Harnessing Context for Budget-Limited Crowdsensing with Massive Uncertain Workers

Feng Li, *Member, IEEE,* Jichao Zhao, Dongxiao Yu, *Member, IEEE,* Xiuzhen Cheng, *Fellow, IEEE,* and Weifeng Lv, *Member, IEEE,*

Abstract—Crowdsensing is an emerging paradigm of ubiquitous sensing, through which a crowd of workers are recruited to perform sensing tasks collaboratively. Although it has stimulated many applications, an open fundamental problem is how to select among a massive number of workers to perform a given sensing task under a limited budget. Nevertheless, due to the proliferation of smart devices equipped with various sensors, it is very difficult to profile the workers in terms of sensing ability. Although the uncertainties of the workers can be addressed by standard *Combinatorial Multi-Armed Bandit* (CMAB) framework through a trade-off between exploration and exploitation, we do not have sufficient allowance to directly explore and exploit the workers under the limited budget. Furthermore, since the sensor devices usually have quite limited resources, the workers may have bounded capabilities to perform the sensing task for only few times, which further restricts our opportunities to learn the uncertainty. To address the above issues, we propose a *Context-Aware Worker Selection* (CAWS) algorithm in this paper. By leveraging the correlation between the context information of the workers and their sensing abilities, CAWS aims at maximizing the expected total sensing revenue efficiently with both budget constraint and capacity constraints respected, even when the number of the uncertain workers are massive. The efficacy of CAWS can be verified by rigorous theoretical analysis and extensive experiments.

Index Terms—Multi-Armed Bandits, worker selection, crowdsensing

1 Introduction

Due to the proliferation of hand-held smart devices (e.g., smart phones, smart glasses, smart watches, etc) which are usually equipped with various sensors [1], the concept of crowdsensing has become a new paradigm for ubiquitous sensing [2]. Thousands or even millions of human crowds (a.k.a., *workers*) may be engaged in a sensing task (e.g., traffic information collection, ambient surveillance, urban business characterization, etc) with their sensor devices, and their collective contributions can be utilized to considerably improve sensing quality across a wide spectrum of applications [3].

By utilizing the crowdsensing paradigm, although there is no need to deploy specialized sensor devices to complete sensing tasks, thereby considerably reducing the overhead of data acquirement, the *requester* of a sensing task is still constrained by budget such that the requester only affords to recruit a limited number of workers. Therefore, how to select a set of high-qualified workers is a very crucial issue for the crowdsensing paradigm. There have been many existing studies exploring the combinatorial nature of the worker selection problem by assuming the workers' sensing abilities are known in advance [4, 5, 6, 7].

Unfortunately, due to the diversities of the sensor devices and human behaviors, the workers may have different sensing abilities to provide data with different qualities even for the same sensing task, while it is usually very difficult to pre-profile the heterogeneous workers to characterize their sensing abilities, especially considering the number of the workers may be huge. To address the uncertainties of the workers, one popular choice is to apply the *Combinatorial Multi-Armed Bandits* (CMAB) framework (e.g.,in [8, 9, 10]) such that the workers are sequentially selected to perform the sensing task under a budget. The essence of the CMAB framework is to leverage a trade-off between exploitation and exploration for each of workers [8, 10] (or each of worker-task combinations [9]). Hence, when there are a huge number of unknown workers, directly exploring and exploiting the workers results in significant overhead, while the total budget is limited. For one extreme example, if we do not have sufficient budget to select each of the workers once, the CMAB-based approaches even cannot be initialized. In addition, the bounded sensing capacities also restrict us from learning the sensing abilities of the workers individually. Specifically, the workers perform the task only limited times such that we do not have sufficient opportunities to explore and exploit them. In a nut shell, in this paper, we focus on addressing the following problem: *given a sensing task with limited budget and a massive number of unknown workers with limited capacities, how to fully utilize the budget and the capacities to explore and exploit the workers for task assignment?*

In this paper, we propose a *Context-Aware Worker Selection* (CAWS) algorithm. Specifically, by utilizing the correlation between workers' context information and their sensing abilities, we innovate in leveraging a CMAB framework to balance the exploitation and exploration in the context space rather than the massive workers. By learning the sensing ability distribution in the context space, we sequentially

- F. Li, J. Zhao, D. Yu and X. Cheng are with School of Computer Science and Technology, Shandong University, Qingdao, China.
E-mail: {fli, dxyu, xzcheng}@sdu.edu.cn, zhaojichao@mail.sdu.edu.cn
- W. Lv is with School of Computer Science and Engineering, Beihang University, Beijing, China.
E-mail: lwf@nlsde.buaa.edu.cn

Manuscript received April 19, 2005; revised August 26, 2015.

select the uncertain workers to efficiently maximize the expected sensing revenue with both strict budget constraint and capacity constraint respected. We conduct solid theoretical analysis to quantify the performance gap (a.k.a. *regret*) between our algorithm and the (nearly) optimal solution with the workers' sensing abilities known as prior. We also perform extensive experiments to verify the efficacy of CAWS. The main contribution of this paper is summarized as follows

- To the best of our knowledge, this is the first work considering the scalability of the uncertain worker selection in large-scale crowdsensing systems.
- We propose a context-aware worker selection algorithm to address the contradiction between the massive uncertain workers and the limit budget as well as the workers' bounded sensing capacities.
- We conduct rigorous theoretic analysis to quantify the regret between our CAWS algorithm and the approximately optimal one and perform extensive experiments (with both synthetic data and real data) to verify the advantages of CAWS over other state-of-the-art methods.

The remaining of our paper is organized as follows. We first introduce the system model and describe our problem in Sec. 2. We then present the details of our CAWS algorithm in Sec. 3. The analysis of our CAWS algorithm is given in Sec. 4. We report our experiment results in Sec. 5. We finally survey related literature and conclude this paper in Sec. 6 and Sec. 7, respectively.

2 System Model and Problem Description

2.1 System Model

We consider a crowdsensing process assigning a sensing task to a set of workers $\mathcal{N} = \{1, 2, \cdots, N\}$ under budget B. For each worker $i \in \mathcal{N}$, let c_i denote the cost to recruit (or select) worker i for one time to collect a data sample. Note that the cost parameters for different workers may be heterogeneous. We also define a capacity attribute τ_i for each worker $i \in \mathcal{N}$, which represents the maximum number of data samples worker i can contribute (or the maximum number of times worker i is selected). Let $c_{min} = \min_{i\in\mathcal{N}} c_i$, $c_{max} = \max_{i\in\mathcal{N}} c_i$ and $\tau_{max} = \max_{i\in\mathcal{N}} \tau_i$.

Each worker $i \in \mathcal{N}$ is associated with contextual information denoted by ϕ_i which is closely related to the worker's sensing ability. We assume that, for $\forall i \in \mathcal{N}$, $\phi_i \in \mathcal{S}$ is an M-dimensional vector, where $\mathcal{S} = [0, 1]^M$ is the so-called "*context space*". The context dimensions could be the proficiency of the workers in some required skills, the personal backgrounds of the workers or the performance parameters of the sensor devices, and we normalize each of the dimensions into a range of $[0, 1]$. We define a *stochastic* reward function $r : \mathcal{S} \to \{0, 1\}$. For $\forall i \in \mathcal{N}$, binary random variable $r(\phi_i) \in \{0, 1\}$ indicates if a data sample provided by worker i is qualified and thus represents the random *reward* obtained by selecting worker i to provide a qualified data sample. We assume $r(\phi_i)$ for each selection (and thus for each data sample) is identically and independently drawn from an *unknown* Bernoulli distribution and let $\mu_i = \mathbb{E}[r(\phi_i)]$ denote the unknown expected value of $r(\phi_i)$ [1]. In fact, μ_i is a measure of worker i's sensing ability. To facilitate our presentation, we suppose $r_i = r(\phi_i)$ and thus $\mu_i = \mathbb{E}[r_i]$ in the following.

2.2 Problem Description

Assuming $x_i \in \{0, 1, \cdots, \tau_i\}$ is the number of times we select worker i (i.e., the number of data samples we recruit worker i to provide), our problem can be formulated as

$$\max \quad f(\{x_i\}_{i=1}^N) = \sum_{i=1}^{N} \mu_i x_i \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} x_i c_i \le B, \tag{2}$$

$$x_i \in \{0, 1, 2, \cdots, \tau_i\}, \ \forall i \in \mathcal{N} \tag{3}$$

Our objective (1) is to maximize the expected cumulative revenue of our task assignment $\{x_i\}_{i=1}^N$, subject to budget constraint (2) and capacity constraints (3). In particular, the total cost of our task assignment cannot exceed the budget and each worker cannot be selected for more than τ_i times. It is apparent that, if μ_i (or $r(\cdot)$) was known as prior knowledge, our problem could be cast to a *Bounded Knapsack Problem* (BKP). Although the BKP is of NP-hardness, it can be addressed by many approximate algorithms efficiently [11]. For example, in the 2-approximation density-order greedy algorithm, we first sort the workers in decreasing order with respect to their densities $\rho_i = \mu_i / c_i$, and then greedily select the workers in the order until we do not have sufficient residual budget to select any available worker with non-zero residual capacity. In fact, as will shown in Sec. 3, we adapt this algorithm as a subroutine in our CAWS algorithm, where we sort the workers according to the estimates of their densities.

Unfortunately, it is usually very difficult to pre-profile the workers due to the huge number of workers as well as the diversity of sensor devices carried by the workers. Consequently, $\{\mu_i\}_{i=1}^N$ may not always be available as a prior, which makes our problem is much more difficult than the BKP. To address such uncertainties, one choice is to apply the CMAB framework. For example, in [12], the workers (corresponding to the arms) are explored and exploited through UCB indexing. Nevertheless, when there are a huge number of workers (and thus arms), leveraging the trade-off between exploration and exploitation directly among the workers results in considerable overhead. For example, in an extreme case where $\sum_{i=1}^N c_i > B$, we even do not have sufficient budget to select each of the workers for one time to initialize the workers' UCB indices. Furthermore, due to the workers' bounded capacities, we may do not have sufficient opportunities to explore and exploit the workers individually. Therefore, the problem is, *with unknown sensing abilities, how to efficiently select among the massive workers to maximize the expected total sensing revenue under limited budget and bounded capacities?* In this paper, we propose to utilize the correlation between context information and

1. Although we hereby assume $r(\phi_i)$ is an i.i.d. random variable obeying an unknown Bernoulli distribution parameterized by $\mu_i = \mathbb{P}(r(\phi_i) = 1)$, our algorithm is readily to work with arbitrary probability distributions with normalized supports in $[0, 1]$.

sensing ability, for the purpose of balancing exploration and exploitation among the workers in the context space.

3 ALGORITHM

Our CAWS algorithm is motivated by a common sense that workers with similar context may have similar sensing abilities for a certain type of sensing tasks (which is the main basis for our later theoretic analysis). We divide the context space $\mathcal{S}$ into d^M disjoint cubic sub-space (which are called "*hypercubes*" in the following). Each of the M-dimensional hypercubes is of identical size $\frac{1}{d} \times \frac{1}{d} \times \cdots \times \frac{1}{d}$ [2]. We denote by Ω the set of all hypercubes and by $Q_i \in \Omega$ the one such that $\phi_i \in Q_i$. As mentioned above, the workers in the same hypercube may have similar sensing abilities. Therefore, the essence of our CAWS algorithm is to leverage the trade-off between exploration and exploitation among the hypercubes rather than the workers. By learning the "sensing abilities" of the hypercubes, we can estimate the ones of the workers.

The pseudo-code of our CAWS algorithm is described in **Algorithm** 1. Our algorithm proceeds in iterations. We denote by $i(t) \in \mathcal{N}$ the worker selected in the t-th iteration and by $r_{i(t)}$ the reward yielded by this selection. For $\forall Q \in \Omega$, it is said that we choose Q in the t-th iteration if $\phi_{i(t)} \in Q$. We then denote by

$$\lambda_Q(t) = \sum_{t'=1}^{t} \mathbb{I}(\phi_{i(t')} \in Q) = \lambda_Q(t-1) + \mathbb{I}(\phi_{i(t)} \in Q) \quad (4)$$

the number of times Q is chosen up to the t-th iteration, where $\mathbb{I} : \{\text{True}, \text{False}\} \rightarrow \{1, 0\}$ is an indicator function. We also denote by

$$\begin{aligned} \bar{r}_Q(t) &= \frac{\sum_{t'=1}^{t} \mathbb{I}(\phi_{i(t')} \in Q) r_{i(t')}}{\lambda_Q(t)} \\ &= \frac{\bar{r}_Q(t-1)\lambda_Q(t-1) + \mathbb{I}(\phi_{i(t)} \in Q) r_{i(t)}}{\lambda_Q(t)} \end{aligned} \quad (5)$$

the average reward obtained up to the t-th iteration by choosing Q. At the beginning of the t-th iteration, we also let $B(t)$ be the residual budget and $\tau_i(t)$ be the residual capacity of worker $i \in \mathcal{N}$, which are initialized by $B(1) = B$ and by $\tau_i(1) = \tau_i$, respectively, as shown in Line 1. Worker $i \in \mathcal{N}$ is said to be *available* in the t-th iteration if $\tau_i(t) \geq 1$. Our algorithm proceeds only if there exists sufficient budget to select at least one available worker (see Line 2). In the first d^M iterations, we randomly choose a worker from each of the hypercubes, so as to initialize $\lambda_Q(t)$ and $\bar{r}_Q(t)$ for $\forall Q$ (see Lines 3 and 4). In the following, we use a density-ordered greedy subroutine (see **Algorithm** 2) to calculate a non-negative integral weight $x_i(t)$ for $\forall i \in \mathcal{N}$, which represents how many times we can (virtually) select worker i using residual budget $B(t)$ in a greedy manner (see Line 6). We then choose worker $i(t)$ with probability $\frac{x_i(t)}{\sum_{i'=1}^{N} x_{i'}(t)}$ (see Line 7) and increase $x_{i(t)}$ by one accordingly (see Line 8). Next, we update $\bar{r}_{Q_{i(t)}}(t)$ and $\lambda_{Q_{i(t)}}(t)$ for the hypercube $Q_{i(t)}$ (see Line 11). We finally renew the residual capacity of $i(t)$ and the residual budget (as shown in Lines 12 and 13, respectively) and proceed to the next iteration (see Line 14).

2. We will introduce how to partition the context space by choosing a proper value for d later in Sec. 4.

Algorithm 1: Our context-aware worker selection algorithm.

Input: $\{\tau_i, c_i, \phi_i\}_{i=1}^N$, B
Output: $\mathbf{x} = \{x_i\}_{i=N}^N$
1 $t = 1$; $B(t) = B$; $\tau_i(t) = \tau_i$ and $x_i = 0$ for $\forall i \in \mathcal{N}$;
2 **while** $B(t) \geq \min\{c_i \mid i \in \mathcal{N}, \tau_i(t) \geq 1\}$ **do**
3   **if** $t \leq d^M$ **then**
4   Randomly choose worker $i(t)$ in the t-th hypercube;
5   **else**
6   Call the density ordered greedy subroutine (see **Algorithm** 2) to calculate $\{x_i(t)\}_{i=1}^N$;
7   Choose worker $i(t) \in \mathcal{N}$ with probability $\frac{x_i(t)}{\sum_{i' \in \mathcal{N}} x_{i'}(t)}$;
8         $x_{i(t)} = x_{i(t)} + 1$;
9   **end**
10   Observe $r_{i(t)}$;
11   Update $\lambda_{Q_{i(t)}}(t)$ and $\bar{r}_{Q_{i(t)}}(t)$ according to (4) and (5), respectively;
12    $\tau_{i(t)}(t+1) = \tau_{i(t)}(t) - 1$;
13    $B(t+1) = B(t) - c_{i(t)}$;
14    $t = t + 1$;
15 **end**

As demonstrated in **Algorithm** 1, a density-ordered greedy subroutine is called in each iteration to calculate $x_i(t)$. The pseudo-code of the subroutine is given in **Algorithm** 2. Specifically, in the t-th iteration, we first calculate UCB index

$$U_i(t) = \bar{r}_{Q_i}(t-1) + \sqrt{\frac{2 \log t}{\lambda_{Q_i}(t-1)}} \quad (6)$$

for each worker i (see Line 1), and the workers are then sorted in decreasing order with respect to $\rho_i(t) = U_i(t)/c_i$. The UCB index $U_i(t)$ actually can be thought as an estimate on worker i's sensing ability. We greedily choose the workers with budget $B(t)$ in the order until there is no available workers or the residual budget is not sufficient for us to select any available workers (see Lines 3∼11).

4 ANALYSIS

As mentioned in Sec. 2.2, the BKP (1)∼(3) is NP-hard when $\{\mu_i\}_{i=1}^N$ are known as a prior. We now introduce a *rounding*-based approximation algorithm which can serve as a baseline to theoretically evaluate our CAWS algorithm. We first fractionalize the (integral) BKP as follows

$$\max \quad f(\{x_i\}_{i=1}^N) = \sum_{i=1}^{N} \mu_i x_i \quad (7)$$

$$\text{s.t.} \quad \sum_{i=1}^{N} x_i c_i \leq B, \; 0 \leq x_i \leq \tau_i, \; \forall i \in \mathcal{N} \quad (8)$$

and then round the fractional solution to an integral one. Compare with the (integral) BKP (1)∼(3), the only difference between them is that the variable x_i is a fractional non-negative number in the *Fractional BKP* (FBKP) rather than

Algorithm 2: Density-ordered greedy subroutine in the t-th iteration.

Input: $\{\tau_i(t), c_i\}_{i=1}^N$, $B(t)$, $\{\bar{r}_Q(t-1), \lambda_Q(t-1)\}_{Q\in\Omega}$
Output: $\mathbf{x}(t) = \{x_i(t)\}_{i\in\mathcal{N}}$
Calculate $U_i(t)$ for $\forall i \in \mathcal{N}$ according to (6);
Sort the workers $\mathcal{N}$ in decreasing order with respect to $\rho_i(t) = U_i(t)/c_i$;
$b = 0$;
for $i = 1, 2, \cdots, N$ **do**
  if $b + c_i \le B(t)$ **then**
    $x_i(t) = \min\left\{\tau_i(t), \left\lfloor \frac{B(t)-b}{c_i} \right\rfloor\right\}$;
    $b = b + c_i \cdot x_i(t)$;
  else
    $x_i(t) = 0$;
  end
end

an integral non-negative number in the BKP. The FBKP problem can be addressed by a density-ordered greedy approach. We first sort the workers in decreasing order with respect to their densities $\rho_i = \mu_i/c_i$ such that $\rho_1 \ge \rho_2 \ge \cdots \ge \rho_N$. Then, the optimal solution to the FBKP can be calculated as

$$x_i^* = \begin{cases} \tau_i, \ \ \forall i = 1, 2, \cdots, k-1 \\ \frac{B - \sum_{j=1}^{k-1} c_j \tau_j}{c_i}, \ \ i = k \\ 0, \ \ \forall i = k+1, k+2, \cdots, N \end{cases} \tag{9}$$

where the k-th worker is continuously "split" such that $\sum_{j=1}^{k-1} c_j\tau_j \le B$ and $\sum_{j=1}^{k} c_j\tau_j > B$. We finally round downward x_i^* for $\forall i \in \mathcal{N}$, and denote by $\lfloor \mathbf{x}^* \rfloor = \{\lfloor x_i^* \rfloor\}_{i=1}^N$ the resulting integral solution to the BKP. Letting f_{BKP}^* and f_{FBKP}^* be the optimal objective value of the BKP and the one of the FBKP, respectively, we have

$$\sum_{i=1}^N \mu_i \lfloor x_i^* \rfloor \le f_{BKP}^* \le f_{FBKP}^* \le \sum_{i=1}^N \mu_i \lfloor x_i^* \rfloor + \mu_k \tag{10}$$

It is shown that the gap between the lower bound of f_{BKP}^* and its upper bound is constrained; hence, it is rational to use the lower bound $\sum_{i=1}^N \mu_i \lfloor x_i^* \rfloor$ as the baseline to evaluate the performance of our algorithm. Specifically, given time horizon T, we are interested in investigating the following regret function

$$\mathsf{Regret}(T, \{i(t)\}_{t=1}^T) = \sum_{j=1}^N \mu_j \lfloor x_j^* \rfloor - \sum_{j=1}^N \mu_j \mathbb{E}_{T,\{i(t)\}_{t=1}^T}[x_i] \tag{11}$$

which indicates the gap between the expected cumulative reward yielded by the (nearly) optimal solution $\lfloor \mathbf{x}^* \rfloor$ and the expected one produced by the solution of our algorithm $\mathbf{x}$. In the following, we will show the upper-bound of the above regret function.

As mentioned in Sec. 3, our CAWS algorithm is based on the natural assumption that the workers with similar context could have similar sensing abilities. This assumption can be formalized by the following Hölder condition.

Assumption 1 (Hölder Condition)**.** *There exist $L > 0$ and $\alpha > 0$ such that for any contexts $s, s' \in \mathcal{S}$, it holds that*

$$|\mathbb{E}[r(s)] - \mathbb{E}[[r(s')]| \le L\|s - s'\|^\alpha \tag{12}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^M$.

It should be noted that our CAWS algorithm still works if the assumption does not strictly hold. However, the regret might not be bounded if the assumption was violated.

Lemma 1. *For $\forall i, i' \in \mathcal{N}$ such that $Q_i = Q_{i'}$, we have*

$$|\mu_i - \mu_{i'}| \le \Delta = L\left(M^{\frac{1}{2}} d^{-1}\right)^\alpha \tag{13}$$

Proof. Since the workers i and i' have their contexts in the same hypercube, we have $\|\phi_i - \phi_{i'}\| \le M^{\frac{1}{2}} d^{-1}$ according to our strategy of evenly partitioning the context space (as depicted in Sec. 3). Then, considering the Hölder condition shown above, we have $|\mu_i - \mu_{i'}| = |\mathbb{E}[r(\phi_i)] - \mathbb{E}[r(\phi_{i'})]| \le L\|\phi_i - \phi_{i'}\|^\alpha = L\left(M^{\frac{1}{2}} d^{-1}\right)^\alpha$ □

For each hypercube $Q \in \Omega$, we denote by μ_Q the expected reward yielded by selecting the workers in Q (i.e., the "sensing ability" of the hypercube Q). It is apparent that $|\mu_i - \mu_{Q_i}| \le \Delta$ for $\forall i \in \mathcal{N}$, which implies μ_{Q_i} can be used as an estimate on μ_i.

The definition of the regret function suggests the key of our analysis should be to quantify the impact of mischoosing workers on the sensing revenue. The reason for the regret is two-fold: on one hand, we leverage the qualities of the contextual hypercubes to estimate the ones of the workers such that we may not be able to make "right" selection decisions even we learn μ_Q exactly; on the other hand, according to MAB theory, we learn the qualities of the contextual hypercubes through a trade-off between exploration and exploitation, while making "wrong" selection decisions is the price we have to pay for the learning process. Therefore, supposing $\lfloor \tilde{\mathbf{x}}^* \rfloor = \{\lfloor \tilde{x}_i^* \rfloor\}_{i=1}^N$ is the solution obtained by applying the rounding-based density-ordered greedy algorithm to BKP instance $(\{i, \mu_{Q_i}, c_i, \tau_i\}_{i=1}^N, B)$ (where we use μ_{Q_i} as an estimate on μ_i), we decompose the regret function as follows

$$\begin{aligned}
&\mathsf{Regret}(T, \{i(t)\}_{t=1}^T) \\
=& \sum_{j=1}^N \mu_j \lfloor x_j^* \rfloor - \sum_{j=1}^N \mu_j \lfloor \tilde{x}_j^* \rfloor \\
&+ \sum_{j=1}^N \mu_j \lfloor \tilde{x}_j^* \rfloor - \sum_{j=1}^N \mu_j \mathbb{E}_{T,\{i(t)\}_{t=1}^T}[x_j] \\
\le& \sum_{j=1}^N \mu_j \lfloor x_j^* \rfloor - \sum_{j=1}^N \mu_j \lfloor \tilde{x}_j^* \rfloor + \sum_{j=1}^N (\mu_{Q_j} + \Delta) \lfloor \tilde{x}_j^* \rfloor \\
&- \sum_{j=1}^N (\mu_{Q_j} - \Delta) \mathbb{E}_{T,\{i(t)\}_{t=1}^T}[x_j] \\
\le& \sum_{j=1}^N \mu_j \lfloor x_j^* \rfloor - \sum_{j=1}^N \mu_j \lfloor \tilde{x}_j^* \rfloor + \sum_{j=1}^N \mu_{Q_j} \lfloor \tilde{x}_j^* \rfloor \\
&- \sum_{j=1}^N \mu_{Q_j} \mathbb{E}_{T,\{i(t)\}_{t=1}^T}[x_j] + \frac{2B\Delta}{c_{min}}
\end{aligned} \tag{14}$$

where we have the second inequality since $|\mu_i - \mu_{Q_i}| \leq \Delta$ holds for $\forall i \in \mathcal{N}$ as mentioned above and the third one due to the fact that $\sum_{j=1}^{N} \lfloor \tilde{x}_j^* \rfloor \leq \frac{B}{c_{min}}$ and $\sum_{j=1}^{N} \mathbb{E}_{T,\{i(t)\}_{t=1}^{T}}[x_j] \leq \frac{B}{c_{min}}$. By defining

$$\mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N}) = \sum_{j=1}^{N} \mu_j \lfloor x_j^* \rfloor - \sum_{j=1}^{N} \mu_j \lfloor \tilde{x}_j^* \rfloor \quad (15)$$

and

$$\begin{aligned} &\mathsf{Regret}(\lfloor \tilde{\mathbf{x}}^* \rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^{N}) \\ &= \sum_{j=1}^{N} \mu_{Q_j} \lfloor \tilde{x}_i^* \rfloor - \sum_{j=1}^{N} \mu_{Q_j} \mathbb{E}_{T,\{i(t)\}_{t=1}^{T}}[x_j] \end{aligned} \quad (16)$$

the regret function $\mathsf{Regret}(T, \{i(t)\}_{t=1}^{T})$ can be re-written as

$$\begin{aligned} &\mathsf{Regret}(T, \{i(t)\}_{t=1}^{T}) \leq \mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N}) \\ &+ \mathsf{Regret}(\lfloor \tilde{\mathbf{x}}^* \rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^{N}) + \frac{2B\Delta}{c_{min}} \end{aligned} \quad (17)$$

$\mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N})$ represents the loss due to our partition of the contextual space, while $\mathsf{Regret}(\lfloor \tilde{\mathbf{x}}^* \rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^{N})$ indicates the one resulting from our learning process. In the following, we first present the main result showing the upper-bound of the regret function (17) in Sec. 4.1 and then report the details of the proof in Sec. 4.2, by bounding the two sub-regret functions (15) and (16), respectively.

To facilitate our analysis, we reuse the notion $\in$ when there is no ambiguity, such that for each worker i, it is said that $i \in \lfloor \mathbf{x}^* \rfloor$ (resp. $i \in \lfloor \tilde{\mathbf{x}}^* \rfloor$) if $\lfloor x_i^* \rfloor \geq 1$ (resp. $\lfloor \tilde{x}_i^* \rfloor \geq 1$). We also give some notions as follows which will be useful to our later analysis.

$$i^* = \arg\max_{i \in \mathcal{N}} \frac{\mu_{Q_i}}{c_i} \quad (18)$$

$$\mathcal{N}_Q = \{i \in \mathcal{N} \mid \phi_i \in Q\} \quad (19)$$

$$\mathcal{N}_Q^+ = \{i \in \mathcal{N} \mid \phi_i \in Q, i \in \lfloor \tilde{\mathbf{x}}^* \rfloor\} \quad (20)$$

$$\mathcal{N}_Q^- = \{i \in \mathcal{N} \mid \phi_i \in Q, i \notin \lfloor \tilde{\mathbf{x}}^* \rfloor\} \quad (21)$$

$$c_{max}(\mathcal{N}_Q^+) = \max_{i \in \mathcal{N}_Q^+} c_i,\ c_{min}(\mathcal{N}_Q^+) = \min_{i \in \mathcal{N}_Q^+} c_i \quad (22)$$

$$c_{max}(\mathcal{N}_Q^-) = \max_{i \in \mathcal{N}_Q^+} c_i,\ c_{min}(\mathcal{N}_Q^-) = \min_{i \in \mathcal{N}_Q^+} c_i \quad (23)$$

$$\delta_{min} = \min_{Q,Q' \in \Omega} \left| \frac{\mu_Q}{c_{min}(\mathcal{N}_Q^-)} - \frac{\mu_{Q'}}{c_{max}(\mathcal{N}_Q^+)} \right| \quad (24)$$

$$\xi = \frac{8}{c_{min}^2 \delta_{min}^2} + \left(\frac{c_{max}}{c_{min}}\right)^2 \quad (25)$$

4.1 Main Result

Theorem 1. *Assuming* $d = \left\lceil B^{\frac{1}{\alpha+M}} \right\rceil$, *the regret function of our CAWS algorithm (11) is upper-bounded by*

$$\left(\tau_{max} + 2^M B^{\frac{M}{\alpha+M}} h(\ln B) + 1\right) \frac{c_{max}}{c_{min}} + \frac{4LM^{\frac{\alpha}{2}} B^{\frac{M}{\alpha+M}}}{c_{min}} + 1 \quad (26)$$

where

$$h(\ln B) = \xi \ln\left(\frac{B}{c_{min}}\right) + \frac{\pi^2}{3} + 1 \quad (27)$$

which implies that the regret for our CAWS algorithm is $\mathcal{O}\left(B^{\frac{M}{\alpha+M}} \ln B\right)$.

4.2 Detailed Proof

We first prove $\mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N})$ is upper-bounded in **Theorem** 2.

Theorem 2. *Recall that* $\lfloor \mathbf{x}^* \rfloor = \{\lfloor x_i^* \rfloor\}_{i=1}^{N}$ *and* $\lfloor \tilde{\mathbf{x}}^* \rfloor = \{\lfloor \tilde{x}_i^* \rfloor\}_{i=1}^{N}$ *be the results we obtain by applying the rounding-based density-ordered greedy algorithm to the two BKP instances* Instance1 $= (\{i, \mu_i, c_i, \tau_i\}_{i=1}^{N}, B)$ *and* Instance2 $= (\{i, \mu_{Q_i}, c_i, \tau_i\}_{i=1}^{N}, B)$, *respectively. Considering* μ_{Q_i} *is an estimate on* μ_i *for* $\forall i \in \mathcal{N}$, *we have*

$$\mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N}) \leq \frac{2\Delta B}{c_{min}} + 1 \quad (28)$$

Proof. We denote by $\mathbf{x}^* = \{x_i^*\}_{i=1}^{N}$ and $\tilde{\mathbf{x}}^* = \{\tilde{x}_i^*\}_{i=1}^{N}$ the fractional solutions to the FBKP versions of Instance1 and Instance2, respectively. Considering the inequality (10),

$$\begin{aligned} &\mathsf{Regret}(\lfloor \mathbf{x}^* \rfloor, \lfloor \tilde{\mathbf{x}}^* \rfloor, \{\mu_i\}_{i=1}^{N}) \\ &\leq \sum_{i=1}^{N} \mu_i x_i^* - \left(\sum_{i=1}^{N} \mu_i \tilde{x}_i^* - \mu_{Q_{\tilde{k}}}\right) \leq \sum_{i=1}^{N} \mu_i x_i^* - \sum_{i=1}^{N} \mu_i \tilde{x}_i^* + 1 \end{aligned} \quad (29)$$

where $\tilde{k}$ is the split worker in Instance2 and $\mu_{\tilde{k}} \leq 1$. According to the procedure of our rounding-based density-ordered greedy algorithm shown in Sec. 4, if there is a worker i with $x_i^* > \tilde{x}_i^*$, there must be at least another worker i' with $x_{i'}^* < \tilde{x}_{i'}^*$ such that $\frac{\mu_{i'}}{c_{i'}} \leq \frac{\mu_i}{c_i}$ and $\frac{\mu_{Q_{i'}}}{c_{i'}} \geq \frac{\mu_{Q_i}}{c_i}$. Therefore,

$$\begin{aligned} \frac{\mu_i}{c_i} - \frac{\mu_{i'}}{c_{i'}} &\leq \frac{\mu_{Q_i} + \Delta}{c_i} - \frac{\mu_{Q_{i'}} - \Delta}{c_{i'}} \\ &= \frac{\mu_{Q_i}}{c_i} - \frac{\mu_{Q_{i'}}}{c_{i'}} + \Delta\left(\frac{1}{c_i} + \frac{1}{c_{i'}}\right) \leq \frac{2\Delta}{c_{min}} \end{aligned} \quad (30)$$

where we have the first inequality due to $|\mu_i - \mu_{Q_i}| \leq \Delta$ for $\forall i \in \mathcal{N}$ (see **Lemma** 1) and the third one by considering the facts that $\frac{\mu_{Q_{i'}}}{c_{i'}} \geq \frac{\mu_{Q_i}}{c_i}$ and $c_i \geq c_{min}$ for $\forall i \in \mathcal{N}$. In other words, if worker i is not (fractionally) selected in Instance2, our algorithm will select some other workers to replace worker i in $\tilde{\mathbf{x}}^*$. Nevertheless, since the workers (selected to replace worker i) have smaller densities than i, these replacements may result in the reward loss, which can be indicated by the difference between the first two terms in (29). Hence, assuming $\widetilde{\mathcal{N}_i}$ denote the set of those workers and $\tilde{i} = \arg\min_{j \in \widetilde{\mathcal{N}_i}} \frac{\mu_{Q_j}}{c_j}$, we have

$$\begin{aligned} &\sum_{i=1}^{N} \mu_i x_i^* - \sum_{i=1}^{N} \mu_i \tilde{x}_i^* \\ \leq\ &\sum_{i: x_i^* > \tilde{x}_i^*} \left((x_i^* - \tilde{x}_i^*)\, \mu_i - \frac{(x_i^* - \tilde{x}_i^*) c_i}{c_{\tilde{i}}} \cdot \mu_{\tilde{i}}\right) \\ =\ &\sum_{i: x_i^* > \tilde{x}_i^*} \left(c_i (x_i^* - \tilde{x}_i^*) \left(\frac{\mu_i}{c_i} - \frac{\mu_{\tilde{i}}}{c_{\tilde{i}}}\right)\right) \end{aligned} \quad (31)$$

Considering $\frac{\mu_i}{c_i} - \frac{\mu_{\tilde{i}}}{c_{\tilde{i}}} \leq \frac{2\Delta}{c_{min}}$ (see (30)), we have

$$\sum_{i=1}^{N} \mu_i x_i^* - \sum_{i=1}^{N} \mu_i \tilde{x}_i^* \leq \frac{2\Delta}{c_{min}} \sum_{i: x_i^* > \tilde{x}_i^*} c_i (x_i^* - \tilde{x}_i^*) \leq \frac{2\Delta B}{c_{min}}$$

by substituting which into (29), we complete the proof. □

In the following, we prove the upper-bound of the sub-regret function $\text{Regret}(\lfloor\tilde{\mathbf{x}}^*\rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N)$. We introduce an redundant term $\mathbb{E}_T[T]\mu_{Q_{i^*}}$ such that

$$\begin{aligned}
&\text{Regret}\left(\lfloor\tilde{\mathbf{x}}^*\rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N\right) \\
&= \mathbb{E}_T\left[\sum_{j=1}^N \mu_{Q_j}\lfloor\tilde{x}_i^*\rfloor - T\mu_{Q_{i^*}} + T\mu_{Q_{i^*}}\right. \\
&\quad \left. - \sum_{Q\in\Omega}\sum_{i\in\mathcal{N}_Q}\mu_Q\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_i \mid T]\right] \\
&= \sum_{j=1}^N \mu_{Q_j}\lfloor\tilde{x}_i^*\rfloor - \mu_{Q_{i^*}}\mathbb{E}_T[T] \\
&+ \mathbb{E}_T\left[T\mu_{Q_{i^*}} - \sum_{Q\in\Omega}\sum_{i\in\mathcal{N}_Q}\mu_Q\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_i \mid T]\right] \qquad (32)
\end{aligned}$$

Apparently, the key to bounding the above sub-regret function is to figure out the lower-bound of the second term and the upper-bound of the third one (see **Lemma** 2).

Theorem 3. *Letting $\lfloor\tilde{\mathbf{x}}^*\rfloor$ be the results we obtain by applying the rounding-based density-ordered greedy algorithm to the BKP instances $(\{i, \mu_{Q_i}, c_i, \tau_i\}_{i=1}^N, B)$ and $\mathbf{x}^*$ be the output of our CAWS algorithm, we have*

$$\begin{aligned}
&\text{Regret}\left(\lfloor\tilde{\mathbf{x}}^*\rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N\right) \\
&\le \frac{c_{max}}{c_{min}}\left(\tau_{max} + d^M h(\ln B) + 1\right) \qquad (33)
\end{aligned}$$

Proof. We first introduce **Lemma** 2 which will be helpful in our later derivation. Due to the space limit, the proofs of **Lemma** 2 can be found in **Appendix A**.

Lemma 2. *Supposing T denotes the total number of the iterations our CAWS proceeds with budget B, we have the following two inequalities holds*

$$\begin{aligned}
&\mathbb{E}_T[T] \\
&\ge \frac{B - c_{max}}{c_{i^*}} - \mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\frac{c_j - c_{i^*}}{c_{i^*}}\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \\
&- \sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}h(\ln B) \qquad (34)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}_T\left[T\mu_{Q_{i^*}} - \sum_{Q\in\Omega}\sum_{i\in\mathcal{N}_Q}\mu_Q\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_i \mid T]\right] \\
&\le \sum_{Q:\mu_{Q_{i^*}}>\mu_Q}(\mu_{Q_{i^*}} - \mu_Q)h(\ln B) \\
&\quad - \mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}} - \mu_{Q_j})\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \qquad (35)
\end{aligned}$$

By substituting the above two inequalities (34) and (35) into (32), we have

$$\begin{aligned}
&\text{Regret}\left(\lfloor\tilde{\mathbf{x}}^*\rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N\right) \\
= &\sum_{Q\in\Omega}\sum_{j\in\Omega}\mu_Q\lfloor\tilde{x}_j^*\rfloor - \mathbb{E}_T[T]\mu_{Q_{i^*}} \\
&+\mathbb{E}_T\left[T\mu_{Q_{i^*}} - \sum_{Q\in\Omega}\sum_{i\in\mathcal{N}_Q}\mu_Q\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_i \mid T]\right] \\
\le &\sum_{Q\in\Omega}\sum_{j\in\Omega}\mu_Q\lfloor\tilde{x}_j^*\rfloor - \frac{\mu_{Q_{i^*}}(B - c_{max})}{c_{i^*}} \\
&+\mu_{Q_{i^*}}\mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\frac{c_j - c_{i^*}}{c_{i^*}}\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \\
&+\mu_{Q_{i^*}}\sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}h(\ln B) \\
&+\sum_{Q:\mu_{Q_{i^*}}>\mu_Q}(\mu_{Q_{i^*}} - \mu_Q)h(\ln B) \\
&+\mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}} - \mu_{Q_j})\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \\
\le &\sum_{Q\in\Omega}\sum_{j\in\Omega}\mu_Q\lfloor\tilde{x}_j^*\rfloor - \frac{B\mu_{Q_{i^*}}}{c_{i^*}} + \frac{c_{max}\mu_{Q_{i^*}}}{c_{i^*}} \\
&+\sum_{Q\in\Omega}g\cdot h(\ln B) \\
&+\mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\left(\frac{\mu_{Q_{i^*}}c_j}{c_{i^*}} - \mu_{Q_j}\right)E_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \\
&\qquad (36)
\end{aligned}$$

where

$$\begin{aligned}
g \;=\; &\mathbb{I}(c_{max}(\mathcal{N}_Q^-) > c_{i^*})\cdot\frac{\mu_{Q_{i^*}}(c_{max}(\mathcal{N}_Q^-) - c_{i^*})}{c_{i^*}} \\
&+\mathbb{I}(\mu_{Q_{i^*}} - \mu_Q > 0)\cdot(\mu_{Q_{i^*}} - \mu_Q) \qquad (37)
\end{aligned}$$

Since $B \ge \sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}c_j\lfloor\tilde{x}_j^*\rfloor$, we have

$$\begin{aligned}
&\sum_{Q\in\Omega}\sum_{j\in\Omega}\mu_Q\lfloor\tilde{x}_j^*\rfloor - \frac{B\mu_{Q_{i^*}}}{c_{i^*}} \\
\le &\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\mu_{Q_j}\lfloor\tilde{x}_j^*\rfloor - \frac{\mu_{Q_{i^*}}\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}c_j\lfloor\tilde{x}_j^*\rfloor}{c_{i^*}} \\
= &\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\left(\mu_{Q_j} - \frac{\mu_{Q_{i^*}}c_j}{c_{i^*}}\right)\lfloor\tilde{x}_j^*\rfloor \qquad (38)
\end{aligned}$$

In addition, since $0 \le \mu_{Q_{i^*}} \le 1$, $c_{max}(\mathcal{N}_Q^-) - c_{i^*} \le c_{max} - c_{min}$ and $\mu_{Q_{i^*}} - \mu_Q \le 1$, we have

$$g \le \frac{c_{max} - c_{min}}{c_{min}} + 1 = \frac{c_{max}}{c_{min}} \qquad (39)$$

Substituting the above two inequalities (38) and (39) into

(36), we have

$$\begin{aligned}
&\mathsf{Regret}\left(\lfloor \tilde{\mathbf{x}}^* \rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N\right) \\
\leq &\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \left(\mu_{Q_j} - \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}}\right) \lfloor \tilde{x}_j^* \rfloor + \frac{c_{max}\mu_{Q_{i^*}}}{c_{i^*}} \\
&+ d^M \frac{c_{max}}{c_{min}} \left(\xi \ln T + \frac{\pi^2}{3} + 1\right) \\
&+ \mathbb{E}_T\Bigg[\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \left(\frac{\mu_{Q_{i^*}}(c_j - c_{i^*})}{c_{i^*}} + (\mu_{Q_{i^*}} - \mu_{Q_j})\right) \\
&\qquad \mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\Bigg] \\
= &\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \left(\mu_{Q_j} - \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}}\right) \lfloor \tilde{x}_j^* \rfloor + \frac{c_{max}\mu_{Q_{i^*}}}{c_{i^*}} \\
&+ d^M \frac{c_{max}}{c_{min}} \left(\xi \ln\left(\frac{B}{c_{min}}\right) + \frac{\pi^2}{3} + 1\right) \\
&- \mathbb{E}_T\left[\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \left(\mu_{Q_j} - \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}}\right) \mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]\right] \\
\leq &\,\mathbb{E}_T\left[\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T] - \lfloor \tilde{x}_j^* \rfloor\right)\right] \\
&+ d^M \frac{c_{max}}{c_{min}} \left(\xi \ln\left(\frac{B}{c_{min}}\right) + \frac{\pi^2}{3} + 1\right) + \frac{c_{max}\mu_{Q_{i^*}}}{c_{i^*}} \qquad (40)
\end{aligned}$$

where we have the last inequality holds by considering $\frac{\mu_{Q_{i^*}}}{c_{i^*}} \geq \frac{\mu_{Q_j}}{c_j}$ for $\forall j \in \mathcal{N}$.

The first term at the right side of the above inequality can be written as

$$\begin{aligned}
&\mathbb{E}_T\left[\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T] - \lfloor \tilde{x}_j^* \rfloor\right)\right] \\
= &\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_T[\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T]] - \lfloor \tilde{x}_j^* \rfloor\right) \\
= &\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_T[\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j - \lfloor \tilde{x}_j^* \rfloor \mid T]]\right) \\
\leq &\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor : x_j > \lfloor \tilde{x}_j^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_T[\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j - \lfloor \tilde{x}_j^* \rfloor \mid T]]\right)
\end{aligned} \qquad (41)$$

Suppose $\tilde{k}$ denotes the split worker in $\lfloor \tilde{\mathbf{x}}^* \rfloor$. For any worker j such that $j \in \lfloor \tilde{\mathbf{x}}^* \rfloor$ and $j \neq \tilde{k}$, we have $\lfloor \tilde{x}_j^* \rfloor = \tau_j$, while $\lfloor \tilde{x}_{\tilde{k}}^* \rfloor \leq \tau_{\tilde{k}}$. Therefore, for $\forall j \in \lfloor \tilde{\mathbf{x}}^* \rfloor$, $x_j - \lfloor \tilde{x}_j^* \rfloor \leq 0$, and the split worker $\tilde{k}$ is the only possible one such that $x_{\tilde{k}} - \lfloor \tilde{x}_{\tilde{k}}^* \rfloor \geq 0$ may hold. Also, since $c_j \leq c_{max}$ for $\forall j \in \mathcal{N}$, continuing the above equation (41), we have

$$\begin{aligned}
&\mathbb{E}_T\left[\sum_{j\in\lfloor \tilde{\mathbf{x}}^* \rfloor} \frac{\mu_{Q_{i^*}} c_j}{c_{i^*}} \left(\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j \mid T] - \lfloor \tilde{x}_j^* \rfloor\right)\right] \\
\leq &\frac{\tau_{max}\mu_{Q_{i^*}} c_{max}}{c_{i^*}}
\end{aligned} \qquad (42)$$

We complete the proof by substituting (42) into (40) as follows

$$\begin{aligned}
&\mathsf{Regret}\left(\lfloor \tilde{\mathbf{x}}^* \rfloor, \mathbf{x}, \{\mu_{Q_i}\}_{i=1}^N\right) \\
\leq &\frac{\tau_{max}\mu_{Q_{i^*}} c_{max}}{c_{i^*}} + \frac{c_{max}\mu_{Q_{i^*}}}{c_{i^*}} + d^M \frac{c_{max}}{c_{min}} h(\ln B) \\
\leq &(\tau_{max} + 1)\frac{c_{max}}{c_{min}} + d^M \frac{c_{max}}{c_{min}} h(\ln B) \qquad (43)
\end{aligned}$$

□

Now, we are ready to prove our main result shown in **Theorem** 1. Combining **Theorem** 2 and **Theorem** 3 into (17), we have

$$\begin{aligned}
&\mathsf{Regret}(T, \{i(t)\}_{t=1}^T) \\
\leq &\left(\tau_{max} + d^M h(\ln B) + 1\right) \frac{c_{max}}{c_{min}} + \frac{4\Delta B}{c_{min}} + 1 \qquad (44)
\end{aligned}$$

Letting $d = \left\lceil B^{\frac{1}{\alpha+M}} \right\rceil$, we have

$$d^M = \left\lceil B^{\frac{1}{\alpha+M}} \right\rceil^M \leq 2^M B^{\frac{M}{\alpha+M}} \qquad (45)$$

and

$$\Delta = L\left(M^{\frac{1}{2}} d^{-1}\right)^\alpha \leq L M^{\frac{\alpha}{2}} B^{-\frac{\alpha}{\alpha+M}} \qquad (46)$$

when $\alpha > 0$ as shown in the **Assumption** 1. We finally complete the proof of our main result by substituting (45) and (46) into (44).

5 EXPERIMENTS

In this section, we evaluate the performance of our CAWS algorithm through extensive experiments. We first introduce the reference algorithms in Sec. 5.1 and then compare them with our CAWS algorithm using both synthetic dataset and real dataset in Sec. 5.2 and Sec. 5.3, respectively.

5.1 Reference Algorithms

We mainly compare our CAWS algorithm with the following ones which can be applied to our problem.

- **Oracle**: Oracle is aware of the sensing abilities of the workers; therefore, it applies the density-ordered greedy algorithm to output a nearly optimal solution.
- **Bounded ϵ-first**: The bounded ϵ-first algorithm is with decoupled exploitation and exploration [13]. Under a ϵ-fraction of the budget, it explores the workers uniformly to estimate their sensing abilities; while with the remaining budget, it assigns the task to the workers according to their estimated sensing abilities according to the density-ordered greedy algorithm.
- **B-KUBE**: B-KUBE is a CMAB-based algorithm to handle BKP, where the workers with unknown expected sensing abilities and bounded sensing capacities are selected under a given budget [12]. It can be considered to be a degeneration of CAWS where the context space is sufficiently partitioned such that each hypercube contains only one worker. Our CAWS algorithm is then degraded to that we estimate the workers' sensing abilities directly by their UCB indices which are calculated according to their historical performances.

- **Random**: The (purely) random algorithm selects an available worker (whose residential capacity is non-zero) uniformly in each iteration until the budget is exhausted or none of the workers have non-zero residual capacity.

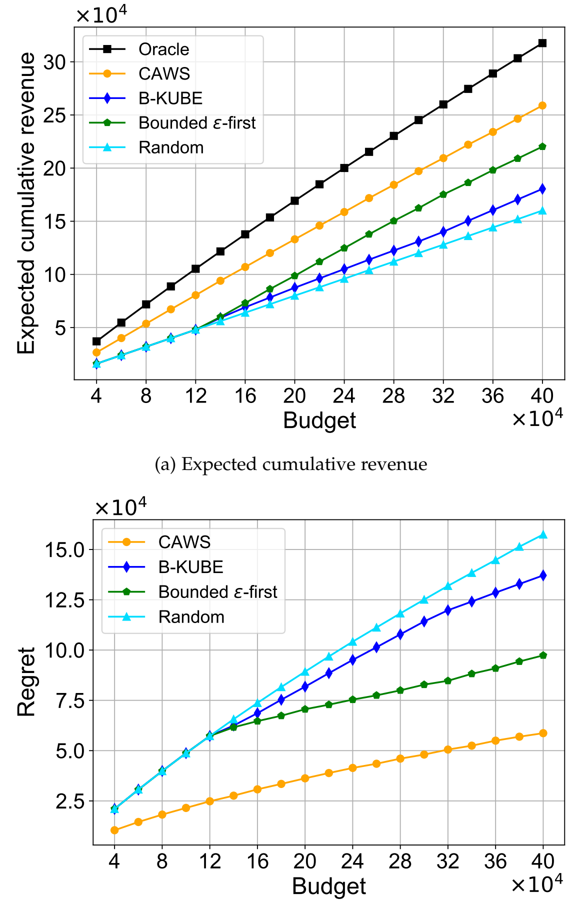
5.2 Evaluation with Synthetic Data

We first quantitatively evaluate the above algorithms in terms of expected cumulative revenue by synthetic data. We conduct our simulations by assuming there are 10^5 workers whose capacities and costs are distributed uniformly in $[20, 40]$ and $[1, 1.5]$, respectively. We suppose the context space $\mathcal{S}$ has $M = 2$ dimensions and each dimension is normalized in $[0, 1]$ as mentioned in Sec. 2.1. The workers have their contexts uniformly distributed in the context space $\mathcal{S}$. We also randomly set the workers' sensing abilities such that the Hölder condition holds for $\alpha = 1$ for the purpose of quantitative analysis (e.g., one choice is to let each worker has its sensing ability being the average of its context values).

We vary the budget from 4×10^4 to 4×10^5 with a step size 4×10^4 and report the results in Fig. 1. Note that in this setting, the budget is only at most 4 times higher (or even smaller) than the number of the workers. In a nutshell, compared with N, B is quite limited. It is shown in Fig. 1 (a) that our CAWS algorithm yields higher expected cumulative revenue than the other ones, since CAWS fully utilizes the context information of the workers such that we can effectively estimate the workers' sensing abilities according to the context information even we do not have sufficient budget to fully exploit them. Furthermore, the performance of our algorithm is very close to the one of the oracle algorithm, especially when the budget is limited.

We also plot the regrets of the different algorithms in Fig. 1(b). Since the regret function of the oracle algorithm is always almost zero, we do not show it in Fig. 1(b). Consistent with what has been shown in Fig. 1(a), our CAWS algorithm has a much lower regret than the other three alternatives. When the budget is increased, our algorithm proceeds more iterations such that the regret is increased but at a very low rate, which is consistent with our main theoretical result in **Theorem** 1.

We then fix the budget $B = 1 \times 10^5$ and vary the number of the workers $N = 4, 6, 8, 10 \times 10^4$ to show the scalabilities of the different algorithms. The results in terms of expected cumulative revenue are presented in Fig. 2. In a given context space, when there are more workers to "fulfill" the context space, our CAWS algorithm has a better utilization of the context information to estimate the worker's sensing abilities more accurately. Therefore, with an increasing number of workers, our CAWS algorithm yields more expected cumulative revenue. In contrast, the performances of the other algorithms are degraded in face of a large number of workers, as they have no sufficient budget to explore and exploit the workers.



(a) Expected cumulative revenue

(b) Regret

Fig. 1. Comparisons of different algorithms under varying budget settings with synthetic data.

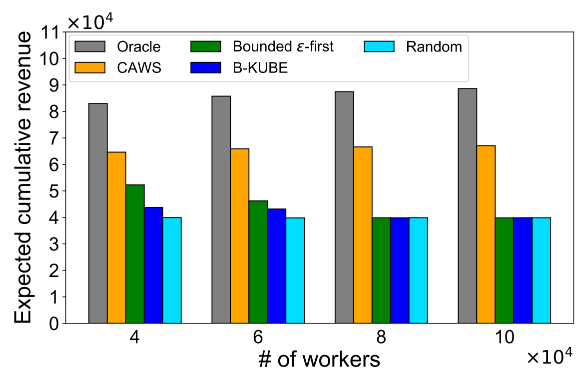


Fig. 2. Comparisons of different algorithms with different numbers of workers.

5.3 Evaluation with Real Data

In this section, we evaluate the performance of our CAWS algorithm in a crowdsensing application based on the dataset published by Yelp [14]. In fact, crowdsensing is a general paradigm for ubiquitous sensing, and the dataset includes abundant real-world traces for emulating spatial crowdsensing where Yelp workers are employed to review (or "sense") local business.

We randomly choose 10^5 workers from the dataset. For each worker, we set the number of his/her reviews as his/her capacities. Since there is no cost parameters for the workers in the dataset, we randomly set the cost parameters $[1, 1.5]$. We choose *number of friends*, *number of fans* and *number of year as elite* as the context dimensions. In our experiments, we gradually increase the dimensionality of the contextual space, to evaluate our algorithm under the differently dimensioned contextual space. In the Yelp dataset, the sensed data (i.e., the reviews on the business) is voted by reviewers. For each of the sensed data, its quality (or reward feedback) can be calculated according to the votes it receives. We hereby assume that we get a unit of reward if the review receives at least three positive votes.

Different from the synthetic dataset where α is controllable, we have to figure out an appropriate value for α when using the Yelp dataset to partition the contextual space, since α is an intrinsic parameter for real data. To quantitatively evaluate our algorithm, we first illustrate in Fig. 3 the impact of different values of α on the performance of our algorithm. In the following, we set $\alpha = 2, 0.75, 0.25$ for $M = 1, 2, 3$, respectively. It is worthy to note that our algorithm still work with an arbitrary value of α and we hereby seek for an appropriate value for α only for the purpose of quantitative evaluation.

Since the dataset does not include the sensing abilities of the workers, we focus on investigating the performances of the algorithms in terms of cumulative revenue (rather than the expected one). In addition, we vary budget B from 2×10^4 to 4×10^5 with a step size 2×10^4 to show the adaptivity of CAWS to different budgets. It is shown by the results in Fig. 4 that, our CAWS algorithm outperforms the others and its performance is very close to the ones of the oracle (for all $M = 1, 2, 3$), especially under limited budget. Furthermore, since our CAWS algorithm adaptively tunes the granularity of partitioning the context space (according to the number of dimensions), it results in similar cumulative revenues in all the three contextual spaces. By taking into account more relevant dimensions (e.g., by increasing M from 1 to 2), our algorithm yields more cumulative revenue. Nevertheless, a higher-dimensional contextual space does not always imply much higher cumulative revenue. For example, the resulting cumulative revenue in the three-dimensional contextual space is very close to the one in the two-dimensional contextual space.

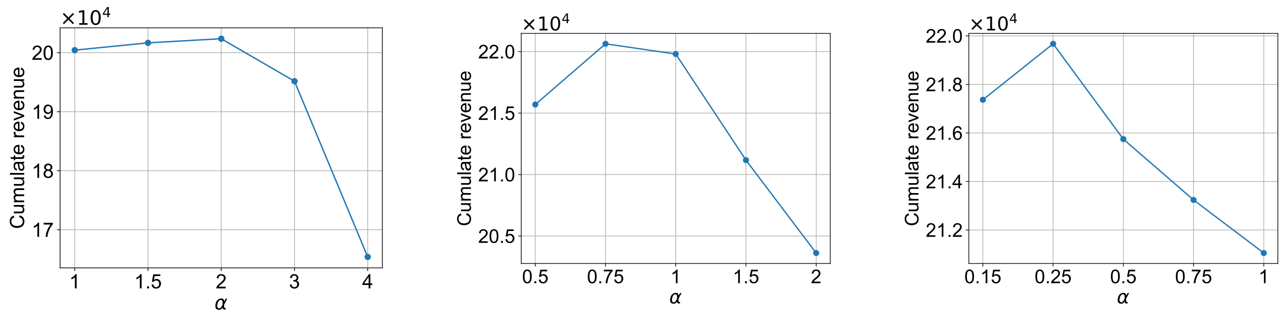
We also evaluate the algorithms under different numbers of workers, by using the Yelp dataset. We vary the number of workers from 4×10^4 to 1×10^5 with a step size 2×10^4. We fix the budget $B = 1 \times 10^5$. As illustrated by the results in Fig. 5, our CAWS algorithm yields much higher than the others in all settings. Similar with our observations in Fig. 4, the cumulative revenue obtained by applying our algorithm in the two-dimensional contextual space is very close to the one yielded by our algorithm in the three-dimensional space. Additionally, since we partition the contextual space with a carefully tuned granularity, the performance of our algorithm still can be ensured when we introduce much more workers with limited budget.

6 Related Work

In the past decades, there have been a vast body of studies on the fundamental problem of worker selection in crowdsensing systems [4, 5, 6, 7]. However, most of the existing proposals assume that the workers' sensing abilities are known as prior, while such an assumption may not be the case in practice. Therefore, there have been a few recent studies considering the uncertain worker selection problems where the worker's sensing abilities are unknown. For example, [8] studies the worker selection problem such that the workers with uncertain sensing abilities are selected sequentially under a limited budget to perform a given sensing task. In [9], a multi-task assignment problem is investigated. Therein, unknown workers are selected to maximize the sensing revenue, such that the resulting total cost does not exceed the budget and all the sensing tasks can be completed. The multi-task assignment problem is also studied in [10]. Each worker first submits its options (i.e., a subset of the tasks), and the crowdsensing platform assigns one of the options to each worker under a given budget, aiming at maximizing the sensing quality. Although these proposals leverage the CMAB framework, none of them takes into account the constraints on the workers' capacities. Moreover, as mentioned in Sec. 1, the arms in the CMAB framework (e.g., the workers in [8, 10] or the combinations of the workers and the tasks in [9]) are exploited and explored individually. Therefore, the standard CMAB framework is of low efficacy especially when the number of arms is huge while the budget is limited, as shown in Sec. 5.

Context information is very useful for crowdsensing systems and has been extensively utilized in designing worker selection algorithms [15, 16]. In [17], the context similarities between tasks and workers are employed to characterize the eligibilities of the workers. The sensing tasks are then assigned to the workers according to the eligibilities, aiming at improving sensing efficiency. In [18], a context-data quality classifier is trained from historical data in an off-line manner. It is then used to estimate the data qualities in an on-line manner, according to which, the workers are selected. Although machine learning methods are applied to train the classifier, it takes into account neither budget constraints nor capacity constraints. In [19], the dependence of workers' sensing abilities on both the workers' and tasks' context information is learned in an on-line manner, such that the tasks can be assigned accordingly with only budget constraints considered.

MAB problem is a typical reinforcement learning problem and has been studied for years. So far, several well-known algorithms, e.g., ϵ-greedy algorithm, UCB algorithm, etc, have been proposed [20, 21]. It is then extended to CMAB problem, to address the uncertainties in combinatorial optimization problem [12, 22, 23, 24, 25]. Motivated by contextual bandit where the context information of the arms is utilized [26, 27], [28] proposes a contextual CMAB framework which inherits properties from both contextual bandit and combinatorial bandit. Specifically, it studies the budget-limited worker selection problem within a given time horizon. In each time slot, it allocates a fixed amount of budget to either exploit or explore a group of workers. Therefore, it cannot be applied to our problem where the total budget is limited such that we have to make full use of the budget to discriminate the workers with uncertain sensing abilities. Also, it cannot handle the capacity constraints of the workers, whereas these constraints are the main concerns of our CAWS algorithm.



(a) $M = 1$ (b) $M = 2$ (c) $M = 3$

Fig. 3. Cumulative revenues under different values of α.

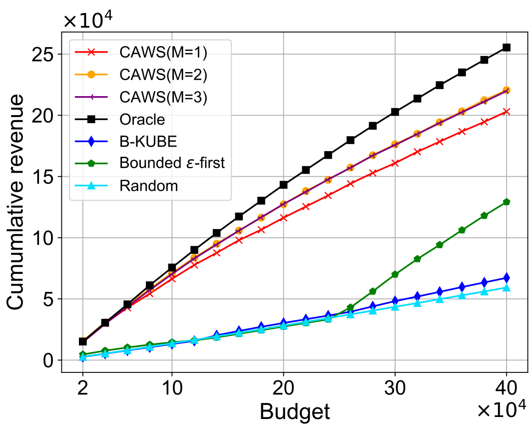


Fig. 4. Cumulative revenue with varying budget in Yelp dataset.

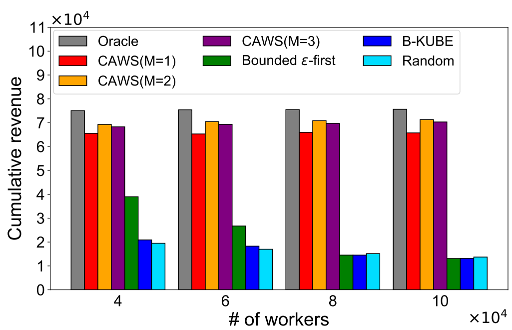


Fig. 5. Cumulative revenue with different numbers of workers.

7 Conclusion

In this paper, we have studied how to select among a massive number of uncertain workers with bounded sensing capacities under a limited budget, such that the expected total sensing revenue can be maximized with both the capacity constraint and budget constraint respected. Although the standard CMAB framework can be applied to address the above problem by exploring and exploiting the workers individually, it is of quite low efficiency when the number of workers is huge while the budget is significantly limited. To address the above issue, we have proposed a worker selection algorithm, i.e., CAWS, which makes a trade-off between exploitation and exploration on the partitions of the context space instead of the individual workers. We have performed rigorous theoretical analysis to prove the regret of our CAWS algorithm is upper-bounded by $\mathcal{O}(B^{\frac{M}{\alpha+M}} \ln B)$ under a properly quantified partition of the context space. We also have conducted extensive experiments with both the synthetic data set and the real one to verify the considerable advantages of our CAWS algorithm over the existing state-of-the-art CMAB-based algorithms.

Appendix A
Proof of Lemma 2

Lemma 3. *Suppose our CAWS algorithm proceeds T iterations and let $B(t)$ denotes the residual budget at the beginning of the t-th iteration. Initially, $B(1) = B$. For each iteration $t = 1, 2, \cdots, T$, we have*

$$\frac{c_{min}}{B(t)} \leq \frac{1}{T - t + 1} \tag{47}$$

Proof. At the beginning of the t-th iteration, the residual budget is $B(t)$. Since we select the workers T times in total, for any $1 \leq t \leq T$,

$$(T - t + 1)c_{min} \leq c_{i(t)} + c_{i(t+1)} + \cdots + c_{i(T)} \leq B(t)$$

based on which, we have the inequality (47). □

Assume $\lfloor \tilde{\mathbf{x}}^*(t) \rfloor$ and $\lfloor \hat{\mathbf{x}}^*(t) \rfloor$ denote the solutions by applying the rounding-based density-ordered greedy algorithm to the BKP instances $(i, \mu_{Q_i}, c_i, \tau_i(t), B(t))_{i=1}^N$ and $(i, U_i(t), c_i, \tau_i(t), B(t))_{i=1}^N$, respectively. By replacing μ_{Q_i} with $U_i(t)$, $\lfloor \hat{\mathbf{x}}^*(t) \rfloor$ can be considered as an estimate on $\lfloor \tilde{\mathbf{x}}^*(t) \rfloor$.

Lemma 4. *If there is a worker j such that $j \notin \lfloor \tilde{\mathbf{x}}^* \rfloor$ and $j \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor$, then there is at least one worker $j' \in \lfloor \tilde{\mathbf{x}}^* \rfloor$ such that*

$$\frac{\mu_{Q_j}}{c_j} \leq \frac{\mu_{Q_{j'}}}{c_{j'}} \tag{48}$$

and

$$\frac{1}{c_j}\left(\bar{r}_{Q_j}(t) + \sqrt{\frac{2\log t}{\lambda_{Q_j}(t)}}\right) \geq \frac{1}{c_{j'}}\left(\bar{r}_{Q_{j'}}(t) + \sqrt{\frac{2\log t}{\lambda_{Q_{j'}}(t)}}\right) \tag{49}$$

where $Q_j \neq Q_{j'}$. Also, the worker j' has non-zero residual capacity to perform sensing tasks.

Proof. If a worker $j \notin \lfloor \tilde{\mathbf{x}}^* \rfloor$, then the worker $j \notin \lfloor \tilde{\mathbf{x}}^*(t) \rfloor$, since $B(t) \leq B$. Moreover, according to the procedures of the rounding-based density-ordered greedy algorithm, if the worker $j \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor$, there exists at least one work $j' \in \lfloor \tilde{\mathbf{x}}^*(t) \rfloor$ such that $\frac{\mu_{Q_j}}{c_j} \leq \frac{\mu_{Q_{j'}}}{c_{j'}}$ and $\frac{U_j(t)}{c_j} \geq \frac{U_{j'}(t)}{c_{j'}}$. Also, $Q_j \neq Q_{j'}$; otherwise, we would have both $c_j \geq c_{j'}$ and $c_j \leq c_{j'}$ hold since $\mu_{Q_j} = \mu_{Q_{j'}}$ and $U_j(t) = U_{j'}(t)$. Also, $j' \in \lfloor \tilde{\mathbf{x}}^*(t) \rfloor$ implies that the worker j' has non-zero residual capacity to perform additional tasks. Furthermore, since $\lfloor \tilde{\mathbf{x}}^*(t) \rfloor \subseteq \lfloor \tilde{\mathbf{x}}^* \rfloor$, $j' \in \lfloor \tilde{\mathbf{x}}^* \rfloor$ and j' can perform more tasks. □

Lemma 5. *Assume our CAWS algorithm proceeds T iterations. For $\forall Q \in \Omega$, we have*

$$\mathbb{P}(i(t) \in \mathcal{N}_Q^- \mid T) \leq \mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \in \mathcal{N}_Q^- \mid T) + (c_{max}/c_{min})^2/(T-t+1) \tag{50}$$

Proof. Since

$$\mathbb{P}(i(t) \in \mathcal{N}_Q^- \mid T) = \mathbb{P}(i(t) \in \mathcal{N}_Q^-, i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T) + \mathbb{P}(i(t) \in \mathcal{N}_Q^-, i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T) \tag{51}$$

we can prove this lemma by deriving the upper bound of $\mathbb{P}(i(t) \in \mathcal{N}_Q^-, i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T)$

Recall that $\lfloor \hat{\mathbf{x}}^*(t) \rfloor$ denotes the solution of the BKP instance $(i, U_i(t), c_i, \tau_i(t), B(t))_{i=1}^N$ by the rounding-based density-ordered greedy algorithm. Let $\hat{k}(t)$ denote the split worker. Then, after selecting the worker $\hat{k}$, the residual budget is less than or equal to $c_{\hat{k}(t)}$; therefore,

$$\sum_{i \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor} x_i(t) \leq \frac{c_{\hat{k}(t)}}{c_{min}} \leq \frac{c_{max}}{c_{min}} \tag{52}$$

Furthermore, considering the selection outputted by our density-ordered greedy subroutine can be bounded as $\sum_{i \in \mathcal{N}} x_i(t) \geq \frac{B(t)}{c_{max}}$, we have

$$\frac{\sum_{i \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor} x_i(t)}{\sum_{i \in \mathcal{N}} x_i(t)} \leq \frac{c_{\hat{k}(t)}}{c_{min}} \leq \frac{c_{max}}{c_{min}} \cdot \frac{c_{max}}{B(t)} \tag{53}$$

By substituting the inequality (47) in **Lemma** 3 into the above inequality, we have

$$\frac{\sum_{i \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor} x_i(t)}{\sum_{i \in \mathcal{N}} x_i(t)} \leq \left(\frac{c_{max}}{c_{min}}\right)^2 \cdot \frac{1}{T-t+1} \tag{54}$$

Then, the upper bound of $\mathbb{P}(i(t) \in \mathcal{N}_Q^-, i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T)$ can be derived as follows

$$\begin{aligned}
&\mathbb{P}(i(t) \in \mathcal{N}_Q^-, i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T) \\
\leq &\mathbb{P}(i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid T) \\
= &\sum_{\{x_i(t)\}_{i=1}^N} \mathbb{P}(i(t) \notin \lfloor \hat{\mathbf{x}}^*(t) \rfloor \mid \{x_i(t)\}_{i=1}^N, T) \cdot \mathbb{P}(\{x_i(t)\}_{i=1}^N) \\
\leq &\sum_{\{x_i(t)\}_{i=1}^N} \left(\frac{c_{max}}{c_{min}}\right)^2 \cdot \frac{1}{T-t+1} \cdot \mathbb{P}(\{x_i(t)\}_{i=1}^N) \\
= &(c_{max}/c_{min})^2/(T-t+1)
\end{aligned} \tag{55}$$

By substituting which into (51), we complete the proof. □

Lemma 6. *For $\forall Q \in \Omega$, let $Y_Q(t) = \sum_{t'=1}^{t} \mathbb{I}(i(t') \in \mathcal{N}_Q^-)$ denotes the number of times the workers in $\mathcal{N}_Q^-$ is selected by our CAWS algorithm up to the t-th iteration. We then have*

$$\mathbb{E}_{\{i(t)\}_{t=1}^T}[Y_Q(T) \mid T] \leq \xi \ln T + \frac{\pi^2}{3} + 1 \tag{56}$$

Proof. According to **Lemma** 5, $\mathbb{E}_{\{i(t)\}_{t=1}^T}[Y_Q(T) \mid T]$ can be written as

$$\begin{aligned}
&\mathbb{E}_{\{i(t)\}_{t=1}^T}[Y_Q(T) \mid T] \\
=&1 + \sum_{t=d^M+1}^{T} \mathbb{P}(i(t) \in \mathcal{N}_Q^- \mid T) \\
\leq&1 + \sum_{t=d^M+1}^{T} \left(\mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \in \mathcal{N}_Q^- \mid T) + \frac{\left(\frac{c_{max}}{c_{min}}\right)^2}{T-t+1} \right) \\
\leq&1 + \sum_{t=d^M+1}^{T} \mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \in \mathcal{N}_Q^- \mid T) + \left(\frac{c_{max}}{c_{min}}\right)^2 \ln T \\
\leq&1 + \sum_{t=d^M+1}^{T} \mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \notin \lfloor \tilde{\mathbf{x}}^* \rfloor \mid T) + \left(\frac{c_{max}}{c_{min}}\right)^2 \ln T
\end{aligned} \tag{57}$$

We then derive the bound of the sum of the first two items by considering **Lemma** 4 as follows.

$$\begin{aligned}
&1 + \sum_{t=d^M+1}^{T} \mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \notin \lfloor \tilde{\mathbf{x}}^* \rfloor \mid T) \\
\leq&1 + \sum_{t=d^M+1}^{T} \mathbb{P}\left(\frac{\bar{r}_Q(t-1) + \sqrt{\frac{2\log t}{\lambda_Q(t-1)}}}{c_{min}(\mathcal{N}_Q^-)} \geq \frac{\bar{r}_{Q'}(t-1) + \sqrt{\frac{2\log t}{\lambda_{Q'}(t-1)}}}{c_{max}(\mathcal{N}_{Q'}^+)} \,\middle|\, T \right) \\
\leq&\ell + \sum_{t=d^M+1}^{T} \mathbb{P}\left(\begin{array}{c} \frac{\bar{r}_Q(t-1) + \sqrt{\frac{2\log t}{\lambda_Q(t-1)}}}{c_{min}(\mathcal{N}_Q^-)} \geq \frac{\bar{r}_{Q'}(t-1) + \sqrt{\frac{2\log t}{\lambda_{Q'}(t-1)}}}{c_{max}(\mathcal{N}_{Q'}^+)}, \\ \lambda_Q(t-1) \geq \ell \end{array} \,\middle|\, T \right) \\
\leq&\ell + \sum_{t=d^M+1}^{T} \mathbb{P}\left(\begin{array}{l} \max_{\ell \leq s_Q < t} \frac{\bar{r}_Q(t-1) + \sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \\ \geq \min_{1 \leq s_{Q'} < t} \frac{\bar{r}_{Q'}(t-1) + \sqrt{\frac{2\log t}{s_{Q'}}}}{c_{max}(\mathcal{N}_{Q'}^+)} \end{array} \,\middle|\, T \right) \\
\leq&\ell + \sum_{t=1}^{T} \sum_{s_{Q'}=1}^{t-1} \sum_{s_Q=\ell}^{t-1} \mathbb{P}\left(\frac{\bar{r}_Q(t-1) + \sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \geq \frac{\bar{r}_{Q'}(t-1) + \sqrt{\frac{2\log t}{s_{Q'}}}}{c_{max}(\mathcal{N}_{Q'}^+)} \,\middle|\, T \right)
\end{aligned} \tag{58}$$

If it holds that $\frac{\bar{r}_Q(t-1) + \sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \geq \frac{\bar{r}_{Q'}(t-1) + \sqrt{\frac{2\log t}{s_{Q'}}}}{c_{max}(\mathcal{N}_{Q'}^+)}$, then at least one of the following three event must happen

$$\text{Event1}: \frac{\bar{r}_Q(t-1)}{c_{min}(\mathcal{N}_Q^-)} - \frac{\sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \geq \frac{\mu_Q}{c_{min}(\mathcal{N}_Q^-)} \tag{59}$$

$$\text{Event2}: \frac{\bar{r}_{Q'}(t-1)}{c_{max}(\mathcal{N}_{Q'}^+)} - \frac{\sqrt{\frac{2\log t}{s_{Q'}}}}{c_{max}(\mathcal{N}_{Q'}^+)} \leq \frac{\mu_{Q'}}{c_{max}(\mathcal{N}_{Q'}^+)} \tag{60}$$

$$\text{Event3}: \frac{\mu_{Q'}}{c_{max}(\mathcal{N}_{Q'}^+)} - \frac{\mu_Q}{c_{min}(\mathcal{N}_Q^-)} \leq \frac{2\sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \tag{61}$$

By applying the Chernoff-Hoeffding inequalities [29], we have

$$\mathbb{P}(\text{Event1}) = \mathbb{P}\left(\bar{r}_Q(t-1) - \sqrt{\frac{2\log t}{s_Q}} \geq \mu_Q\right) \leq t^{-4}$$

Similarly,

$$\mathbb{P}(\text{Event2}) = \mathbb{P}\left(\bar{r}_{Q'}(t-1) - \sqrt{\frac{2\log t}{s_{Q'}}} \le \mu_{Q'}\right) \le t^{-4}$$

When $\ell = \left\lceil \frac{8\ln T}{c_{min}^2 \delta_{min}^2} \right\rceil$, for any $s_Q = \ell, \ell+1, \cdots, t-1$, we have

$$\frac{\mu_{Q'}}{c_{max}(\mathcal{N}_{Q'}^+)} - \frac{\mu_Q}{c_{min}(\mathcal{N}_Q^-)} > \frac{2\sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)}$$

and thus $\mathbb{P}(\text{Event3}) = 0$. Combining the above inequalities, we have

$$\begin{aligned}
&\mathbb{E}_{\{i(t)\}_{t=1}^T}[Y_Q(T) \mid T] \\
\le& 1 + \sum_{t=d^M}^{T} \mathbb{P}(i(t) \in \lfloor \hat{\mathbf{x}}^*(t) \rfloor, i(t) \notin \lfloor \tilde{\mathbf{x}}^* \rfloor \mid T) + \left(\frac{c_{max}}{c_{min}}\right)^2 \ln T \\
\le& \ell + \sum_{t=1}^{T}\sum_{s_{Q'}=1}^{t-1}\sum_{s_Q=\ell}^{t-1} \mathbb{P}\left(\frac{\bar{r}_Q(t-1)+\sqrt{\frac{2\log t}{s_Q}}}{c_{min}(\mathcal{N}_Q^-)} \ge \frac{\bar{r}_{Q'}(t-1)+\sqrt{\frac{2\log t}{s_{Q'}}}}{c_{max}(\mathcal{N}_{Q'}^+)} \,\middle|\, T \right) \\
\le& \ell + \sum_{t=1}^{T}\sum_{s_{Q'}=1}^{t-1}\sum_{s_Q=\ell}^{t-1} (\mathbb{P}(\text{Event1}) + \mathbb{P}(\text{Event2}) + \mathbb{P}(\text{Event3})) \\
\le& \left\lceil \frac{8\ln T}{c_{min}^2\delta_{min}^2} \right\rceil + \sum_{t=1}^{T}\sum_{s_{Q'}=1}^{t-1}\sum_{s_Q=\ell}^{t-1} 2t^{-4} \\
&+ \left(\frac{c_{max}}{c_{min}}\right)^2 \ln T + \frac{\pi^2}{3} + 1 \\
\le& \xi \ln T + \frac{\pi^2}{3} + 1
\end{aligned}$$

where ξ is defined in (25). □

Now, we are ready to prove the inequalities (34) and (35). Our CAWS algorithm proceeds until we have no more residual budget to select any workers such that

$$\mathbb{P}\left(\sum_{t=1}^{T} c_{i(t)} \ge B - c_{max}\right) = 1 \tag{62}$$

therefore, we have

$$\begin{aligned}
&B - c_{max} \\
\le& \mathbb{E}_{T,\{i(t)\}_{t=1}^T}\left[\sum_{t=1}^{T} c_{i(t)}\right] \\
=& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}\sum_{j\in\mathcal{N}_Q} c_j \mathbb{P}(i(t)=j \mid T)\right] \\
=& \mathbb{E}_T\left[\sum_{t=1}^{T}\left(c_{i^*} + \sum_{Q\in\Omega}\sum_{j\in\mathcal{N}_Q}(c_j - c_{i^*})\mathbb{P}(i(t)=j \mid T)\right)\right] \\
\le& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}\left(\sum_{j\in\mathcal{N}_Q^-}(c_j - c_{i^*})\mathbb{P}(i(t)=j \mid T)\right.\right. \\
&\left.\left. + \sum_{j\in\mathcal{N}_Q^+}(c_j - c_{i^*})\mathbb{P}(i(t)=j \mid T)\right)\right] + \mathbb{E}_T[T]c_{i^*} \\
\le& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}\left(\sum_{j\in\mathcal{N}_Q^-}(c_{max}(\mathcal{N}_Q^-) - c_{i^*})\mathbb{P}(i(t)=j \mid T)\right.\right. \\
&\left.\left. + \sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(c_j - c_{i^*})\mathbb{P}(i(t)=j \mid T)\right)\right] + \mathbb{E}_T[T]c_{i^*} \\
\le& \mathbb{E}_T[T]c_{i^*} + \mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(c_j - c_{i^*})\mathbb{E}_{\{i(t)\}_{i=1}^T}[x_j \mid T]\right] \\
&+ \mathbb{E}_T\left[\sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}(c_{max}(\mathcal{N}_Q^-) - c_{i^*})\mathbb{E}_{\{i(t)\}_{i=1}^T}[Y_Q(T)|T]\right]
\end{aligned} \tag{63}$$

and thus

$$\begin{aligned}
&\mathbb{E}_T[T] &(64)\\
\ge& \frac{B - c_{max}}{c_{i^*}} - \mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}\frac{c_j - c_{i^*}}{c_{i^*}}\mathbb{E}[x_j \mid T]\right] \\
&- \mathbb{E}_T\left[\sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}\mathbb{E}_{\{i(t)\}_{i=1}^T}[Y_Q(T)|T]\right] &(65)
\end{aligned}$$

The validity of (34) can be proved by substituting (56) (see **Lemma** 6) into the second item on the right side of the above inequality such that

$$\begin{aligned}
&\mathbb{E}_T\left[\sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}\mathbb{E}_{\{i(t)\}_{i=1}^T}[Y_Q(T) \mid T]\right] \\
\le& \sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}\left(\xi\cdot\mathbb{E}_T[\ln T] + \frac{\pi^2}{3} + 1\right) \\
\le& \sum_{Q:c_{max}(\mathcal{N}_Q^-)>c_{i^*}}\frac{c_{max}(\mathcal{N}_Q^-) - c_{i^*}}{c_{i^*}}\cdot h(\ln B) &(66)
\end{aligned}$$

where the second inequality holds since $T \le \frac{B}{c_{min}}$.

We finally prove that the inequality (35) hold by the derivation as follows

$$\begin{aligned}
&\mathbb{E}_T\left[T\mu_{Q_{i^*}} - \sum_{Q\in\Omega}\sum_{i\in\mathcal{N}_Q}\mu_Q\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_i \mid T]\right] \\
=& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}(\mu_{Q_{i^*}} - \mu_Q)\mathbb{P}(i(t)\in\mathcal{N}_Q \mid T)\right] \\
=& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}(\mu_{Q_{i^*}} - \mu_Q)\mathbb{P}(i(t)\in\mathcal{N}_Q^- \mid T)\right. \\
&\left. - \sum_{t=1}^{T}\sum_{Q\in\Omega}(\mu_{Q_{i^*}} - \mu_Q)\mathbb{P}(i(t)\in\mathcal{N}_Q^+ \mid T)\right] \\
=& \mathbb{E}_T\left[\sum_{t=1}^{T}\sum_{Q\in\Omega}(\mu_{Q_{i^*}} - \mu_Q)\mathbb{P}(i(t)\in\mathcal{N}_Q^- \mid T)\right. \\
&\left. - \sum_{t=1}^{T}\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}} - \mu_{Q_j})\mathbb{P}(i(t)=j \mid T)\right]
\end{aligned}$$

$$\leq \mathbb{E}_T\Big[\sum_{t=1}^{T}\sum_{Q:\mu_{Q_{i^*}}>\mu_Q}(\mu_{Q_{i^*}}-\mu_Q)\mathbb{P}(i(t)\in\mathcal{N}_Q^- \mid T)$$
$$-\sum_{t=1}^{T}\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}}-\mu_{Q_j})\mathbb{P}(i(t)=j \mid T)\Big]$$
$$= \mathbb{E}_T\left[\sum_{Q:\mu_{Q_{i^*}}>\mu_Q}(\mu_{Q_{i^*}}-\mu_Q)\mathbb{E}_{\{i(t)\}_{t=1}^T}[Y_Q(T)\mid T]\right]$$
$$-\mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}}-\mu_{Q_j})\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j\mid T]\right]$$
$$\leq \sum_{Q:\mu_{Q_{i^*}}>\mu_Q}(\mu_{Q_{i^*}}-\mu_Q)h(\ln B)$$
$$-\mathbb{E}_T\left[\sum_{j\in\lfloor\tilde{\mathbf{x}}^*\rfloor}(\mu_{Q_{i^*}}-\mu_{Q_j})\mathbb{E}_{\{i(t)\}_{t=1}^T}[x_j\mid T]\right] \quad (67)$$

where we have last inequality by considering **Lemma** 6 again.

REFERENCES

[1] Cisco Annual Internet Report (20182023) White Paper. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf, 2020. Online; accessed March 9, 2020.

[2] A. Capponi, C. Fiandrino, B. Kantarci, L. Foschini, D. Kliazovich, and P. Bouvry. A Survey on Mobile Crowdsensing Systems: Challenges, Solutions, and Opportunities. *IEEE Communications Surveys & Tutorials*, 21(3):2419–2465, 2019.

[3] W. Khan, Y. Xiang, M. Aalsalem, and Q. Arshad. Mobile Phone Sensing Systems: A Survey. *IEEE Communications Surveys & Tutorials*, 15(1):402–427, 2013.

[4] Z. Song, C. Liu, J. Wu, J. Ma, and W. Wang. QoI-Aware Multitask-Oriented Dynamic Participant Selection With Budget Constraints. *IEEE Trans. on Vehicular Technology*, 63(9):4618–4632, 2014.

[5] H. Li, T. Li, and Y. Wang. Dynamic Participant Recruitment of Mobile Crowd Sensing for Heterogeneous Sensing Tasks. In *Proc. of the 12th IEEE MASS*, pages 136–144, 2015.

[6] L. Pu, X. Chen, J. Xu, and X. Fu. Crowdlet: Optimal Worker Recruitment for Self-Organized Mobile Crowdsourcing. In *Proc. of IEEE INFOCOM*, pages 1–9, 2016.

[7] Y. Yang, W. Liu, E. Wang, and J. Wu. A Prediction-Based User Selection Framework for Heterogeneous Mobile CrowdSensing. *IEEE Trans. on Mobile Computing*, 18(11):2460–2473, 2019.

[8] K. Han, C. Zhang, and J. Luo. Taming the Uncertainty: Budget Limited Robust Crowdsensing Through Online Learning. *IEEE/ACM Trans. on Networking*, 24(3):1462–1475, 2016.

[9] G. Gao, J. Wu, Z. Yan, M. Xiao, and G. Chen. Unknown Worker Recruitment with Budget and Covering Constraints for Mobile Crowdsensing. In *Proc. of the 25th IEEE ICPADS*, pages 539–547, 2019.

[10] G. Gao, J. Wu, M. Xiao, and G. Chen. Combinatorial Multi-Armed Bandit Based Unknown Worker Recruitment in Heterogeneous Crowdsensing. In *Proc. of IEEE INFOCOM*, pages 179–188, 2020.

[11] R. Kohli, R. Krishnamurti, and P. Mirchandani. Average Performance of Greedy Heuristics for The Integer Knapsack Problem. *European Journal of Operational Research*, 154(1):36–45, 2014.

[12] A. Rangi and M. Franceschetti. Multi-Armed Bandit Algorithms for Crowdsourcing Systems with Online Estimation of Workers' Ability. In *Proc. of the 17th AAMAS*, page 13451352, 2018.

[13] L. Tran-Thanh, S. Stein, A. Rogers, and N. Jennings. Efficient Crowdsourcing of Unknown Experts using Bounded Multi-Armed Bandits. *Artificial Intelligence*, 214:89–111, 2014.

[14] Yelp Dataset: www.yelp.com/dataset/challenge.

[15] O. Yurur, C. Liu, Z. Sheng, V. Leung, W. Moreno, and K. Leung. Context-Awareness for Mobile Sensing: A Survey and Future Directions. *IEEE Communications Surveys & Tutorials*, 18(1):68–93, 2016.

[16] H. Vahdat-Nejad, E. Asani, Z. Mahmoodian, and M. Mohseni. Context-aware computing for mobile crowd sensing: A survey. *Future Generation Computer Systems*, 99:321–332, 2019.

[17] A. Hassani, P. Haghighi, and P. Jayaraman. Context-Aware Recruitment Scheme for Opportunistic Mobile Crowdsensing. In *Proc. of the 21st IEEE ICPADS*, pages 266–273, 2015.

[18] S. Liu, Z. Zheng, F. Wu, S. Tang, and G. Chen. Context-Aware Data Quality Estimation in Mobile Crowdsensing. In *Proc. of IEEE INFOCOM*, pages 1–9, 2017.

[19] S. Müller, C. Tekin, M. Schaar, and A. Klein. Context-Aware Hierarchical Online Learning for Performance Maximization in Mobile Crowdsourcing. *IEEE/ACM Trans. on Networking*, 26(3):1334–1347, 2018.

[20] T. Lai and H. Robbins. Asymptotically Efficient Adaptive Allocation Rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.

[21] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time Analysis of the Multiarmed Bandit Problem. *Machine Learning*, 47:235256, 2002.

[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 Trans. on Networking*, 20(5):1466–1478, 2012.

[23] L. Tran-Thanh, A. Chapman, A. Rogers, and N. Jennings. Knapsack Based Optimal Policies for BudgetLimited MultiArmed Bandits. In *Proc. of the 26th AAAI*, pages 1135–1140, 2012.

[24] W. Chen, Y. Wang, and Y. Yuan. Combinatorial Multi-Armed Bandit: General Framework and Applications. In *Proc. of the 30th ICML*, pages 151–159, 2013.

[25] L. Tran-Thanh, S. Stein, A. Rogers, and N.Jennings. Efficient Crowdsourcing of Unknown Experts Using Bounded Multi-Armed Bandits. *Artificial Intelligence*, 214:89–111, 2014.

[26] J. Langford and T. Zhang. The Epoch-Greedy Algorithm for Multi-Armed Bandits with Side Information. In *Proc. of the 22nd NIPS*, pages 817–824, 2008.

[27] L. Li, W. Chu, J. Langford, and R. Schapire. A Contextual-Bandit Approach to Personalized News Ar-

ticle Recommendation. In *Proc. of the 19th WWW*, pages 661–670, 2010.

[28] L. Chen, J. Xu, and Z. Lu. Contextual Combinatorial Multi-armed Bandits with Volatile Arms and Submodular Reward. In *Proc. of the 32nd NIPS*, pages 3247–3256, 2018.

[29] D. Dubhashi and A. Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

PLACE PHOTO HERE

Feng Li received his BS and MS degrees in Computer Science from Shandong Normal University, China, in 2007, and Shandong University, China, in 2010, respectively. He got his PhD degree (also in Computer Science) from Nanyang Technological University, Singapore, in 2015. From 2014 to 2015, he worked as a research fellow in National University of Singapore, Singapore. He then joined School of Computer Science and Technology, Shandong University, China, where he is currently an associate professor. His research interests include distributed algorithms and systems, wireless networking, mobile computing, and Internet of Things.

PLACE PHOTO HERE

Jichao Zhao received the BSc degree in 2019 from the School of Information Science and Engineering, Ocean University of China. He is currently a postgraduate student in Department of Computer Science and Technology, Shandong University. Her research interests include crowdsourcing and multi-armed bandit.

PLACE PHOTO HERE

Dongxiao Yu received the BSc degree in 2006 from the School of Mathematics, Shandong University and the PhD degree in 2014 from the Department of Computer Science, The University of Hong Kong. He became an associate professor in the School of Computer Science and Technology, Huazhong University of Science and Technology, in 2016. He is currently a professor in the School of Computer Science and Technology, Shandong University. His research interests include wireless networks, distributed computing and graph algorithms.

PLACE PHOTO HERE

Xiuzhen Cheng received her M.S. and Ph.D. degrees in computer science from the University of Minnesota Twin Cities in 2000 and 2002, respectively. She is a professor in the School of Computer Science and Technology, Shandong University. Her current research interests include cyber-physical systems, wireless and mobile computing, sensor networking, wireless and mobile security, and algorithm design and analysis. She has served on the editorial boards of several technical journals and the technical program committees of various professional conferences/workshops. She also has chaired several international conferences. She worked as a program director for the US National Science Foundation (NSF) from April to October in 2006 (full time), and from April 2008 to May 2010 (part time). She received the NSF CAREER Award in 2004. She is Fellow of IEEE and a member of ACM.

PLACE PHOTO HERE

Weifeng Lv received his Ph.D. degree in computer science from Beihang University in 1998. His current research interests include massive information system, urban cognitive computing, swarm intelligence, and smart cities. He is a professor of computer science and the dean of the School of Computer Science and Engineering, as well as the vice director of the State Key Laboratory of Software Development Environment, at Beihang University. He also serves as the Vice-President and Secretary-General of the China Software Industry Association and the director of National Engineering Research Center for Science and Technology Resources Sharing Service. He received multiple internationally renowned awards, including the second prize of the 2016 China National Science and Technology Invention Award and the first prize of the 2010 Beijing Science and Technology Award.