Combinatorial Multi-Armed Bandit Based Unknown Worker Recruitment in Heterogeneous Crowdsensing

Guoju Gao†, Jie Wu‡, Mingjun Xiao†*, and Guoliang Chen†
†School of Computer Science and Technology / Suzhou Institute for Advanced Study,
University of Science and Technology of China, Hefei, P. R. China
‡Department of Computer and Information Sciences, Temple University
Email: {gaoguoju@mail.ustc.edu.cn, jiewu@temple.edu, xiaomj@ustc.edu.cn, glchen@ustc.edu.cn}

***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 an unknown worker recruitment process as a *novel combinatorial multi-armed bandit* problem, and propose an extended UCB based worker recruitment algorithm. Moreover, we extend the problem to the case where the workers' costs are also unknown and design the corresponding algorithm. We analyze the regrets of the two proposed algorithms and demonstrate their performance through extensive simulations on real-world traces.**

***Index Terms*—Mobile crowdsensing, multi-armed bandits, online learning, worker recruitment.**

I. Introduction

Mobile CrowdSensing (MCS) 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 [1]–[9]. Owing to users' mobility and the diversity of sensors embedded in their smart phones, MCS 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.

A typical MCS 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 and mobile behaviors, 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 MCS. Much effort has been devoted to designing worker recruitment or task allocation algorithms in recent years [4], [7], [10]–[15]. However, most of 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, [16] studies how to maximize the task completion ratio while considering unknown workers' reliability and dynamic arrivals of tasks; [17] develops a modified Thompson Sampling worker selection algorithm to recruit some unknown workers. Nevertheless, these researches mainly involve homogeneous MCS models in which each task can be completed by all workers, although their sensing qualities might be different.

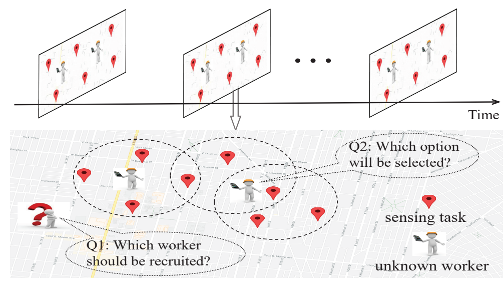


Fig. 1. Illustration of the heterogeneous crowdsensing scenario.

In this paper, we focus on the unknown worker recruitment problems in heterogeneous MCS systems. Consider such a scenario where a requester wants to recruit workers to collect the traffic data (e.g., traffic photos or videos) at some 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. Each worker can provide multiple options, composed of different task combinations and costs, but at most one option will be selected. Moreover, each worker has a sensing quality, following an unknown distribution. Our objective is to design a worker recruitment strategy that can maximize the total task completion quality under a given budget.

In the above unknown worker recruitment problem of heterogeneous MCS systems, the main challenge lies in that the platform does not know 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 [18],

[19], 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, we let a fixed number of arms (i.e., K) be pulled in each round. Our CMAB model has two novel characteristics, 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 [18], [20]. It always selects the arm that has the largest value on the estimated reward and the upper bound of confidence to be pulled. To solve our CMAB problem, 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, due to the reason that 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. In addition, we extend our problem to the scenario where workers' costs are also unknown and devise another algorithm.

Our major contributions are summarized as follows:

- We introduce the unknown worker recruitment problem for heterogeneous MCS 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 worst regret bound of the algorithm: $O(NLK^3 \ln B)$, where B, N, and L are the budget, the number of workers, and the number of options of 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 $O(NLK^3 \ln(NMB))$, where M is the number of tasks.
- We conduct extensive simulations on real-world traces to evaluate the significant performance of our algorithms.

All proofs of theorems and lemmas can be found in Appendix.

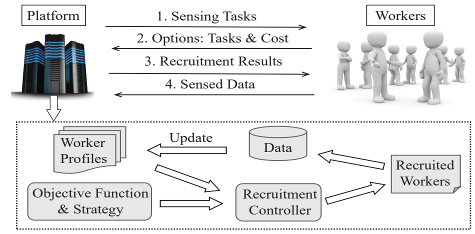


Fig. 2. Illustration of the main procedures in the mobile crowdsensing.

II. System Model & Problem

A. System Overview

Consider an MCS 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 MCS 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, each of which includes a subset of tasks that it can deal with and also attaches 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 MCS 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 evaluated only by the platform after the worker completes some tasks and submits the sensed results. 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. It should be pointed out that we mainly focus on the unknown worker recruitment problem in this paper, and thus we assume that the whole system is secure and truthful by leaving the privacy-preserving and incentive issues to be solved in future works. 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 quality for each worker.

B. Model

We let t denote the index of round, and 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 $\sum_{j\in\mathcal{M}} w_j = 1$.

In our MCS 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,

TABLE I
DESCRIPTION OF COMMONLY-USED NOTATIONS.

Variable	Description
$\mathcal{N}$, $\mathcal{M}$	the sets of workers and sensing tasks, respectively.
i, j, t	the indexes for workers, tasks, and rounds.
K	the number of recruited workers in each round.
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.

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 more 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 camera in the worker's smart phone, the skill, angle, habit of taking photos or videos, etc. 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, which differs from the traditional CMAB model [21], [22].

C. Problem

For the above heterogeneous MCS system, we focus on recruiting K workers in each round so that the weighted sum of the completion qualities (called total weighted completion quality) of all the tasks over all rounds can be maximized under a given budget. 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 $u^j(\mathcal{P}^t)$,

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

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.,

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

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:

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

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

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

$$\textstyle\sum_{l=1}^{L} \mathbb{I}\{p_i^l \in \mathcal{P}^t\} \leq 1 \tag{6}$$

Eqs. (4) and (5) mean the budget and quantity constraint, 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 I.

III. ALGORITHM DESIGN

A. Basic Solution

To address our unknown worker recruitment issue, 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 the 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,

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

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

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, the value of $\overline{q}_i(t)$ 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, i.e.,

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

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 MCS 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. 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 the values of $\{\widehat{q}_i(t-1)|i\in\mathcal{N}\}$, that is,

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

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 $\mathcal{P}\backslash\mathcal{P}^t$ 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 criterion, which can be described as follows:

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

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 $\mathcal{M}_i^l$. 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 $\widehat{q}_i(t)$. 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.

B. 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 of 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 criterion, 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 $\mathcal{P}\backslash\mathcal{P}^{t'}$, 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 $O(NMLK)$.

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

C. 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 [23]. There is no polynomial-time optimal algorithm for this problem. However, 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 $\mathcal{P}^\star\subset\mathcal{P}$. Note that $\mathcal{P}^\star$ satisfies $u_{[q_i]}(\mathcal{P}^\star)\geq\alpha\cdot\max_{\mathcal{P}^t\subset\mathcal{P}}u_{[q_i]}(\mathcal{P}^t)$ where $0<\alpha\leq 1$. 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, denoted by $\mathcal{P}^*\subset\mathcal{P}$, is not fair. Therefore, we introduce the concept of α-*approximation regret* [21], [24] of an algorithm $\mathcal{A}$ under the budget B, that is,

$$R_\alpha^{\mathcal{A}}(B)=\alpha\cdot\sum_{t\geq 1}u_{[q_i]}(\mathcal{P}^*)-\mathbb{E}\Big[\sum_{t\geq 1}u(\mathcal{P}^t)\Big]$$

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 $\forall t\geq 1$, 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 $u_B=u(\mathcal{P}^1)$, $B_t=B-\sum_{p_i^1\in\mathcal{P}^1}\varepsilon_i^1 f(|\mathcal{M}_i^1|)$, $n_i(t)=1$, $\overline{q}_i(t)=(\sum_{j\in\mathcal{M}_i^1}q_{i,j}^t)/|\mathcal{M}_i^1|$ 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 $\forall p_i^l\in\mathcal{P}^t\}$;
    $p_i^l=\underset{p_{i'}^{l'}\in(\mathcal{P}\setminus\mathcal{P}^{t'})}{\operatorname{argmax}}\ u_{[\widehat{r}_i{}^l(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{r}_i{}^l(t-1)]}(\mathcal{P}^t)$;
    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 $\sum_{p_i^l\in\mathcal{P}^t}\varepsilon_i^t f(|\mathcal{M}_i^l|)\geq B_{t-1}$ **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 $\widehat{r}_i{}^l(t)$;
  $B_t=B_{t-1}-\sum_{p_i^l\in\mathcal{P}^t}\varepsilon_i^t f(|\mathcal{M}_i^l|)$, and $u_B=u_B+u(\mathcal{P}^t)$;

$$\leq \textstyle\sum_{t\geq 1}u_{[q_i]}(\mathcal{P}^\star)-\mathbb{E}\Big[\sum_{t\geq 1}u(\mathcal{P}^t)\Big]. \quad (13)$$

According to the selection criterion and the existing work [23], [25], we have $\alpha\geq 1/2$ in our algorithm. Note that in this paper we always let * and * denote the corresponding identifications of the optimal and α-optimal workers, respectively. Then, 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.,

$\Delta_{min}=u_{[q_i]}(\mathcal{P}^\star)-\max_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}u_{[q_i]}(\mathcal{P}'), 0<c_{min}=\min\{c_i^l\}$,
$\Delta_{max}=u_{[q_i]}(\mathcal{P}^\star)-\min_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}u_{[q_i]}(\mathcal{P}'), c_{max}=\max\{c_i^l\}\leq 1$.

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 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 $p_i^l=\operatorname{argmin}_{p_{i'}^{l'}\in\mathcal{P}^t}C_{i'}^{l'}(t-1)$, 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 $\{C_i^l(t)|i\in\mathcal{N},1\leq l\leq L\}$. We first introduce a lemma to analyze the bound of the expected counter $\mathbb{E}[C_i^l(\tau(B))]$ as follows.

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

$$\mathbb{E}[C_i^l(\tau(B))]\leq\tfrac{4K^2(K+1)}{(\Delta_{min}c_{min})^2}\ln(NM\tau(B))+1+\tfrac{K\pi^2}{3}. \quad (14)$$

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 constraint. 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 $c^\star=\sum_{p_i^l\in\mathcal{P}^\star}c_i^l$)

$$\frac{B}{c^\star}-\varphi_3-1-\frac{\varphi_1\varphi_3}{\varphi_2}\ln(\frac{2B}{c^\star}+\varphi_4)\leq\tau(B)\leq\frac{2B}{c^\star}+\varphi_4. \quad (15)$$

Based on Lemmas 1 and 2, we have the following theorem.

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

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

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

IV. Extension

A. Extended Problem

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 $c_i^l=\varepsilon_i f(|\mathcal{M}_i^l|)$ where $f(\cdot)$ is given in our MCS 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 $0<\varepsilon_{min}\leq\varepsilon_i^t\leq 1$. 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 $c_i^l=\varepsilon_i^t f(|\mathcal{M}_i^l|)$. 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 worker profiles. In the extended problem, the platform needs to learn the quality q_i and parameter ε_i simultaneously, and meanwhile maximizes the total weighted qualities of all tasks under a given budget. So it is more challenging to design a suitable recruitment strategy.

B. Basic Solution

Like before, we still let $n_i^l(t)$ denote the number of p_i^l being selected. Differently, when p_i^l is selected in a round, the parameter ε_i is actually learned by 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:

$$\overline{\varepsilon}_i(t)=\begin{cases}\dfrac{\overline{\varepsilon}_i(t-1)\cdot m_i(t-1)+\varepsilon_i^t}{m_i(t-1)+1};\ \text{for}\ \forall 1\leq l\leq L, p_i^l\in\mathcal{P}^t,\\ \overline{\varepsilon}_i(t-1);\qquad\qquad\qquad\ \ \text{for}\ \forall 1\leq l\leq L, p_i^l\notin\mathcal{P}^t.\end{cases} \quad (16)$$

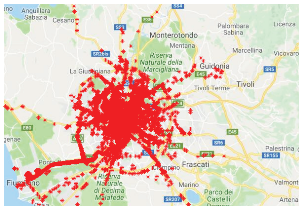


Fig. 3. Exhibition of the dataset

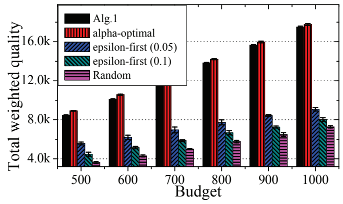
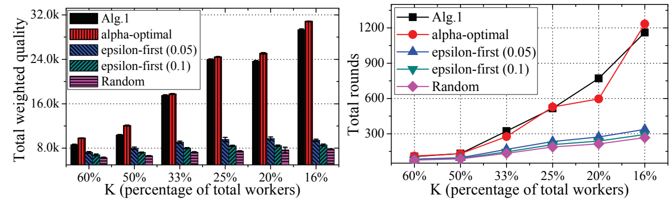


Fig. 4. Alg. 1: quality vs. budget



(a) Total quality vs. K (b) Total rounds vs. K

Fig. 5. Evaluation of Alg. 1 on the parameter K (Gaussian Distribution).

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 criterion, denoted by $\widehat{r}_i(t)$, which takes the obtained quality and cost values into consideration simultaneously. More specifically, we define

$$\widehat{r}_i^{\,l}(t)=f_i^l\cdot\frac{\overline{q}_i(t-1)}{\overline{\varepsilon}_i(t-1)}+f_{max}\cdot\frac{\varepsilon_{min}\cdot Q_{t,i}+C_{t,i}}{\varepsilon_{min}^2},\qquad(17)$$

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,

$$p_i^l=\underset{p_{i'}^{l'}\in(\mathcal{P}\backslash\mathcal{P}^t)}{\operatorname{argmax}}\;u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t\cup\{p_{i'}^{l'}\})-u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t).\quad(18)$$

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.,

$$u_{[\widehat{r}_i^{\,l}(t-1)]}(\mathcal{P}^t)=\sum_{j\in\mathcal{M}}w_j\cdot\max\{\widehat{r}_i^{\,l}(t-1)\cdot\mathbb{I}\{j\in\mathcal{M}_i^l,p_i^l\in\mathcal{P}^t\}\}.(19)$$

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 criterion $\widehat{q}_i(t-1)$ 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 of a worker can be selected in each round. In addition, the computational overhead of Alg. 2 is still $O(NMLK)$.

C. 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,

$$\Delta r_{min}=u_{[r_i^l]}(\mathcal{P}^\star)-\max_{\{\mathcal{P}'\neq\mathcal{P}^\star\}}u_{[r_i^l]}(\mathcal{P}'),\qquad(20)$$

$$\Delta r_{max}=u_{[r_i^l]}(\mathcal{P}^\star)-\min_{\{\mathcal{P}',\mathcal{P}'\neq\mathcal{P}^\star\}}u_{[r_i^l]}(\mathcal{P}').\qquad(21)$$

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. (19) 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 $O(NLK^3\ln(NMB))$. In particular, we give the specific expression:

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

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

V. Performance Evaluation

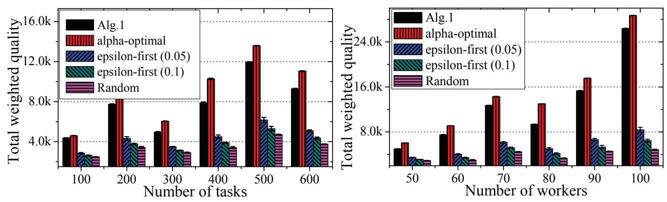
In this section, we evaluate the performance of the proposed algorithms with extensive simulations. We conduct the simulations on a computer with Inter(R) Core(TM) i7-8700 CPU @3.20GHz and 32GB RAM under a Windows platform.

A. Evaluation Methodology

Simulation Settings: We adopt the widely-used trace [26] 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 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 $[5,15]$.

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, and 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, 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 $K=N/3$ in default. In order to estimate the performance of our algorithms adequately, we adopt both Gaussian and Uniform Distributions to generate the quality and cost parameters.

Compared Algorithms: Since our optimization problem involving the budget-limited maximum weighted coverage problem is a novel CMAB problem, there are no existing



(a) Total quality vs. M

(b) Total quality vs. N

Fig. 6. Quality evaluation of Alg. 1 on M and N (Uniform Distribution).

bandit algorithms that can be directly applied in our model. For comparison, we borrow the basic strategy in the existing ϵ-first bandit algorithm [27] 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 $\epsilon = 0.05$ and $\epsilon = 0.1$. Moreover, we implement the Random algorithm which selects K workers in each round randomly. Additionally, we implement the α-optimal algorithm in which all parameters are known in advance.

B. Evaluation Results

First, we display the evaluation results of Alg. 1. To evaluate the effects of budget B, we let B change from 500 to 1000. We see that our algorithm performs much better than the compared algorithms, as shown in Fig. 4. The total weighted quality achieved by Alg. 2 under this setting is at least 75% higher than that of the compared algorithms. Also, we evaluate the performance on the size of K, as shown in Fig. 5. The results indicate that Alg. 1 still outperforms compared algorithms 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). Moreover, we evaluate the performance of Alg. 1 by changing the numbers of tasks and workers, as shown in Fig. 6. Here, in order to demonstrate the applicability of Alg. 1, we use the Uniform Distribution to generate quality values in each round. We observe that our algorithm can obtain more than 90% larger weighted quality than the compared algorithms, and is even going to catch up with the α-optimal algorithm which knows all parameters in advance. These observations exactly validate our theoretical analysis results.

Second, we demonstrate the evaluation results of Alg. 2. As shown in Fig. 7, we first investigate the relationship between the achieved quality and budget. Since the gap among all algorithms is not obvious, we change B in the range $[500, 10^4]$. We get that the total weighted quality achieved by Alg. 2 is at least 81% of that of the α-optimal algorithm. The difference of achieved quality between Alg. 2 and compared algorithms increases when the budget rises. This means Alg. 2 is efficient because the platform has more confidence in the estimation of the unknown parameters. We then evaluate the performance of Alg. 2 by changing the size of the selected workers (i.e., K), as shown in Fig. 8. Here, we let $B = 10^4$ in default. We see that the advantage of Alg. 2 over the compared algorithms is not as overwhelming as that of Alg. 1, due to two unknown parameters existing in the extended problem. Although the

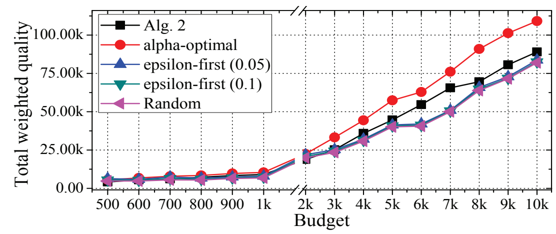


Fig. 7. Evaluation of Alg. 2: quality vs. budget (Gaussian Distribution)

total rounds of the compared algorithms may be higher than that of Alg. 2, the total achieved quality is below that of Alg. 2. In addition, we also evaluate Alg. 2 in terms of M and N under the condition of Uniform Distribution, as shown in Fig. 9. The total quality achieved by Alg. 2 is 21% larger than that of the compared algorithms on average. These observations still remain consistent with our theoretical analysis.

VI. Related Work

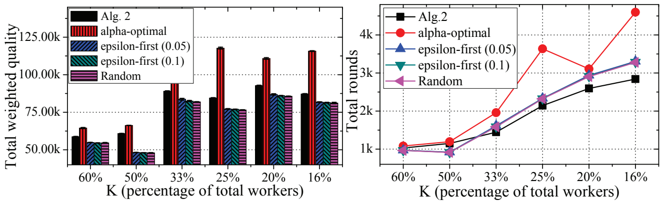
In this paper, we study the combinatorial multi-armed bandit based budget-limited unknown worker recruitment for the heterogeneous MCS. So far, there have been lots of researches on the worker recruitment problem in MCS, such as [4], [10], [11], [28]–[30]. However, most of the existing work assumes that the sensing qualities or costs of workers are known in advance, and then they focus on the quality maximization or cost minimization problems under various constraints.

In fact, only a few researches [16], [17], [24], [31]–[34] consider the unknown sensing qualities or costs in MCS systems. For instance, [33] studies to maximize the total sensing revenue for the budgeted robust MCS; [24] investigates how to select the most informative contributors with unknown costs for budgeted MCS. However, the work [17], [24], [33] either considers the MCS system only contains one task or assumes the sets of tasks for all workers are identical, while other work [16], [32] focuses on the one-to-one matching problem between workers and tasks. Actually, all of them are based on the homogeneous MCS model. Different from the existing work, we study the unknown worker recruitment problem for the heterogeneous MCS system. Especially, it involves a budget-limited maximum weighted coverage problem.

We model our problem as a novel combinatorial multi-armed bandit problem. The existing algorithms for multi-armed bandits [18], [20], [35], [36] cannot be applied in our problem. The most related works are [22], [37], in which they study the top K bandit selection problem. [22] proposes an algorithm which can achieve a good regret bound and only requires linear storage and polynomial computation. [37] designs a UCB-based algorithm with a $O(NK^4 \log B)$ regret bound. Nevertheless, neither of them involves the budget-limited maximum weighted coverage problem and considers that each arm (i.e., a worker) has multiple candidate options.

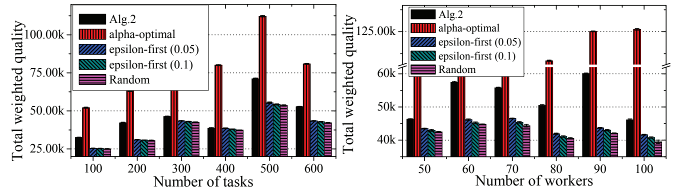
VII. Conclusion

In this paper, we study how to recruit K unknown workers in each round so that the total weighted completion quality of tasks can be maximized under the budget constraint. We model this problem as a novel combinatorial multi-armed bandit problem. It involves a budget-limited maximum weighted coverage problem and each arm has multiple candidate options.



(a) Total quality vs. K (b) Total rounds vs. K

Fig. 8. Evaluation of Alg.2 on the parameter K (Gaussian Distribution).



(a) Total quality vs. M (b) Total quality vs. N

Fig. 9. Quality evaluation of Alg. 2 on M and N (Uniform Distribution).

We propose an unknown worker recruitment algorithm with a good regret bound. Moreover, we study another case where both workers' sensing quality and cost are unknown. A new recruitment algorithm with a provable performance guarantee is designed. Extensive simulations on real-world traces are conducted to show the performance of our algorithms.

ACKNOWLEDGMENTS

This research was supported by the National Key R&D Program of China (Grant No. 2018AAA0101204), the National Natural Science Foundation of China (NSFC) (Grant No. 61872330, 61572457, 61379132, 61936015, U1709217), NSF grants CNS 1824440, CNS 1828363, CNS 1757533, CNS 1629746, CNS 1651947, CNS 1564128, the NSF of Jiangsu Province in China (Grant No. BK20191194, BK20131174, BK2009150), Anhui Initiative in Quantum Information Technologies (Grant No. AHY150300), and China Scholarship Council (CSC) during a visit of 'Guoju Gao' to Temple University (Grant CSC No. 201806340014).

APPENDIX

Proof of Lemma 1: Let $I_i^l(t)$ denote the indicator that $C_i^l(t)$ is incremented at round t. Then, we have

$$\begin{aligned}
C_i^l(\tau) &= \sum_{t=2}^{\tau} \mathbb{I}\{I_i^l(t)=1\} = \lambda + \sum_{t=2}^{\tau} \mathbb{I}\{I_i^l(t)=1, C_i^l(t)\geq\lambda\} \\
&\leq \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{u_{[\widehat{q}_i(t)/c_i^l]}(\mathcal{P}^{t+1}) \geq u_{[\widehat{q}_i(t)/c_i^l]}(\mathcal{P}^\star), C_i^l(t)\geq\lambda\} \\
&= \lambda + \sum_{t=1}^{\tau} \mathbb{I}\{\sum_{p_i^l\in\mathcal{P}^{t+1}} \xi_i^l(t+1)\frac{\widehat{q}_i(t)}{c_i^l} \geq \sum_{p_i^l\in\mathcal{P}^\star} \xi_i^l(\star)\frac{\widehat{q}_i(t)}{c_i^l}, C_i^l(t)\geq\lambda\}, \quad (22)
\end{aligned}$$

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,

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

Obviously, we have $\xi_i^l(t+1) \leq \sum_{j\in\mathcal{M}_i^l} w_j \leq 1$. Then, we continue Eq. (22) and get

$$\begin{aligned}
C_i^l(\tau) &\leq \lambda + \sum_{t=1}^{\tau}\mathbb{I}\Big\{ \max_{\lambda\leq n_{s(1)}\leq\cdots\leq n_{s(K)}\leq t} \sum_{i=1}^{K} \frac{\xi_i^l(t+1)}{c_i^l}\cdot\widehat{q}_{s(i)}(t) \\
&\geq \min_{1\leq n_{s^\star(1)}\leq\cdots\leq n_{s^\star(K)}\leq t} \sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}\cdot\widehat{q}_{s^\star(i)}(t)\Big\} \\
&\leq \lambda + \sum_{t=1}^{\tau}\sum_{n_{s(1)}=\lambda}^{t}\cdots\sum_{n_{s(K)}=\lambda}^{t}\sum_{n_{s^\star(1)}=1}^{t}\cdots\sum_{n_{s^\star(K)}=1}^{t} \\
&\quad \mathbb{I}\Big\{\sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}\cdot\widehat{q}_{s(i)}(t) \geq \sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}\cdot\widehat{q}_{s^\star(i)}(t)\Big\}, \quad (23)
\end{aligned}$$

where $s(i)$ and $s^\star(i)$ mean the i-th element in $\mathcal{P}^{t+1}$ and $\mathcal{P}^\star$, respectively. Here, $\widehat{q}_{s^\star(i)}(t) = \overline{q}_{s^\star(i)}(t) + Q_{t,s^\star(i)}$.

Next, we prove the probability of the following event:

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

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

$$\sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}\overline{q}_{s^\star(i)}(t) \leq \sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}(q_{s^\star(i)} - Q_{t,s^\star(i)}); \quad (24)$$

$$\sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}\overline{q}_{s(i)}(t) \geq \sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}(q_{s(i)} + Q_{t,s(i)}); \quad (25)$$

$$\sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}q_{s^\star(i)} < \sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}(q_{s(i)} + 2Q_{t,s(i)}). \quad (26)$$

Now, we prove the upper bound for Eq. (24), and get

$$\begin{aligned}
&\mathbb{P}\Big\{\sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}\overline{q}_{s^\star(i)}(t) \leq \sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}(q_{s^\star(i)}-Q_{t,s^\star(i)})\Big\} \\
&\leq \sum_{i=1}^{K}\mathbb{P}\Big\{\overline{q}_{s^\star(i)}(t) \leq q_{s^\star(i)} - Q_{t,s^\star(i)}\Big\}. \quad (27)
\end{aligned}$$

After applying the Chernoff-Hoeffding bound introduced in the existing work [20], we have

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

We continue Eq. (27) and get the upper bound, which is

$$\mathbb{P}\{\text{Eq. (27)}\} \leq K\cdot t^{-2(K+1)}.$$

Similarly, we can derive the upper bound for Eq. (25), which is the same as that of the first case. Moreover, if both Eq. (24) and Eq. (25) are false, we can easily infer that Eq. (26) is true. Now, we pick λ such that Eq. (26) becomes impossible.

$$\begin{aligned}
&\sum_{i=1}^{K}\frac{\xi_i^l(\star)}{c_{i^\star}^l}q_{s^\star(i)} - \sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}q_{s(i)} - 2\sum_{i=1}^{K}\frac{\xi_i^l(t+1)}{c_i^l}Q_{t,s(i)} \\
&\geq \Delta_{\mathcal{P}^{t+1}} - K\frac{\sum_{j\in\mathcal{M}} w_j}{c_{min}}\sqrt{\frac{4(K+1)\ln(\sum_{i'\in\mathcal{N}} n_{s(i')}(t))}{n_{s(i)}(t)}} \\
&\geq \Delta_{\mathcal{P}^{t+1}} - \frac{K}{c_{min}}\sqrt{\frac{4(K+1)\ln(NM\tau(B))}{\lambda}} \geq 0. \quad (28)
\end{aligned}$$

Therefore, Eq. (28) always holds, when λ satisfies:

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

Then, we continue Eq. (23) and further have

$$\begin{aligned}
C_i^l(\tau) &\leq \lceil\frac{4(K+1)K^2}{(\Delta_{min}c_{min})^2}\ln(NM\tau(B))\rceil + \sum_{t=1}^{\tau} 2Kt^{-2} \\
&\leq \frac{4(K+1)K^2}{(\Delta_{min}c_{min})^2}\ln(NM\tau(B)) + 1 + \frac{K\pi^2}{3}. \quad (29)
\end{aligned}$$

Proof of Lemma 2: Due to the inequality $\ln\phi < \phi - 1$ for $\forall\phi > 0$, we first have the following expression:

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

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

$B/c^\star-1\leq \tau^\star(B)\leq B/c^\star$.

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

$$\tau(B)\leq \tau^\star(B)+\tau\Big(\sum_{p_i^l\notin\mathcal{P}^\star} n_i^l(\tau(B))c_{max}\Big)$$
$$\leq \tau^\star(B)+NL/(Kc_{min})\mathbb{E}[C_i^l(\tau(B))]. \quad (30)$$

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 $B^-=B-B^\star$ means the budget spent on the non-α-optimal options. Then, we have

$$\tau(B)=\tau(B^\star+B^-)\geq \tau(B^\star)\geq \tau^\star(B^\star)$$
$$\geq \tau^\star(B-\textstyle\sum_{p_i^l\notin\mathcal{P}^\star} n_i^l(\tau(B))c_{max})\geq \frac{B-NL\mathbb{E}[C_i^l(\tau(B))]}{c^\star}-1. \quad (31)$$

According to Eq. (30) and Eq. (30), we further get

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

By substituting the above results into Eq. 31, we get the lower bound on $\tau(B)$ as follows.

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

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

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

Proof of Theorem 2: 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

$$C_i^l(\tau)\leq\lambda+\sum_{t=1}^{\tau}\mathbb{I}\{u_{[\widehat{r}_i^{\,l}(t)]}(\mathcal{P}^{t+1})\geq u_{[\widehat{r}_i^{\,l}(t)]}(\mathcal{P}^\star), C_i^l(t)\geq\lambda\}$$
$$\leq \lambda+\sum_{t=1}^{\tau}\sum_{n_{s(1)}=\lambda}^{t}\cdots\sum_{n_{s(K)}=\lambda}^{t}\sum_{n_{s^\star(1)}=1}^{t}\cdots\sum_{n_{s^\star(K)}=1}^{t}$$
$$\mathbb{I}\Big\{\textstyle\sum_{i=1}^K\xi_i^l(t+1)\cdot\widehat{r}_{s(i)}^{\,l}(t)\geq\sum_{i=1}^K\xi_i^l(\star)\cdot\widehat{r}_{s^\star(i)}^{\,l}(t)\Big\}. \quad (32)$$

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. (24), Eq. (25) and Eq. (26), must hold. Now, we focus on the probability of the following case:

$$\mathbb{P}\Big\{\overline{r}_{s^\star(i)}^{\,l}(t)\leq r_{s^\star(i)}^l-\theta_{t,s^\star(i)}\Big\}$$
$$=\mathbb{P}\Big\{f_{s^\star(i)}^l\cdot\frac{\overline{q}_{s^\star(i)}(t)}{\overline{\varepsilon}_{s^\star(i)}(t)}\leq f_{s^\star(i)}^l\cdot\frac{q_{s^\star(i)}}{\varepsilon_{s^\star(i)}}-\theta_{t,s^\star(i)}\Big\}. \quad (33)$$

Actually, the event in Eq. (33) holds only when at least one of the following events must be true:

$$\overline{q}_{s^\star(i)}(t)\leq q_{s^\star(i)}-Q_{t,s^\star(i)}; \quad (34)$$
$$\overline{\varepsilon}_{s^\star(i)}(t)\geq \varepsilon_{s^\star(i)}+C_{t,s^\star(i)}. \quad (35)$$

We can prove this claim by the counter-evidence. That is, if both Eq. (34) and Eq. (35) are false, we have

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

According to the previous proof, we know $\mathbb{P}\{\text{Eq. (34)}\}\leq t^{-2(K+1)}$ and $\mathbb{P}\{\text{Eq. (35)}\}\leq t^{-2(K+1)}$. We continue Eq. (33) and have $\mathbb{P}\{\text{Eq. (33)}\}\leq 2t^{-2(K+1)}$. 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.,

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

Thus, we choose

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

such that Eq. (36) is impossible. Let's continue Eq. (32):

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

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

$$\tau(B)\leq\tau^\star(B)+\frac{NL}{Kc_{min}}\mathbb{E}[C_i^l(\tau(B))]$$
$$\leq B/c^\star+\tau(B)/2+\varphi_7\leq 2B/c^\star+2\varphi_7,$$

and get the lower bound as follows,

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

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

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

REFERENCES

[1] Y. Wei, Y. Zhu, H. Zhu, Q. Zhang, and G. Xue, "Truthful online double auctions for dynamic mobile crowdsourcing," in *IEEE INFOCOM*, 2015.

[2] 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.

[3] 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.

[4] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.

[5] 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.

[6] 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.

[7] S. He, D. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *IEEE INFOCOM*, 2014.

[8] Y. Chen, B. Li, and Q. Zhang, "Incentivizing crowdsourcing systems with network effects," in *IEEE INFOCOM*, 2016.

[9] 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.

[10] L. Pu, X. Chen, J. Xu, and X. Fu, "Crowdlet: Optimal worker recruitment for self-organized mobile crowdsourcing," in *IEEE INFOCOM*, 2016.

[11] 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.

[12] 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.

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

[14] M. Xiao, J. Wu, L. Huang, Y. Wang, and C. Liu, "Multi-task assignment for crowdsensing in mobile social networks," in *IEEE INFOCOM*, 2015.

[15] 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.

[16] 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.

[17] 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.

[18] L. Tran-Thanh, A. Chapman, A. Rogers, and N. R. Jennings, "Knapsack based optimal policies for budget–limited multi–armed bandits," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[19] S. Wang and L. Huang, "Multi-armed bandits with compensation," in *Advances in Neural Information Processing Systems*, 2018.

[20] 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.

[21] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial multi-armed bandit: General framework and applications," in *International Conference on Machine Learning*, 2013.

[22] 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.

[23] E. L. Lawler, "Fast approximation algorithms for knapsack problems," *Mathematics of Operations Research*, vol. 4, no. 4, pp. 339–356, 1979.

[24] 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.

[25] V. V. Vazirani, *Approximation algorithms*. Springer Science & Business Media, 2013.

[26] 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.

[27] 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.

[28] 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.

[29] H. Li, T. Li, and Y. Wang, "Dynamic participant recruitment of mobile crowd sensing for heterogeneous sensing tasks," in *IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, 2015.

[30] M. Xiao, J. Wu, S. Zhang, and J. Yu, "Secret-sharing-based secure user recruitment protocol for mobile crowdsensing," in *IEEE INFOCOM*, 2017.

[31] 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.

[32] M. Karaliopoulos, I. Koutsopoulos, and M. Titsias, "First learn then earn: Optimizing mobile crowdsensing campaigns through data-driven user profiling," in *ACM MobiHoc*, 2016.

[33] 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.

[34] 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*.

[35] Y. Xia, T. Qin, W. Ma, N. Yu, and T.-Y. Liu, "Budgeted multi-armed bandits with multiple plays." in *IJCAI*, 2016.

[36] 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.

[37] D. P. Zhou and C. J. Tomlin, "Budget-constrained multi-armed bandits with multiple plays," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.