Oasis: Online All-Phase Quality-Aware Incentive Mechanism for MCS

Man Zhang, Xinghua Li, *Member, IEEE*, Yinbin Miao, Bin Luo, Siqi Ma, Kim-Kwang Raymond Choo, *Senior Member, IEEE*, and Robert H. Deng, *Fellow, IEEE*

***Abstract*—To motivate users to submit high quality data for mobile crowdsensing (MCS), some quality-aware incentive mechanisms have been proposed, which recruit and pay users strategically. However, in the existing mechanisms, the recruitment based only on tasks matching degree leads to the ineffective insistent data quality incentive. Meanwhile, the absence of the reasonable payment strategy cannot motivate users to submit high quality data in the current task. To address the above problems, we propose an Online all-phase quality-aware incentive mechanism (Oasis) to realize the quality incentive in both recruitment and payment phases. With the knapsack secretary, Oasis first devises a quality-aware pre-budgeting recruitment strategy, which decides whether the arriving user's long-term data quality and bid satisfy the recruited criterion. Then, in the payment phase, Oasis evaluates and updates the current and long-term data qualities of users. Based on the evaluation results, a two-level payment strategy is devised employing the Myerson theorem, where users submitting higher quality data can obtain more utilities under the budget constraint. Theoretical analysis proves that Oasis satisfies economic feasibility and constant competitiveness while achieving quality incentive in recruitment and payment phases. Extensive experiments using the real-world dataset demonstrate that the sensing result accuracy of Oasis increases 67% compared with the existing works.**

***Index Terms*—Knapsack secretary, mobile crowdsensing, online recruit, quality-aware incentive mechanisms.**

Manuscript received 23 February 2022; revised 11 January 2024; accepted 11 January 2024. Date of publication 16 January 2024; date of current version 10 April 2024. The work of Kim-Kwang Raymond Choo was supported by the Cloud Technology Endowed Professorship. This work was supported in part by the National Natural Science Foundation of China under Grants 62125205 and U23A20303, and in part by the Key Research and Development Program of Shaanxi under Grant 2021ZDLGY05-04. *(Corresponding author: Xinghua Li.)*

Man Zhang, Yinbin Miao, and Bin Luo are with the State Key Laboratory of Integrated Services Networks, and the School of Cyber Engineering, Xidian University, Xi'an 710071, China (e-mail: mzhang0517@outlook.com; ybmiao@xidian.edu.cn; luobin2127@hotmail.com).

Xinghua Li is with the State Key Laboratory of Integrated Service Networks, School of Cyber Engineering, Xidian University, Xi'an 710071, China, and also with the Engineering Research Center of Big data Security, Ministry of Education, Xi'an 710071, China (e-mail: xhli1@mail.xidian.edu.cn).

Siqi Ma is with the School of Information Technology and Electrical Engineering, The University of Queensland, St Lucia, QLD 4072, Australia (e-mail: xdmasiqi@hotmail.com).

Kim-Kwang Raymond Choo is with the Department of Information Systems and Cyber Security and Department of Electrical and Computer Engineering, University of Texas, San Antonio, TX 78249-063 USA (e-mail: Raymond.Choo@utsa.edu).

Robert H. Deng is with the Department of Information Systems, Singapore Management University, Singapore 188065 (e-mail: robertdeng@smu.edu.sg).

Digital Object Identifier 10.1109/TSC.2024.3354240

I. Introduction

With the popularity of personal mobile devices equipped with sensors (e.g., camera, microphone, and accelerometer), mobile crowdsensing (MCS) [1], as a data collection paradigm combining the mobile devices' sensing ability and the idea of crowdsourcing, has brought a lot of convenience to people's daily life. The typical process of MCS is shown in Fig. 1. The service provider (SP) broadcasts sensing tasks to data contributors (DCs). After recruiting DCs and collecting relevant sensing data from them, SP then aggregates the sensing data and extracts the result of sensing tasks. The sensing data collected by MCS are used in many practical applications, such as environmental monitoring [2], traffic monitoring [3], and parking lot querying [4].

In MCS, the valid sensing service depends on a large amount of DCs who submit truthful sensing data [5]. However, due to the high computation, communication and energy costs, DCs are not willing to contribute to MCS without any motivation [5], [6], [7]. Moreover, DCs may submit untruthful data [8], which seriously impairs the accuracy and availability of MCS. For example, an indolent DC may place his mobile device in his pocket rather than carefully take it out to sense surrounding information, leading the inaccurate sensing result. Aiming at these problems, some quality-aware incentive mechanisms for MCS are proposed [8], [9], [10], [11], [12], [13], [14], [15].

The existing quality-aware incentive mechanisms generally assume that SP waits for enough arrival DCs to construct a candidate set. Then, SP compares and recruits DCs from the candidate set and pays recruited DCs based on their data qualities under the certain budget. Specifically, in the recruitment phase, most existing mechanisms select and recruit DCs only based on the matching degree (e.g., bids, locations, sensor accuracy, etc.) between DCs and tasks [12] [16]. However, the matching degree fails to reflect DCs' long-term task completion situation [17]. It leads that DCs who match the current task can still be recruited and make profit even they submitted low quality data in the history tasks. Thus, in the recruitment phase, focusing on the matching degree cannot motivate DCs' insistent high quality sensing data.

In the payment phase, most existing quality-aware incentive mechanisms lack a proper payment strategy to implement quality incentive. They generally determine DCs' payment amounts before DCs submit sensing data [12]. Ignoring DCs' submitted data qualities while allocating payment leads that DCs obtain

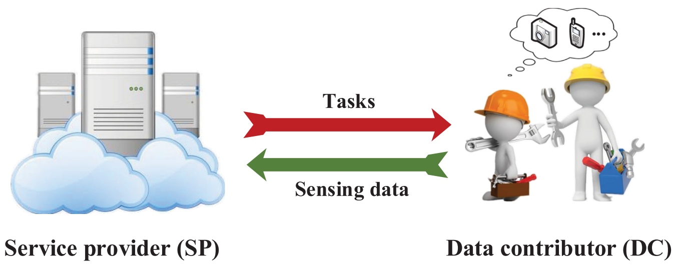


Fig. 1. Typical process of MCS.

equal payment amounts whether they submit high quality data or not. Meanwhile, there're some works paying DCs after receiving DCs' sensing data [14]. However, in their payment strategy, they just consider DC's submitted data quality while ignoring DC's costs for task completion, resulting that DCs submitting high quality data may obtain lower utilities because consuming more costs for the task, and this ruins the data quality incentive goal. Thus, the deficient payment strategy fails to realize the quality incentive in the payment phase, which cannot motivate DCs to submit high quality data in the current task.

From above, some existing works fail to recruit DCs based on their data qualities in the history tasks [12], [13], [14], [16], [18], [19], [20], [21], [22], leading that DCs can be recruited and make profits whether they have submitted high quality data formerly. Moreover, the other works cannot pay DCs according to their data qualities in the current task accurately [12], [16], [18], [19], [20], [21], causing that DCs obtain almost same payments whether they submit high quality data currently. Thus, the existing advanced schemes cannot implement the effective quality incentive in either recruitment or payment phases.

To solve the above problems, we propose an **O**nline **a**ll-pha**s**e quality-aware **i**ncentive mechani**s**m (Oasis). Focusing on the single task's all-phase quality incentive, Oasis claims that only the DCs submitting high quality data can be recruited and obtain relevant payment amounts, which motivates DCs to insistently submit high quality data. The main contributions of this paper are summarized as follows:

1) We implement a recruitment strategy to realize the data quality incentive in the recruitment phase. By the idea of the knapsack secretary, we design a quality-aware pre-budgeting recruitment strategy, which decides whether the arriving DC's long-term data quality and bid satisfy the recruited criterion. The devised recruitment strategy motivates DCs to insistently submit high quality data to maximize the recruited probability, as well as guarantees the following payment quality incentive under the budget constraint.
2) We devise a payment strategy to realize the data quality incentive in the payment phase. Based on the recruitment results, we first evaluate and update DCs' current and long-term data qualities. Then, by referring to Myerson theorem, we design a two-level payment strategy with consideration of both DCs' costs and data qualities, where DCs submitting high quality data can obtain high utilities under the budget constraint.

TABLE I
A COMPARATIVE SUMMARY BETWEEN OUR SCHEMES AND PREVIOUS SCHEMES

Scheme	Recruitment phase			Payment phase	
	Manner	Factor	Quality incentive	Factor	Quality incentive
Xu et al. [13]	offline	MWQ	✗	TCA	✗
Tang et al. [18]	offline	MWQ	✗	–	✗
Wang et al. [19]	offline	MWQ	✗	–	✗
Wang et al. [20]	offline	MWQ	✗	–	✗
Zhao et al. [14]	–	–	✗	IQE	✗
Yang et al. [21]	online	MWQ	✗	–	✗
Wang et al. [16]	online	MWQ	✗	–	✗
Liu et al. [12]	online	MWQ	✗	TCA	✗
Gao et al. [22]	online	MWQ	✗	UQE	✗
Oasis	online	LDQ	✓	CDQ	✓

Note: MWQ: matching degree without data quality; TCA: task completion amount; IQE: inaccurate quality evaluation; UQE: undefined quality evaluation; LDQ: long-term data quality; CDQ: current data quality; –: lack of consideration.

3) Theoretical analysis shows that Oasis can guarantee all-phase quality incentive goals while satisfying economic feasibility and constant competitiveness. Extensive experiments based on the real-world dataset show that Oasis recruits more competitive DCs under the budget constraint. Meanwhile, the sensing result accuracy of Oasis increases 67% compared with the existing works.

The remainder of this paper is organized as follows. Section II reviews the existing related works. Section III presents the problem formulation. Section IV shows the preliminaries. Section V explains Oasis in detail. Sections VI and VII give theoretical analysis and experimental results, respectively. Finally, Section VIII concludes Oasis.

II. Related Works

To motivate DCs' high quality data, many quality-aware incentive mechanisms for MCS have been proposed, which mainly focus on recruitment and payment phases. Table I shows the comparison of various indexes about Oasis and existing schemes.

A. Quality-Aware Incentive Mechanisms in Recruitment Phase

In the recruitment phase, the most existing mechanisms supposed that the DCs who satisfy tasks' demands (temporal-spatial coverage, preference, etc.) can complete more tasks and submit high quality data. Thus, they recruited DCs based on the matching degree between DCs and tasks. Taking the affect of DCs' social relationship on tasks' completion into account, Xu et al. [13] grouped DCs based on their cooperative relationship and assigned sensing task to a group of DCs. Tang et al. [18] matched tasks and DCs based on DCs' preference to gain high quality sensing data. Wang et al. [19] supposed that MCS's data quality can be viewed as the task completion amount, which is related to DCs' spatio-temporal conditions closely. Inspired by this method, Wang et al. [20] proposed an MCS system based on DCs' location diversity. Except for the location awareness, Wang et al. [16] also adopted DCs' expertise in the recruitment phase. Considering that DCs would not wait and build up the candidate set for SP's recruitment, Zhao et al. [23] adopted a multiple-stage sampling-accepting method to online recruit DCs based on the amount of task DCs can complete, which means that they make the recruitment decision as soon as a DC arrives. Supposing that the MCS sensing tasks are heterogeneous, Yang et al. [21] designed a mobility prediction model to forecast the probabilities that DCs complete tasks based on their locations and then online recruited the efficient DCs. Taking the time constraint of sensing tasks and DCs' mobility into consideration, Liu et al. [12] further proposed truthful pricing recruitment to improve the number of completed sensing tasks.

However, in the above works, the DCs who matches the current task can still be recruited and make profits even they submitted low quality data in the history tasks. Due to the ignorance of the quality incentive in the recruitment phase, the above works cannot motivate DCs to submit high quality data insistently.

B. Quality-Aware Incentive Mechanisms in Payment Phase

In the payment phase, some mechanisms determine DCs' payments based on DCs' costs and abilities for the current task completion before DCs submit sensing data [12], [13], [21], [23]. Ignoring DCs' submitted data qualities while allocating payment leads that DCs can obtain equal payment amounts whether they submit high quality data or not. Besides, some mechanisms took DCs' data qualities of the current task into account. Gao et al. [22] paid DCs based on their data qualities, where they failed to proposed the evaluation approach of data quality. Guo et al. [24] and Wang et al. [11] evaluated DCs' data qualities with the distance between DCs' data and the considered truth, which is the mean or voting result of all DCs' data. Peng et al. [8] evaluated DCs' data quality with the help of expectation maximization algorithm and paid DCs based on the evaluation results. Nevertheless, the above aggregation approaches (*e.g.* mean, voting) are incomplete in MCS due to the ignorance of DCs' different reliabilities [25], [26]. To rectify this inaccuracy, Yang *et al.* [27] evaluated DCs' data qualities by adopting the truth discovery that infers the truth among different reliability data sources. Inspired by the scheme in [27], Zhao et al. [14] enabled data requesters to discover the truth and evaluate DCs' data qualities based on the discovered truth, which cannot resist data requesters' slander.

In brief, the existing mechanisms above failed to realize an online all-phase quality incentive, which cannot encourage DCs to submit high quality data effectively.

III. Problem Formulation

In this section, we present the system model, problem definition and design goals of Oasis. Prior to detailed descriptions, we outline main notations in Table II.

TABLE II
Main Notations

Notations	*Descriptions*
$dc_j \in C$	Data contributor
$\mathcal{T} = \{l, r, t, \tau, B\}$	Sensing task $\mathcal{T}$ consisting of task's location, the effective sensing radius, the required time, the type of sensing data, the budget constraint
N_l^t	Estimated number of candidate DCs in location l during time t
$PR_j = \{ID_j, \mathcal{T}, b_j\}$	dc_j's personal resume consisting of his identity certificate, desirable task and bid for task
O, S, Re	Observed/ Sample/ Recruited DC set
Re^{off}	Recruited DC set in offline manner
$q_j^{(c)}$, $q_j^{(l)}$, $\bar{q}_{Re}^{(c)}$	dc_j's current or long-term data quality while completing the task, DCs' average current data quality
c_j, $p_j^{(b)}$, $p_j^{(a)}$, u_j	dc_j's cost for completing sensing task $\mathcal{T}$, his basic and actual payment and his utility
$d_j^{(\mathcal{T})}$, $\tilde{d}^{(\mathcal{T})}$	dc_j's sensing data and SP's aggregated truthful data for task $\mathcal{T}$

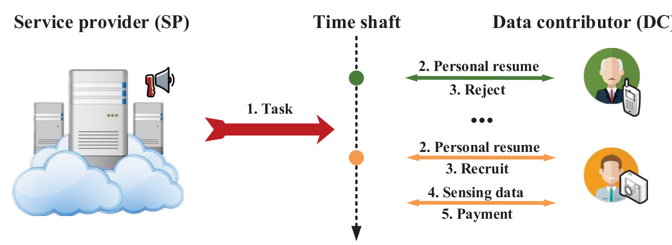


Fig. 2. System model of Oasis.

A. System Model

As depicted in Fig. 2, the system model of Oasis includes two entities, namely service provider (SP) and data contributor (DC). The role of each entity is shown as follows:

- *Service provider (SP):* SP broadcasts a single sensing task and recruits arrival DCs to collect sensing data. After receiving the data from recruited DCs, SP calculates the truthful result of sensing task and pays DCs.
- *Data contributor (DC):* DC decides whether to participate in sensing task according to task demands. If DC is recruited, he submits sensing data to SP to obtain relevant payment.

SP broadcasts a sensing task $\mathcal{T} = \{l, r, t, \tau, B\}$ (step 1), where l, r, t, τ, B indicate the location, the effective sensing radius, the required time, the sensing data type and the budget of task $\mathcal{T}$, respectively. DC decides whether to participate in according to $\mathcal{T}$ and submit his personal resume $PR = \{ID, \mathcal{T}, b\}$ to SP (step 2), where ID, $\mathcal{T}$, and b represent DC's identity certificate, desirable task, and bid for completing the task, respectively. When receiving DC's PR, SP immediately recruits DC or not according to his long-term data quality $q^{(l)}$ (step 3). After being recruited, DC submits sensing data to SP (step 4). Then, SP pays DC $p^{(a)}$ based on DC's current data quality $q^{(c)}$ (step 5).

Assumption 1. Notice that the number of malicious DCs who submit low quality data deliberately is limited, where Oasis can still discover the sensing task's truthful result from the multiple recruited DCs. Besides, the DCs' long-term data qualities and bids are *independent and identically distributed* (*i.i.d.*) sampled from some unknown distributions, which is reasonable and accepted generally.

B. Problem Definition

Oasis supposes that DCs are intelligent enough to determine rationally whether or not to participate in tasks according to his personal condition and tasks' demands. In other words, DC defaults that he can complete the task with high quality in time if he decides to participate in.

In the recruitment phase, considering that DCs submitting higher quality data chronically are more likely to keep performing well in future, Oasis requires SP to broadcast one single task $\mathcal{T} = \{l, r, t, \tau, B\}$ to DCs and maximize the recruited DCs' long-term data qualities $q^{(l)}$ under the budget constraint B in an online manner, which means that

$$Maximize \sum_{dc_j \in Re} q_j^{(l)} \tag{1}$$

$$\text{s.t.} \sum_{dc_j \in Re} p_j^{(a)} \leq B. \tag{2}$$

In the payment phase, to motivate DCs' high quality data, Oasis further requires SP to pay DC based on their current data qualities $q_j^{(c)}$ and bids b_j accurately, which means that

$$p_j^{(a)} = \mathcal{P}(q_j^{(c)}, b_j) \tag{3}$$

$$\text{s.t.} \sum_{dc_j \in Re} p_j^{(a)} \leq B. \tag{4}$$

where $\mathcal{P}$ denotes the strategic payment function.

C. Design Goals

Oasis should first satisfy the classic accepted *economic feasibility* to guarantee the effectiveness of the incentive mechanism, including *individual rationality* and *truthfulness*. The former enables DCs' participation motivation and the latter guarantees DCs' bidding strategy honesty.

- *Individual rationality:* The utility u_j of DC who submits a truthful bid and high quality data should be nonnegative, which means that $u_j \geq 0$;
- *Truthfulness:* The truthful bids $\hat{b}_j$ of DCs should be their dominant strategies, which means that all DCs bidding truthfully (i.e., $b_j = \hat{b}_j$) can maximize their utilities.

After submitting personal resumes, DCs will move casually (online mode) rather than stay and wait for the recruitment decision (offline mode) [23]. Thus, Oasis should also satisfy *online recruitment* and *constant competitiveness* to fit DCs' inconclusive movement and guarantee practicability and feasibility in dynamic MCS [12]. The former requires recruitment decisions are made irrevocably as soon as DCs arrive, and the latter implies effectiveness gap between online and offline mode is acceptable.

- *Online recruitment:* Without the future information, the recruitment decision should be made irrevocably as soon as a DC submits his personal resume. Meanwhile, the recruited DCs' long-term data quality is desired to be as higher as possible.
- *Constant competitiveness:* The ratio of DCs' long-term data qualities amount $\sum q^{(l)}$ between online manner and offline manner should be constant, which means that $\sum_{dc_j \in Re} q_j^{(l)} \geq c \sum_{dc_j \in Re^{off}} q_j^{(l)}$, $c \in \mathbb{C}$, where Re and Re^{off} denote the recruited DCs in online manner and offline manner, respectively.

Meanwhile, Oasis is also required to satisfy the quality incentive in both recruitment and payment phases.

- *Quality incentive in recruitment phase:* DCs should be recruited based on their long-term data qualities by (5)

$$\Pr(dc_j \in Re) \geq \Pr(dc_i \in Re),\ if\ b_j = b_i,\ q_j^{(l)} \geq q_i^{(l)}, \tag{5}$$

  where $\Pr(dc_j \in Re)$ represents the probability that dc_j is recruited.
- *Quality incentive in payment phase:* DCs' payments and utilities should be affected by their current data qualities in a single task by (6)

$$\begin{cases} p_j^{(a)} \geq p_j^{(b)},\ if\ q_j^{(c)} \geq \bar{q}_{Re}^{(c)}, \\ p_j^{(a)} < p_j^{(b)},\ if\ q_j^{(c)} < \bar{q}_{Re}^{(c)}, \end{cases} \tag{6}$$

  where $p_j^{(b)}, p_j^{(a)}$ denote DCs' basic and actual payment respectively, and $\bar{q}_{Re}^{(c)}$ denotes recruited DCs' average current data qualities.

The detailed analysis of all above design goals can be found in Section VI.

IV. Preliminaries

To evaluate DCs' data qualities accurately and realize online recruitment, Oasis makes good use of two techniques including truth discovery and knapsack secretary shown as follows.

A. Truth Discovery

To infer true information among multi-source data, the truth discovery is proposed and studied extensively [28], [29], [30]. The general principle of truth discovery is that the sources providing truth more often are considered to be more reliable, and the information supported by reliable sources is regarded as truth. Here we adopt the *Optimization based methods* [25] of truth discovery. The detailed process is discussed as follows.

For the sensing task $\mathcal{T}$, its truthful result can be deduced from multiple DCs by the optimization goal

$$\min F(W, \tilde{d}^{(\mathcal{T})}) = \sum_{dc_j \in \text{Re}} w_j \cdot dis(d_j^{(\mathcal{T})}, \tilde{d}^{(\mathcal{T})}), \tag{7}$$

where DCs' reliabilities (weight) and sensing data are denoted as W=$\{w_1, w_2, \ldots, w_n\}$ and D=$\{d_1^{(\mathcal{T})}, d_2^{(\mathcal{T})}, \ldots, d_n^{(\mathcal{T})}\}$, respectively. $\tilde{d}^{(\mathcal{T})}$ means the aggregated truthful result of $\mathcal{T}$. $dis(\cdot)$ can measure the distance between DCs' data and the truthful result. The objective is to minimize (7), meaning a DC providing data far from the truthful result is assigned a low weight as well as the truth $\tilde{d}^{(\mathcal{T})}$ is closer to the data of DC with high

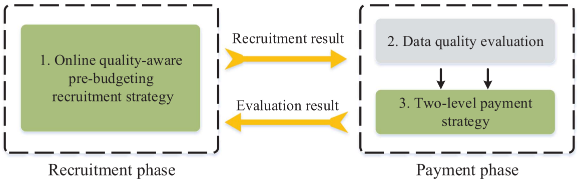


Fig. 3. Technical overview of Oasis.

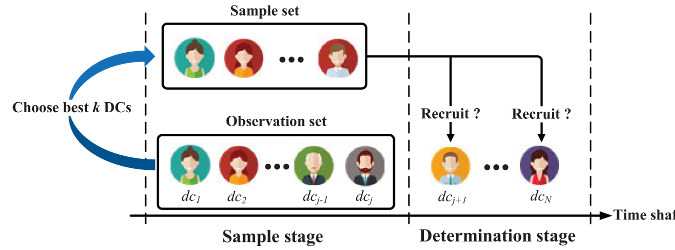


Fig. 4. Recruitment phase: online quality-aware pre-budgeting recruitment strategy.

weight. We can leverage coordinate descent to calculate $\tilde{d}^{(\mathcal{T})}$ and $W=\{w_1, w_2, \ldots, w_n\}$, where one variable is fixed to obtain the other variable and conduct iteratively until convergence.

B. Knapsack Secretary

Under the budget constrain and without knowing the condition of the future candidates, employers must decide whether or not to recruit a coming candidate who brings a certain value. Aiming to maximize the value brought by candidates under the budget constrain in an online manner, knapsack secretary [31] has been proposed to solve the knapsack problem as follows.

$$Maximize \sum_{c_i \in C} v(c_i) \tag{8}$$

$$\text{s.t.} \quad \sum_{c_i \in C} b(c_i) \leq B, \tag{9}$$

where B is budget constrain, $c_i \in C$ denotes candidates, $v(c_i)$, $b(c_i)$ represent c_i's value and bid. To recruit better candidates, the algorithm observes and collects candidates' information in first several steps, but does not recruit any candidates. After obtaining the general information, the algorithm recruits candidates in last several steps.

V. Detailed Design: Oasis

In the all phases of MCS, the data quality incentive is an indispensable factor, which motivates DCs' high quality data and further guarantees the distinguished MCS service. To solve the above problem, we propose Oasis that guarantees the quality incentive in both recruitment and payment phases under the budget constraint. This section first gives the technical overview of Oasis and then describes its construction in detail.

A. Technical Overview

Oasis consists of two tightly coupled phases including *recruitment phase* and *payment phase*, which are shown in Fig. 3. Technically, by the idea of *knapsack secretary*, an online quality-aware pre-budgeting recruitment strategy is devised, which not only realizes the quality incentive but also constrains the budget in advance for the following payment phase. Besides, *truth discovery* is adopted to evaluate DCs' data qualities, which affect DCs' actual payments that calculated from the *Myerson theorem* as well as the recruited probability of the next round recruitment phase. Both the *recruitment phase* and *payment phase* can implement data quality incentive and motivates DCs to submit high quality data effectively.

In the recruitment phase, referring to the *knapsack secretary*, we design a quality-aware pre-budgeting recruitment strategy. It decides whether the arriving DC's long-term data quality and bid satisfy the recruited criterion. Specifically, SP first observes the general level of DCs' long-term data qualities. Meanwhile, SP samples the information of k DCs whose long-term data qualities are largest in the observation process. Based on the sample set S, SP confirms the recruited criterion and decides whether recruit a arriving DC dc_j based on long-term data quality $q_j^{(l)}$ and bid b_j. In addition, the sample set S is dynamically updated in the recruitment phase, which adapts the market changes. Until the budget B is exhausted, the recruitment phase is terminate.

In the payment phase, with the result of *recruitment phase*, SP calculates the task's truthful result $\tilde{d}^{(\mathcal{T})}$ and further evaluates and updates DCs' current and long-term data qualities, $q^{(c)}$ and $q^{(l)}$. Moreover, the current data quality $q^{(c)}$ guides the two-level payment strategy to realize the quality incentive in payment phase and the long-term data quality $q^{(l)}$ affects the next round of *recruitment phase*. Specifically, receiving the data $d^{(\mathcal{T})}$ from all recruited DCs, SP first aggregates them and calculates the truth $\tilde{d}^{(\mathcal{T})}$ of sensing task. Based on the method of *truth discovery*, the truth is depended on the multiple iterative aggregation processes where the recruited DCs' different reliabilities are considered by Oasis. Subsequently, SP further evaluates DCs' data qualities in the current task $q^{(c)}$ by the distance between DCs' data $d^{(\mathcal{T})}$ and the truth $\tilde{d}^{(\mathcal{T})}$. Referring to the *Myerson theorem*, the evaluation results determine the recruited DCs' actual payment $p^{(a)}$ under the budget constraint as well as their long-term data qualities $q^{(l)}$, which reflect DCs' utilities in the current task and the recruited possibility in the future tasks.

B. Recruitment Phase: Online Quality-Aware Pre-Budgeting Recruitment Strategy

Unlike the existing works that also adopt the secretary method, Oasis devises the online quality-aware pre-budgeting recruitment strategy. It recruits DCs based on their long-term data qualities and bids, which motivates DCs' high quality data and supports the latter budget constraint payment strategy.

To clarify the devised online quality-aware pre-budgeting recruitment strategy, we assume that there's DC set $C = \{dc_1, dc_2, \ldots dc_N\}$. Each DC dc_j with personal resume $PR_j = \{ID_j, \mathcal{T}, b_j\}$ arrives to require for sensing task $\mathcal{T} = \{l, r, t, \tau, B\}$ in an online manner. Based on this assumption, the recruitment strategy is shown in Fig. 4. Specifically, SP just observes DCs' standard but not recruits any DCs in the sample stage. Then, with the sample set consisting of k DCs in the sample stage, the eligible DCs are recruited in the determination stage.

- *Sample Stage*

The hardest problem in the online recruitment is that SP cannot know the future information, especially when the budget constraint exists. To make a wise decision in this situation, using knapsack secretary, Oasis requires that SP only observes and samples but not recruits any DCs in the first δ steps to guide the subsequent recruitment phase.

Because of the periodical distribution of DCs' participant time that has been recognized widely [12], SP can first forecast the number N_l^t of candidate DCs in location l during time slot t according to the history check-in records of Apps. Then, SP observes first δ DCs ($\delta = \frac{N_l^t}{e}$ concluded by Vaze et al. [32]) and calculates their *value-per-buck* $\rho_j = \frac{q_j^{(l)}}{b_j}$ indicating the ratio of DC's long-term data quality to bid. Sorting the first δ DCs in non-increasing order based on their value-per-buck, SP gets the observed set

$$O = \{dc_1, dc_2, \ldots dc_\delta\}, \tag{10}$$

where $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_\delta$. SP picks up the first k DCs in observed set O to build up the sample set S, which satisfies

$$\sum_{j=1}^{k} p_j^{(b)} \leq B \, and \sum_{j=1}^{k+1} p_j^{(b)} > B, \tag{11}$$

$$\text{s.t.} \quad \begin{cases} p_j^{(b)} = b_j \cdot \frac{\rho_j}{\rho_*} \\ \rho_* = \frac{q_*^{(l)}}{b_*} \leq \rho_j, \forall dc_j \in S \end{cases} \tag{12}$$

where $p_j^{(b)}$ denotes the basic payment of dc_j, dc_* represents the DC whose value-per-buck ρ_* is the smallest in sample set. It's worth noting that the worst DC in sample set can be adjusted in recruitment phase to adapt to market changes.

- *Determination Stage*

Based on the sample set S and the worst DC dc_* in S, SP recruits DCs in the remaining steps.

When dc_j ($j \geq \delta + 1$) arrives, SP calculates his value-per-buck ρ_j and verifies that whether dc_j satisfies the recruited criterion defined as (13)

$$\begin{cases} \rho_j > \rho_*, \\ b_j < b_*, \end{cases} \tag{13}$$

note that dc_j's value-per-buck should be larger than ρ_* and bid is less than b_*. If dc_j satisfies conditions above, he can be recruited, where

$$Re = Re \cup dc_j. \tag{14}$$

No matter dc_j is recruited or not, SP adjusts sample set S and dc_* if $\rho_j > \rho_*$ for adapting market changes. SP first removes dc_* with ρ_* from S and joins dc_j into S, which means that

$$S = (S \backslash dc_*) \cup dc_j. \tag{15}$$

Then, SP sorts DCs of S in non-increasing order based on their value-per-buck again and updates dc_* to meet the next round of recruitment. This online recruitment phase can continue until the budget is exhausted, which means that

$$\sum_{dc_j \in Re} p_j^{(b)} \leq B. \tag{16}$$

It's worth noting that the basic payment can only imply the universal level of dc_j's payment, which assists the sum of the recruited DCs' actual payment to be under the budget constraint B. The actual payment of dc_j is determined by his basic payment and current data quality simultaneously, which is discussed in the latter payment phase.

The above two stages make up the integrated process of the online quality-aware pre-budgeting recruitment strategy that is shown in Algorithm V-B. The sample stage is discussed in the line 1∼7 and the determination stage is discussed in the line 8∼18.

Algorithm 1: Online Quality-Aware Pre-Budgeting Recruitment Strategy.

Input: dc_j with $PR_j = \{ID_j, \mathcal{T}, b_j\}$, N_l^t
Output: *Re*
1 **Sample stage**
2 $O = \{$first $\frac{N_l^t}{e}-th$ arriving $dc_j\}$ // *The first* $\frac{N_l^t}{e}$ *arrival DCs set up the observation set.*
3 Calculate $dc_j \in O$'s value-per-buck $\rho_j = \frac{q_j^{(l)}}{b_j}$
4 Sort $dc_j \in O$ in decreasing order based on their ρ_j
5 Select first k $dc_j \in O$ into S if $\sum_{j=1}^{k} p_j^{(b)} \leq B$ and $\sum_{j=1}^{k+1} p_j^{(b)} > B$
6 $S = \{dc_1, \ldots, dc_k\}$ // *The k DCs with largest* ρ *set up the sample set.*
7 $\rho_* \leq \rho_j, \forall dc_j \in S$ // ρ *of DC in sample set is larger than* ρ_*.
8 **Determination stage**
9 **for** *every new arrival* dc_j **do**
10   **if** $\rho_j > \rho_*$ **then**
11   **if** $b_j < b_*$ **then**
12   dc_j is recruited $Re = Re \cup dc_j$ if $\sum_{dc_j \in Re} p_j^{(b)} \leq B$
13   **else**
14   Removes dc_* with ρ_* from S and joins dc_j into S
15   Sort $dc_j \in S$ in decreasing order based on their ρ_j
16   **end**
17   **end**
18 **end**
19 Return *Re*

C. Payment Phase: Two-Level Payment Strategy Based on Data Quality Evaluation

After the recruitment phase, to motivate high quality data, Oasis pays recruited DCs $Re = \{dc_1, dc_2, \ldots dc_{|Re|}\}$ discriminatively based on their data quality relied on multiple rounds quality evaluation. The payment phase is shown in Fig. 5. Specifically, referring to the truth discovery, the truth of task is first aggregated from multiple recruited DCs' data. Then, each dc_j's data quality can be evaluated according to the distance

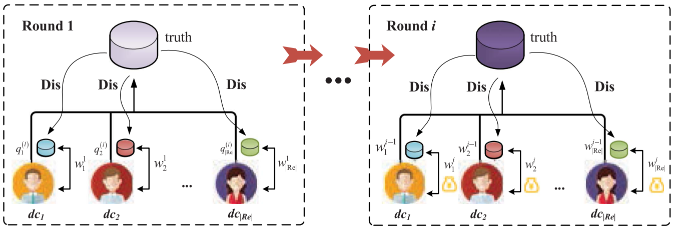


Fig. 5. Payment phase: two-level payment strategy based on data quality evaluation.

between his submitted data and the aggregated truth while dc_j's payment is determined relied on his data quality accurately.

1) Data Quality Evaluation: SP first calculates the task's truth by utilizing the optimization framework in (7), which can be realized by the following two iterative processes.

- *Truth update*

SP calculates truth $[\tilde{d}^{(\mathcal{T})}]^{(i)}$ in i-th round of truth calculation by (17)

$$[\tilde{d}^{(\mathcal{T})}]^{(i)} = \frac{\sum_{dc_j \in \mathrm{Re}} d_j^{(\mathcal{T})} \cdot w_j^{(i)}}{\sum_{dc_j \in \mathrm{Re}} w_j^{(i)}}. \tag{17}$$

- *Weight update*

Based on $[\tilde{d}^{(\mathcal{T})}]^{(i)}$, SP then updates DCs' weights in the i-th rounds by (18)

$$w_j^{(i)} = -\ln\left(\frac{dis(d_j^{(\mathcal{T})}, [\tilde{d}^{(\mathcal{T})}]^{(i)})}{\sum_{dc_{j'} \in \mathrm{Re}} dis(d_{j'}^{(\mathcal{T})}, [\tilde{d}^{(\mathcal{T})}]^{(i)})}\right). \tag{18}$$

Until these two processes are convergent in the i-th round, the aggregated truth is concluded by (19)

$$\tilde{d}^{(\mathcal{T})} = [\tilde{d}^{(\mathcal{T})}]^{(i)}, \; where \; dis([\tilde{d}^{(\mathcal{T})}]^{(i)}, [\tilde{d}^{(\mathcal{T})}]^{(i-1)}) \leq o. \tag{19}$$

$o \in \mathbb{R}$ is the small positive value that fulfills the quality requirement of $\mathcal{T}$.

With the truth $\tilde{d}^{(\mathcal{T})}$, SP evaluates dc_j's current data quality

$$q_j^{(c)} = e^{-dis(d_j^{(\mathcal{T})}, \tilde{d}^{(\mathcal{T})})}, \tag{20}$$

where $q_j^{(c)} = 1$ means that dc_j's sensing data $d_j^{(\mathcal{T})}$ is identical with the truth $\tilde{d}^{(\mathcal{T})}$, and $q_j^{(c)} \to 0$ implies that dc_j's sensing data $d_j^{(\mathcal{T})}$ is more deviant. To eliminate the effect of unit and scale differences between current data qualities, we further normalize $q_j^{(c)}$ and its value equals

$$[q_j^{(c)}]^{(nor)} = \frac{q_j^{(c)} - q_{\min}^{(c)}}{q_{\max}^{(c)} - q_{\min}^{(c)}}, \tag{21}$$

where $q_{\min}^{(c)}$ denotes the minimum current data quality and $q_{\max}^{(c)}$ denotes the max one of all DCs.

Then, to observe and reflect DCs' long-term data qualities, Oasis requires that the weight of DCs' data qualities decreases exponentially over time. That is, the more recent value is, the heavier weight is, but the older value is also given a certain weight. Thus, SP leverages EWMA (exponentially weighted moving average) to aggregate DCs' current data qualities in every round of tasks shown as follows

$$[q_j^{(l)}]^{(new)} = \varphi[q_j^{(l)}]^{(old)} + (1 - \varphi)q_j^{(c)}, \tag{22}$$

where $0 < \varphi < 0.5$, $\varphi \in \mathbb{R}$ denotes an adjustable parameter. It's worth noting that each DC's long-term data qualities are recorded by SP. We do not discuss the privacy preservation of data quality because this problem can be summarized as anonymous reputation question which has been extensively discussed [33].

Algorithm 2: Quality Evaluation and Payment Allocation.

Input: Re, $d_j^{(\mathcal{T})}$
Output: $\tilde{d}^{(\mathcal{T})}$, $q_j^{(c)}$, $q_j^{(l)}$, $p_j^{(b)}$, $p_j^{(a)}$

Current and long-term data quality evaluation
Calculate the task $\mathcal{T}$'s truthful result $\tilde{d}^{(\mathcal{T})}$ by truth discovery
for $dc_j \in Re$ **do**
  dc_j's current data quality $q_j^{(c)} = e^{-dis(d_j^{(\mathcal{T})}, \tilde{d}^{(\mathcal{T})})}$
  $[q_j^{(c)}]^{(nor)} = \frac{q_j^{(c)} - q_{\min}^{(c)}}{q_{\max}^{(c)} - q_{\min}^{(c)}}$ // *Eliminate the effect of unit and scale differences between current data qualities.*
  $[q_j^{(l)}]^{(new)} = \varphi[q_j^{(l)}]^{(old)} + (1-\varphi)q_j^{(c)}$
end
Two-level payment strategy
for $dc_j \in Re$ **do**
  dc_j's basic payment $p_j^{(b)} = b_j \cdot \frac{\rho_j}{\rho_*}$ // *Basic payment is affected by bid and value-per-buck.*
  dc_j's actual payment $p_j^{(a)} = \frac{[q_j^c]^{(nor)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)}$ // *Actual payment is affected by current data quality.*
end

2) Two-Level Payment Strategy: The two-level payment strategy consists of the basic payment and the actual payment. Specifically, the basic payment in the first level guarantees Oasis's economic feasibility, which is previously determined in the recruitment phase, and the actual payment in the second level implements Oasis's data quality incentive in the payment phase.

- *Basic payment*

To guarantee Oasis's economic feasibility, including truthfulness and individual rationality, we refer to the method in [34] and determine dc_j's basic payment $p_j^{(b)}$ by (23)

$$p_j^{(b)} = b_j \cdot \frac{\rho_j}{\rho_*}, \tag{23}$$

where b_j and ρ_j represent dc_j's bid and value-per-buck, ρ_* represents the recruitment threshold. With (23), Oasis's truthfulness and individual rationality can be proved, where the detailed proof process is shown in Section VI. Meanwhile, this basic payment strategy previously considers the budget constraint in the recruitment phase, which guarantees the budget feasibility in the payment phase.

- *Actual payment*

The nuclear idea of the actual payment is that the DCs who submit higher quality data may receive more payment than his basic payment and those who submit lower one is paid less by

(24)

$$\begin{cases} p_j^{(a)} \geq p_j^{(b)}, \; if \; [q_j^{(c)}]^{(nor)} \geq \bar{q}_{Re}^{(c)}, \\ p_j^{(a)} < p_j^{(b)}, \; if \; [q_j^{(c)}]^{(nor)} < \bar{q}_{Re}^{(c)}, \end{cases} \tag{24}$$

where $\bar{q}_{Re}^{(c)}$ indicates the mean value of current data quality among the recruited DCs

$$\bar{q}_{Re}^{(c)} = \frac{\sum_{dc_j \in Re} [q_j^{(c)}]^{(nor)}}{|Re|}. \tag{25}$$

Apparently, here Oasis assumes that the DC whose current data quality is larger than the mean level among his peer can be regarded as submitting high quality data. Thus, based on the basic payment, we further determine DCs' actual payment according to their current data qualities by (26)

$$p_j^{(a)} = \frac{[q_j^c]^{(nor)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)}. \tag{26}$$

Though there's a little distance between sensing data and truth, DC whose current data quality is higher than average level can also be paid more than what he expects. It means that Oasis allows DCs' mistakes within the certain range, which is helpful for motivation on DCs' subsequent participation. With (26), Oasis's quality incentive in the payment phase can be proved, where the detailed proof process is shown in Section VI.

The integrated process of the payment phase is shown in Algorithm V-C1. Specifically, DCs' current and long-term data qualities are evaluated in the line 1∼7 and their payments are determined in the line 8∼12.

D. Discussion

Here we demonstrate the practicability of our scheme, which means that Oasis is suitable and practical for MCS where multiple heterogeneous sensing tasks, mass data and participants exist.

Firstly, although focusing on the single task's quality incentive, Oasis can also solve this motivation problem under multiple heterogeneous sensing tasks situation, which is related works in Table I cannot realize. This scalability is based on the fact that the multi-tasks incentive problem can be transformed to the parallel single-task incentive in Oasis. Specifically, different sensing tasks' quality-aware incentives rely on different discovered sensing truth results that are certainly diverse, leading that Oasis cannot simultaneously realize multiple tasks' quality incentive with only one discovered sensing truth However, with the powerful computation and storage ability, SP is able to simultaneously calculate different sensing truth of different tasks. Based on these diverse sensing truth, SP can evaluate DCs' data quality and further realize quality-aware incentive on multiple tasks respectively. However, due to the absence of effective quality-aware incentive design and the non-quality-guided aiming setting, related works in Table I cannot realize the quality motivation in neither single-task's situation nor multi-tasks'.

Secondly, Oasis is able to handle the mass sensing data and DCs effectively, which is rooted in Oasis's ingenious technique adoption. In recruitment phase, Oasis first adopts knapsack secretary to devise an online quality-aware recruitment strategy where a dynamically mobile DC does not need to wait for all other DCs' arrival to get the recruitment decision. Meanwhile, because of the adoption of knapsack secretary, our online recruitment effectiveness is considerable compared with the offline mode when facing a large scale participants. In payment phase, Oasis evaluates DCs' data quality with the method of truth discovery and then pay for them based on the evaluation results according to Myerson theorem. Specifically, truth discovery is a mature and generally discussed technique inferring true information among multi-source mass data, and the Myerson-based payment calculation process is linear operation, meaning that SP can obtain result efficiently even facing a large scale sensing data and DCs. Thus, based on the above analysis, Oasis can fit in a large scale situation, realizing the practicability in mass data MCS.

TABLE III
COMPARISON OF VARIOUS INDEXES

Indexes \ *Schemes*	ON-SEG [12]	PACE [14]	Oasis
QIR	✗	✗	✓
QIP	✗	✗	✓
RM	online	offline	online
BF	✗	✓	✓

Note: QIR: quality incentive in recruitment; QIP: quality incentive in payment; RM: recruitment manner; BF: budget feasibility.

VI. Scheme Analysis

In this section, we prove that Oasis satisfies *economic feasibility* including *truthfulness* and *individual rationality*, *constant competitiveness*, *quality incentive* in recruitment and payment phases. Moreover, Table III shows the comparison of various indexes about Oasis and existing schemes [12], [14].

Theorem 1: If Oasis's payment strategy satisfies *bid-monotone*, meaning that a winning player can still win the auction if he declares a false bid that is less than his truthful bid, and *critical value*, meaning that a winning player will lose the auction if he declares a false bid that is more than his truthful payment, Oasis satisfies *truthfulness* [34].

Proof: Here, we first prove that Oasis's online recruitment process is *bid-monotone*. If dc_j is recruited, we can naturally conclude that

$$\rho_j = \frac{q_j^{(l)}}{b_j} > \rho_*. \tag{27}$$

Therefore, if dc_j falsely declares his bid ${b_j}' < b_j$, we can find that

$${\rho_j}' = \frac{q_j^{(l)}}{{b_j}'} > \frac{q_j^{(l)}}{b_j} = \rho_j > \rho_* \tag{28}$$

With the false bid ${b_j}' < b_j < b_*$, dc_j is still be recruited. Thus, Oasis's online recruitment process is *bid-monotone*.

To realize the data quality motivation, Oasis proposes the basic payment to guide the fundamental level of DCs' reward, which determines a large part of DCs' actual payment. Thus, here we analyze basic payment is the *critical value*.

Assuming that dc_j falsely declares his bid $b_j'' > p_j^{(b)}$, we can find that dc_j's false value-per-buck ρ_j'' satisfies

$$\rho_j'' = \frac{q_j^{(l)}}{b_j''} < \frac{q_j^{(l)}}{p_j^{(b)}}. \tag{29}$$

The truthful basic payment of dc_j if he bids truthfully satisfies

$$p_j^{(b)} = b_j \cdot \frac{\rho_j}{\rho_*}. \tag{30}$$

So the upper bound of dc_j's buck-per-bang value is

$$\rho_j'' < \frac{q_j^{(l)}}{p_j^{(b)}} = \frac{q_j^{(l)}}{b_j \cdot \frac{\rho_j}{\rho_*}}, \tag{31}$$

where

$$\frac{q_j^{(l)}}{b_j \cdot \rho_j} = \frac{q_j^{(l)}}{b_j \cdot \frac{q_j^{(l)}}{b_j}} = 1. \tag{32}$$

Thus, we can conclude that

$$\rho_j'' < \rho_*, \tag{33}$$

which means that dc_j with false bid $b_j'' > p_j^{(b)}$ cannot be recruited anymore, proving that the basic payment of Oasis is the *critical value*.

Based on the above proof that Oasis guarantees both *bid-monotone* and *critical value*, which meets Theorem 1, we can conclude that Oasis satisfies truthfulness.

Theorem 2: If each DC $dc_j \in Re$ who behaves truthfully can obtain the non-negative utility, where $u_j > 0$, Oasis satisfies *individual rationality*.

Proof: Here, based on Theorem 1, we can learn that each DC bids truthfully and his utility

$$u_j = p_j^{(a)} - c_j = p_j^{(a)} - b_j. \tag{34}$$

If dc_j is recruited, we can conclude that

$$p_j^{(b)} = b_j \cdot \frac{\rho_j}{\rho_*} > b_j. \tag{35}$$

Then, if dc_j submits high quality data in this task, which means that $[q_j^{(c)}]^{(nor)} \geq \bar{q}_{Re}^{(c)}$, his actual payment satisfies

$$p_j^{(a)} = \frac{[q_j^{(c)}]^{(nor)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)} \geq p_j^{(b)} > b_j = c_j. \tag{36}$$

We can conclude that

$$u_j = p_j^{(a)} - c_j > 0, \tag{37}$$

which means that dc_j can get a nonnegative utility as long as he can be considered to submit high quality data, meeting Theorem 2.

Based on the above proof that Oasis meets Theorem 2, we can conclude that Oasis satisfies individual rationality.

Remark 1: Oasis cannot guarantee the individual rationality of DC whose current data quality is lower than the peer mean level, which also motivates DC's high quality sensing data.

Theorem 3: If the ratio of DCs' long-term data qualities amount $\sum q^{(l)}$ between online manner and offline manner is constant, where $\sum_{dc_j \in Re} q_j^{(l)} \geq c \sum_{dc_j \in Re^{off}} q_j^{(l)}$, $c \in \mathbb{C}$, Oasis satisfies *constant competitiveness*.

Proof: Here, to prove Theorem 3, we first relax Algorithm V-B to the offline scenario, where all of the arriving DCs wait and set up a candidate set. Using the recruitment method in Algorithm V-B, SP compares and recruits DCs from the candidate set as many as possible under the budget constraint, where the recruited DCs hold the biggest value-per-buck in the candidate set. Let Re^{off} be the recruited DCs set of Algorithm V-B in the offline scenario and the value-per-buck threshold is ρ_*^{off}. Running in an offline manner, ρ_*^{off} is concluded after observing all DCs, and we can obtain that $\rho_*^{off} \geq \rho_*$. Based on this conclusion, we can assume that all DCs in Re^{off} are eligible for Re if they appear in the recruitment phase in Algorithm V-B.

Based on Algorithm V-B, a new dc_j arriving at step s $(\delta < s \leq N_l^t)$ has the chance to be recruited if dc_i is sampled in S with the smallest value-per-buck and $\rho_i < \rho_j$. Because the permutations are uniformly random, the probability that at step s, the DC with biggest value in sample set S is sampled at or before step δ is $\frac{\delta}{s-1}$. Meanwhile, the probability of any $dc_j \in Re^{off}$ arriving at step s is $\frac{1}{N_l^t}$. Thus, the probability of recruiting dc_j when he arrives at step s without any other condition is

$$\Pr(dc_j \in Re^{off}) = \sum_{s=\delta+1}^{N_l^t} \frac{1}{N_l^t} \cdot \frac{\delta}{s-1}, \tag{38}$$

where

$$\sum_{s=\delta+1}^{N_l^t} \frac{1}{N_l^t} \cdot \frac{\delta}{s-1} = \frac{\delta}{N_l^t} \cdot \sum_{s=\delta+1}^{N_l^t} \frac{1}{s-1} = \frac{\delta}{N_l^t} \cdot \int_{\delta+1}^{N_l^t} \frac{1}{s-1} ds \tag{39}$$

$$> \frac{\delta}{N_l^t} \cdot \int_{\delta}^{N_l^t} \frac{1}{x} dx = \frac{\delta}{N_l^t} \cdot \ln\left(\frac{N_l^t}{\delta}\right).$$

Following the conclusion of Vaze *et al* [32], we have $\delta = \frac{N_l^t}{e}$. Thus, we can conclude that the probability of $dc_j \in Re^{off}$ is recruited in Algorithm V-B is more than $\frac{1}{e}$. By linearity of expectation, we have the expected value of Re which is at least

$$E\left(\sum\nolimits_{dc_j \in Re} q_j^{(l)}\right) \geq \frac{1}{e} \sum\nolimits_{dc_j \in Re^{off}} q_j^{(l)}. \tag{40}$$

Then we discuss the second recruitment constraint $b_j < b_*$. There's no doubt that for any two DCs dc_* and dc_j, if $\rho_* < \rho_j$, $\Pr(b_* > b_j) = \frac{1}{2}$ under the assumption that $\rho_j = \frac{q_j^{(l)}}{b_j}$ [32]. Thus, the probability that dc_j who satisfies $\rho_* < \rho_j$ is recruited is $\frac{1}{2}$. Based on this conclusion, we can further discuss the lower bound of the expected value of Re which is at least

$$E\left(\sum\nolimits_{dc_j\in Re} q_j^{(l)}\right) \geq \frac{1}{2e}\sum\nolimits_{dc_j\in Re^{off}} q_j^{(l)}. \tag{41}$$

$\frac{1}{2e}$ is a constant that is independent of the other parameters, meeting Theorem 3.

Based on the above proof that Oasis meets Theorem 3, we can conclude that Oasis satisfies constant competitiveness.

Theorem 4: If a DC dc_j with the higher long-term data quality has a greater probability being recruited, where $\Pr(dc_j \in Re) \geq \Pr(dc_i \in Re),\ if\ b_j = b_i,\ q_j^{(l)} \geq q_i^{(l)}$, Oasis satisfies *quality incentive in recruitment phase*.

Proof: Here, we assume that there're two DCs dc_i, dc_j arriving and requiring for task simultaneously with $q_i^{(l)} < q_j^{(l)}$ and $b_i = b_j$. Based on Algorithm V-B, SP calculates their value-per-buck

$$\rho_i = \frac{q_i^{(l)}}{b_i} < \frac{q_j^{(l)}}{b_j} = \rho_j. \tag{42}$$

- If value-per-buck threshold ρ_* satisfies

$$\rho_* < \rho_i < \rho_j, \tag{43}$$

  then dc_j is recruited first and it is uncertain whether dc_i can be recruited after the recruitment of dc_j because of the change of ρ_*.
- If value-per-buck threshold ρ_* satisfies

$$\rho_i < \rho_* < \rho_j, \tag{44}$$

  then dc_j is recruited and dc_i has no chance.
- If value-per-buck threshold ρ_* satisfies

$$\rho_i < \rho_j < \rho_*, \tag{45}$$

  neither of them can be recruited.

Thus, with the same bid and arrival time, the probability of recruitment

$$\Pr(dc_j \in Re) > \Pr(dc_i \in Re), \tag{46}$$

which satisfies Theorem 4.

Thus, we can conclude that Oasis satisfies quality incentive in recruitment phase.

Theorem 5: If a DC dc_i obtains more payment than his expectation when he submit higher quality data than his peer, where $p_j^{(a)} > p_j^{(b)}, q_j^{(c)} > q_{Re}^{(c)}$, and vice versa, where $p_j^{(a)} < p_j^{(b)}, q_j^{(c)} < q_{Re}^{(c)}$, Oasis satisfies *quality incentive in payment phase*.

Proof: In Oasis, each DC $dc_j \in Re$ can obtain the actual payment $p_j^{(a)} = \frac{q_j^{(c)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)}$. It means that the DC whose current data quality is lower than the mean level cannot receive his desired payment, where

$$p_j^{(a)} = \frac{q_j^{(c)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)} < p_j^{(b)}. \tag{47}$$

On the contrary, the DC whose current data quality is no less than the mean value can receive his desired payment or more than it, where

$$p_j^{(a)} = \frac{q_j^{(c)}}{\bar{q}_{Re}^{(c)}} \cdot p_j^{(b)} \geq p_j^{(b)}. \tag{48}$$

From above, (47) and (48) meet Theorem 4. Thus, we can conclude that Oasis satisfies quality incentive in payment phase.

In addition to the above theorems, we also notice that Oasis's recruitment strategy inevitably makes it impossible for first-arriving DCs to be recruited, which roots in the knapsack secretary method. Meanwhile, there're several works that attempt to settle this problem [23]. While, this is not the focus of our work, therefore, we do not discuss this issue anymore.

VII. Performance Evaluation

In this section, we discuss the effectiveness of Oasis on recruitment result and data quality motivation, respectively.

A. Experimental Setup

Experimental environment: Python language programming was used to achieve the proposed scheme and relevant contrast schemes. Furthermore, the experiment environment is: Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz, 8192 MB RAM, and the Operating System is Windows 7.

Dataset and parameters setup: To make the comparison experiments more equitable and acceptable,

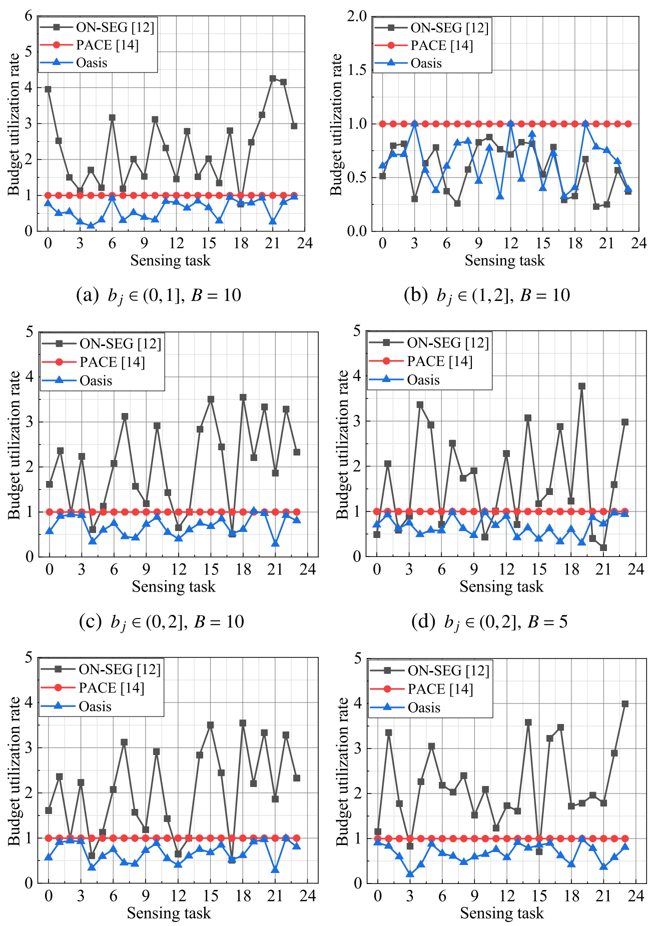
we adopt two real-world datasets, including temperature dataset[1] and movie rating dataset MovieLens [35]. Firstly, the former dataset is utilized generally in MCS [14], [26] and includes 5030 temperature data sensed by 367 taxis running in Rome. This data structure and large scale can effectively and practically fit experimental demands in MCS. Specifically, we divide February 2, 2014, into 24 timeslots, with each slot being one hour, and set up 24 sensing tasks requiring each hour's temperature in Rome. Meanwhile, we assume that the taxis with relevant temperature data are willing to participate in sensing tasks. Secondly, the latter dataset MovieLens contains 1000209 rating records (1-5 point) of 3900 movies made by 6040 Movielens users. To simulate the MCS process, we first add the random noise into the rating records based on the differential privacy, which can be viewed as the varying-quality sensing data. Subsequently, we model the sensing task as collecting 10 certain movies' mean rating records where Oasis's and comparison works' sensing result accuracy can be tested. Moreover, we dynamically set up the parameters as shown in Table IV.

Comparison indexes: To demonstrate Oasis's superiority in recruitment result and data quality motivation, we adopt ON-SEG [12] and PACE [14] as comparisons, representative schemes for online incentive mechanism and quality-aware incentive mechanism, respectively. On the one hand, we explore the budget feasibility and the recruited DCs' long-term data qualities to testify the recruitment result of three schemes. On the other hand, we also test DCs' data qualities, utilities and sensing result accuracy to verify three schemes' data quality motivation.

[1]This outdoor temperature sensing traces dataset can be found in http://crawdad.org/queensu/crowd-temperature/20151120.

TABLE IV
THE PARAMETERS SETUP

Parameter	Value
B	$5, 10, 15$
b_j	$(0, 1]$, $(1, 2]$, $(0, 2]$
$q_j^{(l)}$	0.5
φ	$0.4, 0.3$



(a) $b_j \in (0, 1]$, $B = 10$ (b) $b_j \in (1, 2]$, $B = 10$

(c) $b_j \in (0, 2]$, $B = 10$ (d) $b_j \in (0, 2]$, $B = 5$

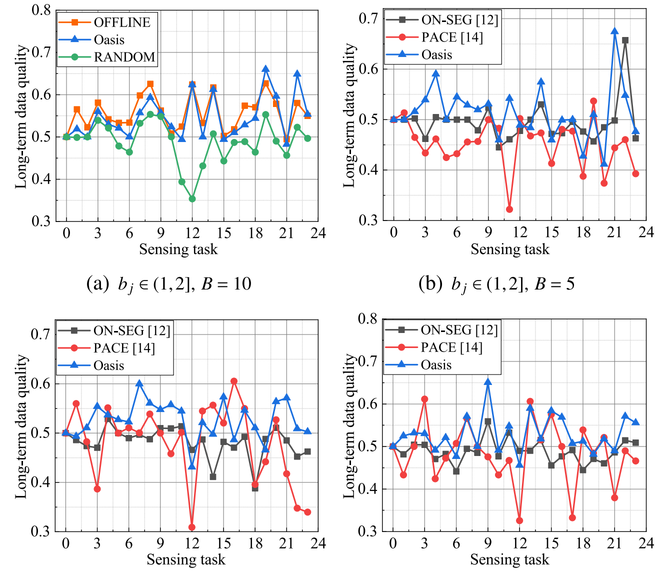
(e) $b_j \in (0, 2]$, $B = 10$ (f) $b_j \in (0, 2]$, $B = 15$

Fig. 6. Budget utilization rates.

B. Recruitment Result

1) Budget Feasibility: If the budget feasibility cannot be satisfied, the recruitment strategy is hard to implement chronically. Thus, we test the budget utilization rate $UR = \frac{\sum_{dc_j \in Re} p_j{}^{(a)}}{B}$ under ON-SEG [12], PACE [14], and Oasis to explore whether the budget feasibility of three schemes can be satisfied.

The experiment results under different DCs' bids are shown in Fig. 6(a), (b), and (c), respectively. We can find that PACE can always guarantees that $UR = 1$, which means that it uses all budget adequately to reward DCs. When $b_j \in (0, 2]$ and $b_j \in (0, 1]$, ON-SEG overspends generally because of the unreasonable recruitment and payment strategies. Like ON-SEG, Oasis runs in an online manner similarly but Oasis can always



(a) $b_j \in (1, 2]$, $B = 10$ (b) $b_j \in (1, 2]$, $B = 5$

(c) $b_j \in (1, 2]$, $B = 10$ (d) $b_j \in (1, 2]$, $B = 15$

Fig. 7. Recruited DCs' long-term data qualities.

ensure $UR \leq 1$. Moreover, under the different budget constraint, the budget utilization rates of ON-SEG, PACE and Oasis are shown in Fig. 6(d), (e), and (f), respectively. There's obvious that PACE and Oasis can always satisfy $UR \leq 1$ and implement chronically. However, ON-SEG cannot cover the total payment for all DCs under the budget constraint in most case.

Even though recruiting DCs in an online manner, Oasis previously considers the budget constraint, which prevents SP from recruiting more DCs than his can afford. Thus, thanks to the pre-budgeting recruitment strategy and the reasonable payment strategy, Oasis can obtain abundant data from the recruited DCs under the budget constraint. Therefore, Oasis enables SP to be profitable, which guarantees Oasis's implementation chronically.

2) Recruited DCs' Long-Term Data Quality: Oasis aims to recruit DCs with higher long-term data qualities $q^{(l)}$ in an online manner, which is different from the existing works. Thus, to test Oasis's effectiveness, we first compare the recruited DCs' long-term data qualities of Oasis with the OFFLINE recruitment strategy (Algorithm V-B in the offline scenario) and the RANDOM recruitment strategy, which recruits DCs randomly. Then, to demonstrate Oasis's superiority, we further evaluate recruited DCs' long-term data qualities of ON-SEG [12], PACE [14], Oasis, respectively. Due to the difference in recruited DCs' number, we evaluate the average (not total) long-term data qualities $\bar{q}^{(l)}$ of different schemes, which is more impartial.

The comparison result of Oasis, OFFLINE, and RANDOM is shown in Fig. 7(a). The effectiveness of RANDOM is unstable and unpromising obviously. Meanwhile, Oasis can always maintain a little distance between OFFLINE, which is ascribed to Oasis's ignorance of future information. Due to the evaluated average value, the effectiveness of Oasis is better than OFFLINE casually, including sensing tasks 19, 22, where the recruited DCs' number of OFFLINE is more than Oasis.

Overall, due to the observation on DCs' general long-term data quality level in the sample set, although online recruiting DCs, rather than selecting DCs with the candidate set in the greedy manner, Oasis still achieves superior recruitment results. Thus, Oasis can keep an outstanding superiority invariably compared with RANDOM. Meanwhile, Oasis can approach OFFLINE's performance, which can be viewed as the optimal result.

Fig. 7(b), (c), and (d) indicate the recruited DCs' long-term data qualities in ON-SEG, PACE, Oasis under different budget constraints, respectively. Because of ignoring DCs' data qualities while recruiting DCs, PACE cannot recruit DCs with high long-term data qualities, which leads to low quality data purchase and damage to MCS service. Despite recruiting DCs based on matching degree between DCs and task, ON-SEG can only match DCs and task well, rather than recruiting DCs with high long-term data qualities. However, recruiting DCs based on their long-term data qualities and bids simultaneously, Oasis's recruited DCs keeps preferable qualities under the different budget constraint.

Because Oasis's recruitment factor is DC's long-term data quality rather than the others, which means that only the DCs with high long-term data qualities have the chance to be recruited, Oasis generally stays ahead compared with the other works. It implies that the recruited DCs of Oasis are more likely to submit high quality data in the current task. Meanwhile, Oasis prefers to recruit DCs with high long-term data qualities, which motivates DCs to submit high quality data insistently to maximize their recruited probability, while the other works cannot. Thus, Oasis guarantees that the recruited DCs' submitted data qualities as well as realizes the quality incentive in recruitment phase.

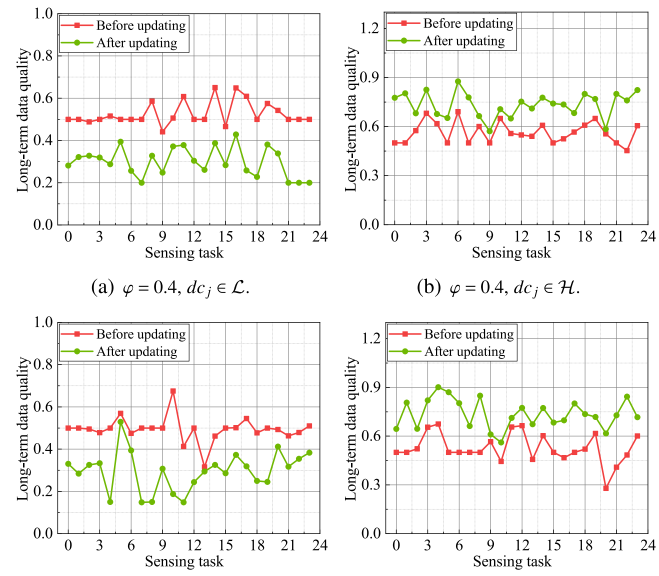
C. Data Quality Motivation

1) Quality Incentive in Recruitment Phase: The long-term data quality $q^{(l)}$ can affect the DC's recruited possibility. Only DCs' $q^{(l)}$ can be changed dynamically based on their data qualities at each round of sensing task can incentive mechanism realize the quality-aware in recruitment phase. However, ON-SEG [12] and PACE [14] are failed to adjust DCs' long-term data qualities. Thus, we only test $q^{(l)}$ of DCs before and after a sensing task in Oasis. Here we suppose that a DC can be viewed as submitting high quality data, $dc_j \in \mathcal{H}$, if he satisfies

$$dis(d_j^{(\mathcal{T})}, \tilde{d}^{(\mathcal{T})}) \leq \frac{\sum_{dc_j \in \text{Re}} dis(d_j^{(\mathcal{T})}, \tilde{d}^{(\mathcal{T})})}{|\,\text{Re}\,|}, \tag{49}$$

where $d_j^{(\mathcal{T})}$ represents dc_j's sensing data and $\tilde{d}^{(\mathcal{T})}$ denotes the ground truth. Otherwise, a DC is viewed as submitting low quality data, $dc_j \in \mathcal{L}$.

Fig. 8(a) and (c) shows the contrast between the average long-term data quality $\bar{q}^{(l)}$ of $dc_j \in \mathcal{L}$ before and after a sensing task. We can find that DCs' $\bar{q}^{(l)}$ drops sharply, which leads to the loss of the next round of recruitment because of their poor long-term data qualities. On the contrary, due to the less distance between sensing data and truth, $\bar{q}^{(l)}$ of $dc_j \in \mathcal{H}$ rises steadily in Fig. 8(b) and (d). This growth makes DCs more competitive in future recruitment.



(a) $\varphi = 0.4$, $dc_j \in \mathcal{L}$. (b) $\varphi = 0.4$, $dc_j \in \mathcal{H}$.

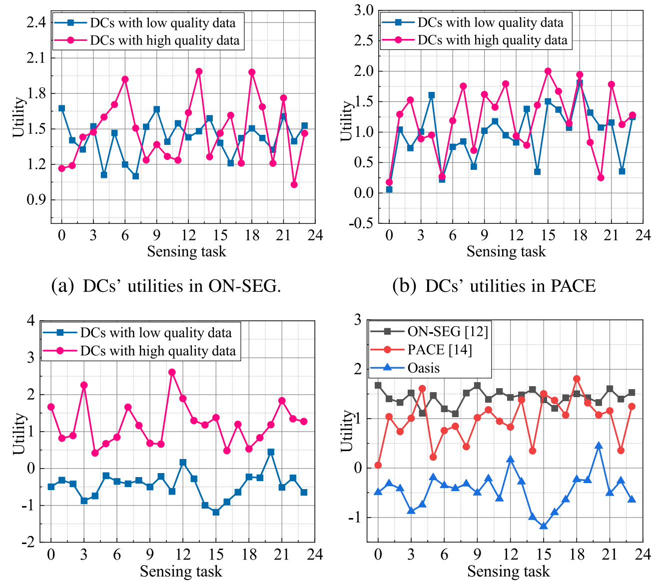
(c) $\varphi = 0.3$, $dc_j \in \mathcal{L}$. (d) $\varphi = 0.3$, $dc_j \in \mathcal{H}$.

Fig. 8. Long-term data quality's change.

Because long-term data qualities are dynamically changed in Oasis, but not static, the DCs' submitted data qualities in the current task severely affect his long-term data qualities. Specifically, the DCs submitting low quality data are unlikely to be recruited due to the decreasing long-term data qualities. On the contrary, thanks to the increasing long-term data qualities, the DCs with high quality data have a higher chance of being recruited in the future. Therefore, considering the sustainable utilities, the rational DCs are willing to submit high quality sensing data for the long term in Oasis. However, because DCs' data quality adjustment is ignored, ON-SEG and PACE cannot guarantee the data quality motivation in recruitment phase.

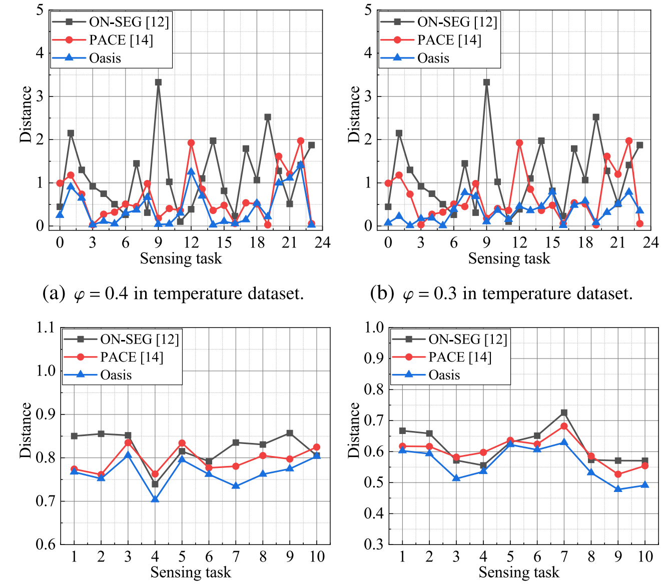
2) Quality Incentive in Payment Phase: If all DCs get equal utilities roughly whether they collect high quality data or not, the quality incentive in the payment phase is ineffective. Thus, to test ON-SEG [12], PACE [14], and Oasis's quality incentive in the payment phase, we evaluate the utilities of DCs submitting high quality data or not, respectively.

Fig. 9(a) shows the average utilities of DCs in ON-SEG. We can find that the utilities of all DCs are equal roughly, which means that whether DCs submit high quality data or not can always obtain similar utilities compared with others. Fig. 9(b) indicates the average utilities of DCs in PACE. There's evident that the average utilities of DCs submitting high quality data is more than that of DCs submitting low quality data generally. However, because of the inaccurate data quality evaluation, the situation exists where DCs submitting low quality data obtain more utilities than DCs submitting high quality data. On the other hand, Oasis pays DCs based on their data quality strictly, and the distance between utilities of DCs is shown in Fig. 9(c). In addition, the average utilities of DCs submitting low quality data in three schemes are shown in Fig. 9(d). We can find that



(a) DCs' utilities in ON-SEG. (b) DCs' utilities in PACE (c) DCs' utilities in Oasis (d) DCs' utilities with low quality data.

Fig. 9. DCs' utilities in payment phase.



(a) $\varphi = 0.4$ in temperature dataset. (b) $\varphi = 0.3$ in temperature dataset. (c) $\varphi = 0.4$ in MovieLens. (d) $\varphi = 0.3$ in MovieLens.

Fig. 10. Sensing result accuracy.

the results of Oasis are even negative. On the contrary, ON-SEG and PACE can always permit DCs to obtain positive utilities.

From above, on account of the accurate evaluation on DCs' data quality and the reasonable quality-aware payment strategy, Oasis guarantees that the higher quality data DCs submit, the more payment they obtain, vice versa. In other words, there's no doubt that only DCs submitting high quality data can obtain a higher utility than his peer, which means that Oasis guarantees quality-aware in payment phase effectively. Meanwhile, compared with the other works, Oasis has tougher penalties (negative utilities) for the DCs with low quality data. However, the DCs with low quality data in the other works can still obtain positive utilities, leading that DCs arbitrarily submit low quality data with out any punishment. Therefore, we deem that Oasis can motivate DCs to submit high quality data and punish DCs submitting low quality data in payment phase fiercely, but ON-SEG and PACE fail to.

3) Sensing Result Accuracy: Improving the sensing result accuracy is the most essential aim of the quality-aware incentive mechanisms. Thus, based on two real-world datasets, here we compare the sensing result accuracy of ON-SEG [12], PACE [14], and Oasis. To unify standard, we first view the sensing result accuracy as the distance between the ground truth and the scheme's sensing result [17]

$$dis(t^{(g)}, \tilde{d}^{(\mathcal{T})}) = |t^{(g)} - \tilde{d}^{(\mathcal{T})}|, \tag{50}$$

where $t^{(g)}$ denotes the ground truth[2] and $\tilde{d}^{\mathcal{T}}$ represents the scheme's sensing result. Then, we test $dis(t^{(g)}, \tilde{d}^{(\mathcal{T})})$ of three schemes.

[2]The ground truths in Fig. 10(a) and (b) mean the real temperature in Rome, which can be found in https://www.worldweatheronline.com/. The ground truths in Fig. 10(c) and (d) mean the accuracy mean rating calculated by the original rating data without disturbance noise.

The test results based on two different datasets are shown in Fig 10. Viewing the task completion amount as the data quality, ON-SEG can only focus on DCs' locations but ignore DCs' sensing data, which is the reason why ON-SEG's sensing result is far from the ground truth both in temperature dataset and MovieLens. Though focusing on the DCs' data qualities, PACE supposes that the average value is the truth, which is unreasonable in MCS due to the difference of DCs' reliabilities. Thus, the distance between the ground truth and PACE's sensing result is big relatively in all experimental results. However, with the change of φ, Oasis's sensing result can always maintain a less distance with the ground truth due to the all-phase data quality motivation and reasonable data aggregation.

In brief, due to the quality-incentive in both recruitment phase and payment phase, Oasis can guarantee the higher sensing results qualities, where the average distance of Oasis is 62.7% and 36% less than ON-SEG's and PACE's, respectively. Thus, Oasis is always able to improve sensing result accuracy significantly.

VIII. Conclusion and Future Work

In this article, we propose an online all-phase quality-aware incentive mechanism (Oasis) to realize the quality incentive in both recruitment and payment phases. Referring to knapsack secretary method, we design a quality-aware pre-budgeting recruitment strategy, which decides whether the arriving DC's long-term data quality and bid satisfy the recruited criterion. It realizes the quality incentive in the recruitment phase and guarantees the payment phase's quality incentive under the budget constraint. Then, with the task's truthful result calculated by truth discovery, we evaluate and update users' current and long-term data qualities. The evaluation results guide the design of two-level payment strategy, implementing Oasis's economic feasibility and quality incentive in payment phase. Meanwhile, the updated results affect the next round of the online recruitment. Extensive experiments demonstrate that Oasis has more outstanding recruitment effectiveness and data quality motivation compared with the existing mechanisms.

However, Oasis only focuses on the single task's quality incentive but cannot adapt to multiple heterogeneous tasks simultaneously. As part of future work, we will consider multiple heterogeneous tasks quality-aware recruitment and payment, realizing data quality motivation.

References

[1] A. Capponi, C. Fiandrino, B. Kantarci, L. Foschini, D. Kliazovich, and P. Bouvry, "A survey on mobile crowdsensing systems: Challenges, solutions, and opportunities," *IEEE Commun. Surv. Tut.*, vol. 21, no. 3, pp. 2419–2465, Third Quarter 2019.

[2] Y. Cheng et al., "Aircloud: A cloud-based air-quality monitoring system for everyone," in *Proc. ACM Conf. Embedded Netw. Sensor Syst.*, 2014, pp. 251–265.

[3] WAZE, "Waze," [EB/OL], Accessed: Jul. 14, 2021. [Online]. Available: www.waze.com

[4] PARKNET, "Parknet," [EB/OL], Accessed: Jul. 14, 2021. [Online]. Available: https://sourceforge.net/projects/parknet/

[5] X. Zhang et al., "Incentives for mobile crowd sensing: A survey," *IEEE Commun. Surv. Tut.*, vol. 18, no. 1, pp. 54–67, First Quarter 2016.

[6] J. Liu, S. Huang, D. Li, S. Wen, and H. Liu, "Addictive incentive mechanism in crowdsensing from the perspective of behavioral economics," *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 5, pp. 1109–1127, May 2022.

[7] Y. Lin, Z. Cai, X. Wang, F. Hao, L. Wang, and A. M. V. V. Sai, "Multi-round incentive mechanism for cold start-enabled mobile crowdsensing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 993–1007, Jan. 2021.

[8] D. Peng, F. Wu, and G. Chen, "Data quality guided incentive mechanism design for crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 307–319, Feb. 2018.

[9] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 167–176.

[10] Y. Wen et al., "Quality-driven auction-based incentive mechanism for mobile crowd sensing," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4203–4214, Sep. 2015.

[11] H. Wang, S. Guo, J. Cao, and M. Guo, "MeLoDy: A long-term dynamic quality-aware incentive mechanism for crowdsourcing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 4, pp. 901–914, Apr. 2019.

[12] W. Liu, Y. Yang, E. Wang, and J. Wu, "Dynamic user recruitment with truthful pricing for mobile crowdsensing," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 1113–1122.

[13] J. Xu, Z. Rao, L. Xu, D. Yang, and T. Li, "Incentive mechanism for multiple cooperative tasks with compatible users in mobile crowd sensing via online communities," *IEEE Trans. Mobile Comput.*, vol. 19, no. 7, pp. 1618–1633, Jul. 2020.

[14] B. Zhao, S. Tang, X. Liu, and X. Zhang, "PACE: Privacy-preserving and quality-aware incentive mechanism for mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 20, no. 5, pp. 1924–1939, May 2021.

[15] K. Xue, B. Zhu, Q. Yang, N. Gai, D. S. L. Wei, and N. Yu, "InPPTD: A lightweight incentive-based privacy-preserving truth discovery for crowdsensing systems," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4305–4316, Mar. 2021.

[16] X. Wang et al., "Online spatial crowdsensing with expertise-aware truth inference and task allocation," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 412–427, Jan. 2022.

[17] F. Restuccia, N. Ghosh, S. Bhattacharjee, S. Das, and T. Melodia, "Quality of information in mobile crowdsensing: Survey and research challenges," *ACM Trans. Sensor Netw.*, vol. 13, no. 4, pp. 1–43, 2017.

[18] W. Tang, K. Zhang, J. Ren, Y. Zhang, and X. Shen, "Privacy-preserving task recommendation with win-win incentives for mobile crowdsourcing," *Inf. Sci.*, vol. 527, pp. 477–492, 2020.

[19] J. Wang et al., "Multi-task allocation in mobile crowd sensing with individual task quality assurance," *IEEE Trans. Mobile Comput.*, vol. 17, no. 9, pp. 2101–2113, Sep. 2018.

[20] X. Wang, R. Jia, X. Tian, X. Gan, L. Fu, and X. Wang, "Location-aware crowdsensing: Dynamic task assignment and truth inference," *IEEE Trans. Mobile Comput.*, vol. 19, no. 2, pp. 362–375, Feb. 2020.

[21] Y. Yang, W. Liu, E. Wang, and J. Wu, "A prediction-based user selection framework for heterogeneous mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2460–2473, Nov. 2019.

[22] H. Gao, C. H. Liu, J. Tang, D. Yang, P. Hui, and W. Wang, "Online quality-aware incentive mechanism for mobile crowd sensing with extra bonus," *IEEE Trans. Mobile Comput.*, vol. 18, no. 11, pp. 2589–2603, Nov. 2019.

[23] D. Zhao, X. Li, and H. Ma, "Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 647–661, Apr. 2016.

[24] B. Guo et al., "TaskMe: Toward a dynamic and quality-enhanced incentive mechanism for mobile crowd sensing," *Int. J. Hum.- Comput. Stud.*, vol. 102, pp. 14–26, 2017.

[25] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2014, pp. 1187–1198.

[26] C. Zhao, S. Yang, and J. A. McCann, "On the data quality in privacy-preserving mobile crowdsensing systems with untruthful reporting," *IEEE Trans. Mobile Comput.*, vol. 20, no. 2, pp. 647–661, Feb. 2021.

[27] S. Yang, F. Wu, S. Tang, X. Gao, B. Yang, and G. Chen, "On designing data quality-aware truth estimation and surplus sharing method for mobile crowdsensing," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 4, pp. 832–847, Apr. 2017.

[28] Y. Li et al., "A survey on truth discovery," *ACM SIGKDD Explorations Newslett.*, vol. 17, no. 2, pp. 1–16, 2015.

[29] X. Pang, Z. Wang, D. Liu, J. C. S. Lui, Q. Wang, and J. Ren, "Towards personalized privacy-preserving truth discovery over crowdsourced data streams," *IEEE/ACM Trans. Netw.*, vol. 30, no. 1, pp. 327–340, Feb. 2022.

[30] D. Wang, J. Ren, Z. Wang, X. Pang, Y. Zhang, and X. S. Shen, "Privacy-preserving streaming truth discovery in crowdsourcing with differential privacy," *IEEE Trans. Mobile Comput.*, vol. 21, no. 10, pp. 3757–3772, Oct. 2022.

[31] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg, "A knapsack secretary problem with applications," in *Proc. Approximation, Randomization, Combinatorial Optimization. Algorithms Techniques, 10th Int. Workshop, APPROX, 11th Int. Workshop, RANDOM*, 2007, pp. 16–28.

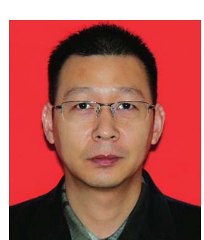
[32] R. Vaze, "Online knapsack problem under expected capacity constraint," in *Proc. IEEE Conf. Comput. Commun.*, 2018, pp. 2159–2167.

[33] X. O. Wang, W. Cheng, P. Mohapatra, and T. F. Abdelzaher, "ARTSense: Anonymous reputation and trust in participatory sensing," in *Proc. IEEE Conf. Comput. Commun.*, 2013, pp. 2517–2525.

[34] R. B. Myerson, "Optimal auction design," *Math. Oper. Res.*, vol. 6, no. 1, pp. 58–73, 1981.

[35] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Trans. Interact. Intell. Syst.*, vol. 5, no. 4, pp. 19:1–19:19, 2016.

Man Zhang received the BS degree in computer science from the Xi'an University of Technology, in 2018. She is currently working toward the PhD degree in Security of Cyberspace in Xidian University. Her research interests include privacy protection and intelligent mobile computing.



Xinghua Li (Member, IEEE) received the MS and PhD degrees in computer science from Xidian University, in 2004 and 2007, respectively. He is currently a professor in the School of Cyber Engineering, Xidian University, China. His research interests include wireless networks security, privacy protection, cloud computing, and security protocol formal methodology.

Yinbin Miao received the BS degree from the School of Telecommunication Engineering, Jilin University, Changchun, China, in 2011, and the PhD degree from the School of Telecommunication Engineering, Xidian University, Xi'an, China, in 2016. He is currently an associated professor with Xidian university. His current research interests include network security and cryptography.

Bin Luo received the BS degree in computer science from Xidian University, in 2015. She is currently working toward the PhD degree in Security of Cyberspace in Xidian University. Her research interests include privacy protection and trust management.

Siqi Ma received the PhD degree in information system from Singapore Management University, in 2018. She is currently a lecturer in the School of Information Technology and Electrical Engineering with the University of Queensland. Her research interests are mobile security, web security, IoT security.

Kim-Kwang Raymond Choo (Senior Member, IEEE) received the PhD degree in Information Security from the Queensland University of Technology, Australia, in 2006, and currently holds the Cloud Technology Endowed Professorship with the University of Texas at San Antonio. He is the founding co-editor-in-chief of ACM Distributed Ledger Technologies: Research & Practice, and founding Chair of IEEE TEMS Technical Committee on Blockchain and Distributed Ledger Technologies.

Robert H. Deng (Fellow, IEEE) is AXA chair professor of Cybersecurity and professor of Information Systems in the School of Information Systems, Singapore Management University since 2004. His research interests include data security and privacy, multimedia security, network and system security. He served/is serving on the editorial boards of many international journals, including TIFS, TDSC. He has received the Distinguished Paper Award (NDSS 2012), Best Paper Award (CMS 2012), Best Journal Paper Award (IEEE Communications Society 2017).