Budgeted Unknown Worker Recruitment for Heterogeneous Crowdsensing Using CMAB

Guoju Gao, He Huang, *IEEE Member*, Mingjun Xiao, *IEEE Member*, Jie Wu, *IEEE Fellow*, Yu-E Sun, *IEEE Member*, and Yang Du

Abstract—Mobile crowdsensing, through which a requester can coordinate a crowd of workers to complete some sensing tasks, has attracted significant attention recently. In this paper, we focus on the unknown worker recruitment problem in mobile crowdsensing, where workers' sensing qualities are unknown a priori. We consider the scenario of recruiting workers to complete some continuous sensing tasks. The whole process is divided into multiple rounds. In each round, every task may be covered by more than one recruited workers, but its completion quality only depends on these workers' maximum sensing quality. Each recruited worker will incur a cost and each task is attached a weight to indicate its importance. Our objective is to determine a recruiting strategy to maximize the total weighted completion quality under a limited budget. We model such unknown worker recruitment process as a *novel combinatorial multi-armed bandit* (CMAB) problem, and propose an unknown worker recruitment algorithm based on the modified upper confidence bound (UCB). Moreover, we extend the problem to the case where the workers' costs are also unknown and design the corresponding algorithm. We analyze the regret bounds of the two proposed algorithms through rigorous proofs. In addition, we also study the unknown worker recruitment problem with fairness constraints. Here, the term "fairness" means that the platform must guarantee a minimum selection fraction for each registered worker, so that the platform can avoid the scenario where some workers are over-recruited but some others might be under-recruited. For this problem, we devise a fairness-aware unknown worker recruitment algorithm. Finally, we demonstrate the performance of the proposed algorithms through extensive simulations on real-world traces.

Index Terms—Heterogeneous crowdsensing, combinatorial multi-armed bandits, online learning, worker recruitment, fairness.

1 Introduction

Mobile Crowdsensing (MC) is a newly-emerging paradigm where a crowd of mobile users can be recruited to cooperatively complete some sensing tasks by using their carried smart phones [2]–[12]. Owing to users' mobility and the diversity of sensors embedded in their smart phones, MC can deal with various sensing tasks distributed in a large-scale area. Consequently, it has stimulated many applications that a single user cannot cope with, such as traffic information collection, noise pollution collection, water pollution monitoring, and urban WiFi characterization, etc, which has provided a great convenience for our daily life.

A typical MC system includes a platform residing on a cloud. Through the platform, service requesters can publicize their sensing tasks and recruit mobile users (a.k.a., *workers*) to complete these tasks. Generally, due to the diverse smart phones (e.g., camera pixel, software version, storage capacity, etc.) and worker behaviors (e.g., skill level, engagement level, etc.), workers might result in different sensing qualities, even for the same task. Thus, recruiting workers to achieve higher completion qualities or lower costs is one of the most important problems in MC. Much effort has been devoted to designing worker recruitment or task allocation algorithms in recent years [5], [8], [13]–[16]. However, most of the existing work assumes that workers' sensing qualities are known in advance, which is not true in practice. So far, only a few researches have investigated the scenario where workers' sensing qualities are unknown a priori, i.e., the so-called unknown worker recruitment problem. For example, [17] studies how to maximize the task completion ratio while considering unknown workers' reliability and the dynamic arrivals of tasks; [18] develops a modified Thompson Sampling worker selection algorithm to recruit some unknown workers; [19] investigates how to select the most informative contributors with unknown costs for budgeted crowdsensing; [20] designs a context-aware hierarchical online learning algorithm for performance maximization of MC. Nevertheless, these researches either neglect the requester's budget constraint or mainly involve homogeneous MC models in which each task can be completed by all workers, although their sensing qualities might be different.

In this paper, we focus on the unknown worker recruitment problems in heterogeneous MC systems. Consider a scenario where a requester wants to recruit workers to collect the traffic data (e.g., traffic photos or videos) at some

- *G. Gao, H. Huang and Y. Du are with the School of Computer Science and Technology, Soochow University, Suzhou, P. R. China. Correspondence to: huangh@suda.edu.cn*
- *M. Xiao is with the School of Computer Science and Technology / Suzhou Institute for Advanced Study, University of Science and Technology of China, Hefei, P. R. China. Email: xiaomj@ustc.edu.cn*
- *J. Wu is with the Department of Computer and Information Sciences, Temple University, 1805 N. Broad Street, Philadelphia, PA 19122. E-mail: jiewu@temple.edu*
- *Y. Sun is with the Rail Transportation Department, Soochow University, Suzhou, P. R. China. E-mail: sunye12@suda.edu.cn*

This paper is an extended version of the conference paper [1] published in IEEE INFOCOM 2020. The work of Guoju Gao, He Huang, Yu-E Sun, and Yang Du was supported by the National Natural Science Foundation of China (NSFC) under Grant 61873177, 62072322, and U20A20182. The work of Mingjun Xiao was supported in part by the NSFC under Grant 61872330, 61936015, and U1709217, in part by Anhui Initiative in Quantum Information Technologies under Grant AHY150300, and in part by the NSF of Jiangsu in China under Grant BK20191194. The work of Jie Wu was supported by NSF Grants CNS 1824440, CNS 1828363, CNS 1757533, CNS 1629746, CNS 1651947, and CNS 1564128.

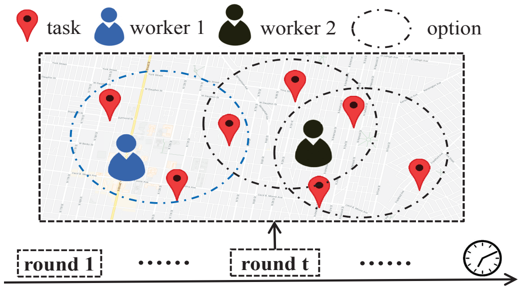

Fig. 1: Illustration of the heterogeneous crowdsensing.

urban intersections for a period of time. The whole data collection is divided into multiple rounds. In each round, it consists of many location-related sensing tasks, each of which corresponds to a traffic intersection, as shown in Fig. 1. Here, each task is attached with a weight to indicate its importance. Each worker can complete (a.k.a., *cover*) one or more tasks. The tasks that each worker can deal with might be different, i.e., the sets of tasks covered by different workers are heterogeneous. All workers will tell the platform the tasks they want to perform and the costs they expect to charge. In order to increase the probability of being recruited in each round, each worker can provide multiple options at the beginning, where each option is composed of different task combinations and costs, but at most one option will be selected in each round. Moreover, each worker has a sensing quality, following an unknown distribution. Note that only when the workers complete these tasks and return the sensing results, the platform or requester is able to measure the sensing qualities for each worker. Our objective is to design a worker recruitment strategy that can maximize the total task completion quality (i.e., the weighted sum of the completion qualities of all tasks over all rounds) under a given budget.

In the above unknown worker recruitment problem of heterogeneous MC systems, the main challenge lies in that the platform does not know the workers' sensing qualities in advance, so it needs to learn their quality values by tentatively recruiting workers to complete some tasks and then selects the best group of workers according to the learned results. Generally, the two processes are called exploration and exploitation [21]–[23], respectively. We need to balance the two processes so as to maximize the total task completion quality under a given budget. To address this challenge, we model the unknown worker recruitment process as a novel Combinatorial Multi-Armed Bandit (CMAB) problem, where each worker is seen as an arm, its sensing quality is seen as the corresponding reward, and recruiting workers is equivalent to pulling arms. Moreover, in order to balance the achieved total completion quality and the required number of total rounds under the budget, we let a fixed number of arms (i.e., K) be pulled in each round. Our CMAB model has two novel characteristics, which is different from all the existing CMAB models. First, each arm has multiple options, each of which corresponds to a set of covered tasks and a cost. The platform needs to not only select arms but also determine the option for each arm. Second, it contains a budget-limited maximum weighted coverage problem (i.e., maximizing the total task completion quality, which involves a weighted sum function on some maximum sensing qualities), making it very challenging.

As we know, Upper Confidence Bound (UCB) is a widely-used arm-pulling strategy, designed for the traditional multi-armed bandit problem [21], [24]. It always selects the arm that has the largest value on the estimated reward and the upper bound of confidence to be pulled. However, the simple application of the existing algorithms cannot solve our problem efficiently. To that end, we extend the UCB strategy by adding two extra designs. First, when estimating the reward and computing the confidence for each arm, we consider that workers' sensing qualities might be learned multiple times in one round, since each worker has multiple options and covers multiple tasks. Second, we adopt the greedy strategy to solve the budget-limited maximum weighted coverage problem, when determining which arms should be pulled. Next, according to the extended UCB arm-pulling strategy, we design an unknown worker recruitment algorithm. Furthermore, we extend our problem to the scenario where workers' costs are also unknown and devise another algorithm. In addition, we also study the unknown worker recruitment problem by involving the fairness constraint, which can guarantee a minimum selection fraction for each registered worker. In such a way, the MC system can avoid the scenario that some workers are over-recruited but some others might be under-recruited. We propose the fairness-aware unknown worker recruitment algorithm for this problem.

Our major contributions are summarized as follows:

- We introduce the unknown worker recruitment problem for heterogeneous MC systems and turn it into a novel K-arm CMAB problem. Unlike existing researches, this CMAB model contains a budget-limited maximum weighted coverage problem and each arm has multiple candidate options.
- We propose an extended UCB based arm-pulling strategy to solve our CMAB problem and design the corresponding unknown worker recruitment online algorithm. Moreover, we derive the regret bound of the proposed algorithm, that is, $O(NLK^3 \ln(B + NLK^2 \ln(N^2K^2ML)))$, in which B, N, M, and L denote the budget, the number of workers, the number of tasks, and the number of options for each worker, respectively.
- We also study an extended case where both the sensing qualities and the costs of workers are unknown, and devise another algorithm with a provable regret guarantee $O(NLK^3 \ln(NMB + N^2K^2ML \ln(N^2K^2ML)))$.
- Moreover, we consider another case by involving the fairness constraint of workers. That is, the platform must guarantee a minimum selection fraction for each registered worker under the given budget. By introducing the virtual queue technique in the CMAB problem, we design a fairness-aware unknown worker recruitment algorithm.
- We conduct extensive experimental simulations on real-world traces to evaluate the significant performance of the three proposed algorithms, and the results show that our policies outperform the compared algorithm.

The remainder of the paper is organized as follows. We first introduce the crowdsensing model and the optimization problem in Section 2. Next, we design an unknown worker recruitment algorithm in Section 3. We then propose another algorithm for the extended problem in Section 4. In Section 5, we devise the fairness-aware unknown worker recruitment algorithm to address the fairness constraint of workers. In Section 6, we evaluate the performances of the proposed algorithms. After reviewing the related work in Section 7, we conclude the paper in Section 8.

2 System Model & Problem

In this section, we first introduce the system overview, and then present the model. Finally, we formalize the problem.

2.1 System Overview

Consider an MC system, composed of a platform and a crowd of workers. A requester wants to collect some traffic data (e.g., photos, videos, etc.) for a period of time via the MC system, but constrained by a budget. The whole data collection consists of some location-related sensing tasks and is also divided into multiple rounds, each of which lasts a certain time interval. Each task here is attached with a weight to indicate its importance. First, the requester publicizes these tasks to all workers via the platform. Then, each worker will tell the platform which sensing tasks it is willing to perform. Moreover, the worker can provide multiple options at first, each of which consists of a subset of tasks that it can deal with and a cost that it wants to charge. Next, the platform will recruit some workers to perform the tasks round by round according to some strategy, until the budget is exhausted. Fig. 2 illustrates the main procedures.

For generality, we assume that the MC system is heterogeneous, where each task can be completed by multiple workers and each worker can also cover multiple tasks in each round. Moreover, each worker has a sensing quality when performing tasks. The quality value can be measured by the platform after the worker completes some tasks and submits the sensed results. For example, the platform may take the photo clarity, photographing angle, the number of photos, and some other factors into consideration when determining workers' sensing quality in terms of the photos. If a task is completed by more than one workers, we will only select the best sensing data and let the completion quality of this task be the maximum sensing quality of these workers. We assume that workers' sensing qualities follow some unknown distributions. The platform can learn and estimate these distributions after the workers complete some tasks. A profile is used to record the learned qualities of each worker. In this paper, we consider that all workers are truthful about their costs. This is reasonable because some auction mechanisms [25]–[30] can be introduced to solve the incentive issues.

2.2 Model

We let t denote the index of the round; let $\mathcal{N}=\{1,\cdots,i,\cdots,N\}$ and $\mathcal{M}=\{1,\cdots,j,\cdots,M\}$ denote N workers and M sensing tasks in the system, respectively. We use B to denote the requester's budget. Since each task has a different level of importance for the requester, we use w_j to denote the weight of the j-th task, and let $\sum_{j\in\mathcal{M}} w_j = 1$.

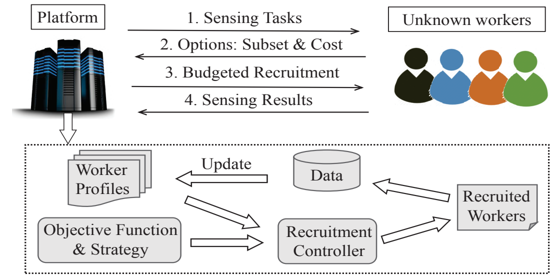

Fig. 2: Illustration of the main procedures in MC.

In our MC system, each worker $i \in \mathcal{N}$ would submit L (≥ 1) candidate options to the platform. We use $p_i^l = \langle \mathcal{M}_i^l, c_i^l \rangle$ to denote the l-th $(1 \leq l \leq L)$ option submitted by the worker i, where $\mathcal{M}_i^l \subseteq \mathcal{M}$ means the subset of tasks he can perform and c_i^l denotes the corresponding cost. Note that for each worker, at most one option can be selected in each round. Moreover, we suppose $c_i^1 \leq c_i^2 \leq \cdots \leq c_i^L$ for $\forall i \in \mathcal{N}$. In reality, the corresponding cost c_i^l is highly related to the number of tasks, i.e., the value of $|\mathcal{M}_i^l|$. We consider that the cost is proportional to the function of the number of tasks for simplicity. That is, we let $c_i^l = \varepsilon_i \cdot f(|\mathcal{M}_i^l|)$, where $f(\cdot)$, as a monotonically increasing function (i.e., performing more tasks must result in a higher cost), is given in our model. The values of ε_i (called cost parameter) for different workers are heterogeneous. For example, a worker carrying the smart phone with the advanced configurations (e.g., high-resolution camera, 5G network, etc.) generally has a large cost parameter. Moreover, ε_i is assumed to be known a priori here. In this paper, we will also consider an extended case where ε_i is unknown. Note that the value of c_i^l is normalized to $(0, 1]$. We let $\mathcal{P}_i = \{p_i^l | 1 \leq l \leq L\}$ denote the set of options submitted by worker i for simplicity, and further use $\mathcal{P} = \cup_{i\in\mathcal{N}} \mathcal{P}_i$ to denote the set of all options.

On the other hand, we use a normalized nonnegative random variable $q_{i,j}^t \in (0, 1]$ to denote the sensing quality of the worker i completing the task $j \in \mathcal{M}_i^l$ in the t-th round. In fact, for a particular worker (e.g., i), the values of $\{q_{i,j}^t | j \in \mathcal{M}_i^l, \forall t \geq 1\}$ follow an unknown independent and identically distribution with an unknown (unique) expectation q_i. This is because the sensing quality is mainly determined by the configurations of the workers' smart phones (e.g., camera pixel, software version, storage capacity, etc.) and their sensing behaviors (e.g., skill level, engagement level, worker habit, etc.). In other words, the options submitted by the same worker share the same quality distribution. If the l-th option submitted by the worker i (i.e., p_i^l) is selected in round t, i must perform all tasks in $\mathcal{M}_i^l$, and the quality values $\{q_{i,j}^t | j \in \mathcal{M}_i^l\}$ will be revealed. This indicates that the expected quality (i.e., q_i) would be learned $|\mathcal{M}_i^l|$ times by the platform. Moreover, due to the selection constraint (i.e., at most one of the worker's options can be selected in each round) and the shared quality distribution among the options submitted by a worker, one option cannot be simply seen as a separate worker. These factors make our problem differ from the traditional CMAB model [31], [32].

2.3 Problem

For the above heterogeneous MC system, we focus on recruiting K workers in each round so that the total weighted completion quality of all the tasks over all rounds can be maximized under a given budget. Here, the value of K has a vital impact on the performance of the MC system. We will verify this in the simulation section. We let $\mathcal{P}^t \subset \mathcal{P}$ denote the selected options in round t, in which $p_i^l \in \mathcal{P}^t$ means that the l-th option of worker i will be selected in round t. Since at most one option for a worker can be selected in each round, we have $\sum_{l=1}^{L} \mathbb{I}\{p_i^l \in \mathcal{P}^t\} \leq 1$ for $\forall i \in \mathcal{N}$, where $\mathbb{I}\{true\} = 1$ and $\mathbb{I}\{false\} = 0$. Moreover, we define the final completion quality of a task according to $\mathcal{P}^t$ in round t, denoted by $u^j(\mathcal{P}^t)$,

$$u^j(\mathcal{P}^t) = \begin{cases} 0; & j \notin (\cup_{p_i^l \in \mathcal{P}^t} \mathcal{M}_i^l), \\ \max\{q_{i,j}^t \mid p_i^l \in \mathcal{P}^t\}; & j \in (\cup_{p_i^l \in \mathcal{P}^t} \mathcal{M}_i^l). \end{cases} \quad (1)$$

We further use $u(\mathcal{P}^t)$ to denote the total achieved weighted completion quality of all tasks based on $\mathcal{P}^t$ in round t, i.e.,

$$u(\mathcal{P}^t) = \sum_{j \in \mathcal{M}} (w_j \cdot u^j(\mathcal{P}^t)). \quad (2)$$

Our objective is to determine $\{\mathcal{P}^1, \cdots, \mathcal{P}^t, \cdots\}$ in each round, such that the total expected weighted completion quality of all tasks is maximized under the budget constraint. We formulate our optimization problem as follows:

$$Maximize: \quad \mathbb{E}\Big[\sum_{t \geq 1} u(\mathcal{P}^t)\Big] \quad (3)$$

$$Subject\ to: \quad \sum_{t \geq 1} \sum_{p_i^l \in \mathcal{P}^t} c_i^l \leq B \quad (4)$$

$$|\mathcal{P}^t| = K \quad \text{for} \quad \forall t > 1 \quad (5)$$

$$\sum_{l=1}^{L} \mathbb{I}\{p_i^l \in \mathcal{P}^t\} \leq 1 \quad (6)$$

Eqs. (4) and (5) represent the budget and quantity constraints, while Eq. (6) indicates that at most one option of each worker can be selected in each round. Additionally, we summarize the commonly used notations throughout the paper in Table 1.

TABLE 1: Description of commonly-used notations.

Variable	Description
$\mathcal{N}$, $\mathcal{M}$	the sets of workers and sensing tasks.
i, j, t	the indexes for workers, tasks, and rounds.
K	the number of recruited workers in each round.
w_j	the weight of the j-th task for the requester.
ε_i	the cost parameter of i and $c_i^l = \varepsilon_i f(\lvert\mathcal{M}_i^l\rvert)$.
p_i^l	the l-th option submitted by the worker i.
$\mathcal{P}$	the sets of all options.
$\mathcal{P}^t$	the set of selected options in round t.
L	the number of options that a worker submits.
B	the budget given by the requester.
$q_{i,j}^t$	the quality value of i conducting j in round t.
$\overline{q}_i(t)$	the average quality of i until the t-th round.
$\widehat{q}_i(t)$	the UCB-based quality value of worker i.
q_i	the mean of the distribution $\{q_{i,j}^t \mid t \geq 1, j \in \mathcal{M}_i\}$.
$n_i(t)$	the number of i being learned until round t.
$n_i^l(t)$	the number of p_i^l being selected until round t.
$\mathbb{E}[\cdot]$	the expected function.
η_i	the required minimum fraction of rounds of i.
$\tau(B)$	the total rounds under the given budget B.
$V_i(t)$	the virtual queue length of V_i in round t.
ϱ	the tradeoff parameter for the fairness.

3 Algorithm Design

We first introduce the basic solution and then present the detailed algorithm. After that, we analyze the performance guarantee of the proposed algorithm.

3.1 Basic Solution

To address our unknown worker recruitment issue for the heterogeneous MC system under a given budget, we model it as a budgeted-limited K-arm CMAB problem, where each worker is seen as an arm, sensing quality is seen as the corresponding reward, and recruiting workers is treated as pulling arms. In this model, K workers are recruited in each round and each recruited worker's sensing quality would be learned multiple times in a round. We first extend the Upper Confidence Bound (UCB) to denote the learned quality values (called UCB-based quality). Then, we propose a UCB-based quality function by taking the maximum weighted coverage problem into consideration. Based on this, we adopt a greedy strategy to recruit K unknown workers in each round, that is, we always select the worker with the maximum ratio of the marginal UCB-based quality function value and recruitment cost. We introduce our detailed solution as follows.

When an option of a worker is selected in round t (e.g., $p_i^l \in \mathcal{P}^t$), the worker i must perform all sensing tasks in $\mathcal{M}_i^l$. In other words, the number of times of i being learned by the platform in round t is actually $|\mathcal{M}_i^l|$. Based on this, we first introduce $n_i^l(t)$ and $n_i(t)$ for $i \in \mathcal{N}, 1 \leq l \leq L$ to record the number of times that p_i^l is selected and the number of times that i is learned. That is,

$$n_i^l(t) = \begin{cases} n_i^l(t-1) + 1; & p_i^l \in \mathcal{P}^t, \\ n_i^l(t-1); & p_i^l \notin \mathcal{P}^t. \end{cases} \quad (7)$$

$$n_i(t) = \sum_{l=1}^{L} (n_i^l(t) \cdot |\mathcal{M}_i^l|) \text{ for } \forall i \in \mathcal{N}. \quad (8)$$

Next, we introduce the notation $\overline{q}_i(t)$ to record the average quality value (learned) for i until the t-th round. After $\mathcal{P}^t$ is determined, $\overline{q}_i(t)$ will be updated as follows:

$$\overline{q}_i(t) = \begin{cases} \dfrac{\overline{q}_i(t-1) n_i(t-1) + \sum_{j \in \mathcal{M}_i^l} q_{i,j}^t}{n_i(t-1) + |\mathcal{M}_i^l|}; & p_i^l \in \mathcal{P}^t, 1 \leq l \leq L, \\ \overline{q}_i(t-1); & p_i^l \notin \mathcal{P}^t, 1 \leq l \leq L. \end{cases} \quad (9)$$

In order to balance the relationship between exploitation and exploration, we extend the traditional UCB to propose the concept of UCB-based sensing quality. Concretely speaking, we use $\widehat{q}_i(t)$ to denote the UCB-based quality value:

$$\widehat{q}_i(t) = \overline{q}_i(t) + Q_{t,i}; \quad Q_{t,i} = \sqrt{\frac{(K+1)\ln(\sum_{i' \in \mathcal{N}} n_{i'}(t))}{n_i(t)}}. \quad (10)$$

In this paper, the values of $n_i^l(t)$, $n_i(t)$, $\overline{q}_i(t)$ and $\widehat{q}_i(t)$ make up the *worker profiles* in the platform. Next, we introduce the UCB-based quality function which considers the maximum weighted coverage problem in our MC system. When a task is covered by multiple workers, we let the maximum sensing quality value of these workers denote the final result of this task in a round. In order to determine the solution $\mathcal{P}^t$ in round t, we would refer to the UCB-based quality values revealed in the first t–1 rounds, i.e., the values of $\{\widehat{q}_i(t-1) | i \in \mathcal{N}\}$. More specifically, we let $u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)$ denote the UCB-based quality function for the solution $\mathcal{P}^t$ according to $\{\widehat{q}_i(t-1) | i \in \mathcal{N}\}$, that is,

$$u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t) = \sum_{j \in \mathcal{M}} w_j \cdot \max\{\widehat{q}_i(t-1) \cdot \mathbb{I}\{j \in \mathcal{M}_i^l, p_i^l \in \mathcal{P}^t\}\}. \quad (11)$$

Algorithm 1 The UWR Algorithm

Require: $\mathcal{N}, \mathcal{M}, \mathcal{P}=\{p_i^l|i\in\mathcal{N}, 1\leq l\leq L\}$, $\{w_j|j\in\mathcal{M}\}$, B, K
Ensure: $\mathcal{P}^t\subseteq\mathcal{P}$ for $\forall t\geq 1$, u_B and $\tau(B)$.

1: **Initialization**: $t=1$, recruit all workers, i.e., $\mathcal{P}^1=\{p_i^1|i\in\mathcal{N}\}$, and obtain the quality $q_{i,j}^1$ for $p_i^1\in\mathcal{P}^1$.
2: Let $u_B=u(\mathcal{P}^1)$, $B_t=B-\sum_{p_i^1\in\mathcal{P}^1}c_i^1$, $n_i(t)=|\mathcal{M}_i^1|$ and $\overline{q}_i(t)=(\sum_{j\in\mathcal{M}_i^1}q_{i,j}^t)/|\mathcal{M}_i^1|$ for $\forall i\in\mathcal{N}$;
3: **while** 1 **do**
4: $t\Leftarrow t+1$, $\mathcal{P}^t=\phi$;
5: **while** $|\mathcal{P}^t|<K$ **do**
6: Let $\mathcal{P}^{t'}=\{p_i^{l'}|$ for $\forall p_i^l\in\mathcal{P}^t\}$;
7: Get $p_i^l=\underset{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^{t'})}{\operatorname{argmax}}\frac{u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)}{c_{i'}^{l'}}$;
8: Add p_i^l into $\mathcal{P}^t$, i.e., $\mathcal{P}^t=\mathcal{P}^t+\{p_i^l\}$;
9: **if** $\sum_{p_i^l\in\mathcal{P}^t}c_i^l\geq B_{t-1}$ **then**
10: **return** Terminate and output u_B and $\tau(B)=t$;
11: Obtain the qualities $q_{i,j}^t$ for $\forall p_i^l\in\mathcal{P}^t$;
12: Update worker profiles: $n_i^l(t)$, $n_i(t)$, $\overline{q}_i(t)$ and $\widehat{q}_i(t)$;
13: $B_t=B_{t-1}-\sum_{p_i^l\in\mathcal{P}^t}c_i^l$, and $u_B=u_B+u(\mathcal{P}^t)$;

Based on this, we introduce the greedy recruitment strategy as follows. In the initialization period, the platform will select the first option of each worker (with the minimum cost) to explore the quality values, that is, $\mathcal{P}^1=\{p_i^1|i\in\mathcal{N}\}$. Then, $n_i^l(t)$, $n_i(t)$ and $\overline{q}_i(t)$ will be initialized. In any round $t>1$, the set $\mathcal{P}^t$ is first initialized to be empty. Then, when $|\mathcal{P}^t|<K$, we find the element in $\mathcal{P}\setminus\mathcal{P}^t$ which can increase the UCB-based quality function $u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)$ most quickly with unit cost. That is to say, we let the ratio of the marginal value of the function $u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)$ and cost be the selection criteria, which can be described as follows:

$$p_i^l=\underset{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^t)}{\operatorname{argmax}}\frac{u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)}{c_{i'}^{l'}}. \quad (12)$$

Note that at most one option of a worker can be selected in each round. Thus, if $p_i^l\in\mathcal{P}^t$, $p_i^{l'}$ for $1\leq l'\leq L, l'\neq l$ will not be considered in this round. After K workers are recruited in round t (i.e., $|\mathcal{P}^t|=K$), each worker i (here $p_i^l\in\mathcal{P}^t$) is required to perform all tasks in $\mathcal{M}_i^l$. In this paper, we use the concepts of worker and option exchangeably when no ambiguity exists. Then, the specific completion quality (i.e., $\{q_{i,j}^t|j\in\mathcal{M}_i^l\}$) is obtained by the platform. Based on this information, the platform will update the worker profiles, i.e., the values of $n_i^l(t)$, $n_i(t)$, $\overline{q}_i(t)$ and $\widehat{q}_i(t)$, according to Eq. (7), Eq. (8), and Eq. (9), respectively. At the same time, the total achieved weighted quality, i.e., the value of $u_B=u(\mathcal{P}^1)+\cdots+u(\mathcal{P}^t)$, is updated. Based on the remaining budget, the platform decides whether to continue the recruitment process.

3.2 Detailed Algorithm

According to the above solution, we propose an Unknown Worker Recruitment (UWR) algorithm, as shown in Alg. 1. In steps 1-2, the platform will select the first options from all workers with the minimum cost to initialize several parameters, such as $\overline{n}_i(t)$ and $\overline{q}_i(t)$. In steps 3-8, the platform selects K workers according to the UCB-based qualities and the proposed selection criteria, i.e., Eq. (12). To meet the constraint that at most one option of a worker can be selected in a round, we let $\mathcal{P}^{t'}$ denote the set of not satisfying options, in step 6. Then, the option with the largest ratio of the marginal UCB-based quality function value and cost is selected from the set $\mathcal{P}\setminus\mathcal{P}^{t'}$, in step 7. In steps 9-10, the platform decides whether to terminate the algorithm based on the remaining budget. If the remaining budget is enough, the recruited workers in this round will perform the corresponding tasks, and send the sensing results to the platform in step 11. The platform updates the worker profiles in step 12. The remaining budget and total achieved weighted quality are updated in step 13. Moreover, the computation complexity of the algorithm is dominated by step 7, which is denoted by $O(NMLK)$.

3.3 Performance Analysis

Assume that the platform knows the quality distributions of all workers, i.e., q_i for $\forall i\in\mathcal{N}$. In such a case, the worker recruitment problem is actually a special 0-1 knapsack problem in terms of all rounds, which is NP-hard [33]. There is no polynomial-time optimal algorithm for this problem. For simplicity of following descriptions, we let $\mathcal{P}^*\subset\mathcal{P}$ denote the optimal solution. Here, by recruiting the workers with high ratios of marginal weighted quality value and cost in each round, the platform can output an approximately optimal solution, which is denoted by $\mathcal{P}^\star\subset\mathcal{P}$. According to the existing work [33], [34], we get that the ratio of the approximately optimal solution and the optimal one, denoted as α, is greater than $1/2$. That is, $\alpha=\frac{\sum_{t\geq1}u_{[q_i]}(\mathcal{P}^\star)}{\sum_{t\geq1}u_{[q_i]}(\mathcal{P}^*)}\geq\frac{1}{2}$. Note that in this paper we always let * and * denote the corresponding identifications of the optimal and α-optimal workers, respectively. Here, $u_{[q_i]}(\mathcal{P}^t)$ is defined as follows:

$$u_{[q_i]}(\mathcal{P}^t)=\sum_{j\in\mathcal{M}}w_j\cdot\max\{q_i\cdot\mathbb{I}\{j\in\mathcal{M}_i^l, p_i^l\in\mathcal{P}^t\}\},$$

According to this, directly comparing our unknown worker recruitment results with the optimal solution is not fair. Therefore, we introduce the concept of *α-approximation regret* [19], [31] of an algorithm $\mathcal{A}$ under the budget B, that is, the difference for the total completion qualities achieved by the approximately optimal solution and our solution. More specifically, we let $R_\alpha^\mathcal{A}(B)$ denote the α-approximation regret and have the following equation:

$$\begin{aligned}R_\alpha^\mathcal{A}(B)&=\alpha\cdot\sum_{t\geq1}u_{[q_i]}(\mathcal{P}^*)-\mathbb{E}\Big[\sum_{t\geq1}u(\mathcal{P}^t)\Big]\\&\leq\sum_{t\geq1}u_{[q_i]}(\mathcal{P}^\star)-\mathbb{E}\Big[\sum_{t\geq1}u(\mathcal{P}^t)\Big]. \quad (13)\end{aligned}$$

At the same time, we define the smallest/largest possible difference of the quality values among non-α-optimal workers $\mathcal{P}'\neq\mathcal{P}^\star$, and the minimum/maximum recruitment cost values, i.e.,

$$\Delta_{min}=\frac{u_{[q_i]}(\mathcal{P}^\star)}{\sum_{c_i^l\in\mathcal{P}^\star}c_i^l}-\max_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}\frac{u_{[q_i]}(\mathcal{P}')}{\sum_{c_i^l\in\mathcal{P}'}c_i^l}; \quad (14)$$

$$\Delta_{max}=\frac{u_{[q_i]}(\mathcal{P}^\star)}{\sum_{c_i^l\in\mathcal{P}^\star}c_i^l}-\min_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}\frac{u_{[q_i]}(\mathcal{P}')}{\sum_{c_i^l\in\mathcal{P}'}c_i^l}; \quad (15)$$

$$\nabla_{max}=u_{[q_i]}(\mathcal{P}^\star)-\min_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}u_{[q_i]}(\mathcal{P}'); \quad (16)$$

$$0<c_{min}=\min\{c_i^l\};\ c_{max}=\max\{c_i^l\}\leq1. \quad (17)$$

Then, we introduce $C_i^l(t)$ as the counters after the initialization period, which is updated as follows. In each round, one of the following cases must happen: 1) the α-optimal set of workers is selected; 2) a non-α-optimal set of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3064324, IEEE Transactions on Mobile Computing

workers is recruited. In the former, $C_i^l(t)$ will not change; in the latter, we denote the non-α-optimal set of workers as $\mathcal{P}^t$. Then, there must exist one option $p_i^l \in \mathcal{P}^t$ such that $p_i^l = \mathrm{argmin}_{p_{i'}^{l'} \in \mathcal{P}^t} C_{i'}^{l'}(t-1)$, and we let $C_i^l(t) = C_i^l(t-1)+1$. Here, if there are multiple such options, we arbitrarily choose one. Since exactly one element in $C_i^l(t)$ is increased by 1 when a non-α-optimal set of workers is selected, the total number of non-α-optimal sets of workers is equal to the sum of the values in $\{C_i^l(t) | i \in \mathcal{N}, 1 \le l \le L\}$. Before analyzing the α-approximation regret of our algorithm, we first introduce two lemmas to analyze the bounds of the expected counters $\mathbb{E}[C_i^l(\tau(B))]$ and the stopping round $\tau(B)$, respectively, which are shown as follows.

Lemma 1. We have $\mathbb{E}[C_i^l(\tau(B))] \le \varphi_1 \ln \tau(B) + \varphi_2$ for any $p_i^l \in \mathcal{P}$, where φ_1 and φ_2 are two constants given below. More specifically, we have

$$\mathbb{E}[C_i^l(\tau(B))] \le \frac{4K^2(K+1)}{(\Delta_{min} c_{min})^2} \ln(NM\tau(B)) + 1 + \frac{K\pi^2}{3} \quad (18)$$

Proof: Let $I_i^l(t)$ denote the indicator that $C_i^l(t)$ is incremented at round t. Then, we have

$$\begin{aligned}C_i^l(\tau) &= \sum_{t=2}^{\tau} \mathbb{I}\{I_i^l(t)=1\} = \lambda + \sum_{t=2}^{\tau} \mathbb{I}\{I_i^l(t)=1, C_i^l(t) \ge \lambda\}\\ &= \lambda + \sum_{t=2}^{\tau} \mathbb{I}\{u_{[\widehat{q}_i(t-1)/c_i^l]}(\mathcal{P}^t) \ge u_{[\widehat{q}_i(t-1)/c_i^l]}(\mathcal{P}^\star), C_i^l(t) \ge \lambda\}\\ &\le \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{u_{[\widehat{q}_i(t)/c_i^l]}(\mathcal{P}^{t+1}) \ge u_{[\widehat{q}_i(t)/c_i^l]}(\mathcal{P}^\star), C_i^l(t) \ge \lambda\}\\ &= \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{\sum_{p_i^l \in \mathcal{P}^{t+1}} \xi_i^l(t+1)\frac{\widehat{q}_i(t)}{c_i^l} \ge \sum_{p_i^l \in \mathcal{P}^\star} \xi_i^l(\star)\frac{\widehat{q}_i(t)}{c_i^l}, C_i^l(t) \ge \lambda\}, \quad (19)\end{aligned}$$

where $\xi_i^l(t+1)$ means the product of the effective number of sensing tasks that worker $p_i^l \in \mathcal{P}^{t+1}$ can contribute and the total weight of these effective tasks, that is,

$$\xi_i^l(t+1) = \sum_{j \in \mathcal{M}_i^l} \mathbb{I}\Big\{p_i^l = \underset{p_{i'}^{l'} \in \mathcal{P}^{t+1}}{\mathrm{argmax}}\{\widehat{q}_{i',j}(t+1)\}\Big\} \cdot w_j.$$

Obviously, we have $\xi_i^l(t+1) \le \sum_{j \in \mathcal{M}_i^l} w_j \le 1$.

Then, we continue Eq. (19) and get

$$\begin{aligned}C_i^l(\tau) \le \lambda + \sum_{t=1}^{\tau} \mathbb{I}\Big\{ &\max_{\lambda \le n_{s(1)} \le \cdots \le n_{s(K)} \le t} \sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l} \cdot \widehat{q}_{s(i)}(t)\\ &\ge \min_{1 \le n_{s^\star(1)} \le \cdots \le n_{s^\star(K)} \le t} \sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} \cdot \widehat{q}_{s^\star(i)}(t)\Big\}\\ \le \lambda + \sum_{t=1}^{\tau} \sum_{n_{s(1)}=\lambda}^{t} \cdots \sum_{n_{s(K)}=\lambda}^{t} &\sum_{n_{s^\star(1)}=1}^{t} \cdots \sum_{n_{s^\star(K)}=1}^{t}\\ \mathbb{I}\Big\{\sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l} \cdot \widehat{q}_{s(i)}(t) &\ge \sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} \cdot \widehat{q}_{s^\star(i)}(t)\Big\}, \quad (20)\end{aligned}$$

where $s(i)$ and $s^\star(i)$ mean the i-th element in $\mathcal{P}^{t+1}$ and $\mathcal{P}^\star$, respectively. Here, $\widehat{q}_{s^\star(i)}(t) = \overline{q}_{s^\star(i)}(t) + Q_{t,s^\star(i)}$, in which $Q_{t,i} = \sqrt{\frac{(K+1)\ln(\sum_{i' \in \mathcal{N}} n_{i'}(t))}{n_i(t)}}$.

Next, we prove the probability of the following event:

$$\sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l}(\overline{q}_{s(i)}(t) + Q_{t,s(i)}) \ge \sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l}(\overline{q}_{s^\star(i)}(t) + Q_{t,s^\star(i)}),$$

which means that at least one of the following must hold:

$$\sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} \overline{q}_{s^\star(i)}(t) \le \sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l}(q_{s^\star(i)} - Q_{t,s^\star(i)}); \quad (21)$$

$$\sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l} \overline{q}_{s(i)}(t) \ge \sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l}(q_{s(i)} + Q_{t,s(i)}); \quad (22)$$

$$\sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} q_{s^\star(i)} < \sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l}(q_{s(i)} + 2Q_{t,s(i)}). \quad (23)$$

Now, we prove the upper bound for Eq. (21), and get

$$\begin{aligned}&\mathbb{P}\Big\{\sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} \overline{q}_{s^\star(i)}(t) \le \sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l}(q_{s^\star(i)} - Q_{t,s^\star(i)})\Big\}\\ &\le \sum_{i=1}^{K} \mathbb{P}\Big\{\overline{q}_{s^\star(i)}(t) \le q_{s^\star(i)} - Q_{t,s^\star(i)}\Big\}. \quad (24)\end{aligned}$$

After applying the Chernoff-Hoeffding bound introduced in the existing work [24], [35], we have

$$\begin{aligned}&\mathbb{P}\Big\{\overline{q}_{s^\star(i)}(t) \le q_{s^\star(i)} - Q_{t,s^\star(i)}\Big\}\\ &\le e^{-2n_{s^\star(i)}(t)((K+1)\ln(\sum_{i' \in \mathcal{N}} n_{s^\star(i')}(t))/n_{s^\star(i)}(t))}\\ &\le e^{-2(K+1)\ln(N|\mathcal{M}|_{min} t)} \le t^{-2(K+1)},\end{aligned}$$

where $|\mathcal{M}|_{min} = \min\{|\mathcal{M}_i^l| \text{ for } \forall i \in \mathcal{N} \text{ and } 1 \le l \le L\}$.

We continue Eq. (24) and get the upper bound, which is presented as follows:

$$\mathbb{P}\{\text{Eq. (24)}\} \le K \cdot t^{-2(K+1)}.$$

Similarly, we can derive the upper bound for Eq. (22), which is the same as that of the first case. Moreover, if both Eq. (21) and Eq. (22) are false, we can easily infer that Eq. (23) is true. Now, we pick λ such that Eq. (23) becomes impossible.

$$\begin{aligned}&\sum_{i=1}^{K} \frac{\xi_i^l(\star)}{c_{i^\star}^l} q_{s^\star(i)} - \sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l} q_{s(i)} - 2\sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l} Q_{t,s(i)}\\ &\ge \Delta_{\mathcal{P}^{t+1}} - K\frac{\sum_{j \in \mathcal{M}} w_j}{c_{min}} \sqrt{\frac{4(K+1)\ln(\sum_{i' \in \mathcal{N}} n_{s(i')}(t))}{n_{s(i)}(t)}}\\ &\ge \Delta_{\mathcal{P}^{t+1}} - \frac{K}{c_{min}} \sqrt{\frac{4(K+1)\ln(NM\tau(B))}{\lambda}} \ge 0. \quad (25)\end{aligned}$$

Therefore, Eq. (25) always holds, when λ satisfies:

$$\lambda \ge \frac{4(K+1)K^2}{(\Delta_{min} c_{min})^2} \ln(NM\tau(B)).$$

Then, we continue Eq. (20) and further have

$$\begin{aligned}C_i^l(\tau) &\le \lceil \frac{4(K+1)K^2}{(\Delta_{min} c_{min})^2} \ln(NM\tau(B)) \rceil + \sum_{t=1}^{\tau} 2Kt^{-2}\\ &\le \frac{4(K+1)K^2}{(\Delta_{min} c_{min})^2} \ln(NM\tau(B)) + 1 + \frac{K\pi^2}{3}. \quad (26)\end{aligned}$$

Lemma 1 holds. ■

Based on this, we get that the total number of non-α-optimal sets is at most $O(NLK^3 \ln \tau(B))$. Additionally, since the recruitment cost in each round is uncertain, the stopping round is indeterminate. We let $\tau(B)$ denote the stopping round of Alg. 1 under the budget B. Then, we introduce another lemma to prove the bound on $\tau(B)$.

Lemma 2. The stopping round of our algorithm $\tau(B)$ under the budget B is bounded as follows (here $c^\star = \sum_{p_i^l \in \mathcal{P}^\star} c_i^l$)

$$\frac{B}{c^\star} - \varphi_3 - 1 - \frac{\varphi_1 \varphi_3}{\varphi_2} \ln(\frac{2B}{c^\star} + \varphi_4) \le \tau(B) \le \frac{2B}{c^\star} + \varphi_4. \quad (27)$$

Proof: Due to the inequality $\ln \phi < \phi - 1$ for $\forall \phi > 0$, we first have the following expression:

$$\ln \tau(B) \le \frac{Kc_{min}}{2NL\varphi_1}\tau(B) + \ln(\frac{2NL\varphi_1}{Kc_{min}}) - 1, \quad (28)$$

where Kc_{min} indicates the minimum cost in each round.

Next, we derive the stopping round of the α-optimal algorithm: $\tau^\star(B) = \lfloor \frac{B}{c^\star} \rfloor$, in which $c^\star = \sum_{p_i^l \in \mathcal{P}^\star} c_i^l$. Then, we have $B/c^\star - 1 \le \tau^\star(B) \le B/c^\star$.

In order to derive the upper bound on $\tau(B)$, we have

$$\begin{aligned}\tau(B) &\le \tau^\star(B) + \tau\Big(\sum_{p_i^l \notin \mathcal{P}^\star} n_i^l(\tau(B)) c_{max}\Big)\\ &\le \tau^\star(B) + NL/(Kc_{min})\mathbb{E}[C_i^l(\tau(B))]. \quad (29)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3064324, IEEE Transactions on Mobile Computing

Before proving the lower bound on $\tau(B)$, we first let $0\leq B^\star\leq B$ denote the budget spent on the α-optimal options $\mathcal{P}^\star$, while $B^-=B-B^\star$ represents the budget spent on the non-α-optimal options. Then, we have

$$\tau(B)=\tau(B^\star+B^-)\geq\tau(B^\star)\geq\tau^\star(B^\star)$$
$$\geq\tau^\star(B-\textstyle\sum_{p_i^l\notin\mathcal{P}^\star}n_i^l(\tau(B))c_{max})\geq\frac{B-NL\mathbb{E}[C_i^l(\tau(B))]}{c^\star}-1. \quad (30)$$

According to Eq. (28) and Eq. (29), we further get

$$\tau(B)\leq\tau^\star(B)+\frac{NL}{Kc_{min}}\Big(\varphi_1(\frac{Kc_{min}}{2NL\varphi_1}\tau(B)+\ln(\frac{2NL\varphi_1}{Kc_{min}})-1)+\varphi_2\Big)$$
$$\leq\frac{B}{c^\star}+\frac{\tau(B)}{2}+\frac{NL}{Kc_{min}}(\varphi_1\ln(\frac{2NL\varphi_1}{Kc_{min}})-\varphi_1+\varphi_2)$$
$$\leq\frac{2B}{c^\star}+\frac{2NL}{Kc_{min}}(\varphi_1\ln(\frac{2NL\varphi_1}{Kc_{min}})-\varphi_1+\varphi_2)=\frac{2B}{c^\star}+\varphi_4.$$

By substituting the above results into Eq. 30, we get the lower bound on $\tau(B)$ as follows.

$$\tau(B)\geq B/c^\star-NL\varphi_2/c^\star-1-NL\varphi_1\ln(\tau(B))/c^\star$$
$$\geq B/c^\star-NL\varphi_2/c^\star-1-NL\varphi_1\ln(2B/c^\star+\varphi_4)/c^\star$$
$$=B/c^\star-\varphi_3-1-\ln(2B/c^\star+\varphi_4)\varphi_1\varphi_3/\varphi_2$$

Thus, the lemma holds. ■

Based on Lemmas 1 and 2, we prove the regret bound in the following theorem.

Theorem 1. The worst α-approximate regret of Alg. 1, denoted by $R_\alpha^{\mathcal{A}1}(B)$, is bounded as $O(NLK^3\ln(B+NLK^2\ln(MLN^2K^2)))$, that is,

$$R_\alpha^{\mathcal{A}1}(B)\leq(NL\nabla_{max}\varphi_1+u^\star\varphi_1\varphi_3/\varphi_2)(\ln(\frac{2B}{c^\star}+\varphi_4))+\varphi_5,$$

$$\text{where}\begin{cases}u^\star=u_{[q_i]}(\mathcal{P}^\star),\ \ c^\star=\sum_{p_i^l\in\mathcal{P}^\star}c_i^l\\ \varphi_1=\dfrac{4(K+1)K^2}{(\Delta_{min}c_{min})^2},\ \ \varphi_2=\ln(NM)\varphi_1+1+\dfrac{K\pi^2}{3}\\ \varphi_3=\dfrac{NL\varphi_2}{c^\star},\ \varphi_4=\dfrac{2NL}{Kc_{min}}(\varphi_1\ln(\dfrac{2NL\varphi_1}{Kc_{min}})-\varphi_1+\varphi_2)\\ \varphi_5=NL\nabla_{max}+u^\star(1/c^\star+\varphi_3+1)\end{cases}$$

Proof: According to Lemmas 1 and 2, we get that the α-approximate regret of our algorithm satisfies

$$R_\alpha^{\mathcal{A}1}(B)\leq\textstyle\sum_{t=1}^{\tau^\star(B)}u_{[q_i]}(\mathcal{P}^\star)-\mathbb{E}[\sum_{t=1}^{\tau(B)}u_{[q_i]}(\mathcal{P}^t)]$$
$$\leq\frac{(B+1)u^\star}{c^\star}-\tau(B)u^\star+\tau(B)u^\star-\mathbb{E}[\textstyle\sum_{t=1}^{\tau(B)}u_{[q_i]}(\mathcal{P}^t)]$$
$$\leq u^\star(\frac{B+1}{c^\star}-\tau(B))+\textstyle\sum_{i\in\mathcal{N}}\sum_{l=1}^{L}C_i^l(\tau(B))\nabla_{max}$$
$$\leq u^\star(\frac{B+1}{c^\star}-(\frac{B}{c^\star}-\varphi_3-1-\frac{\varphi_1\varphi_3}{\varphi_2}\ln(\frac{2B}{c^\star}+\varphi_4))$$
$$+NL\nabla_{max}(\varphi_1\ln(\frac{2B}{c^\star}+\varphi_4)+\varphi_2)$$
$$=(NL\nabla_{max}\varphi_1+u^\star\varphi_1\varphi_3/\varphi_2)(\ln(\frac{2B}{c^\star}+\varphi_4))+\varphi_5$$
$$=O\Big(NLK^3\ln(B+NLK^2\ln(MLN^2K^2))\Big).$$

φ_1, φ_2, φ_3 and φ_4 are constants. Theorem 1 holds. ■

Algorithm 2 The EUWR Algorithm

Require: $\mathcal{N}$, $\mathcal{M}$, $\mathcal{P}=\{p_i^l=\langle\mathcal{M}_i\rangle\}$, $\{w_j|j\in\mathcal{M}\}$, B, K, $f(\cdot)$
Ensure: $\mathcal{P}^t\subseteq\mathcal{P}$ for $\forall t\geq1$, u_B and $\tau(B)$.
Initialization: $t=1$, let $\mathcal{P}^1=\{p_i^1|i\in\mathcal{N}\}$ and obtain the quality $q_{i,j}^1$ and cost parameter ε_i^1 for $p_i^1\in\mathcal{P}^1$.
Let $u_B=u(\mathcal{P}^1)$, $B_t=B-\sum_{p_i^1\in\mathcal{P}^1}\varepsilon_i^1f(|\mathcal{M}_i^1|)$, $n_i(t)=1$, $\overline{q}_i(t)=(\sum_{j\in\mathcal{M}_i^1}q_{i,j}^t)/|\mathcal{M}_i^1|$ and $\overline{\varepsilon}_i(t)=\varepsilon_i^1$ for $\forall i\in\mathcal{N}$;
while 1 **do**
  $t\Leftarrow t+1$, $\mathcal{P}^t=\phi$;
  while $|\mathcal{P}^t|<K$ **do**
    Let $\mathcal{P}^{t'}=\{p_i^{l'}|$ for $\forall p_i^l\in\mathcal{P}^t\}$;
    $p_i^l=\operatorname{argmax}_{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^{t'})}u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t)$;
    Add p_i^l into $\mathcal{P}^t$, i.e., $\mathcal{P}^t=\mathcal{P}^t+\{p_i^l\}$;
  Each recruited worker i where $p_i^l\in\mathcal{P}^t$ obtains ε_i^t;
  if $\sum_{p_i^l\in\mathcal{P}^t}\varepsilon_i^tf(|\mathcal{M}_i^l|)\geq B_{t-1}$ **then**
    return Terminate and output u_B and $\tau(B)=t$;
  Perform tasks and obtain the qualities $q_{i,j}^t$ for $\forall p_i^l\in\mathcal{P}^t$;
  Update $n_i^l(t)$, $n_i(t)$, $m_i(t)$, $\overline{q}_i(t)$, $\overline{\varepsilon}_i(t)$, and $\widehat{r}_i^{\,l}(t)$;
  $B_t=B_{t-1}-\sum_{p_i^l\in\mathcal{P}^t}\varepsilon_i^tf(|\mathcal{M}_i^l|)$, and $u_B=u_B+u(\mathcal{P}^t)$;

4 Unknown Quality and Cost

In this section, we consider a more general case where the cost of each worker is unknown too. We first introduce the extended problem and further present the basic solution. Then, we design an Extended Unknown Worker Recruitment (EUWR) algorithm and analyze the corresponding performance guarantee.

4.1 Problem Description

We consider an extended case where both workers' sensing qualities and costs are unknown a priori. Note that the cost of p_i^l is determined by $c_i^l=\varepsilon_if(|\mathcal{M}_i^l|)$ where $f(\cdot)$ is given in our MC system. The unknown cost here means that the cost parameter ε_i is unknown. In each round, when an option of a worker is selected, the worker would estimate the cost parameter according to the current state including battery, network, environment factors, etc. We use ε_i^t to denote the cost parameter in round t. Here, we let $0<\varepsilon_{min}\leq\varepsilon_i^t\leq1$. Note that the values of $\{\varepsilon_i^1,\cdots,\varepsilon_i^t\}$ follow an independent and identically distribution with the unknown expectation ε_i. After receiving the cost parameter ε_i^t in round t, the platform calculates the recruitment cost for $\forall p_i^l\in\mathcal{P}^t$ based on the formula $c_i^l=\varepsilon_i^tf(|\mathcal{M}_i^l|)$. Here, all values of c_i^l will be normalized into $(0,1]$. When the remaining budget cannot cover the total cost of $\mathcal{P}^t$ in round t, the recruitment algorithm will terminate; else, the recruited workers perform the corresponding tasks and return the sensing quality to the platform. The platform then updates the parameters in the worker profiles. In the extended problem, the platform needs to learn the quality q_i and parameter ε_i simultaneously, and meanwhile to maximize the total weighted qualities of all tasks under a given budget. So it is more challenging to design a suitable recruitment strategy.

4.2 Basic Solution

Like before, we still let $n_i^l(t)$ denote the number of p_i^l being selected. However, when p_i^l is selected in a round, the parameter ε_i is actually learned only one time. Thus, we define another notation to record the total number of cost parameter ε_i being learned, denoted by $m_i(t)$, i.e., $m_i(t)=\sum_{l=1}^{L}n_i^l(t)$. Then, the average cost parameter up to round t, denoted by $\overline{\varepsilon}_i(t)$, is calculated as follows:

$$\overline{\varepsilon}_i(t)=\begin{cases}\dfrac{\overline{\varepsilon}_i(t-1)\cdot m_i(t-1)+\varepsilon_i^t}{m_i(t-1)+1}; & \text{for }\forall1\leq l\leq L,p_i^l\in\mathcal{P}^t,\\ \overline{\varepsilon}_i(t-1); & \text{for }\forall1\leq l\leq L,p_i^l\notin\mathcal{P}^t.\end{cases} \quad (31)$$

Similar to the UCB-based expression $\widehat{q}_i(t)=\overline{q}_i(t)+Q_{t,i}$, we also define another UCB-based cost value, which is denoted by $C_{t,i}=\sqrt{(K+1)\ln t/m_i(t)}$.

Before determining $\mathcal{P}^t$, we can use $\overline{\varepsilon}_i(t-1)f(|\mathcal{M}_i^l|)$ to denote the recruitment cost of p_i^l, and the values of $\overline{\varepsilon}_i(t-1)f(|\mathcal{M}_i^l|)$ are finally normalized to $(0,1]$. That is, we have $0 < c_{min} \leq \overline{\varepsilon}_i(t-1)f(|\mathcal{M}_i^l|) \leq 1$. Then, we design another selection criteria, denoted by $\widehat{r}_i(t)$, which takes the obtained quality and cost values into consideration simultaneously. More specifically, we define

$$\widehat{r}_i^{\,l}(t) = f_i^l \cdot \frac{\overline{q}_i(t-1)}{\overline{\varepsilon}_i(t-1)} + f_{max} \cdot \frac{\varepsilon_{min} \cdot Q_{t,i} + C_{t,i}}{\varepsilon_{min}^2}, \quad (32)$$

where $f_i^l = |\mathcal{M}_i^l|/f(|\mathcal{M}_i^l|)$ and $f_{max} = \max_{\mathcal{M}_i^l} f_i^l$.

Next, the platform focuses on determining $\mathcal{P}^t$ in round t by referring to the value of $\widehat{r}_i^{\,l}(t)$. More specifically, in the initialization period, the platform will select the first options of all workers to explore the quality and cost values, that is, $\mathcal{P}^1 = \{p_i^1 | i \in \mathcal{N}\}$. Then, $\overline{q}_i(t)$ and $\overline{\varepsilon}_i(t)$ will be initialized. In any round $t > 1$, the set $\mathcal{P}^t$ is first initialized to be empty. When $|\mathcal{P}^t| < K$, we find the element in $\mathcal{P} \backslash \mathcal{P}^t$ which can increase the function $u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t)$ most quickly, that is,

$$p_i^l = \underset{p_{i'}^{l'} \in (\mathcal{P} \backslash \mathcal{P}^t)}{\operatorname{argmax}} \; u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t \cup \{p_{i'}^{l'}\}) - u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t). \quad (33)$$

Here, $u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t)$ means the total UCB-based ratios of quality and cost for $\mathcal{P}^t$, i.e.,

$$u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t) = \sum_{j \in \mathcal{M}} w_j \cdot \max\{\widehat{r}_i^{\,l}(t-1) \cdot \mathbb{I}\{j \in \mathcal{M}_i^l, p_i^l \in \mathcal{P}^t\}\}. \quad (34)$$

Also, the constraint that at most one option from a worker can be selected in each round still exists. Based on the basic solution, we propose an Extended Unknown Worker Recruitment (EUWR) algorithm, as shown in Alg. 2. The main procedures are similar to that of Alg. 1. The key difference is that the selection criteria $\widehat{q}_i(t-1)$ is replaced by $\widehat{r}_i(t-1)$ in the extended algorithm. Also, each recruited worker in round t would first estimate his cost parameter ε_i^t to determine his recruitment cost in step 9. Note that steps 6-7 cooperate to remove the constraint that at most one option from a worker can be selected in each round. In addition, the computational overhead of Alg. 2 is still $O(NMLK)$.

4.3 Performance Analysis

Before analyzing the regret guarantee of Alg. 2, we define the smallest/largest possible difference of the ratio values among non-α-optimal set of workers $\mathcal{P}' \neq \mathcal{P}^\star$, that is,

$$\Delta r_{min} = u_{[r_i^l]}(\mathcal{P}^\star) - \max_{\{\mathcal{P}' \neq \mathcal{P}^\star\}} u_{[r_i^l]}(\mathcal{P}'), \quad (35)$$

$$\Delta r_{max} = u_{[r_i^l]}(\mathcal{P}^\star) - \min_{\{\mathcal{P}', \mathcal{P}' \neq \mathcal{P}^\star\}} u_{[r_i^l]}(\mathcal{P}'). \quad (36)$$

where $r_i^l = |\mathcal{M}_i^l| q_i / (\varepsilon_i f(|\mathcal{M}_i^l|))$. The calculation of $u_{[r_i^l]}(\mathcal{P}^t)$ is similar to Eq. (34) in which $\widehat{r}_i^{\,l}(t-1)$ is replaced by r_i^l. Then, we have the following theorem.

Theorem 2. The worst α-approximate regret of Alg. 2, denoted by $R_\alpha^{\mathcal{A}2}(B)$, is bounded as $O(NLK^3 \ln(NMB + N^2K^2ML\ln(N^2K^2ML)))$. More precisely, we have:

$$R_\alpha^{\mathcal{A}2}(B) \leq (NL\varphi_6)(\tfrac{u^\star}{c^\star} + \Delta r_{max}) \ln(NM(\tfrac{2B}{c^\star} + 2\varphi_7)) + \varphi_8,$$

$$\text{where} \begin{cases} \varphi_6 = (K+1)\Big(\dfrac{2Kf_{max}(\varepsilon_{min}+1)}{\Delta r_{min}\varepsilon_{min}^2}\Big)^2 \\ \varphi_7 = \dfrac{NL}{Kc_{min}}(\varphi_6 \ln(\dfrac{2N^2ML\varphi_6}{Kc_{min}}) - \varphi_6 + 1 + \dfrac{2K\pi^2}{3}) \\ \varphi_8 = \dfrac{u^\star(1 + c^\star + NL(1 + \frac{2K\pi^2}{3}))}{c^\star} \\ \qquad + NL\Delta r_{max}(1 + \dfrac{2K\pi^2}{3}) \end{cases}$$

Proof: The proof is similar to that of Theorem 1. When the quality and cost distributions are known in advance, we can get the α-optimal solution $\mathcal{P}^\star$ by selecting the workers with high ratios of marginal weighted quality and cost in each round. We also let $C_i^l(t)$, $\tau(B)$ and $I_i^l(t)$ denote the counters, stopping round, and the indicator. In order to prove that $C_i^l(t)$ is bounded, we get

$$\begin{aligned} C_i^l(\tau) \leq\; & \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{u_{[\widehat{r}_i^{\,l}(t)]}(\mathcal{P}^{t+1}) \geq u_{[\widehat{r}_i^{\,l}(t)]}(\mathcal{P}^\star), C_i^l(t) \geq \lambda\} \\ =\; & \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{\sum_{p_i^l \in \mathcal{P}^{t+1}} \xi_i^l(t+1)\widehat{r}_i^{\,l}(t) \geq \sum_{p_i^l \in \mathcal{P}^\star} \xi_i^l(\star)\widehat{r}_i^{\,l}(t), C_i^l(t) \geq \lambda\} \\ \leq\; & \lambda + \sum_{t=1}^{\tau} \sum_{n_{s(1)}=\lambda}^{t} \cdots \sum_{n_{s(K)}=\lambda}^{t} \sum_{n_{s^\star(1)}=1}^{t} \cdots \sum_{n_{s^\star(K)}=1}^{t} \\ & \mathbb{I}\Big\{ \textstyle\sum_{i=1}^K \xi_i^l(t+1) \cdot \widehat{r}_{s(i)}^{\,l}(t) \geq \sum_{i=1}^K \xi_i^l(\star) \cdot \widehat{r}_{s^\star(i)}^{\,l}(t) \Big\}. \end{aligned} \quad (37)$$

where $\xi_i^l(t+1) \leq \sum_{j \in \mathcal{M}_i^l} w_j \leq 1$ and $\xi_i^l(\star) \leq 1$ have the same meanings as before. Here, we use r_i^l to denote the ratio of quality and cost in which all parameters are known, i.e., $r_i^l = f_i^l q_i/\varepsilon_i$ where $f_i^l = |\mathcal{M}_i^l|/f(|\mathcal{M}_i^l|)$. Also, the notation $\overline{r}_i^{\,l}(t) = f_i^l \overline{q}_i(t-1)/\overline{\varepsilon}_i(t-1)$ means the average ratio of quality and cost for p_i^l up to round t. After letting $\theta_{t,i} = f_{max}(\varepsilon_{min}Q_{t,i} + C_{t,i})/\varepsilon_{min}^2$, we get that at least one of the three cases, which are similar to Eq. (21), Eq. (22) and Eq. (23), must hold. Now, we focus on the probability of the following case:

$$\begin{aligned} & \mathbb{P}\Big\{\overline{r}_{s^\star(i)}^{\,l}(t) \leq r_{s^\star(i)}^l - \theta_{t,s^\star(i)}\Big\} \\ &= \mathbb{P}\Big\{f_{s^\star(i)}^l \cdot \tfrac{\overline{q}_{s^\star(i)}(t)}{\overline{\varepsilon}_{s^\star(i)}(t)} \leq f_{s^\star(i)}^l \cdot \tfrac{q_{s^\star(i)}}{\varepsilon_{s^\star(i)}} - \theta_{t,s^\star(i)}\Big\}. \end{aligned} \quad (38)$$

Actually, the event in Eq. (38) holds only when at least one of the following events must be true:

$$\overline{q}_{s^\star(i)}(t) \leq q_{s^\star(i)} - Q_{t,s^\star(i)}; \quad (39)$$

$$\overline{\varepsilon}_{s^\star(i)}(t) \geq \varepsilon_{s^\star(i)} + C_{t,s^\star(i)}. \quad (40)$$

We can prove this claim by the counter-evidence. That is, if both Eq. (39) and Eq. (40) are false, we have

$$\begin{aligned} & f_{s^\star(i)}^l \cdot \Big(\tfrac{q_{s^\star(i)}}{\varepsilon_{s^\star(i)}} - \tfrac{\overline{q}_{s^\star(i)}(t)}{\overline{\varepsilon}_{s^\star(i)}(t)}\Big) \\ &= f_{s^\star(i)}^l \cdot \tfrac{(q_{s^\star(i)} - \overline{q}_{s^\star(i)}(t))\varepsilon_{s^\star(i)} - q_{s^\star(i)} \cdot (\varepsilon_{s^\star(i)} - \overline{\varepsilon}_{s^\star(i)}(t))}{\varepsilon_{s^\star(i)} \cdot \overline{\varepsilon}_{s^\star(i)}(t)} \\ &< f_{s^\star(i)}^l \cdot \big(\tfrac{Q_{t,s^\star(i)}}{\overline{\varepsilon}_{s^\star(i)}(t)} + \tfrac{q_{s^\star(i)}C_{t,s^\star(i)}}{\varepsilon_{s^\star(i)}\overline{\varepsilon}_{s^\star(i)}(t)}\big) \\ &\leq f_{max} \cdot \tfrac{\varepsilon_{min}Q_{t,s^\star(i)} + C_{t,s^\star(i)}}{\varepsilon_{min}^2} = \theta_{t,s^\star(i)}. \end{aligned}$$

According to the previous proof, we know $\mathbb{P}\{\text{Eq. } (39)\} \leq t^{-2(K+1)}$ and $\mathbb{P}\{\text{Eq. } (40)\} \leq t^{-2(K+1)}$. We continue Eq. (38) and have $\mathbb{P}\{\text{Eq. } (38)\} \leq 2t^{-2(K+1)}$. Next, we analyze

$$\textstyle\sum_{i=1}^K \xi_i^l(\star) r_{s^\star(i)}^l < \sum_{i=1}^K \xi_i^l(t+1)(r_{s(i)}^l + 2\theta_{t,s(i)}).$$

We choose λ to make the above inequality false, i.e.,

$$\begin{aligned} & \sum_{i=1}^K \xi_i^l(\star) r_{s^\star(i)}^l - \sum_{i=1}^K \xi_i^l(t+1) r_{s(i)}^l - 2\sum_{i=1}^K \xi_i^l(t+1)\theta_{t,s(i)} \\ &= \Delta r - 2\sum_{i=1}^K \xi_i^l(t+1) f_{max} \tfrac{\varepsilon_{min}Q_{t,s(i)} + C_{t,s(i)}}{\varepsilon_{min}^2} \\ &\geq \Delta r - 2Kf_{max} \frac{\varepsilon_{min}\sqrt{\frac{(K+1)\ln(NM\tau(B))}{\lambda|\mathcal{M}|_{min}}} + \sqrt{\frac{(K+1)\ln\tau(B)}{\lambda}}}{\varepsilon_{min}^2} \\ &\geq \Delta r - 2Kf_{max} \frac{(\varepsilon_{min}+1)\sqrt{\frac{(K+1)\ln(NM\tau(B))}{\lambda}}}{\varepsilon_{min}^2} \geq 0. \end{aligned}$$

Algorithm 3 The FAUWR Algorithm

Require: $\mathcal{N}$, $\mathcal{M}$, $\mathcal{P}=\{p_i^l|i\in\mathcal{N},1\le l\le L\}$, $\{w_j|j\in\mathcal{M}\}$, $\{\eta_i|i\in\mathcal{N}\}$, B, K, ϱ

Ensure: $\mathcal{P}^t\subseteq\mathcal{P}$ for $\forall t\ge 1$, u_B and $\tau(B)$.

Initialization: $t=1$, select the first option of all workers, i.e., $\mathcal{P}^1=\{p_i^1|i\in\mathcal{N}\}$, and obtain the quality $q_{i,j}^1$ for $p_i^1\in\mathcal{P}^1$, and set the virtual length as $V_i(t)=0$ for $i\in\mathcal{N}$.
Let $u_B=u(\mathcal{P}^1)$, $B_t=B-\sum_{p_i^1\in\mathcal{P}^1}c_i^1$, $n_i(t)=|\mathcal{M}_i^1|$ and $\overline{q}_i(t)=(\sum_{j\in\mathcal{M}_i^1}q_{i,j}^t)/|\mathcal{M}_i^1|$ for $\forall i\in\mathcal{N}$;
while 1 **do**
  $t\Leftarrow t+1$, $\mathcal{P}^t=\phi$;
  while $|\mathcal{P}^t|<K$ **do**
    Let $\mathcal{P}^{t'}=\{p_i^{l'}|$ for $\forall p_i^l\in\mathcal{P}^t\}$;
    Get $p_i^l=\underset{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^{t'})}{\operatorname{argmax}}\Big\{\frac{u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{q}_i(t-1)]}(\mathcal{P}^t)}{c_{i'}^{l'}}$ $+\varrho\cdot V_i(t-1)\Big\}$ based on Eq. (43);
    Add p_i^l into $\mathcal{P}^t$, i.e., $\mathcal{P}^t=\mathcal{P}^t+\{p_i^l\}$;
  if $\sum_{p_i^l\in\mathcal{P}^t}c_i^l\ge B_{t-1}$ **then**
    return Terminate and output u_B and $\tau(B)=t$;
  Obtain the qualities $q_{i,j}^t$ for $\forall p_i^l\in\mathcal{P}^t$;
  Update the worker profiles: $n_i^l(t)$, $n_i(t)$, $\overline{q}_i(t)$, $\widehat{q}_i(t)$, and $V_i(t)$ based on Eq. (42);
  $B_t=B_{t-1}-\sum_{p_i^l\in\mathcal{P}^t}c_i^l$, and $u_B=u_B+u(\mathcal{P}^t)$;

Thus, we choose

$$\lambda\ge(K+1)\ln(NM\tau(B))\Big(\frac{2Kf_{max}(\varepsilon_{min}+1)}{\Delta r_{min}\varepsilon_{min}^2}\Big)^2,$$

such that Eq. (41) is impossible. Let's continue Eq. (37):

$$\begin{aligned}C_i^l(\tau)&\le(K+1)\ln(NM\tau(B))\Big(\frac{2Kf_{max}(\varepsilon_{min}+1)}{\Delta r_{min}\varepsilon_{min}^2}\Big)^2\\&\quad+1+\textstyle\sum_{t=1}^{\tau}t^{2K}2K(2t^{-2(K+1)})\\&\le(K+1)\ln(NM\tau(B))\Big(\frac{2Kf_{max}(\varepsilon_{min}+1)}{\Delta r_{min}\varepsilon_{min}^2}\Big)^2+1+\frac{2K\pi^2}{3}\\&=\varphi_6\ln\tau(B)+\varphi_6\ln(NM)+1+2K\pi^2/3.\end{aligned}$$

Similarly, we have $\ln\tau(B)\le Kc_{min}\tau(B)/(2NL\varphi_6)+\ln(2NL\varphi_6/(Kc_{min})-1$. Then, we get the bound of $\tau(B)$,

$$\begin{aligned}\tau(B)&\le\tau^\star(B)+\tfrac{NL}{Kc_{min}}\mathbb{E}[C_i^l(\tau(B))]\\&\le B/c^\star+\tau(B)/2+\varphi_7\le 2B/c^\star+2\varphi_7,\end{aligned}$$

and get the lower bound as follows,

$$\begin{aligned}\tau(B)&\ge(B-NL\mathbb{E}[C_i^l(\tau(B))])/c^\star-1\\&\ge\frac{B}{c^\star}-1-\frac{NL(1+\frac{2K\pi^2}{3})}{c^\star}-\frac{NL\varphi_6\ln(NM(\frac{2B}{c^\star}+2\varphi_7))}{c^\star}.\end{aligned}$$

Finally, we prove the expected regret of Alg. 2 as follows:

$$\begin{aligned}R_\alpha^{\mathcal{A}2}(B)&\le\textstyle\sum_{t=1}^{\tau^\star(B)}u_{[r_i^l]}(\mathcal{P}^\star)-\mathbb{E}[\sum_{t=1}^{\tau(B)}u_{[r_i^l]}(\mathcal{P}^t)]\\&\le u^\star(\tfrac{B+1}{c^\star}-\tau(B))+\textstyle\sum_{i\in\mathcal{N}}\sum_{l=1}^{L}C_i^l(\tau(B))\Delta r_{max}\\&\le(NL\varphi_6)(\tfrac{u^\star}{c^\star}+\Delta r_{max})\ln(NM(\tfrac{2B}{c^\star}+2\varphi_7))+\varphi_8\\&=O\Big(NLK^3\ln(NMB+N^2K^2ML\ln(N^2K^2ML))\Big).\end{aligned}$$

Theorem 2 holds. ■

5 Fairness Constraint of Workers

In this section, we consider an extended case where the fairness constraint of workers is involved. The term "fairness" means that the crowdsensing platform must guarantee a minimum selection fraction for each registered worker, so that the MC system can avoid the scenario that some workers are over-recruited but some others may be under-recruited. **This is because the under-recruited workers may leave the crowdsensing system forever, while fewer workers will certainly harm the sensing results for other requesters. That is, the fairness constraint can potentially guarantee the overall and long-term sensing performance for all requesters.** We first introduce the extended optimization problem and then propose the detailed solution.

TABLE 2: Evaluation Settings

parameter name	default	range
the budget B	$3*10^3$	$5*10^2$–10^4
the number of tasks M	300	100–500
the number of workers N	50	50–100
the parameter K	$N/3$	$N/6$–$3N/5$
the cardinality of $\mathcal{M}_i^l$		5–15
the parameter ϱ		0.1–10

Fig. 3: Exhibition of the real-world dataset.

5.1 Extended Problem

First, we introduce a parameter η_i to denote the required minimum fraction of rounds in which the worker i would be recruited. Based on the definition of $n_i(t)$, i.e., the number of the worker i being selected up to round t, we can formalize the fairness constraint as follows:

$$\liminf_{B\to\infty}\mathbb{E}[n_i(\tau(B))]\ge\tau(B)\cdot\eta_i,\ \text{for}\ \forall i\in\mathcal{N}. \tag{41}$$

where $\tau(B)$ means the total rounds under the budget B, $n_i(\tau(B))$ indicates the total number of worker i being selected under the budget B.

Here, we suppose that there exists at least one worker recruitment solution so that the fairness constraint, i.e., Eq. (41), can be satisfied. That is to say, the MC system would compute the parameter η_i for $i\in\mathcal{N}$ in advance. After considering the fairness during the process of unknown worker recruitment, the extended problem becomes more complicated. Note that the extended problem has the same objective as the original optimization problem, that is, maximizing the total weighted completion qualities of all tasks under the budget. The difference lies in that the fairness constraint of workers is taken into account. Next, we will introduce the detailed solution to the extended problem.

5.2 Extended Solution

In order to solve this extended problem, we introduce a concept of virtual queue [36] to handle the fairness constraint. More specifically, we first use V_i to denote the virtual queue for the worker i, and then let $V_i(t)$ denote the queue length of V_i in round t. Note that we initialize $V_i(0)=0$ for $\forall i\in\mathcal{N}$. Furthermore, $V_i(t)$ is updated as follows:

$$V_i(t)=\max\Big\{0,\ V_i(t-1)+\eta_i-\mathbb{I}\{i_{t-1}=i\}\Big\}, \tag{42}$$

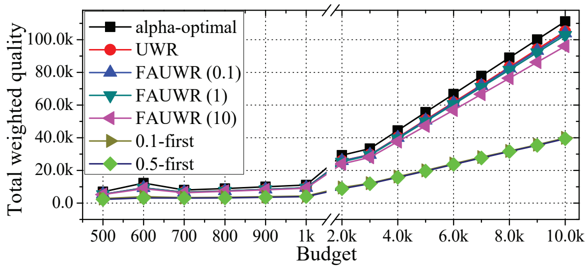

Fig. 4: UWR and FAUWR: total qualities vs. budget

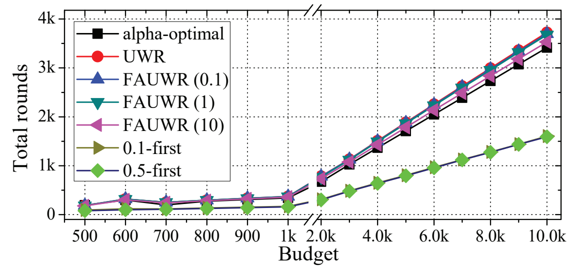

Fig. 5: UWR and FAUWR: total rounds vs. budget

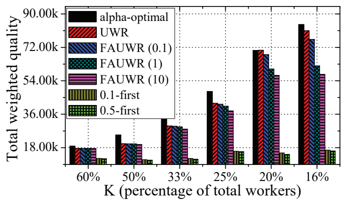

Fig. 6: Qualities vs. K

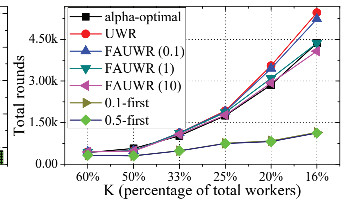

Fig. 7: Rounds vs. K

Fig. 8: Qualities vs. N

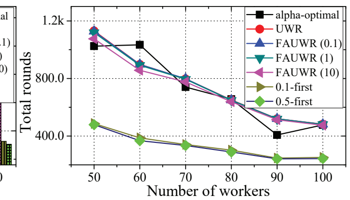

Fig. 9: Rounds vs. N

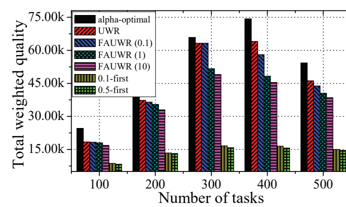

Fig. 10: Qualities vs. M

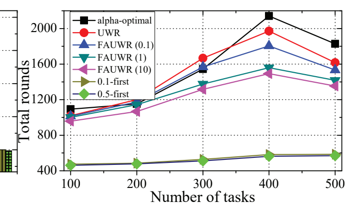

Fig. 11: Rounds vs. M

where $\mathbb{I}\{\text{true}\}=1$ and $\mathbb{I}\{\text{false}\}=0$. Moreover, i_t represents the index for the unknown worker which is selected in round t. Here, for each option p_i^l $(1\le l\le L)$ of the worker i, if p_i^l is selected in the round t, it indicates $i_t=i$ in Eq. (42). Thus, the index $\mathbb{I}\{i_{t-1}=i\}=1$ is obtained.

Based on the virtual queue technique, we design a Fairness-Aware Unknown Worker Recruitment algorithm, called FAUWR. The FAUWR algorithm has the same structure and procedures as the UWR algorithm. The difference lies in the computation of the option index at the beginning of each round. Concretely speaking, the computation of option index p_i^l will involve the virtual queue length to address the fairness constraint. That is, at the beginning of each round t, the platform would choose the option according to the following equation:

$$p_i^l = \underset{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^t)}{\operatorname{argmax}} \Big\{ \frac{u_{[\hat{q}_i(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\hat{q}_i(t-1)]}(\mathcal{P}^t)}{c_{i'}^{l'}} + \varrho\cdot V_i(t)\Big\}. \quad (43)$$

Here, $\varrho>0$ denotes the controlling parameter, which is used to manage the balance between the total weighted quality and the virtual queue lengths. Also, each option p_i^l $(1\le l\le L)$ of a same worker $i\in\mathcal{N}$ shares the same virtual queue lengths, i.e., $V_i(t)$. Actually, when the parameter ϱ is set as 0, we get that the option index computation here is the same as that of the UWR algorithm.

Then, we present the FAUWR algorithm in detail, as shown in Alg. 3. Same as the original settings, we consider that the recruitment cost can be acquired in advance. Therefore, the extended algorithm has the same structure as Alg. 1. In fact, the FAUWR algorithm can also deal with the case where the recruitment cost is unknown a priori. The only difference is the index computation for each option at the beginning of each round. Moreover, the computation complexity of Alg. 3 is denoted by $O(NMLK)$. Note that as long as the number of total rounds (i.e., $\tau(B)$) is large enough, the practical selection fraction for each worker will be higher than that of the required values in Alg. 3. We will verify the fairness performance of the FAUWR algorithm through lots of experimental simulations based on the real-world datasets in Section 6.

6 Performance Evaluation

In this section, we evaluate the performance of the proposed algorithms with extensive simulations. We first introduce the evaluation methodology which mainly includes the detailed simulation settings (such as the real-world datasets, parameter values, etc.) and two compared algorithms. Then, we present and analyze the simulation results in detail.

6.1 Evaluation Methodology

Simulation Settings: We adopt the widely-used trace [37] in our simulations. The trace consists of the GPS coordinates of approximately 320 taxi cabs collected over 30 days in Rome, Italy. The trace on a day (i.e., 2014-02-01) is shown in Fig. 3. We first select M locations from the trace as the task locations, in which M is produced from $[100,600]$. Then, we choose N vehicles from the trace as workers, where N is selected from $[50,100]$. Here, we exclude those vehicles that visit the selected locations with low frequency in our simulations. The default values are $M=300$ and $N=50$. Next, we determine the subset of tasks that a worker can perform, and we let the number of tasks in a subset (i.e., $|\mathcal{M}_i^l|$) be randomly selected from $[5,15]$. Also, we change the budget from $[500,10^4]$ and let $B=3000$ in default.

Now, we focus on determining two parameters: the expected quality q_i and the expected cost parameter ε_i. First, we generate the sensing area for each task. For each task, we use a geographic region with radius $200m$ within its location to denote the sensing area. The workers within this region can perform this task. Based on this, we use the frequency value of a worker i visiting these areas to denote the expected mean q_i, in which q_i is normalized into $(0,1]$. Then, we generate ε_i randomly from $(0,1)$. We directly use the function $f(x)=x$ to determine the cost of each subset,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3064324, IEEE Transactions on Mobile Computing

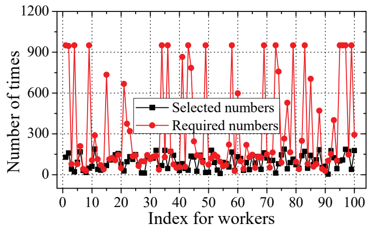

Fig. 12: $\varrho = 0.1$

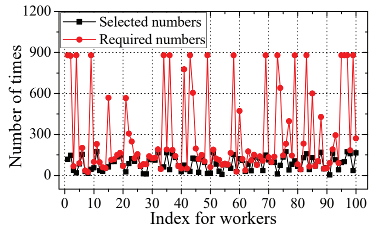

Fig. 13: $\varrho = 1$

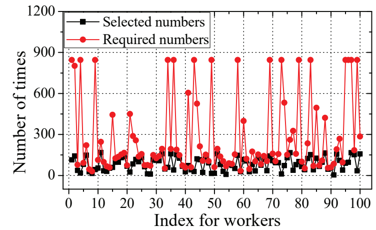

Fig. 14: $\varrho = 10$

and $\varepsilon_i f(|\mathcal{M}_i^l|)$ is normalized to $(0,1]$. Moreover, we let the values of w_j be uniform and let $K = N/3$ by default. In the experimental simulations, we adopt the Gaussian Distributions to generate the quality and cost parameters. Here, in order to ensure that the generated quality values are located in $(0,1]$, we let the variance of the Gaussian Distribution for the worker i, denoted as σ_i, be selected from the range $(0, \min\{\frac{q_i}{2}, \frac{1-q_i}{2}\}]$. In such settings, the generated quality values in each round (i.e., $q_{i,j}^t$) is located in the range $(0,1]$ with the probability of at least 95.4%.

Compared Algorithms: Since our optimization problem involving the budget-limited maximum weighted coverage problem is a novel CMAB problem, there are no existing bandit algorithms that can be directly applied in our model. For comparison, we borrow the basic strategy in the existing ϵ-first bandit algorithm [38] to design a compared algorithm. That is, we randomly selected K workers in each round under the first $\epsilon \cdot B$ budget. In the remaining $(1-\epsilon) \cdot B$ budget, we always recruit the K workers who perform best under the previous $\epsilon \cdot B$ budget. We evaluate the ϵ-first algorithm by choosing $\epsilon = 0.1$, $\epsilon = 0.5$, and $\epsilon = 1$. Note that $\epsilon = 1$ indicates that the ϵ-first algorithm will select K workers in each round randomly. Additionally, we implement the α-optimal algorithm in which all parameters are known in advance.

Moreover, since the UWR and FAUWR algorithms are designed for the case where only the sensing quality of workers is unknown in advance, while the EUWR algorithm is applied to the settings in which both the sensing quality and recruitment cost of workers are unknown a priori, we divide the experimental simulations into two groups. In the first group, we compare the performance of UWR and FAUWR with other algorithms; in the second group, we verify the performance of EUWR in other settings. The main simulation metrics include the total achieved weighted qualities and the rounds. Additionally, we also evaluate the fairness constraint of workers in the FAUWR algorithm. More specifically, when we change the controlling parameter ϱ, we compare the required minimum selection fraction for each worker with the corresponding achieved values. Here, we set the parameter ϱ as $\varrho=0.1$, $\varrho=1$, and $\varrho=10$ in the FAUWR algorithm.

6.2 Evaluation Results

First, we display the evaluation results of the UWR and FAUWR algorithms. To evaluate the effects of budget B, we let B change from 500 to 10000. Clearly, we find that the achieved total weighted qualities and total rounds of all four algorithms rise along with the increase of the budget B, as shown in Fig. 4 and Fig. 5. Moreover, we see that our algorithms perform much better than the compared ϵ-first algorithm. At the same time, the total weighted qualities and the total rounds achieved by the UWR and FAUWR algorithms almost catch up with the α-optimal algorithm. Here, we calculate that the achieved weighted qualities of UWR and FAUWR are at least 139.34% and 132.57% higher than that of the 0.1-first algorithm on average, respectively. Also, we compute that UWR and FAUWR can achieve about 87.86% and 85.41% of the total weighted qualities of the α-optimal algorithm on average, respectively. Meanwhile, the performance for total rounds has a similar conclusion to the quality performance. These observations exactly validate our theoretical analysis results.

Moreover, we also evaluate the performance on the size of K, as shown in Fig. 6 and Fig. 7. The results indicate that the UWR and FAUWR algorithms still outperform the compared ϵ-first algorithm in terms of the achieved total quality and rounds. The smaller K is, the higher total quality can be achieved. However, this will also result in higher recruitment rounds (i.e., more running time). From the simulation results, we get that the selection of the parameter K has a big impact on the performance of our algorithms. Also, we have that the total weighted quality achieved by the α-optimal algorithm is about 10.07% and 18.35% higher than that of the UWR and FAUWR algorithms on average, respectively. The total rounds of UWR and FAUWR are even higher than that of the α-optimal algorithm. This indicates that UWR and FAUWR perform well in terms of the total weighted qualities and total rounds. Then, we evaluate the performance of UWR and FAUWR by changing the numbers of sensing tasks and crowd workers, as shown in Fig. 8, Fig. 9, Fig. 10, and Fig. 11. According to the simulation results, we observe that our proposed UWR and FAUWR algorithms can obtain more than 173% and 168% larger weighted completion quality than the compared ϵ-first algorithm on average, respectively, and is even going to catch up to the α-optimal algorithm that knows all parameters in advance. Furthermore, the total rounds achieved by our algorithms and the α-optimal algorithm almost have the same trends. Along with the increase in the number of workers, the total rounds of all algorithms decrease, and meanwhile, the total weighted quality of all tasks has a slight downward trend. Additionally, when we increase the number of tasks from 100 to 500, the total weighted quality and total rounds of all algorithms show a slight upward trend. This is because the total weighted quality is highly related to the distribution of all tasks. These observations are still consistent with our theoretical analysis results.

On the other hand, the total weighted completion quality of the FAUWR algorithm is a bit smaller than that of the UWR algorithm. This is because FAUWR must handle the fairness constraint of workers by sacrificing a part of the total weighted quality. Here, along with the increase of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3064324, IEEE Transactions on Mobile Computing

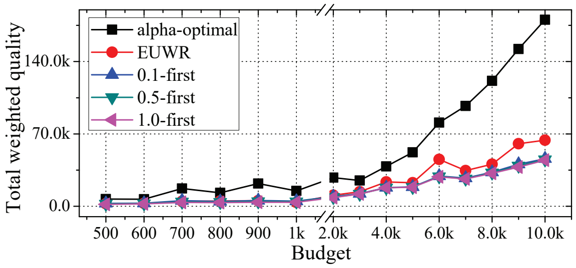

Fig. 15: EUWR: total qualities vs. budget

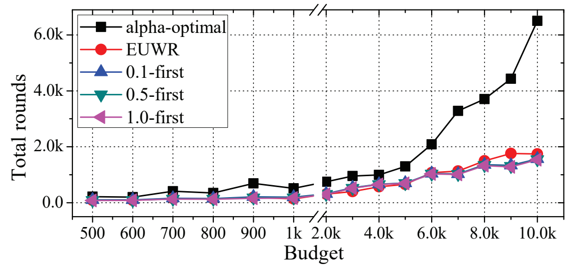

Fig. 16: EUWR: total rounds vs. budget

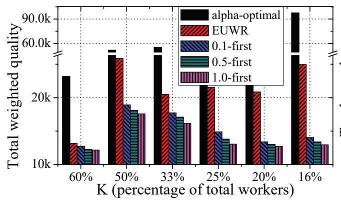

Fig. 17: Qualities vs. K

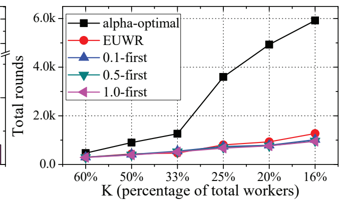

Fig. 18: Rounds vs. K

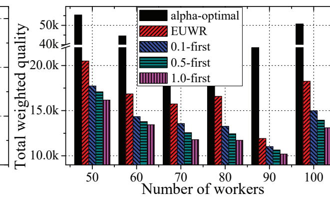

Fig. 19: Qualities vs. N

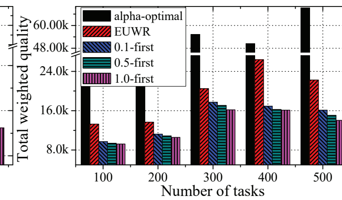

Fig. 20: Qualities vs. M

the parameter ϱ, the FAUWR algorithm would achieve less weighted completion quality. This indicates that FAUWR would spend more budget on the fairness constraint. In order to evaluate the fairness constraint of workers in the FAUWR algorithm, we present the number of times each worker has been selected under the given budget, as shown in Fig. 12, Fig. 13, and Fig. 14. In the simulations, we set the number of workers as $N = 100$. For each worker i, we use the uniform distribution to generate the minimum selection fraction, i.e., η_i. Then, in the experimental simulations for FAUWR, we record the total rounds (i.e., $\tau(B)$) and the numbers of times each worker i has been recruited (i.e., $n_i(\tau(B))$) under the budget B. Note that we set the budget to $B = 5000$ here. After we get the values of η_i, $\tau(B)$, and $n_i(\tau(B))$ for $\forall i \in \mathcal{N}$, we compare $\eta_i \times \tau(B)$ with $n_i(\tau(B))$ to verify the fairness constraint of workers. Actually, only when $n_i(\tau(B)) \geq \eta_i \times \tau(B)$, we can say that the fairness constraint is satisfied. We first let $\varrho = 0.1$ and then we display the simulation results in Fig. 12. We see that the numbers of times some workers have been selected are larger than 900. This means that the workers with good completion quality will be recruited first. However, if we increase the parameter $\varrho = 1$ and $\varrho = 10$, we observe that the number of times the "good" workers have been recruited decreases, while the number of times other ("not good") workers are recruited raises. This demonstrates that the fairness constraint of workers dominates in the worker selection process. These results still remain consistent with our theoretical analysis.

Second, we demonstrate the evaluation results of the EUWR algorithm in other settings where both the sensing quality and recruitment cost are unknown a priori. As shown in Fig. 15, we first investigate the relationship between the achieved quality and budget. We get that the total weighted quality achieved by EUWR does not have an overwhelming advantage over the compared ϵ-first algorithm. More precisely, we calculate that the total weighted quality of EUWR is about 17.61%, 26.44%, and 35.79% higher than that of the 0.1-first, 0.5-first, and 1.0-first algorithms on average, respectively. Also, the difference in achieved quality between EUWR and the compared ϵ-first algorithm increases when the budget rises. This means the EUWR algorithm is efficient because the platform has more confidence in the estimation of the unknown parameters. Moreover, we present the total rounds of all algorithms in Fig. 16. With the increase in the given budget B, the total rounds of all algorithms raise. The simulation results have the similar trends with the quality performance.

We then evaluate the performance of EUWR by changing the size of the selected workers (i.e., K), as shown in Fig. 17 and Fig. 18. We also find that, in these settings, the advantage of EUWR over the compared ϵ-first algorithm is not as overwhelming as that of UWR and FAUWR, due to two unknown parameters existing in the general problem. In all simulation settings, the total weighted completion quality achieved by EUWR is about 44.98% higher than that of the ϵ-first algorithm on average. Although the total rounds of the compared ϵ-first algorithm may be higher than that of EUWR, the total achieved quality is less than that of the EUWR algorithm. In addition, we also evaluate the EUWR algorithm in terms of the number of sensing tasks (i.e., M) and the number of workers (i.e., N), as shown in Fig. 19 and Fig. 20. The total quality achieved by EUWR is about 33.62%, 39.54%, and 44.53% higher than that of the 0.1-first, 0.5-first, and 1.0-first algorithms on average, respectively. Here, when we change the numbers of workers and sensing tasks, the total weighted quality has the same trend as before. These observations are also consistent with our theoretical analysis.

7 Related Work

In this paper, we study the combinatorial multi-armed bandit (CMAB) based budget-limited unknown worker recruitment for the heterogeneous mobile crowdsensing (MC). So far, there have been lots of researches on the worker recruitment problem in MC, such as [5], [13], [14], [39]–[43]. More precisely, [39] investigates how to optimally recruit crowd workers in opportunistic network based crowdsening such that the required space-time paths across the network for collecting data from a set of fixed locations can be generated, where workers are seen as the nodes in the space-time paths. For the case of deterministic node mobility, the authors formulate the worker recruitment as a minimum cost set cover problem with a submodular objective function.

For the more general settings with uncertainty about the worker mobility, they translate the statistics of individual worker mobility to statistics of space-time path formation and feed them to the set cover problem formulation. [13] proposes an optimal worker recruitment mechanism, called Crowdlet, for the self-organized MC paradigm. In the Crowdlet system, a task requester can proactively exploit a massive crowd of encountered workers in real-time for quick and high-quality results. By combining the factors of worker ability, real-timeness and task reward, the authors design an optimal online worker recruitment policy through the dynamic programming principle, so that the expected sum of service quality can be maximized. [43] studies the high quality worker recruitment problem in vehicle-based crowdsourcing by involving the predictable mobility of vehicles, that is, the proposed worker recruitment strategy can guarantee that the vehicle-based crowdsourcing system can perform well using the currently recruited worker for a period of time in the future. However, most of the existing work assumes that the sensing qualities or costs of workers are known in advance, thus focusing on the quality maximization or cost minimization problems under various constraints. Unfortunately, the sensing qualities or costs of workers are generally unknown a priori in real life.

In fact, only a few researches [17]–[20], [44]–[48] consider the unknown sensing qualities or costs in MC systems. For instance, [17] proposes a distance-reliability ratio algorithm to maximize the task completion ratio by considering the unknown reliability of workers and dynamic arrivals of tasks, which is based on a combinatorial fractional programming approach; [45] studies to maximize the total sensing revenue for the budget limited robust mobile crowdsensing, and further proves that the logarithmic regret bound can be achieved in the proposed framework; [20] designs a context-aware hierarchical online learning algorithm for performance maximization of MC, in which the workers' acceptance rate and sensing quality are taken into consideration. Also, the authors analyze that their proposed algorithm can converge to the optimal task assignment strategy; [19] investigates how to select the most informative contributors with unknown costs for budgeted MC and a budgeted multi-armed bandit based worker recruitment algorithm with theoretically proven low-regret guarantee is devised. However, the works [18], [19] either assume that the MC system only contains one task or that the sets of tasks for all workers are identical, while other works [17], [44] focus on the one-to-one matching problem between workers and tasks. Actually, all of them are based on the homogeneous MC model. Differing from the existing work, we study the unknown worker recruitment problem for the heterogeneous MC system. Particularly, our research involves a budget-limited maximum weighted coverage problem.

We model our problem as a novel combinatorial multi-armed bandit problem. The existing algorithms for traditional multi-armed bandits [21], [24], [25], [36], [49], [50] cannot be applied to our problem. The most related works are [32], [51], in which they study the top K bandit selection problem. The authors in [32] consider the CMAB model, where multiple random arms with unknown means can be chosen under the constraints of weights associated with the selected arms in each round. Although the reward for each selected individual arm would be observed by the player, a linearly weighted combination of these selected arms is yielded as the final reward at each round. To this end, the authors propose an efficient algorithm that can achieve a good regret bound, denoted as $O(N^4 \ln T)$, where T represents the total rounds. Also, the proposed algorithm only requires linear storage and polynomial computation. On the other hand, the authors in [51] focus on the budget-limited CMAB problem for both the stochastic and the adversarial settings. In addition to observing the selected individual arms' rewards in each round, the player also needs to learn the cost vector of all selected arms. For the stochastic setting, the authors design a UCB-based algorithm with a $O(NK^4 \ln B)$ regret bound that is actually on the same order of magnitude as ours. For the adversarial setting where the entire sequences of rewards and costs for all arms are fixed in advance, they devise an algorithm based on the well-known $Exp3$ algorithm. The upper and lower bounds on the magnitude of regret are given: $O(\sqrt{NB\ln(N/K)})$ and $\Omega((1–K/N)^2\sqrt{NB/K})$, respectively. Nevertheless, neither of them involves the budget-limited maximum weighted coverage problem or considers that each arm (i.e., a worker) has multiple candidate options.

In contrast, we model our unknown worker recruitment problem as a novel CMAB problem, in which each worker is seen as an arm, the sensing quality of a worker completing a task is seen as the reward of pulling arms, and the worker recruitment is equivalent to the arm-pulling action. Compared to the traditional CMAB models, our proposed CMAB model has two novel characteristics. The first one is that each arm has multiple options and each option is composed of a subset of all sensing tasks and the corresponding cost. The platform (i.e., the player) needs to select the arms and at the same time to determine the option for each arm. The second characteristic is that the objective of our CMAB problem is to maximize the total weighted completion qualities of all tasks under the given budget, i.e., so-called the budget-limited maximum weighted coverage problem. Here, each arm is attached with a weight to indicate its importance for the player. In addition, one sensing task might be covered by several workers, so the computation of the total rewards in a round is more challenging. To solve this problem, we first extend the upper confidence bound to denote the sensing quality of workers, and further propose an unknown worker recruitment algorithm. We prove that the proposed algorithm can achieve a good bound on the regret, i.e., $O(NLK^3\ln(B+NLK^2\ln(MLN^2K^2)))$, in which B, N, M, and L denote the budget, the number of workers, the number of tasks, and the number of options for each worker, respectively. Also, we study an extended case where both the sensing quality and cost parameters of all workers are unknown a priori. We devise an extended unknown worker recruitment algorithm with a provable performance guarantee $O(NLK^3\ln(NMB+N^2K^2ML\ln(N^2K^2ML)))$.

8 Conclusion & Future Work

We study the unknown worker recruitment problem in heterogeneous crowdsensing, where workers' sensing qualities are unknown in advance but follow an independent and identically distribution. In our problem, each task may be performed by multiple workers, but its completion quality only depends on these workers' maximum sensing quality. In future works, we will consider a new function that can combine the sensing results from multiple workers to improve the completion quality, and will refer to the truth discovery method when calculating each worker's completion quality. In this paper, we focus on how to recruit K unknown workers in each round so that the total weighted completion quality of all tasks can be maximized under the budget constraint. We model this problem as a novel combinatorial multi-armed bandit problem. To this end, we propose an extended UCB based unknown worker recruitment algorithm with a good regret bound. Furthermore, we study another case where both workers' sensing quality and cost are unknown a priori. A new algorithm with a provable performance guarantee is designed. Additionally, we investigate the unknown worker recruitment problem with fairness constraints and further propose a fairness-aware unknown worker recruitment algorithm. Extensive simulations on real-world traces are conducted to show the significant performance of the proposed algorithms.

REFERENCES

[1] G. Gao, J. Wu, M. Xiao, and G. Chen, "Combinatorial multi-armed bandit based unknown worker recruitment in heterogeneous crowdsensing," in *IEEE INFOCOM*, 2020.
[2] Y. Wei, Y. Zhu, H. Zhu, Q. Zhang, and G. Xue, "Truthful online double auctions for dynamic mobile crowdsourcing," in *IEEE INFOCOM*, 2015.
[3] X. Wang, J. Zhang, X. Tian, X. Gan, Y. Guan, and X. Wang, "Crowdsensing-based consensus incident report for road traffic acquisition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2536–2547, 2018.
[4] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao, "Incentives for mobile crowd sensing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 54–67, 2015.
[5] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.
[6] J. Liu, H. Shen, H. S. Narman, W. Chung, and Z. Lin, "A survey of mobile crowdsensing techniques: A critical component for the internet of things," *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 3, p. 18, 2018.
[7] D. Zhao, H. Ma, S. Tang, and X. Li, "Coupon: A cooperative framework for building sensing maps in mobile opportunistic networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 2, pp. 392–402, 2015.
[8] S. He, D. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *IEEE INFOCOM*, 2014.
[9] Y. Chen, B. Li, and Q. Zhang, "Incentivizing crowdsourcing systems with network effects," in *IEEE INFOCOM*, 2016.
[10] W. Jin, M. Xiao, M. Li, and L. Guo, "If you do not care about it, sell it: Trading location privacy in mobile crowd sensing," in *IEEE INFOCOM*, 2019.
[11] D. Zhao, X.-Y. Li, and H. Ma, "Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully," *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 647–661, 2016.
[12] G. Gao, M. Xiao, J. Wu, L. Huang, and C. Hu, "Truthful incentive mechanism for nondeterministic crowdsensing with vehicles," *IEEE Transactions on Mobile Computing*, vol. 17, no. 12, pp. 2982–2997, 2018.
[13] L. Pu, X. Chen, J. Xu, and X. Fu, "Crowdlet: Optimal worker recruitment for self-organized mobile crowdsourcing," in *IEEE INFOCOM*, 2016.
[14] Z. Song, C. H. Liu, J. Wu, J. Ma, and W. Wang, "Qoi-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4618–4632, 2014.
[15] S. Yang, K. Han, Z. Zheng, S. Tang, and F. Wu, "Towards personalized task matching in mobile crowdsensing via fine-grained user profiling," in *IEEE INFOCOM*, 2018.
[16] H. Jin, H. Guo, L. Su, K. Nahrstedt, and X. Wang, "Dynamic task pricing in multi-requester mobile crowd sensing with markov correlated equilibrium," in *IEEE INFOCOM*, 2019.
[17] U. ul Hassan and E. Curry, "Efficient task assignment for spatial crowdsourcing: A combinatorial fractional optimization approach with semi-bandit learning," *Expert Systems with Applications*, vol. 58, pp. 36–56, 2016.
[18] Y. Wu, F. Li, L. Ma, Y. Xie, T. Li, and Y. Wang, "A context-aware multi-armed bandit incentive mechanism for mobile crowd sensing systems," *IEEE Internet of Things Journal*, pp. 1–1, 2019.
[19] S. Yang, F. Wu, S. Tang, T. Luo, X. Gao, L. Kong, and G. Chen, "Selecting most informative contributors with unknown costs for budgeted crowdsensing," in *IEEE/ACM IWQoS*, 2016.
[20] S. K. née Müller, C. Tekin, M. van der Schaar, and A. Klein, "Context-aware hierarchical online learning for performance maximization in mobile crowdsourcing," *IEEE/ACM Transactions on Networking*, vol. 26, no. 3, pp. 1334–1347, 2018.
[21] L. Tran-Thanh, A. Chapman, A. Rogers, and N. R. Jennings, "K-napsack based optimal policies for budget–limited multi–armed bandits," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.
[22] J. Vermorel and M. Mohri, "Multi-armed bandit algorithms and empirical evaluation," in *European conference on machine learning*. Springer, 2005.
[23] S. Wang and L. Huang, "Multi-armed bandits with compensation," in *Advances in Neural Information Processing Systems*, 2018.
[24] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.
[25] A. Biswas, S. Jain, D. Mandal, and Y. Narahari, "A truthful budget feasible multi-armed bandit mechanism for crowdsourcing time critical tasks," in *International Conference on Autonomous Agents and Multiagent Systems*, 2015, pp. 1101–1109.
[26] D. Yang, G. Xue, X. Fang, and J. Tang, "Incentive mechanisms for crowdsensing: Crowdsourcing with smartphones," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1732–1744, 2016.
[27] D. Peng, F. Wu, and G. Chen, "Pay as how well you do: A quality based incentive mechanism for crowdsensing," in *ACM MobiHoc*, 2015.
[28] J. Li, Z. Cai, J. Wang, M. Han, and Y. Li, "Truthful incentive mechanisms for geographical position conflicting mobile crowdsensing systems," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 324–334, 2018.
[29] G. Yang, S. He, Z. Shi, and J. Chen, "Promoting cooperation by the social incentive mechanism in mobile crowdsensing," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 86–92, 2017.
[30] J. Lin, D. Yang, M. Li, J. Xu, and G. Xue, "Frameworks for privacy-preserving mobile crowdsensing incentive mechanisms," *IEEE Transactions on Mobile Computing*, vol. 17, no. 8, pp. 1851–1864, 2018.
[31] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial multi-armed bandit: General framework and applications," in *International Conference on Machine Learning*, 2013.
[32] Y. Gai, B. Krishnamachari, and R. Jain, "Combinatorial network optimization with unknown variables: Multi-armed bandits with linear rewards and individual observations," *IEEE/ACM Transactions on Networking*, vol. 20, no. 5, pp. 1466–1478, 2012.
[33] E. L. Lawler, "Fast approximation algorithms for knapsack problems," *Mathematics of Operations Research*, vol. 4, no. 4, pp. 339–356, 1979.
[34] V. V. Vazirani, *Approximation algorithms*. Springer Science & Business Media, 2013.
[35] J. P. Schmidt, A. Siegel, and A. Srinivasan, "Chernoff–hoeffding bounds for applications with limited independence," *SIAM Journal on Discrete Mathematics*, vol. 8, no. 2, pp. 223–250, 1995.
[36] F. Li, J. Liu, and B. Ji, "Combinatorial sleeping bandits with fairness constraints," in *IEEE INFOCOM*, 2019.
[37] L. Bracciale, M. Bonola, P. Loreti, G. Bianchi, R. Amici, and A. Rabuffi, "CRAWDAD dataset roma/taxi (v. 2014-07-17)," Downloaded from http://crawdad.org/roma/taxi/20140717.
[38] L. Tran-Thanh, A. Chapman, E. M. de Cote, A. Rogers, and N. R. Jennings, "Epsilon–first policies for budget–limited multi-armed bandits," in *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.
[39] M. Karaliopoulos, O. Telelis, and I. Koutsopoulos, "User recruitment for mobile crowdsensing over opportunistic networks," in *IEEE INFOCOM*, 2015.

[40] N. B. Truong, G. M. Lee, T. Um, and M. Mackay, "Trust evaluation mechanism for user recruitment in mobile crowd-sensing in the internet of things," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2705–2719, 2019.
[41] E. Wang, Y. Yang, J. Wu, W. Liu, and X. Wang, "An efficient prediction-based user recruitment for mobile crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 16–28, 2017.
[42] M. Xiao, J. Wu, S. Zhang, and J. Yu, "Secret-sharing-based secure user recruitment protocol for mobile crowdsensing," in *IEEE INFOCOM*, 2017.
[43] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *IEEE INFOCOM*, 2015.
[44] M. Karaliopoulos, I. Koutsopoulos, and M. Titsias, "First learn then earn: Optimizing mobile crowdsensing campaigns through data-driven user profiling," in *ACM MobiHoc*, 2016.
[45] K. Han, C. Zhang, and J. Luo, "Taming the uncertainty: Budget limited robust crowdsensing through online learning," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1462–1475, 2016.
[46] G. Gao, J. Wu, Z. Yan, M. Xiao, and G. Chen, "Unknown worker recruitment with budget and covering constraints for mobile crowdsensing," in *IEEE ICPADS*, 2019.
[47] H. Zhao, M. Xiao, J. Wu, Y. Xu, H. Huang, and S. Zhang, "Differentially private unknown worker recruitment for mobile crowdsensing using multi-armed bandits," *IEEE Transactions on Mobile Computing*, 2020.
[48] X. Gao, S. Chen, and G. Chen, "Mab-based reinforced worker selection framework for budgeted spatial crowdsensing," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–1, 2020.
[49] Y. Xia, T. Qin, W. Ma, N. Yu, and T.-Y. Liu, "Budgeted multi-armed bandits with multiple plays." in *IJCAI*, 2016.
[50] S. Tang, Y. Zhou, K. Han, Z. Zhang, J. Yuan, and W. Wu, "Networked stochastic multi-armed bandits with combinatorial strategies," in *IEEE ICDCS*, 2017.
[51] D. P. Zhou and C. J. Tomlin, "Budget-constrained multi-armed bandits with multiple plays," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Guoju Gao is an assistant professor in the School of Computer Science and Technology at the Soochow University. He received his Ph.D. degree in computer science and technology with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China. He visited the Temple University, USA, from Jan. 2019 to Jan. 2020. His research interests include incentive mechanisms, mobile crowdsensing, edge computing and reinforcement learning.

He Huang is currently a professor in the School of Computer Science and Technology at Soochow University, P. R. China. He received his Ph.D. degree in School of Computer Science and Technology from University of Science and Technology of China (USTC), China, in 2011. From 2019 to 2020, he was a Visiting Research Scholar with Florida University, Gainesville, USA. He has authored over 100 papers in related international conference proceedings and journals. His current research interests include traffic measurement, computer networks, and algorithmic game theory. Dr. Huang is a Member of both IEEE and the Association for Computing Machinery (ACM). He received the best paper awards from Bigcom 2016, IEEE MSN 2018, and Bigcom 2018. He has served as the Technical Program Committee Member of several conferences, including IEEE INFOCOM, IEEE MASS, IEEE ICC, and IEEE Globecom.

Mingjun Xiao is a professor in the School of Computer Science and Technology at the University of Science and Technology of China (USTC). He received his Ph.D. from USTC in 2004. His research interests include crowdsourcing, mobile social networks, mobile cloud computing, blockchain, data security and privacy. He has published more over 80 papers in referred journals and conferences, including TMC, TC, TPDS, TKDE, TSC, INFOCOM, ICNP, etc. He served as the TPC member of INFOCOM'20, DASFAA'20, INFOCOM'19, ICDCS'19, DASFAA'19, etc. He is on the reviewer board of several top journals such as TMC, TON, TPDS, TSC, TVT, TCC, etc.

Jie Wu is the Director of the Center for Networked Computing and Laura H. Carnell professor at Temple University. He also serves as the Director of International Affairs at College of Science and Technology. He served as Chair of Department of Computer and Information Sciences from the summer of 2009 to the summer of 2016 and 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 at the National Science Foundation and was a distinguished professor at Florida Atlantic University. His current research interests include mobile computing and wireless net- works, routing protocols, cloud and green computing, network trust and security, and social network applications. Dr. Wu regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including IEEE Transactions on Service Computing and the Journal of Parallel and Distributed Computing. Dr. Wu was general co-chair for IEEE MASS 2006, IEEE IPDPS 2008, IEEE ICDCS 2013, ACM MobiHoc 2014, ICPP 2016, and IEEE CNS 2016, as well as program co-chair for IEEE INFOCOM 2011 and CCF CNCC 2013. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). Dr. Wu is a CCF Distinguished Speaker and a Fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.

Yu-E Sun is a professor of School of Rail Transportation, Soochow University, P.R. China. She received her Ph.D. degree in Shenyang Institute of Computing Technology from Chinese Academy of Science. From 2019 to 2020, she was a Postdoc Research Scholar with Florida University, Gainesville, USA. She has authored over 80 papers in related international conference proceedings and journals. Her current research interests span traffic measurement, privacy preserving, algorithm design and analysis for wireless networks. She is a Member of both IEEE and ACM. She received the best paper awards from Bigcom 2016, IEEE MSN 2018, and Bigcom 2018. She has served as the Technical Program Committee Co-Chairs of ACM MSCC-MobiHoc 2016.

Yang Du is currently a postdoctoral fellow in Soochow University. He received his B.E. degree from the Soochow University in 2015 and Ph.D. degree from University of Science and Technology of China in 2020. His research interests include network traffic measurement, mobile crowdsensing, and truth discovery.