[49] Y. Wang, Z. Cai, Z. Zhan, Y. Gong, and X. Tong, "An optimization and auction-based incentive mechanism to maximize social welfare for mobile crowdsourcing," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 3, pp. 414–429, Jun. 2019.
[50] G. Gao, M. Xiao, J. Wu, S. Zhang, L. Huang, and G. Xiao, "DPDT: A differentially private crowd-sensed data trading mechanism," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 751–762, Jan. 2020.



Yin Xu received the BS degree from the School of Computer Science and Technology, Anhui University, Hefei, China, in 2019. She is currently working toward the PhD degree with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China. Her research interests include spatial crowdsourcing, game theory, reinforcement learning, federated learning, mobile computing, privacy preservation, and incentive mechanism.

Mingjun Xiao (Member, IEEE) received the PhD degree from the University of Science and Technology of China in 2004. He is currently a professor with the School of Computer Science and Technology, University of Science and Technology of China. He has authored or coauthored more than 100 papers in referred journals and conferences, including TMC, TC, TPDS, TON, TKDE, TSC, INFOCOM, ICDE, and ICNP. His research interests include mobile crowdsensing, edge computing, federated learning, auction theory, data security and privacy. He was a TPC member of INFOCOM'22, IJCAI'22, INFOCOM'21, IJCAI'21, INFOCOM'20, INFOCOM'19, ICDCS'19, DASFAA'19, and INFOCOM'18. He is on the reviewer board of several top journals, such as TMC, TON, TPDS, TSC, TVT, and TCC.

Jie Wu (Fellow, IEEE) is currently the director of the Center for Networked Computing and Laura H. Carnell professor with Temple University. He is also the director of International Affairs with the College of Science and Technology. He was the chair of the Department of Computer and Information Sciences from the summer of 2009 to the summer of 2016 and an associate vice provost for International Affairs from the fall of 2015 to the summer of 2017. Prior to joining Temple University, he was a program director with National Science Foundation and was a distinguished professor with Florida Atlantic University. He has regularly authored or coauthored in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including the *IEEE Transactions on Mobile Computing*, *IEEE Transactions on Service Computing*, *Journal of Parallel and Distributed Computing*, and *Journal of Computer Science and Technology*. His research interests include mobile computing and wireless networks, routing protocols, network trust and security, distributed algorithms, applied machine learning, and cloud computing. He is/was the general chair or co-chair for IEEE IPDPS'08, IEEE DCOSS'09, IEEE ICDCS'13, ACM MobiHoc'14, ICPP'16, IEEE CNS'16, WiOpt'21, and ICDCN'22, and the program chair or co-chair of IEEE MASS'04, IEEE INFOCOM'11, CCF CNCC'13, and ICCCN'20. He was an IEEE Computer Society distinguished visitor, ACM distinguished speaker, and the chair of IEEE Technical Committee on Distributed Processing (TCDP). He is a fellow of AAAS. He was the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.

Sheng Zhang (Member, IEEE) received the BS and PhD degrees from Nanjing University in 2008 and 2014, respectively. He is currently an associate professor with the Department of Computer Science and Technology, Nanjing University. He is also a member of the State Key Lab. for Novel Software Technology. To date, he has authored or coauthored more than 80 papers, including those appeared in JSAC, TMC, TON, TPDS, TC, MobiHoc, ICDCS, INFOCOM, SECON, IWQoS, and ICPP. His research interests include cloud computing and edge computing. He was the recipient of the Best Paper Award of IEEE ICCCN 2020 and Best Paper Runner-Up Award of IEEE MASS 2012. He was the recipient of 2020 ACM Nanjing Rising Star Award and 2015 ACM China Doctoral Dissertation Nomination Award. He is a member of ACM and a senior member of CCF.

Guoju Gao received the PhD degree in computer science and technology from the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China. He is currently an assistant professor with the School of Computer Science and Technology, Soochow University. From January 2019 to January 2020, he visited to Temple University, USA. He has authored or coauthored 30 articles in referred journals and conferences, including IEEE TMC, ToN, TPDS, TSC, INFOCOM, and ICPP. His research interests include mobile computing, network traffic measurement, and reinforcement learning.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.