

An Online Reinforcement Learning Approach to Quality-Cost-Aware Task Allocation for Multi-Attribute Social Sensing

Yang Zhang, Daniel Zhang, Nathan Vance, Dong Wang

*Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556*

Abstract

Social sensing has emerged as a new sensing paradigm where humans (or devices on their behalf) collectively report measurements about the physical world. This paper focuses on a quality-cost-aware task allocation problem in multi-attribute social sensing applications. The goal is to identify a task allocation strategy (i.e., decide when and where to collect sensing data) to achieve an optimized tradeoff between the data quality and the sensing cost. While recent progress has been made to tackle similar problems, three important challenges have not been well addressed: (i) “online task allocation”: the task allocation schemes need to respond quickly to the potentially large dynamics of the measured variables in social sensing; (ii) “multi-attribute constrained optimization”: minimizing the overall sensing error given the dependencies and constraints of multiple attributes of the measured variables is a non-trivial problem to solve; (iii) “nonuniform task allocation cost”: the task allocation cost in social sensing often has a nonuniform distribution which adds additional complexity to the optimized task allocation problem. This paper develops a **Quality-Cost-Aware Online Task Allocation** (QCO-TA) scheme to address the above challenges using a principled on-

line reinforcement learning framework. We evaluate the QCO-TA scheme through a real-world social sensing application and the results show that our scheme significantly outperforms the state-of-the-art baselines in terms of both sensing accuracy and cost.

Keywords: Social Sensing; Multi-Attribute Optimization;
Quality-Cost-Aware Task Allocation; Online Reinforcement Learning

1. Introduction

This paper presents an online reinforcement learning framework to solve the quality-cost-aware task allocation problem in multi-attribute social sensing applications. Social sensing has emerged as a new sensing paradigm in pervasive and mobile computing applications where humans (or devices on their behalf) collectively report measurements about the physical world [1, 2]. Examples of social sensing applications include air quality and environment monitoring in smart cities using mobile devices [3], malfunctioning urban infrastructures reporting using geotagging [4], and damage assessment in disaster response using online social media [5]. In social sensing applications, participants perform sensing tasks at assigned locations to collect different attributes of the measured variables that are of interests to the application [6]. For example, in an urban air quality sensing application, participants are tasked to measure various air quality attributes (e.g., $PM_{2.5}$, PM_{10} , CO_2) at different locations of the city to estimate the overall air quality and identify potential health risks. We refer to this category of applications as *multi-attribute social sensing applications*.

In multi-attribute social sensing applications, there exists a fundamental tradeoff between *data quality* and *sensing (task allocation) cost* [3, 7].

In particular, it is essential to obtain comprehensive and accurate measurements to ensure the desired data quality of the social sensing applications. Such a dedicated data collection process often also encounters a high sensing cost (e.g., more incentives to recruit participants to perform sensing tasks) [3]. However, the high sensing cost may not always be affordable to the applications with a finite budget [8]. Therefore, a key challenge for social sensing applications is to find a task allocation strategy (i.e., decide when and where to collect sensing data) that achieves an optimized trade-off between the data quality and sensing cost. We refer to this problem as the *quality-cost-aware task allocation* problem. The current solutions to address this problem primarily focus on identifying an *optimal set* of sensing locations (i.e., cells) to collect measurements to minimize the overall sensing error [3, 9, 10, 11, 12, 13, 14, 15, 16]. However, these solutions cannot be directly adapted to solve our task allocation problem due to three challenges that have not been fully addressed: *online task allocation*, *multi-attribute constrained optimization* and *nonuniform task allocation cost*. We elaborate them below.

Online Task Allocation. Many social sensing applications are delay sensitive and require timely response to meet the application requirement [12]. For example, during a hurricane, it is crucial for the application to decide when and where the data should be collected to provide real-time situation awareness about the disaster. However, online task allocation in social sensing is challenging due to the large spatial-temporal dynamics of the measured variables and opportunistic nature of social sensing participants [12, 9]. This problem becomes more challenging in a multi-attribute sensing scenario where the values of all attributes change simultaneously. Several task allocation methods have been developed to address similar problems [3, 17].

However, a few important limitations exist. First, existing models largely ignore the high dynamics of the social sensing applications and allocate the sensing task to cells *one by one* until the data quality requirement is met [3]. Second, current solutions do not explicitly consider the correlation between different attributes, thus leading to sub-optimal task allocation solutions.

Multi-attribute Constrained Optimization. We observe that different sensing attributes often have different spatio-temporal distributions that will affect the task allocation decisions [9]. For example, the local optimized task allocation strategy for a particular sensing attribute may not be the global optimized task allocation strategy for all attributes. Furthermore, different sensing attributes may have inherent and complex dependencies. For example, the PM_{2.5} and CO₂ are often found to be correlated in a social sensing application that measures the air quality of a city [9]. It is not a trivial job to design a task allocation strategy that can effectively identify the optimal set of sensing cells that can minimize the sensing error across multiple interdependent sensing attributes with diversified distributions.

Nonuniform Task Allocation Cost. The task allocation cost in social sensing is often related to the incentives to motivate a participant to travel from one sensing location to another [18]. Such task allocation cost often has a non-uniform distribution (e.g., different travel distances will lead to different amount of incentives), which adds additional complexity to the optimized task allocation problem [19]. For example, in a social sensing application as shown in Figure 1, a participant at location *A* may be assigned to collect air quality readings at two possible locations: location *B* and *C*. The sensing measurements collected at *B* will reduce the overall sensing error more significantly. However, the task allocation cost at *B* is also higher than *C* because the travel distance between *A* and *B* is larger than the one

between A and C (i.e., $d_2 > d_1$). The question is which location we should send participant to perform the sensing task. To answer this question, the task allocation scheme needs to carefully explore the tradeoff between the data quality and the nonuniform task allocation cost.

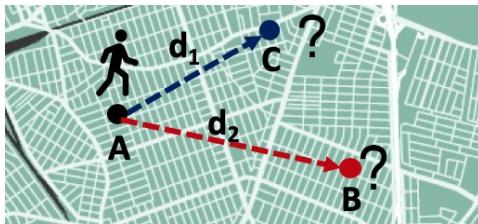


Figure 1: Tradeoff Between Quality and Cost

In this paper, we develop a Quality-Cost-Aware Online Task Allocation (QCO-TA) scheme to address above challenges under a principled online reinforcement learning framework. To address the *online task allocation* challenge, we develop an online learning algorithm that dynamically estimates the priority of each sensing cell for different sensing attributes at each cycle. To address the *multi-attribute constrained optimization* challenge, we develop a Bayesian inference scheme that judiciously combines the priority estimations of all sensing attributes into a comprehensive priority score that identifies the cells to effectively reduce the overall sensing error across different sensing attributes. To address the *nonuniform task allocation cost* challenge, we develop a principled reinforcement learning method to explicitly consider the nonuniform task allocation cost and learns the optimal set of sensing cells for the task allocation. Finally, we evaluate the QCO-TA scheme on a real-world social sensing dataset: Piemonte Air Quality Dataset. The results show that our scheme significantly outperforms the state-of-the-art baselines in both sensing accuracy and cost.

We choose the online and reinforcement learning framework to address the quality-cost-aware task allocation problem in multi-attribute social sensing applications for two main reasons. First, the sensing measurements in social sensing applications are often collected in real time [20]. The online learning technique is a great fit for such application scenarios because it is capable of dynamically adjusting the task allocation decisions based on the streaming sensing measurements. This is in contrast to the batch based learning techniques which often require a large amount of high-quality training data *a priori* that are not available in our problem setting. Second, our quality-cost-aware task allocation problem aims to find a task allocation strategy that achieves an optimized trade-off between data quality and sensing cost given incomplete sensing measurements (i.e., due to a finite sensing budget). The reinforcement learning technique is a goal-oriented learning technique that fits nicely into our problem. In particular, it provides a data-driven solution that learns to achieve a complex objective (i.e., Equation (2)) given a set of partially available sensing measurements. The reinforcement learning solution sharply contrasts with classical optimization techniques (e.g., linear programming, dynamic programming) which often require a complete set of sensing measurements to learn the optimized trade-off between data quality and sensing cost. However, such a complete sensing dataset is often not available due to limited sensing resources and coverage in social sensing applications [3].

A preliminary version of this work was published in the [21]. We refer to the scheme developed in the previous work as the online optimized multi-attribute task allocation (OO-MTA) scheme. The current paper is a significant extension of the previous work in the following aspects. First, we extend our previous model by explicitly exploring the optimized tradeoff

between the data quality and sensing cost in multi-attribute social sensing applications. In contrast, the OO-MTA only focuses on optimizing data quality and does not take the sensing cost into consideration. Second, we develop a novel reinforcement learning algorithm that explicitly addresses the nonuniform task allocation cost challenge identified in this paper. Third, we add a new set of experiments to explicitly evaluate the performance of all compared schemes in terms of both data quality and sensing cost. Fourth, we compare our scheme with more recent task allocation schemes including OO-MTA and demonstrate the performance gains achieved by the QCO-TA scheme compared to all baselines. Finally, we extend the related work by adding a new discussion on the cost-aware task allocation schemes and the difference between QCO-TA and those schemes (Section 2).

2. Background and Related Work

2.1. Social Sensing

Social sensing has emerged as a new application paradigm due to the proliferation of portable devices and ubiquitous Internet connectivity [2]. A recent survey of social sensing can be found in [1]. Social sensing has been widely used in environment sensing [12], traffic monitoring [22], emergence and disaster response [23], social sensor profiling [24], point-of-interest (POI) identification [25], clickbait video detection [26], and abnormal event identification [27]. Quality-cost-aware optimal task allocation in social sensing remains to be an open challenge that has not been fully addressed [3]. This paper addresses the quality-cost-aware task allocation problem in a more challenging scenario where the measured variables have multiple dependent attributes and nonuniform sensing costs.

2.2. Task Allocation

Task allocation with sparse resources has been well studied in mobile crowdsensing literature [3, 9, 10, 28, 29, 11, 12, 13, 14, 15, 30, 31]. Those techniques can be classified into two main categories based on their primary objectives: 1) *Resource and Cost Reduction*: For example, Wang *et al.* developed a data quality aware task allocation scheme that leverages active learning and Bayesian inference techniques to allocate sensing tasks to a limited number of crowd sensors to reduce the overall sensing cost [3]. Zhang *et al.* developed a bottom-up task allocation scheme where mobile sensors bid for tasks to minimize energy cost using a game-theoretic approach [9]. Tong *et al.* proposed a two-phased-based online task allocation scheme to reduce the task allocation cost in real-time crowdsourcing systems [10]. 2) *Quality-of-Service (QoS) Improvement*: For example, Zhou *et al.* developed a budget-aware task allocation scheme that maximizes the quality of sensing data under the constraints imposed by the physical distance between tasks [11]. Hsieh *et al.* developed a greedy task allocation scheme to allocate sensors to cells that would generate minimum entropy to improve the inference accuracy [12]. Zhang *et al.* proposed an expertise-aware task allocation scheme to ensure the quality of the collected data in mobile crowdsourcing systems by inferring the expertise of task participants through truth analysis [13]. There also exist a few solutions that explore both cost and QoS. For example, Yu *et al.* proposed a quality and budget aware task allocation scheme to improve the data quality for spatial crowdsourcing systems given the application specific budget limitations [14]. Liu *et al.* developed an information distribution aware task allocation framework to minimize the task allocation cost of mobile crowdsourcing applications while ensuring the social fairness (e.g., task load balancing) of each participant [15].

Our work is clearly different from the above solutions in several important aspects. First, we explicitly consider the multi-attribute sensing problem which is more challenging than the single attribute problem addressed in the above literature. In the multi-attribute sensing problem, different sensing attributes often have different and correlated spatial-temporal distributions that can lead to inconsistent or even conflicting task allocation decisions [32]. For example, the sensing measurements that can significantly improve the estimation accuracy of a specific sensing attribute may not be equally helpful for other sensing attributes. Moreover, it is not a trivial task to model the complex dependencies between different sensing attributes and understand how such dependencies would affect the global optimized task allocation solution. Second, the goal of our solution is to achieve a complex optimization objective (i.e., jointly optimize the trade-off between the nonuniform task allocation cost and the sensing data quality). The above objective becomes more challenging when we consider the incomplete sensing measurements introduced by a finite sensing budget in our problem. Such incomplete sensing measurements often provide inadequate evidence to explore the aforementioned trade-off between the data quality and sensing cost in multiple-attribute social sensing applications.

2.3. Online Learning

Our work is also related to online learning techniques which have been applied in solving decision making problems in social sensing applications [33, 34, 35, 36, 37, 38]. In particular, online learning learns to make sequential decisions to achieve the desired quality-of-service of an application and dynamically adjust the learning process based on the streaming data received at each step. For example, Rajan *et al.* developed an online learning

task scheduling scheme that coordinates the real-time execution of crowd tasks through the learning of crowd performance [35]. Feng *et al.* developed an online learning algorithm to detect abnormal behavior patterns in crowds using the online self-organization mapping technique [33]. Zhang *et al.* proposed an online learning framework to improve the efficiency of competence-based knowledge compression in machine learning [34]. Xu *et al.* developed an efficient online learning algorithm for dynamic workload offloading (to the centralized cloud) to minimize the system delay and operation cost [37]. To the best of our knowledge, the QCO-TA scheme is one of the first approaches to leverage online learning techniques to address the quality-cost-aware multi-attribute task allocation problem in social sensing.

2.4. Reinforcement Learning

Finally, our work also bears some relevance with the reinforcement learning techniques that have been applied in recommendation systems, intelligent transportation systems, computer vision, natural language processing, and control theory [39, 40, 41, 42, 43]. In particular, reinforcement learning learns to optimize the desired objective of an application by maximizing the *cumulative reward* received from the environment when exploring the application-specific search space. For instance, Zhang *et al.* developed a scholar collaboration and recommendation system via competitive multi-agent reinforcement learning [39]. Xu *et al.* applied a deep reinforcement learning approach in intelligent transportation systems to improve the control robustness for autonomous driving vehicles [40]. Supancic III *et al.* presented a reinforcement learning based decision making framework to continuously track the objects in streaming videos [41]. Branavan *et al.* developed a new reinforcement learning approach to effectively map natural

language instructions to executable actions [42]. Abbeel *et al.* proposed a reinforcement learning based autonomous helicopter flight control system using differential dynamic programming [43]. To the best of our knowledge, the QCO-TA scheme is among the first frameworks to leverage reinforcement learning techniques to explore the optimized tradeoff between the data quality and sensing cost in multi-attribute social sensing applications.

3. Problem Statement

In this section, we formulate the problem of quality-cost-aware task allocation in multi-attribute social sensing applications. We first define a few terms that will be used in the problem statement.

Definition 1. *Sensing Cell*: *We divide the target area for multi-attribute social sensing task into disjoint cells where each cell represents a subarea of interest. In particular, we define S to represent the set of sensing cells in the target area, X to be the total number of sensing cells, and x to be the x^{th} sensing cell in the target area.*

Definition 2. *Sensing Cycle*: *A sensing cycle is a period of time where participants perform one round of the sensing tasks. We define Y to be the total number of sensing cycles, and y to be the y^{th} sensing cycle.*

Definition 3. *Sensing Attribute*: *In social sensing applications, the measured variables often have multiple sensing attributes. We define A to be the total number of sensing attributes, and a to be the a^{th} attribute.*

Consider a social sensing application where the goal is to monitor the air quality index of a city by tasking people to collect sensing measurements

at different locations. In this case, a sensing cell is a neighborhood where the sensing values stay relatively stable spatially [3]. A sensing cycle reflects the frequency of sensing data updates (e.g., hourly, daily). In order to estimate the air quality index, participants collect a set of sensing attributes (e.g., NO₂, CO, PM2.5 and PM10) that are associated with the measured variables (i.e., the air quality index at different cells).

Definition 4. *Real Sensing Value (RS)*: *We define the RS matrix to represent the ground-truth sensing value of measured variables. In particular, RS^a is the ground truth sensing value of attribute a , and $RS_{x,y}^a$ is the ground truth sensing value of attribute a in cell x at cycle y .*

Definition 5. *Collected Sensing Value (CS)*: *We define a CS matrix to represent the collected sensing values of measured variables. In particular, CS^a is the collected sensing values of attribute a , and $CS_{x,y}^a$ is the collected sensing value of attribute a in cell x at cycle y .*

Definition 6. *Inferred Sensing Value (IS)*: *We define an IS matrix to represent the inferred sensing values of measured variables from the inference algorithms by leveraging the collected sensing value. In particular, IS^a is the inferred sensing value of attribute a , and $IS_{x,y}^a$ is the inferred sensing value of attribute a in cell x at cycle y .*

Definition 7. *Sensing Error (SE)*: *we define the Sensing Error to be the mean absolute error between the inferred sensing value and the real sensing value. In particular, we have $SE_{x,y}^a = |IS_{x,y}^a - RS_{x,y}^a|$, where $SE_{x,y}^a$ is the inference error of attribute a in cell x at cycle y .*

Definition 8. Number of Participants (\mathcal{P}): we define \mathcal{P} to be the total number of participants that can be assigned for task allocation in the application. In social sensing, we observe that the number of available participants is often much smaller than the number of sensing cells X due to the budget and resource limitations [3], i.e., $\mathcal{P} \ll X$. We denote p as the p^{th} participant. In addition, we assume the participants in our multi-attribute social sensing applications to be collaborative, i.e., participants agree to perform all assigned task during the application period as long as compensation is provided.

Definition 9. Task Allocation Cost (\mathcal{C})¹: we define the task allocation cost to be the compensation to cover the cost of a participant to travel from one sensing cell to another. We consider the cost to be proportional to the travel distance of a participant in this paper. In particular, we have

$$\mathcal{C}_y^p \propto \text{Distance}_y^p \quad (1)$$

where \mathcal{C}_y^p and Distance_y^p are the task allocation cost and travel distance of participant p at sensing cycle y . In this paper, we choose the above simplified cost model, however, we observe that the above cost function can be readily extended by considering additional cost/incentive design and mobility modeling [44] to accommodate specific requirements of a social sensing application. For example, we can extend the Equation (1) by modeling the diversified response time and transportation expense of different sensing participants, e.g., participants may choose different means of transportation (e.g., walk, bike, drive, bus) to travel between different sensing cells depending on personal preference or availability of different transportation options.

¹We use the term *task allocation cost* and *sensing cost* interchangeably in the paper.

The goal of our online quality-cost-aware task allocation in multi-attribute social sensing applications is to make real-time task allocation decisions that optimize the tradeoff between the overall sensing error for all sensing attributes of the measure variables and the task allocation costs over all participants. Formally our problem is defined as:

$$\begin{aligned}
& \text{select } S_y \text{ from } S, \forall 1 < y < Y_c \\
& \text{minimize } \sum_{a=1}^A \left(\Gamma_a \left(\frac{1}{Y_c} \cdot \sum_{y=1}^{Y_c} \frac{1}{X} \cdot \sum_{x=1}^X SE_{x,y}^a \right) \right) \\
& \text{minimize } \frac{1}{Y_c} \cdot \sum_{y=1}^{Y_c} \left(\frac{1}{\mathcal{P}} \cdot \sum_{p=1}^{\mathcal{P}} \mathcal{C}_y^p \right)
\end{aligned} \tag{2}$$

where $|S_y| \leq \mathcal{P}$

where Y_c denotes the total number of sensing cycles of interest, and S_y represents the cells that are allocated to the participants for sensing tasks at cycle y . Considering that different sensing attributes have different ranges of values, we define Γ_a as the normalization function to normalize the sensing error for attribute a .

The above problem is NP-hard since each of its two objectives (i.e., minimizing sensing error or minimizing sensing cost) can be reduced to the Knapsack problem (i.e., one of the Karp's 21 NP-hard problems) ² [45]. In this paper, we develop a QCO-TA scheme that judiciously explores the tradeoff between the data quality and sensing cost and identifies an optimized task allocation strategy to jointly optimize the sensing error and cost. The details of the QCO-TA scheme are discussed in the next section.

²The detailed proof of NP-hardness of the proposed optimization problem can be found in the Appendix.

4. Solution

In this section, we present the Quality-Cost-Aware Online Task Allocation (QCO-TA) scheme to address the problem formulated in the previous section using a principled online reinforcement learning framework. The QCO-TA scheme consists of three modules: 1) a Single-Attribute Priority Estimation (SPE) module, 2) a Multi-Attribute Priority Integration (MPI) module, and 3) a Nonuniform-Cost-Aware Task Selection (NTS) module. The overall architecture of the QCO-TA scheme is shown in Figure 2.

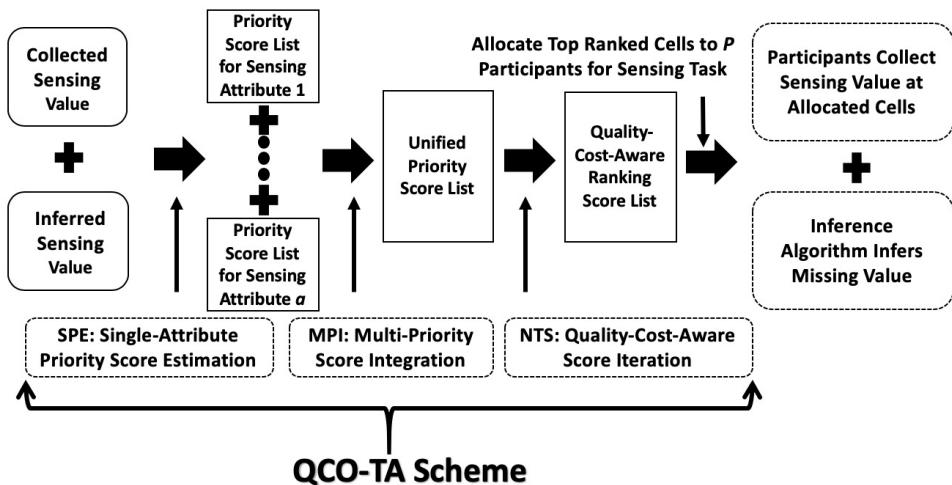


Figure 2: Overview of QCO-TA Scheme

4.1. Single-Attribute Priority Estimation (SPE)

In this subsection, we present the single-attribute priority estimation module that addresses the *online task allocation* challenge discussed in the introduction. We first define a few terms that will be used in this module.

Definition 10. Task Priority: *We define the task priority as the order in which the sensing cells are selected for task allocation given a particular*

sensing attribute ³. A task with the highest priority will be selected first.

Definition 11. Priority Score: We further define the priority score as a scalar to quantify the task priority defined above.

In particular, the SPE module estimates the priority score of each cell for a given sensing attribute and dynamically updates the estimations based on the collected sensing values from the previous cycles using an online learning algorithm. To compute the task priority of each cell for a given attribute in real time, we need to know which cell's sensing value, if collected, would be the most helpful one to reduce the sensing error. This problem has been proven to be an NP-hard without knowing the real sensing value of cells in advance [3]. To solve this problem, we develop an efficient approximation algorithm that considers two factors directly related to the sensing error of a cell: 1) *uncertainty*: the estimation confidence of the sensing values in a cell from a given inference algorithm (e.g., KNN, SVR) [3]; 2) *representativeness*: how accurately the sensing value of the target cell can be used to represent the values of its neighboring cells [46]. In QCO-TA, we use *temporal entropy* and *spatial mutual information* to estimate the uncertainty and representativeness of a sensing cell, respectively.

Temporal Entropy (TE): We define temporal entropy to quantify the uncertainty of the inferred sensing value of a sensing cell as follows:

$$TE_{x,y}^a = \Theta_\Lambda(\{IS_{x,1}^a, IS_{x,2}^a, \dots, IS_{x,y}^a\}) \quad (3)$$

where $TE_{x,y}^a$ is the temporal entropy of cell x at cycle y for attribute a . $IS_{x,y}^a$ is the inferred sensing value of cell x at cycle y for attribute a (defined in

³Since each task is associated with a particular sensing cell, we use task and cell priority interchangeably in the rest of the paper.

Definition 6). Λ is the distribution (e.g., normal distribution) of the inferred sensing value $IS_{x,y}^a$. Θ_Λ is the function to calculate the differential entropy for the distribution Λ [47]. Intuitively, a high temporal entropy of a cell indicates that the inference algorithm is uncertain about its inferred sensing values of that cell and vice versa.

Spatial Mutual Information (SMI): We define the spatial mutual information of a sensing cell to be the aggregated mutual information between the target cell and the rest of cells:

$$SMI_{x,y}^a = \sum_{i=1, i \neq x}^X I(\{IS_{x,1}^a, IS_{x,2}^a \dots IS_{x,y}^a\}; \{IS_{i,1}^a, IS_{i,2}^a \dots IS_{i,y}^a\}) \quad (4)$$

where $SMI_{x,y}^a$ represents the spatial mutual information of cell x at cycle y for attribute a . I is the function to calculate the mutual information of inferred sensing values between different cells [48]. In particular, $I(\{IS_{x,1}^a, IS_{x,2}^a \dots IS_{x,y}^a\}; \{IS_{i,1}^a, IS_{i,2}^a \dots IS_{i,y}^a\})$ is the mutual information between cell x and i . Intuitively, a high spatial mutual information of a cell indicates that the sensing values of the cell, if selected, can be used to reduce the inference error significantly.

We then combine the temporal entropy (TE) and spatial mutual information (SMI) to compute the *priority score* (PS) that determines the task priority of each sensing cell to be selected for task allocation as follows:

$$PS_{x,y}^a = \alpha_{TE}^y \cdot TE_{x,y}^a + \alpha_{SMI}^y \cdot SMI_{x,y}^a \quad (5)$$

where $PS_{x,y}^a$ represents the priority score of cell x at sensing cycle y given sensing attribute a . α_{TE}^y and α_{SMI}^y are the weights for temporal entropy and spatial mutual information at cycle y , respectively. The values of α_{TE}^y and α_{SMI}^y are tuned based on the requirements of specific applications.

4.2. Multi-Attributes Priority Integration (MPI)

In this subsection, we describe the Multi-Attribute Priority Integration (MPI) module to address the *multi-attribute constrained optimization* challenge. First, we formally define a comprehensive ranking score as follows.

Definition 12. Unified Priority Score (UPS): *We define the UPS to be the weighted summation of the priority score PS of all sensing attributes generated by the SPE module as follows:*

$$UPS_{x,y} = \sum_{a=1}^A w_y^a \cdot PS_{x,y}^a \quad (6)$$

where $UPS_{x,y}$ is the unified priority score for cell x at cycle y and A is the number of sensing attributes. w_y^a is the weight for attribute a in cycle y . $PS_{x,y}^a$ is the priority score of cell x at cycle y for attribute a .

The key question is how to dynamically compute the weights for all attributes at each cycle so that the aggregated sensing error (defined in Equation (2)) can be minimized by exploiting the dependencies between attributes. To solve this problem, we develop an exponential weighted online learning algorithm that dynamically updates the weights for all attributes based on collected sensing values in real time. In particular, we have the following updating rule for the weight:

$$w_{y+1}^a = w_y^a \cdot \exp(-\eta \cdot \ell(IS_y^a, RS_y^a)) \quad (7)$$

where w_y^a and w_{y+1}^a are the weights for attribute a at cycle y and $y+1$, respectively. η is the learning rate parameter that directly controls the scale of the weight assigned to each sensing attribute. $\ell(IS_y^a, RS_y^a)$ is the loss function that measures the sensing error between the inferred sensing value

(IS_y^a , defined in Definition 6) and the real sensing value (RS_y^a , defined in Definition 4) in the current cycle y . The intuition of this weight updating function is that it increases the weights of the sensing attributes that contribute less sensing error.

The challenging part of computing the above weight updating function is how to calculate the loss function since we do not have the real sensing value for all cells due to the budget limitation. To address this problem, we apply Bayesian inference to estimate the loss function as follows:

$$\ell(IS_y^a, RS_y^a) \approx \ell(IS_y^a, RS_y^a | IS_y^a, CS_y^a) = \Phi_{\mathcal{N}_a}(\delta) \quad (8)$$

where $\ell(IS_y^a, RS_y^a | IS_y^a, CS_y^a)$ is the estimated loss given the collected sensing values CS_y^a and inferred sensing values IS_y^a . Φ is the inverse of the cumulative distribution function given the distribution \mathcal{N}_a . \mathcal{N}_a is the distribution of the mean absolute error (MAE) between CS_y^a and IS_y^a for attribute a in current cycle y (i.e., $|CS_y^a - IS_y^a|$). We assume \mathcal{N}_a follows the normal distribution, which is a common assumption for MAE in social sensing applications [3]. In addition, δ is a probability threshold to determine the level of approximation between the loss $\ell(IS_y^a, RS_y^a)$ and estimated loss $\ell(IS_y^a, RS_y^a | IS_y^a, CS_y^a)$. It is usually set to be higher than 0.95.

4.3. Nonuniform-Cost-Aware Task Allocation (NTS)

In this subsection, we describe the Nonuniform-Cost-Aware Task Allocation (NTS) module to address the *nonuniform task allocation cost* challenge. In particular, the NTS module judiciously integrates the nonuniform task allocation costs with the unified ranking score generated by the MPI module to explore the optimized tradeoff between the data quality and sensing cost through a principled reinforcement learning framework. We first define a quality-cost-aware ranking score as follows.

Definition 13. *Quality-Cost-Aware Ranking Score (QRS)*: We define the *QRS* to be the overall ranking score that jointly considers both *UPS* score (defined in Definition 12) generated by the *MPI* module and the nonuniform sensing cost (defined in Definition 9) as follows:

$$QRS_{x,y} = \Gamma(UPS_{x,y}, \mathcal{C}_{[x',x],y}) \quad (9)$$

where $QRS_{x,y}$ is the quality-cost-aware ranking score for cell x at cycle y . $UPS_{x,y}$ is the unified priority score of cell x at cycle y . $\mathcal{C}_{[x',x],y}$ is the cost for participants to move from current cell x' to cell x to perform the sensing task at cycle y . Γ is the mapping function that integrates the *UPS* score and nonuniform sensing cost. Intuitively, a high *QRS* value indicates that a cell, if selected, would most likely to reduce the overall sensing errors with the minimal sensing costs, and vice versa.

The key challenge now is how to design an effective mapping function Γ to compute the *QRS* score so that the aggregated sensing error and overall sensing cost can be jointly optimized as indicated in Equation (2). To solve this problem, we develop a principled reinforcement learning algorithm that iteratively explores the optimal tradeoff between the data quality and the cost for task allocation. We first define a few key terms that will be used in our reinforcement learning framework.

Definition 14. *State (\mathcal{S})*: We define a state (\mathcal{S}_x) as a sensing cell (i.e., $\mathcal{S}_x = x$) that is considered as a candidate for task allocation. Each state carries a state value (\mathcal{V}), which indicates the priority of the corresponding cell to be selected for a sensing task allocation. In particular, we define \mathcal{V}_x^y to be the state value for state \mathcal{S}_x at sensing cycle y .

We initiate the state value for each sensing cell with the *UPS* score generated

by MPI at each sensing cycle (i.e., $\mathcal{V}_x^y = UPS_x^y$). The state value will be dynamically updated by our reinforcement learning algorithm to explicitly consider the nonuniform task allocation cost, which will be elaborated in this subsection.

Definition 15. Action (\mathcal{A}): *We define an action $\mathcal{A}_{[x',x'']}$ as the move of a participant p who travels from the sensing cell x' to x'' to perform the assigned sensing task in two consecutive sensing cycles.*

A participant can take an action to move from current cell x' to a new cells x'' or stay at current cell x' (i.e., $x'' = x'$) to perform the sensing task. The goal of our reinforcement learning algorithm is to learn the optimal action for each participant so the aggregated sensing error and overall sensing cost can be jointly optimized.

Definition 16. Reward (\mathcal{R}): *We define the reward $\mathcal{R}_{[x',x'']}$ for action $\mathcal{A}_{[x',x'']}$ to be inversely proportional to the task allocation cost (i.e., travel distance $Distance_{[x',x'']}$ between the two cells) defined in Definition 9 as:*

$$\mathcal{R}_{[x',x'']} = \gamma \cdot \frac{1}{Distance_{[x',x'']}} \quad (10)$$

where γ is the scaling parameter, the value of γ is tuned based on the specific requirements of the applications. The rewards will be used to dynamically update the state values in our reinforcement learning algorithm.

Using the above definitions, we leverage a Bellman optimality equation [49] to integrate the nonuniform task allocation cost with the unified ranking score generated by MPI as follows:

$$\mathcal{V}_{x'} = \max_{\mathcal{A}_{[x',x'']}} \mathbb{E}[\beta \cdot (\mathcal{V}_{x''}) + R_{[x',x'']} | S = S_{x'}, \mathcal{A} = \mathcal{A}_{[x',x'']}] \quad (11)$$

where $\mathcal{A}_{[x',x'']}$ is the action of moving from the sensing cell x' to x'' for a sensing task at sensing cell x'' . $R_{[x',x'']}$ is the reward for the action. β is the discount parameter to control the updating rate of the state values, which is usually set to be a small value (i.e., less than 1) to ensure a desired performance of the reinforcement learning algorithm. The intuition of the above equation is to assign a higher priority score (i.e., state value) to a cell that leads to lower overall sensing error with minimal sensing cost.

Using the Bellman optimality equation, we can learn the optimized state value V_x^* for each sensing cell x by iteratively updating each state value V_x until all state values are converged as follows:

$$\begin{aligned} \text{abs}(\mathcal{V}'_x - \mathcal{V}_x) &< \Theta, \forall 1 < x < X \\ \mathcal{V}_x^* &= \mathcal{V}_x, \forall 1 < x < X \end{aligned} \tag{12}$$

where \mathcal{V}'_x is the state value for sensing cell x in the previous iteration, and \mathcal{V}_x is the updated state value for sensing cell x in the current iteration. Θ is the threshold to stop the learning process, which is usually set to be a small value (e.g., less than 0.1) to ensure the convergence of the learning process and the accuracy of the learned model. We take the learned optimized state value \mathcal{V}_x^* as the overall ranking score QRS_x^y for a sensing cell x at the sensing cycle y as follows:

$$QRS_{x,y} = \mathcal{V}_x^*, \forall 1 < x < X \tag{13}$$

Finally, each participant is allocated to move from the current cell x' to the new cell x'' that leads to the highest $QRS_{x'',y}$ value as follows:

$$\begin{aligned} \arg \max_{x''} (ORS_{x'',y} | \mathcal{A} = \mathcal{A}_{[x',x'']}, x'' \notin S_y) \\ \text{add } x'' \text{ to } S_y \end{aligned} \tag{14}$$

where S_y is set of sensing cells that has been allocated to the participants in the current sensing cycle.

5. Evaluation

In this section, we evaluate the performance of the QCO-TA scheme through a real world social sensing application. We compare the performance of QCO-TA with state-of-the-art task allocation baselines. The evaluation results show that QCO-TA significantly outperforms the baselines in terms of both sensing accuracy and task allocation cost.

5.1. Datasets

Piemonte Air Quality Dataset: In our evaluation, we use a social sensing dataset published by Blangiardo *et al.* [50]⁴. This dataset consists of daily air quality measurements across 24 locations (i.e., cells) in Piemonte, Italy (as shown in Figure 3(a)). The measurements include the following air quality related attributes: wind speed, temperature, emission rates of primary aerosols, and particulate matter (PM) 10. We choose this dataset because i) it contains multiple sensing attributes of the measured variable (i.e., the air quality); ii) the measurements have large spatial-temporal dynamics (Figure 3(b)), which make our problem more challenging to solve. The sensing cycle is set to be one day for this application.

5.2. Inference Algorithm

In the experiment, we select the following inference algorithms to work with the task allocation schemes to estimate the sensing values of the cells that are not selected for sensing in each cycle.

⁴<https://sites.google.com/a/r-inla.org/stbook/datasets>

⁵The dark colors in each horizontal line indicate larger temporal dynamics and the various color patterns across different lines indicate large spatial variations.

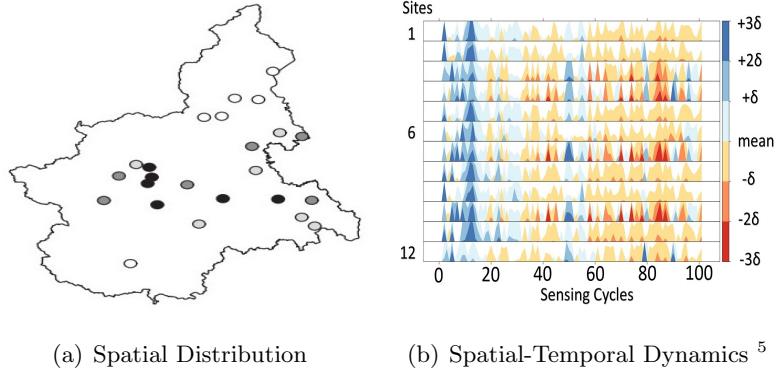


Figure 3: Piemonte Air Quality Dataset

1. **K-Nearest Neighbour (KNN):** KNN estimates the missing value of a cell by averaging the collected sensing values from the k nearest cells of the target cell.
2. **Inverse Distance Weighting (IDW):** IDW estimates the missing value of a cell by calculating the weighted average value of the collected sensing values from its n closest neighbors, where the weights are proportional to the reciprocal of the spatial distances between the target cell and its neighbors.
3. **Support Vector Regression (SVR):** SVR first establishes a prediction model with the collected sensing values from the selected cells using a support vector machine and then applies the prediction model to infer the missing value of the target cell [51].

5.3. Baseline algorithms

We choose several representative task allocation schemes as the baselines.

1. **OO-MTA:** OO-MTA scheme is a simplified version of QCO-TA scheme that selects the sensing cells solely based on the unified priority score

(defined in Definition 12) from our previous work [21]. The OO-MTA scheme does not consider the nonuniform task allocation cost.

2. **GPS-TA:** GPS-TA (Greedy Priority Selection Task Allocation) is a greedy task allocation algorithm that selects the d cells with highest priorities for each sensing attribute for task allocation [12]. d is equal to the number of participants divided by the number of attributes.
3. **EWA-TA:** EWA-TA (Equal Weighted Aggregation Task Allocation) is a task allocation scheme that generates the overall priority of a cell by calculating the mean of the priorities of the cell across all sensing attributes and then selects the top \mathcal{P} ranked cells for task allocation.
4. **UNS-TA:** UNS-TA (Uniform Sampling Task Allocation) is a task allocation scheme that uniformly samples \mathcal{P} cells from all cells for sensing task allocation in each cycle, where each cell has an equal probability to be selected for task allocation [52].

5.4. Evaluation Metrics

In our evaluation, we define the following metrics to evaluate the task allocation performance of all compared schemes.

- **Aggregated Sensing Error (ε):** We define Aggregated Sensing Error (ε) to be the aggregated sensing errors for all the sensing attributes of the measured variable. Specifically, we define:

$$\varepsilon = \sum_{a=1}^A \left(\Gamma_a \left(\frac{1}{Y} \cdot \sum_{y=1}^Y \frac{1}{X} \cdot \sum_{x=1}^X SE_{x,y}^a \right) \right) \quad (15)$$

where A is the number of the sensing attributes, X is the number of sensing cells, and Y is the the number of sensing cycles. $SE_{x,y}^a$ is the sensing error for attribute a of cell x at cycle y as we defined in

Definition 7. Γ_a is the normalization function to normalize the sensing error for attribute a as we defined in Equation (2).

- **Average Task Allocation Cost (φ):** We define Average Task Allocation Cost (φ) as follows:

$$\varphi = \frac{1}{Y} \cdot \sum_{y=1}^Y \left(\frac{1}{P} \cdot \sum_{p=1}^P \mathcal{C}_y^p \right) \quad (16)$$

where P is the number of participants and \mathcal{C}_y^p is the task allocation cost for participant p at sensing cycle y as we defined in Definition 9. In our evaluation, the task allocation cost is measured by the physical distance between the source and destination cells a participant travels.

5.5. Evaluation Results

In this subsection, we present the results of our QCO-TA scheme and the compared baselines on the real world social sensing dataset. In the experiment, we evaluate the performance of all compared schemes by varying the number of sensing attributes of the measured variable. In particular, we change the number of attributes from two to four in our experiment based on the number of attributes available in the dataset (i.e., we have four attributes in total). For a given number of sensing attributes, we evaluate the performance of all compared schemes using the *aggregated sensing error* and *average task allocation cost* metrics defined in Equation (15) and Equation (16), respectively. In our experiment, we evaluate the performance of all schemes by changing the *number of participants*. Specifically, we vary the number of participants P from 8 to 14 by considering the number of sensing cells in our dataset (i.e., $X=24$).

Table 1: Evaluation Results on Aggregated Sensing Error

TA-Scheme	P-Num	2 Sensing Attributes			3 Sensing Attributes			4 Sensing Attributes		
		KNN	IDW	SVR	KNN	IDW	SVR	KNN	IDW	SVR
QCO-TA	$\mathcal{P}=8$	0.153	0.149	0.141	0.184	0.180	0.172	0.232	0.228	0.217
	$\mathcal{P}=10$	0.139	0.138	0.134	0.170	0.171	0.166	0.214	0.215	0.211
	$\mathcal{P}=12$	0.133	0.131	0.130	0.164	0.163	0.161	0.210	0.209	0.207
	$\mathcal{P}=14$	0.129	0.128	0.127	0.160	0.159	0.158	0.207	0.206	0.204
OO-MTA	$\mathcal{P}=8$	0.201	0.197	0.175	0.329	0.321	0.263	0.245	0.244	0.231
	$\mathcal{P}=10$	0.173	0.170	0.163	0.257	0.252	0.234	0.230	0.230	0.222
	$\mathcal{P}=12$	0.163	0.162	0.158	0.230	0.230	0.221	0.226	0.225	0.221
	$\mathcal{P}=14$	0.160	0.164	0.159	0.221	0.220	0.214	0.222	0.224	0.219
GPS-TA	$\mathcal{P}=8$	0.239	0.240	0.221	0.231	0.241	0.226	0.293	0.286	0.275
	$\mathcal{P}=10$	0.229	0.232	0.223	0.227	0.228	0.224	0.275	0.275	0.273
	$\mathcal{P}=12$	0.235	0.242	0.229	0.228	0.232	0.227	0.274	0.277	0.276
	$\mathcal{P}=14$	0.238	0.245	0.234	0.232	0.234	0.230	0.276	0.276	0.277
EWA-TA	$\mathcal{P}=8$	0.292	0.275	0.245	0.344	0.333	0.289	0.377	0.376	0.334
	$\mathcal{P}=10$	0.239	0.234	0.227	0.291	0.290	0.271	0.335	0.335	0.319
	$\mathcal{P}=12$	0.234	0.226	0.223	0.268	0.264	0.251	0.319	0.318	0.306
	$\mathcal{P}=14$	0.216	0.210	0.202	0.245	0.240	0.232	0.301	0.299	0.290
UPS-TA	$\mathcal{P}=8$	0.245	0.246	0.223	0.248	0.249	0.229	0.320	0.312	0.288
	$\mathcal{P}=10$	0.230	0.232	0.223	0.238	0.242	0.232	0.294	0.298	0.289
	$\mathcal{P}=12$	0.227	0.229	0.224	0.236	0.239	0.234	0.292	0.294	0.289
	$\mathcal{P}=14$	0.226	0.226	0.222	0.235	0.237	0.233	0.292	0.293	0.290

5.5.1. Evaluation Results on Data Quality

The results on *aggregated sensing error* are shown in Table 1. We observe that the QCO-TA scheme outperforms all of the baselines by achieving the smallest sensing error. The performance gain achieved by the QCO-TA scheme is consistent over different inference algorithms and different numbers of sensing attributes. The performance gain achieved by QCO-TA compared to the best performing baseline on $\mathcal{P}=8, \mathcal{P}=10, \mathcal{P}=12, \mathcal{P}=14$ are 4.8%, 4.4%, 3.0%, 3.1% under the KNN inference algorithm when the number of sensing attributes is 2. Such performance gains are achieved by two key designs in the proposed QCO-TA scheme. First, the SPE module (de-

scribed in Section 4.1) judiciously uses temporal entropy and spatial mutual information to estimate the priority of cells for different sensing attributes. Two types of sensing cells are often selected for sensing tasks to reduce the overall sensing error: i) the sensing cells with the high uncertainty of their inferred sensing values, and ii) the representative sensing cells whose sensing value can be used to represent the sensing values of their neighboring cells. Second, the MPI module (described in Section 4.2) unitizes principled exponential weighted online learning to integrate the priority estimations of different sensing attributes. In particular, the MPI module explicitly exploits the dependencies between the attributes and carefully increases the weights of the sensing attributes that contribute less to the overall sensing error. Additionally, we observe that the sensing error of all schemes generally decreases when the number of participants increases. This is because a larger number of participants allow the schemes to collect sensing values from more cells, which reduces the errors of the inference algorithms of all compared schemes. These results demonstrate that the QCO-TA scheme can minimize the sensing error of social sensing applications with multiple sensing attributes in comparison with the state-of-the-art baselines.

5.5.2. Evaluation Results on Task Allocation Cost

The results on *average task allocation cost* are shown in Table 2. We observe that the QCO-TA scheme outperforms all baselines by achieving the lowest task allocation cost. The performance gain achieved by QCO-TA compared to the best performing baseline on $\mathcal{P}=8$, $\mathcal{P}=10$, $\mathcal{P}=12$, $\mathcal{P}=14$ are 6.458 km, 3.645 km, 2.633 km, 6.279 km in terms of travel distance under the KNN inference algorithm when the number of sensing attributes is 2. Similar performance gains are also observed for different inference algorithms and

Table 2: Evaluation Results on Average Task Allocation Cost (Unit: km)

TA-Scheme	P-Num	2 Sensing Attributes			3 Sensing Attributes			4 Sensing Attributes		
		KNN	IDW	SVR	KNN	IDW	SVR	KNN	IDW	SVR
QCO-TA	$\mathcal{P}=8$	2.872	1.551	5.528	3.043	2.136	6.918	5.696	5.846	6.751
	$\mathcal{P}=10$	2.978	5.615	3.136	4.619	4.056	3.787	4.235	5.671	6.008
	$\mathcal{P}=12$	3.924	3.789	2.471	3.114	2.779	2.226	3.620	4.259	4.863
	$\mathcal{P}=14$	3.281	3.401	2.798	2.563	2.759	3.139	3.116	3.324	4.283
OO-MTA	$\mathcal{P}=8$	9.338	5.920	11.758	13.810	8.673	7.261	8.310	10.321	8.045
	$\mathcal{P}=10$	6.234	6.111	6.572	4.916	4.389	8.969	6.585	7.634	6.394
	$\mathcal{P}=12$	6.557	6.041	5.073	6.303	5.587	6.536	7.378	4.683	5.876
	$\mathcal{P}=14$	9.560	5.284	4.290	5.571	5.602	5.790	5.619	4.473	4.632
GPS-TA	$\mathcal{P}=8$	16.798	17.436	32.693	29.173	23.686	39.436	20.151	18.421	31.053
	$\mathcal{P}=10$	13.847	11.883	27.726	16.796	13.283	27.182	20.879	18.253	37.555
	$\mathcal{P}=12$	10.527	8.250	21.368	14.060	15.727	20.971	17.964	13.922	23.505
	$\mathcal{P}=14$	7.196	9.840	19.905	15.128	20.066	19.356	13.139	11.631	22.636
EWA-TA	$\mathcal{P}=8$	11.939	14.584	35.633	14.345	7.202	31.720	15.190	15.925	35.674
	$\mathcal{P}=10$	9.061	11.515	28.204	10.146	8.822	23.193	10.557	11.685	30.714
	$\mathcal{P}=12$	6.244	8.726	21.129	5.944	5.577	16.901	8.068	7.868	24.725
	$\mathcal{P}=14$	7.069	5.664	14.829	5.920	4.226	12.756	5.773	7.736	18.110
UPS-TA	$\mathcal{P}=8$	39.036	36.221	38.862	39.256	36.996	35.553	35.346	35.450	36.655
	$\mathcal{P}=10$	33.646	32.604	31.180	34.079	31.568	31.678	31.699	35.127	31.713
	$\mathcal{P}=12$	28.999	28.256	28.083	30.627	30.966	28.356	28.650	28.392	31.266
	$\mathcal{P}=14$	24.816	23.877	25.082	23.430	25.758	24.251	25.179	25.890	25.301

different numbers of sensing attributes. Such performance gains of the QCO-TA scheme are achieved by the key design in the principled reinforcement learning framework (NTS module as described in Section 4.3), which directly learns to achieve the optimized tradeoff between the data quality and the nonuniform task allocation cost. In particular, the NTS module iteratively updates the proposed Bellman optimality equation (defined in Equation 11) to assign a higher priority score to the cell that leads to a lower overall sensing error with the minimal sensing cost. The reinforcement learning process ensures that the QCO-TA scheme always selects the sensing cells which lead to the lowest sensing cost under the premise of sensing quality

assurance. In summary, the above results demonstrate the capability of QCO-TA to achieve the goal of quality-cost-aware task allocation (defined in Equation 2). We also would like to acknowledge that we only consider the task allocation cost to be the travel distances of participants to perform the sensing tasks in our evaluation. However, in real-world applications, there can be additional factors that affect the task allocation cost (e.g., the time consumed by the participants and the way of commute). We plan to further validate our QCO-TA scheme with a more complex and comprehensive cost model in our future work.

5.5.3. Affordability of QCO-TA scheme

Finally, we study the affordability of the QCO-TA scheme by examining the performance of QCO-TA scheme over multiple sensing cycles. The performance of QCO-TA scheme is shown in Figure 4. We observe that the QCO-TA scheme only requires a small number of sensing cycles at the beginning to quickly learn the optimal task allocation strategy that optimizes the performance of the system (in terms of data quality). The results demonstrate that our scheme can efficiently learn the optimal task allocation strategy within a limited number of sensing cycles.

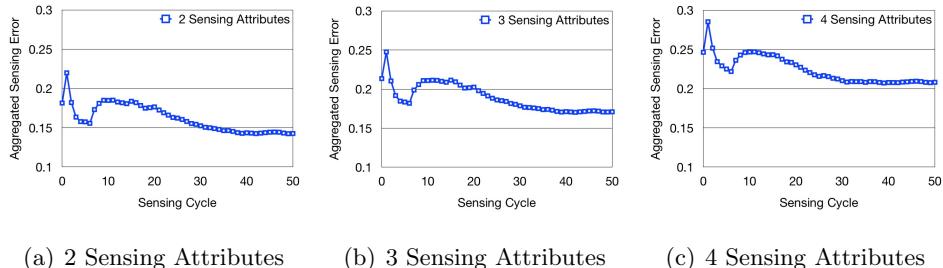


Figure 4: Affordability of QCO-TA Scheme

6. Conclusion

This paper develops a QCO-TA scheme to solve the quality-cost-aware task allocation problem in multi-attribute social sensing applications. In the QCO-TA scheme, we develop a *single-attribute priority estimation* module to estimate the priority score of each cell for a given sensing attribute, a *multi-attribute priority integration* module to integrate the priority scores from all sensing attributes into a unified ranking score for the task allocation, and a *task-cost-aware task allocation* module to explore the optimized tradeoff between the data quality and sensing cost. The evaluation results on a real-world data trace demonstrate that the QCO-TA scheme achieves significant performance gains in terms of both data quality and sensing cost compared to the state-of-the-art baselines in various application scenarios.

Acknowledgment

This research is supported in part by the National Science Foundation under Grant No. CNS-1845639, CNS-1831669, CBET-1637251, IIS-1447795, Army Research Office under Grant W911NF-17-1-0409. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

- [1] D. Wang, B. K. Szymanski, T. Abdelzaher, H. Ji, L. Kaplan, The age of social sensing, *Computer* 52 (1) (2019) 36–45 (2019).

- [2] D. Wang, T. Abdelzaher, L. Kaplan, Social sensing: building reliable systems on unreliable data, Morgan Kaufmann, 2015 (2015).
- [3] L. Wang, D. Zhang, A. Pathak, C. Chen, H. Xiong, D. Yang, Y. Wang, Ccs-ta: Quality-guaranteed online task allocation in compressive crowdsensing, in: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2015 (2015).
- [4] Y.-C. Yu, A mobile social networking service for urban community disaster response, in: Semantic Computing (ICSC), 2015 IEEE International Conference on, IEEE, 2015, pp. 503–508 (2015).
- [5] D. Wang, L. Kaplan, H. Le, T. Abdelzaher, On truth discovery in social sensing: A maximum likelihood estimation approach, in: Information Processing in Sensor Networks (IPSN), 2012 ACM/IEEE 11th International Conference on, IEEE, 2012, pp. 233–244 (2012).
- [6] D. Zhang, Y. Zhang, Q. Li, N. Vance, D. Wang, Robust state prediction with incomplete and noisy measurements in collaborative sensing, in: 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), IEEE, 2018, pp. 460–468 (2018).
- [7] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, X. Mao, Incentives for mobile crowd sensing: A survey, IEEE Communications Surveys & Tutorials 18 (1) (2016) 54–67 (2016).
- [8] L. G. Jaimes, I. Vergara-Laurens, M. A. Labrador, A location-based incentive mechanism for participatory sensing systems with budget constraints, in: Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on, IEEE, 2012 (2012).

- [9] D. Zhang, Y. Ma, Y. Zhang, S. Lin, X. S. Hu, D. Wang, A real-time and non-cooperative task allocation framework for social sensing applications in edge computing systems, in: 2018 IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS), IEEE, 2018, pp. 316–326 (2018).
- [10] Y. Tong, J. She, B. Ding, L. Wang, L. Chen, Online mobile micro-task allocation in spatial crowdsourcing, in: 2016 IEEE 32Nd international conference on data engineering (ICDE), IEEE, 2016, pp. 49–60 (2016).
- [11] C. Zhou, C.-K. Tham, M. Motani, Qoata: Qoi-aware task allocation scheme for mobile crowdsensing under limited budget, in: Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on, IEEE, 2015, pp. 1–6 (2015).
- [12] H.-P. Hsieh, S.-D. Lin, Y. Zheng, Inferring air quality for station location recommendation based on urban big data, in: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2015, pp. 437–446 (2015).
- [13] X. Zhang, Y. Wu, L. Huang, H. Ji, G. Cao, Expertise-aware truth analysis and task allocation in mobile crowdsourcing, in: 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), IEEE, 2017, pp. 922–932 (2017).
- [14] H. Yu, C. Miao, Z. Shen, C. Leung, Quality and budget aware task allocation for spatial crowdsourcing, in: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 1689–1690 (2015).

- [15] Q. Liu, T. Abdessalem, H. Wu, Z. Yuan, S. Bressan, Cost minimization and social fairness for spatial crowdsourcing tasks, in: International Conference on Database Systems for Advanced Applications, Springer, 2016, pp. 3–17 (2016).
- [16] J. Liu, H. Shen, H. S. Narman, W. Chung, Z. Lin, A survey of mobile crowdsensing techniques: A critical component for the internet of things, *ACM Transactions on Cyber-Physical Systems* 2 (3) (2018) 18 (2018).
- [17] A. Ahmed, K. Yasumoto, Y. Yamauchi, M. Ito, Distance and time based node selection for probabilistic coverage in people-centric sensing, in: Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on, IEEE, 2011, pp. 134–142 (2011).
- [18] L. Kazemi, C. Shahabi, Geocrowd: enabling query answering with spatial crowdsourcing, in: Proceedings of the 20th international conference on advances in geographic information systems, ACM, 2012 (2012).
- [19] Y. Zhao, Q. Han, Spatial crowdsourcing: current state and future directions, *IEEE communications magazine* 54 (7) (2016) 102–107 (2016).
- [20] C. C. Aggarwal, T. Abdelzaher, Social sensing, in: Managing and mining sensor data, Springer, 2013, pp. 237–297 (2013).
- [21] Y. Zhang, D. Zhang, N. Vance, D. Wang, Optimizing online task allocation for multi-attribute social sensing, in: 2018 27th International Conference on Computer Communication and Networks (ICCCN), IEEE, 2018, pp. 1–9 (2018).

- [22] Y. Zhang, Y. Lu, D. Zhang, L. Shang, D. Wang, Risksens: A multi-view learning approach to identifying risky traffic locations in intelligent transportation systems using social and remote sensing, in: 2018 IEEE International Conference on Big Data (Big Data), IEEE, 2018, pp. 1544–1553 (2018).
- [23] D. Y. Zhang, D. Wang, Y. Zhang, Constraint-aware dynamic truth discovery in big data social media sensing, in: 2017 IEEE International Conference on Big Data (Big Data), IEEE, 2017, pp. 57–66 (2017).
- [24] Y. Zhang, N. Vance, D. Zhang, D. Wang, On opinion characterization in social sensing: A multi-view subspace learning approach, in: 2018 14th International Conference on Distributed Computing in Sensor Systems (DCOSS), IEEE, 2018, pp. 155–162 (2018).
- [25] D. Y. Zhang, D. Wang, H. Zheng, X. Mu, Q. Li, Y. Zhang, Large-scale point-of-interest category prediction using natural language processing models, in: 2017 IEEE International Conference on Big Data (Big Data), IEEE, 2017, pp. 1027–1032 (2017).
- [26] L. Shang, D. Y. Zhang, M. Wang, S. Lai, D. Wang, Towards reliable on-line clickbait video detection: A content-agnostic approach, Knowledge-Based Systems (2019) 104851 (2019).
- [27] P. Giridhar, M. T. Amin, T. Abdelzaher, D. Wang, L. Kaplan, J. George, R. Ganti, Clarisense+: An enhanced traffic anomaly explanation service using social network feeds, Pervasive and Mobile Computing 33 (2016) 140–155 (2016).
- [28] N. Vance, M. T. Rashid, D. Zhang, D. Wang, Towards reliability in on-

line high-churn edge computing: A deviceless pipelining approach, in: 2019 IEEE International Conference on Smart Computing (SMART-COMP), IEEE, 2019, pp. 301–308 (2019).

[29] D. Zhang, Y. Ma, C. Zheng, Y. Zhang, X. S. Hu, D. Wang, Cooperative-competitive task allocation in edge computing for delay-sensitive social sensing, in: 2018 IEEE/ACM Symposium on Edge Computing (SEC), IEEE, 2018, pp. 243–259 (2018).

[30] D. Y. Zhang, D. Wang, An integrated top-down and bottom-up task allocation approach in social sensing based edge computing systems, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, IEEE, 2019, pp. 766–774 (2019).

[31] D. Y. Zhang, T. Rashid, X. Li, N. Vance, D. Wang, Heteroedge: Taming the heterogeneity of edge computing system in social sensing, in: Proceedings of the International Conference on Internet of Things Design and Implementation, ACM, 2019, pp. 37–48 (2019).

[32] Y. Zheng, F. Liu, H.-P. Hsieh, U-air: When urban air quality inference meets big data, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2013, pp. 1436–1444 (2013).

[33] J. Feng, C. Zhang, P. Hao, Online learning with self-organizing maps for anomaly detection in crowd scenes, in: Pattern Recognition (ICPR), 2010 20th International Conference on, IEEE, 2010, pp. 3599–3602 (2010).

[34] Y. Zhang, S. Zhang, D. Leake, Maintenance for case streams: A stream-

ing approach to competence-based deletion, in: International Conference on Case-Based Reasoning, Springer, 2017, pp. 420–434 (2017).

- [35] V. Rajan, S. Bhattacharya, L. E. Celis, D. Chander, K. Dasgupta, S. Karanam, Crowdcontrol: An online learning approach for optimal task scheduling in a dynamic crowd platform, in: Proceedings of ICML Workshop: Machine Learning Meets Crowdsourcing, Vol. 2, 2013 (2013).
- [36] Y. Zhang, D. Zhang, N. Vance, Q. Li, D. Wang, A light-weight and quality-aware online adaptive sampling approach for streaming social sensing in cloud computing, in: 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), IEEE, 2018, pp. 1–8 (2018).
- [37] J. Xu, L. Chen, S. Ren, Online learning for offloading and autoscaling in energy harvesting mobile edge computing, IEEE Transactions on Cognitive Communications and Networking 3 (3) (2017) 361–373 (2017).
- [38] D. Zhang, D. Wang, N. Vance, Y. Zhang, S. Mike, On scalable and robust truth discovery in big data social media sensing applications, IEEE Transactions on Big Data (2018).
- [39] Y. Zhang, C. Zhang, X. Liu, Dynamic scholarly collaborator recommendation via competitive multi-agent reinforcement learning, in: Proceedings of the Eleventh ACM Conference on Recommender Systems, ACM, 2017, pp. 331–335 (2017).
- [40] Z. Xu, C. Tang, M. Tomizuka, Zero-shot deep reinforcement learning

driving policy transfer for autonomous vehicles based on robust control, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2018, pp. 2865–2871 (2018).

- [41] J. S. Supancic III, D. Ramanan, Tracking as online decision-making: Learning a policy from streaming videos with reinforcement learning., in: ICCV, 2017, pp. 322–331 (2017).
- [42] S. R. Branavan, H. Chen, L. S. Zettlemoyer, R. Barzilay, Reinforcement learning for mapping instructions to actions, in: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1, Association for Computational Linguistics, 2009, pp. 82–90 (2009).
- [43] P. Abbeel, A. Coates, M. Quigley, A. Y. Ng, An application of reinforcement learning to aerobatic helicopter flight, in: Advances in neural information processing systems, 2007, pp. 1–8 (2007).
- [44] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, X. Mao, Incentives for mobile crowd sensing: A survey, IEEE Communications Surveys & Tutorials 18 (1) (2015) 54–67 (2015).
- [45] D. Pisinger, A minimal algorithm for the bounded knapsack problem, in: International Conference on Integer Programming and Combinatorial Optimization, Springer, 1995, pp. 95–109 (1995).
- [46] F. Pan, W. Wang, A. K. Tung, J. Yang, Finding representative set from massive data, in: Data Mining, Fifth IEEE International Conference on, IEEE, 2005, pp. 8–pp (2005).

- [47] J. Friedman, T. Hastie, R. Tibshirani, *The elements of statistical learning*, Vol. 1, Springer series in statistics New York, 2001 (2001).
- [48] B. C. Ross, Mutual information between discrete and continuous data sets, *PloS one* 9 (2) (2014) e87357 (2014).
- [49] S. J. Bradtke, M. O. Duff, Reinforcement learning methods for continuous-time markov decision problems, in: *Advances in neural information processing systems*, 1995, pp. 393–400 (1995).
- [50] M. Blangiardo, M. Cameletti, *Spatial and spatio-temporal Bayesian models with R-INLA*, John Wiley & Sons, 2015 (2015).
- [51] K. Chen, J. Yu, Short-term wind speed prediction using an unscented kalman filter based state-space support vector regression approach, *Applied Energy* 113 (2014) 690–705 (2014).
- [52] C.-J. Ho, J. W. Vaughan, Online task assignment in crowdsourcing markets., in: *AAAI*, Vol. 12, 2012, pp. 45–51 (2012).

7. Appendix

7.1. *The proof of NP-hardness of proposed task allocation problem*

We use the *reduction* technique to prove the quality-cost-aware task allocation problem is NP-hard by reducing it to a well-known NP-hard problem, i.e, the bounded knapsack problem. Let us consider a simplified problem of the quality-cost-aware task allocation problem, where the goal is to only optimize the overall sensing error (given the predefined task allocation cost constraint).

The proof of NP-hard: Bounded Knapsack problem is a known NP-complete problem [45]. We do the transformation as follows:

1. For a bounded knapsack problem with a set of M items, each item m comes with a weight of w^m and a value of v^m . The goal is to select a subset of at most N items from the M items (i.e., $N < M$) to maximize the overall value of the selected items $\sum_{i=0}^N v^i$ while ensuring the overall weight of the selected items is under the knapsack's weight capacity $\sum_{i=0}^N W^i < \Theta$ (where Θ is the knapsack's weight capacity).
2. Let us convert the bounded knapsack problem to the simplified task allocation problem as follows: considering a set of M sensing cells, selecting each sensing cell m requires a sensing cost of w^m and reduces the sensing error by v^m . The goal is to select a subset of at most N sensing cells from the M sensing cells to maximize the reduced sensing error $\sum_{i=0}^N v^i$ while ensuring the overall sensing cost is under the predefined threshold $\sum_{i=0}^N w^i < \Theta$.
3. By solving the task allocation problem, we can get the answer for the bounded knapsack problem. In particular, if we can allocate N sensing cells to maximize the reduced sensing error while ensuring the overall sensing cost is under the predefined threshold, then we must be able to find N items in the bounded knapsack problem so that the overall value can be maximized while ensuring the overall weight is under the knapsack's weight capacity.

The above reduction process proves that the simplified task allocation problem is NP-hard. In other words, we know that the simplified task allocation problem is at least as hard as all problems in NP, so our proposed quality-cost-aware task allocation problem is also at least as hard as all problems

in NP. Therefore, our proposed quality-cost-aware task allocation problem is NP-hard.