



# Quality aware cost efficient reward mechanism in mobile crowdsensing system with uncertainty constraints

Sanjoy Mondal<sup>1</sup> · Abhishek Das<sup>2</sup>

Received: 6 February 2023 / Accepted: 14 June 2023 / Published online: 1 July 2023  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

Mobile Crowdsensing (MCS) has shown the greatest potential that allows smart devices to collect and share different sensing data. Mobile users (or participants) send the desired sensing data to the service providers and collect rewards. However, the reward needs to be given such as, it does not increase platform costs. On the other hand, the unsatisfactory reward may reduce the interest of the participant which may degrade the quality data. Therefore, increasing sensing data quality with a constrained budget is a crucial challenge. There has been extensive research on the reward mechanism for MCS, but, most of the work is on the basic assumption that participant will complete their assigned task positively. In this paper, we propose an efficient user selection mechanism for Mobile Crowdsensing System (MCS) by considering the Probability of Success (PoS) of users (i.e. participant may fail to complete the assigned task with some probability). For the selection of an efficient user, the proposed mechanism also accounts the parameters like data quality and platform cost. We also propose a reward calculation model for the selected users. Minimizing the platform cost with a constrained budget is an NP-hard problem. To provide a sub-optimal solution to this problem Chaotic Krill-Herd optimization algorithm is used. The extensive simulation results reveal that the proposed method outperforms the existing work by considerable margins.

## List of symbols

$n$	Number of mobile users (or participants)
$m$	Number of tasks
$b_j$	Allocated budget for task $j$
$S_i$	Set of tasks covered by user $i$
$p_i^j$	Probability of success of user $i$ for the task $j$
$r_i^s$	Sensing range of user $i$
$\mathcal{P}$	Set of locations
$c^j$	Platform cost to complete task $j$
$d_i^j$	Required minimum data quality for task $j$
$\mathcal{D}_Q$	Total data quality
$\omega_n$	Inertia weight of the induced motion
$\omega_f$	Inertia weight of the foraging motion
$\alpha_i$	Direction of the induced motion
$V^f$	Foraging speed
$\rho_i^{best}$	Best fitness effect of the $i$ th Krill

PoI Point of Interest

## 1 Introduction

This is the era of smart mobile devices embedded with a broad range of sophisticated sensors like smartphones, smart wearable devices, GPS devices, digital compasses and so on Xiao et al. (2018). By utilizing the expeditious proliferation of smart mobile devices, MCS (Ganti et al. 2011) has become a large-scale sensing paradigm to facilitate monitoring and sensing the surrounding environment. It acquires real-world conditions including environmental pollution, traffic congestion detection, health care condition, hotspot identification etc (Yang et al. 2017). Traditional static sensing applications depend on the expensive and specialized sensing infrastructure (Yick et al. 2008). However, utilizing this immoderate infrastructure only to monitor the surrounding environment is not a smart solution due to its high maintenance cost and lack of reusability. In contrast to that, MCS provides this infrastructure to monitor by leveraging ubiquitous mobile devices and their advanced embedded sensors. In recent years, a lot of magnificent applications of MCS have been

✉ Sanjoy Mondal  
sanjoymondal@soa.ac.in

<sup>1</sup> Department of Computer Science and Information Technology, ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India

<sup>2</sup> Department of Computer Science and Engineering, Aliah University, Kolkata, India

proposed but note that in all applications a 'large number of mobile users' is a very important parameter in the MCS platform. Therefore to motivate, the users to involve and selecting appropriate participants are two challenging tasks in MCS (Gao et al. 2019). Over and above, sensing tasks deplete the computational resources of mobile devices, require human effort, and incur privacy issues. Therefore users are reluctant to participate in the MCS unless they get sufficient rewards (Luo et al. 2017).

Researchers have studied extensively incentive mechanisms in MCS. However, they ignore the possible unreliability in mobile crowdsensing, the majority of the existing incentive schemes always assume that when a user is assigned a sensing assignment, she would successfully execute it. Here we argue that in a more realistic MCS scenario, due to the uncertainty of mobile devices or others parameters, users may fail to complete the assigned tasks.

Acquiring high-quality data at a reasonable sensing cost is the key factor of a successful MCS application (Peng et al. 2018). Participants may not provide trusted data due to malicious intent or malfunctioning devices (Jiang et al. 2017). That's why service providers assign multiple participants for a specific data sensing task, which may incur the sensing cost. Thus, these two factors - data quality and sensing cost have certain intrinsic collisions in MCS. Receiving high-quality data with a limited budget is the prime concern of all organizations.

A lot of research has already been done on the effective incentive mechanism to motivate the participants (Ogie 2016) or to measure the data quality. Authors in Wang et al. (2018d) proposed a framework to optimize the overall utility within a reasonable incentive budget. Here utility defines the quality of sensing data. In Li et al. (2019) authors optimize overall quality utility and overall incentive payment, the two most important objectives of MCS with a limited resource pool.

To deal with the above-mentioned problem in this paper we have an MCS environment as shown in Fig. 1. Service providers (SP) broadcast some tasks to mobile users. Engrossed mobile users do response and send their profiles to the SP through Wi-fi or cellular network. Among the involved users it's a very challenging task to select efficient

users. In this paper, our objective is to design a reward mechanism for the selected users that minimizes the MCS platform cost while guaranteeing the desired data quality for each task. To design such a mechanism we incorporate the probability to complete a task by the corresponding users. Selecting the users that minimize the platform cost with maximum data quality can be proved to be an NP-Hard problem. Thus to search for a sub-optimal and computationally efficient solution we use Chaotic Krill Herd (CKH) Optimization algorithm.

The Chaotic Krill Herd (CKH) Optimization algorithm (Saremi et al. 2014), is considered one of the most advanced heuristic optimization algorithms developed to solve the global numerical optimization problem. The basic Krill Herd (KH) algorithm (Wang et al. 2019) was designed inspired by the behavior of Krill Herd. Next, the Chaos properties were introduced into KH to increase its global search ability. The objective function of the Krill movement is designed based on the minimum distance of each Krill individual and the high density of the herd. We summarize our contributions in this work as follows:

- We formulate the problem of incentive mechanism in a Mobile Crowdsensing System with platform budget constraints and user uncertainty. Which can approximately maximize the data quality with allocated budgets by the service providers.
- For large-scale mobile users, we propose an efficient user selection mechanism. The selection process considers various parameters like data quality, platform cost of the task, the budget allocated for the task, and the probability of success of the users.
- For the selected user we propose a reward calculation algorithm based on the user's uncertainty.
- We perform extensive simulations and experimental results that validate our proposed mechanism, which can effectively diminish the platform cost while maintaining the desired data quality and satisfactory reward for the users.

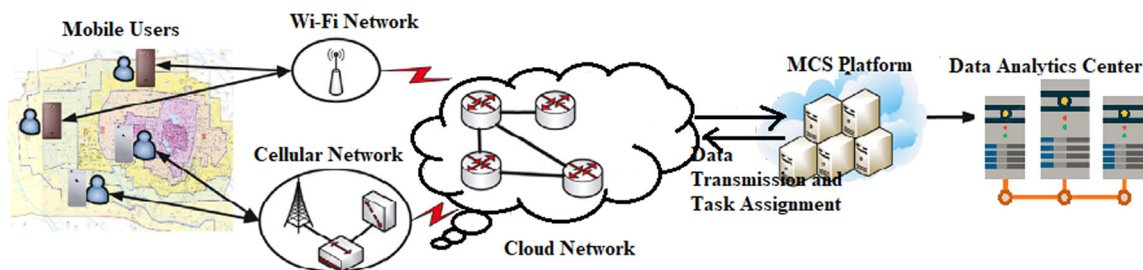


Fig. 1 A typical mobile crowdsensing system environment

## 1.1 Paper's organization

The rest of the paper is organized as follows. Section 2 describes a brief survey on contemporary research works whereas Sect. 3 describes the MCS system and formulate the problem. Section 4 briefly describes the background of CKHO algorithms. In Sect. 5 proposed framework and algorithms are discussed. Section 6 presents the simulation environment and parameters. The simulated results are also validated in the same section followed by the Conclusion and future scope described in Sect. 7.

## 2 Related works

As Crowd sensed Data quality is an important factor of MCS and IOT, a lot of research work has been done on this topic. Also to motivate the participants for delivering high-quality data, a lot of reward mechanism has been proposed. So in this section, we survey different proposed algorithms for improving crowd sensed data quality and about different reward mechanisms in the domain of MCS. Nowadays the Quality of Data is a major research interest. In 2017, S. Ding, X. He, and J. Wang proposed a model for improving the quality of service in the paper (Ding et al. 2017), where they gave attention to the selection of appropriate service nodes with good coverage and transmission performances. In this paper, they mapped the service node selection problem into a multiobjective optimization problem. Although the multitasking capability of the sensor nodes increases the performance of the system the drawback is the sensing quality of individual nodes may become poor. In pape Wang et al. (2017) proposed a framework for optimizing the overall utility while the multiple tasks are distributed among the limited number of participants but the problem is this framework does not concern with the data quality. To deal with the data quality in paper (Yang et al. 2017). integrate the incentive mechanism with the data mining and machine learning-based data analysis technique. They have designed an unsupervised learning model to evaluate participants' data qualities and exploited an outlier detection technique to cut off anomalous data items. Again in the paper (Guo et al. 2017) the proposed mechanism measures and improves the data quality by a visual crowdsensing system. It emphasizes data diversity rather than the resolution of data. There are some proposed mechanisms for quality-aware incentives to motivate the participants for delivering high-quality data. In Peng et al. (2018) D. Peng, F. Wu and G. Chen designed a quality-based incentive mechanism that analyzes the data reading to estimate the quality of sensing data and depending on the quality of the data participants

are rewarded. In the paper (Wang et al. 2016) the authors designed a fine-grained MCS model based on a quality-driven actuation that predicts the possibility of participant's incorrectness. Based on this analysis the submitted sensed data are considered highly reliable or not reliable. In Jin et al. (2015) authors designed incentive mechanisms depending on the reverse actuation model to maximize social welfare while maintaining individual rationality and computational efficiency. In the paper (Pouryazdan et al. 2017) Pouryazdan et al. analyzed data trustworthiness based on three existing data quantifying approaches which are statistical, recommendation-based, and anchor-based approaches under collaborative reputation scores in a crowd-sensing system Also, in Jin et al. (2017) Jin et al. proposed a payment mechanism based on game theory that measures the effort of all participants and the participant who gives high effort will be highly paid. All the above-discussed incentive mechanisms are offline. In the paper (Gao et al. 2019) the authors proposed an online incentive mechanism model by formulating an optimization problem that maximizes the amount of high-quality data under a restricted budget. Also, it introduces a bonus system by measuring the task completion level of each participant that motivates each participant to deliver their best, and also this model solves the problem for online scenarios of random movement of participants.

We can see from Table 1 that authors have suggested various participant selection algorithms by taking into account various things including energy efficiency, geographical coverage, budgetary constraints, and data quality, among other things. The MCS system has a variety of circumstances, and the same participant selection algorithm could not work well in all of them. A single participant may be allowed to upload data for each target in some applications where data quality may be affected, for instance, whereas MCS applications for quality assurance may not accept a single data packet for each target. As a result, numerous participants may be permitted to submit data packets for each target in this scenario. Therefore, it is necessary to build several algorithms to choose efficient ones for various application circumstances.

## 3 System model and problem formulation

We consider a mobile crowdsensing environment, where the service provider distributes some task or provides some point of interest (PoI) (also known as target) such as hospitals, restaurants, etc. to the participants (or) mobile users. We assume that mobile users have sensor-enabled mobile devices through which they can collect data for the corresponding PoI and transmit it to the service provider. The set of PoI is denoted by  $\mathcal{P} = \{p_1, p_2, p_3, \dots, p_m\}$ , where  $m$

**Table 1** Existing works with various MCS parameters

Authors	Mobile crowdsensing parameters			
	Target coverage	Energy efficiency	Budget constraint	Data quality
Yu et al. (2016)	✗	✗	✓	✗
Chen and Yang (2019)	✓	✗	✗	✗
Ko et al. (2019)	✓	✓	✗	✗
Wang et al. (2018b)	✗	✗	✗	✓
Song et al. (2020)	✓	✗	✓	✓
Wang et al. (2018c)	✓	✗	✓	✗
Wang et al. (2018a)	✓	✗	✗	✓
Zhang et al. (2016)	✓	✗	✓	✓
El Khatib Rawan et al. (2018)	✓	✗	✓	✗
Li et al. (2017)	✓	✗	✓	✓
Mondal et al. (2022)	✓	✓	✗	✗
Mondal et al. (2018)	✓	✓	✗	✗

denotes the number of targets. The locations of  $\mathcal{P}$  can be denoted as  $\vec{\mathcal{P}} = \{\vec{p}_1, \vec{p}_2, \vec{p}_3, \dots, \vec{p}_m\}$ . Each  $\vec{p}_i$  consisting of 2D location information and denoted as  $\vec{p}_i = \{x_i, y_i\}$ . The corresponding set of task  $\mathcal{T} = \{t_1, t_2, t_3, \dots, t_m\}$ . MCS platform consisting of  $n$  participants/mobile users  $\mathcal{M} = \{m_1, m_2, \dots, m_n\}$ . The data quality of user  $m_i$  for the task  $j$  is denoted as  $d_i^j$ . The maximum budget for each task is denoted by the budget set  $\mathcal{B} = \{b_1, b_2, b_3, \dots, b_n\}$ . Here  $b_j$  denotes the maximum budget that can be allocated for the task  $t_j$ . The location of each user can be defined as  $\vec{l}_{m_j}$ . We consider an arbitrary movement for each mobile user. If a mobile user is close to the PoI, we consider that the PoI is fully covered by the mobile user. Based on the coverage of the task,  $S_i$  denotes a set of tasks covered by user  $i$ . To upload the sensed data, a mobile user can use the cellular network and for exceptional cases, if the area is not covered by the cellular network, for short-range communication, a mobile user can use Bluetooth or WiFi. The sensing range and communication range of the  $j^{th}$  mobile user are denoted by  $r_j^s$  and  $r_j^c$  respectively. A PoI  $p_j$  is covered by the mobile user  $m_j$  iff,  $\|\vec{l}_{m_j} - \vec{p}_i\| \leq r_j^s$ . Whether a mobile user  $m_j$  covers the PoI  $p_i$  is defined as

$$\eta_{m_j} = \begin{cases} 1 & \text{if } \|\vec{l}_{m_j} - \vec{p}_i\| \leq r_j^s \\ 0 & \text{otherwise} \end{cases}$$

Due to network issues or other users may fail to complete the task. The MCS platform includes a factor probability of success (PoS) of the user for each task. PoS of user  $i$  to the task is defined as  $P_i^j$ . Each task  $t_j$  should be assigned to the user  $i$  having PoS greater than  $T_j$ . The cost for completion of the task  $j$  by the user  $i$  is defined as  $c_i^j$ . In contrast to minimizing the platform cost, our objective is to reduce the

platform cost while ensuring the data quality of each task with satisfactory rewards received by the users. So our problem can be formulated as,

$$\text{Maximize : } \mathcal{D}^Q \tag{1}$$

$$\text{Subjectto : } d^j \leq d_j \quad \forall j \in \mathcal{T} \tag{2}$$

$$b_j \geq c_j + r_j \quad \forall j \in \mathcal{T} \tag{3}$$

The constraints in eq. (2) ensure that data quality must be maintained within the allocated budgets for each task. Here  $d'_m$  denotes the final rewards of  $m^{th}$  user.

### 4 Priliminary of CKHO algorithm

Krill Herd is an innovative meta-heuristic swarm intelligence optimization algorithm that can be used to solve various complex problems. The Algorithm performs based on krill's individuals. At the time of food, hunting krill swarm are communicate with another member of the swarm and this is the basis of the KHO algorithm. Three movements are implemented for this KHO which repeats multiple times. To set the position of the krill's individual three main activities are considered and they are as follows:

- (1) The movement induced by the other krill (known as induced movement).
- (2) Foraging actions.
- (3) Random diffusion.

KHO follows the lagrangian model given in the equation

$$\frac{dX_i}{dt} = N_i + F_i + D_i \tag{4}$$

where  $N_i$  determines the motion produced by other krill individuals,  $F_i$  specifies the foraging motion, and  $D_i$

indicates the physical diffusion of the  $i^{th}$  krill individuals. The motion induced by the  $i^{th}$  krill can be determined as follows

$$N_i^{new} = N^{max} \alpha_i + \omega_n N_i^{old} \tag{5}$$

where  $N^{max}$  represents the maximal induced speed,  $\omega_n$  indicates the inertia weight of the induced motion, and  $N_i^{old}$  is the last induced motion

Foraging motion can be formulated in terms of two important factors such as current food location and previous experience with the food locations. The motion for the  $i$ th krill individuals can be expressed as follows:

$$F_i = V^f \beta_i + \omega_f F_i^{old} \tag{6}$$

$$\beta_i = \beta_i^{food} + \beta_i^{best} \tag{7}$$

Here  $V^f$  indicates the foraging speed and  $\omega_f$  represents the inertia of the foraging motion. The last foraging motion is indicated by  $F_i^{old}$  whereas  $\beta_i^{food}$  represents the attraction of the food. The parameter  $\beta_i^{best}$  determines the best fitness effect of the  $i^{th}$  krill till now.

Random diffusion is a process consisting of two components such as the highest diffusion speed and a random directional vector. It can be determined by the following equation:

$$D_i = D_{max} \delta \tag{8}$$

Where  $D_{max}$  indicates the maximum diffusion speed and the random directional vector is represented by  $\delta$ .

Along with these three motion actions, the position of the  $i$ th krill individuals during the time interval  $t$  and  $t + \Delta t$  can be expressed by the following equation

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \tag{9}$$

Here  $\Delta$  can be calculated from the equation

$$\Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j) \tag{10}$$

Where  $NV$  indicated the total number of variables.  $UB_j$  and  $LB_j$  are the upper and the lower bound respectively for the variable  $j$ .  $C_t$  is the constant number.

Chaos has the properties of randomness, regularity, and ergodicity. The integration of basic KH and chaos phenomenon, known as CKH can execute iterative searches at a higher speed than the other standard ones. Fig. 2 explain different steps of CKH. We know the main parameters of the KH algorithm is inertia weights  $\omega_f$  and  $\omega_n$ . These two values are important for determining the convergence speed and changing the position of the krill in KH. To improve this algorithm CKH algorithm introduces a chaotic map random variable  $c_0$  for tuning the inertia

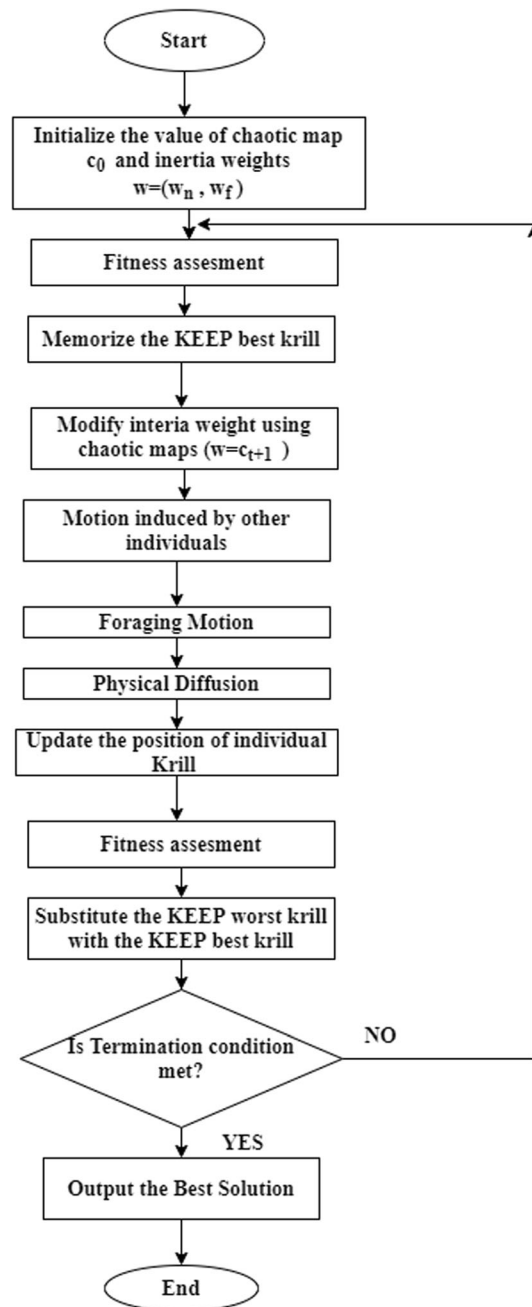


Fig. 2 Flowchart of chaotic krill herd optimization

weights  $\omega_f$  and  $\omega_n$  which gives the better convergence speed with the best solution.

### 5 Proposed methodology

In this section we propose an efficient user selection mechanism for large-scale mobile users. After the selection of efficient users, we propose a reward mechanism because the user may fail to complete the assigned tasks.

## 5.1 Efficient user selection

We have discussed the NP-Hardness of the problem of selecting efficient users among large-scale mobile users with minimizing platform cost and maximizing data quality. To provide the sub-optimal solution we designed a CKHO-based user selection mechanism by considering the important parameters data quality of the mobile users, Cost to complete the task, and Probability of Success (PoS) to complete the task.

### 5.1.1 Data quality of mobile user

Mobile crowd sensing platform expects the maximum data quality  $\mathcal{D}^Q$  with the limited budget  $\mathcal{B}$  at every stage. The Mobile Crowdsensing platform will select high-quality sensing data among the participants. Therefore our optimization goal can be formulated as:

$$\text{Maximize } \mathcal{F}_1 = \mathcal{D}^Q = \sum_{i \in \mathcal{M}} d_i^j \quad \forall j \in \mathcal{T} \quad d_i^j \leq d_j \quad (11)$$

To choose the high-quality sensing data, we can define the data quality as a utility function given below:

$$\mathcal{D}^Q = \sum_{t \in \mathcal{T}} \min \left( d_t \sum_{m \in \mathcal{M}} x_{m,t} \right) \quad (12)$$

Here  $x_{m,t} = 1$ , if the mobile crowdsensing platform successfully picked up its desired data quality for the task  $t$  by the participant  $m$ , and  $x_{m,t} = 0$  otherwise.

### 5.1.2 Cost function

The mobile users belonging to the corresponding region can only participate in that task. In this paper, the cost function consists of three components *viz.* platform cost to complete a task, reward received by the user to complete a task, and penalty factor to selecting insufficient user. Based on the factors mentioned, the cost function for assigning all the tasks by the MCS system can be defined as follows:

$$\text{Minimize } \mathcal{F}_2 = C = \sum_{j \in \mathcal{T}} c^j + \sum_{i=1}^{\mathcal{T}} \sum_{j=1}^{\mathcal{M}} \mathcal{R}_{i,j} \times \mathcal{T}_{i,j} + \sum_{j=1}^{\mathcal{M}} \mathcal{K}_j \quad (13)$$

Here  $\mathcal{R}_{i,j}$  represents the number of rewards received by the  $i^{\text{th}}$  participant for accomplishing the given  $j^{\text{th}}$  task.  $\mathcal{T}_{i,j}$  is a matrix containing binary values. If  $i^{\text{th}}$  participant is eligible (i.e. maintains the desired data quality) for  $j^{\text{th}}$ , then  $\mathcal{T}_{i,j}$  is set to 1, and otherwise  $\mathcal{T}_{i,j} = 0$ .

$\mathcal{K}_j$  is defined as a penalty for the chosen solution due to the insufficient number of participants for the task  $j$ . Here  $c^j$  defines the platform cost for the task  $j$ . For each task  $j \in \mathcal{T}$ ,  $C_j \geq b_j$ , here  $C_j$  denotes the cost for the task  $j$  and  $b_j$  is the allocated budget by the MCS platform for the task  $j$ .

### 5.1.3 Probability of success (PoS)

To deal with faulty or incomplete transmission, we consider the parameter PoS. Our objective is to select the user having PoS greater than a given threshold value. It can be defined as:

$$\mathcal{F}_3 = \sum_{i \in \mathcal{M}} p_i^j | j \in S_i \quad (14)$$

Here,  $S_i$  denotes the set of tasks that can be allocated for user  $i$ .

By consolidating the above factors we formulate objective function as follows:

$$F = \alpha \times \mathcal{F}_1 + \beta \times 1/\mathcal{F}_2 + \gamma \times \mathcal{F}_3 \quad (15)$$

Here  $\alpha$ ,  $\beta$  and  $\gamma$  are weight factors. The weighted factors are optimized by the CKHO algorithm.

## 5.2 Reward calculation for the selected users

After the selection of users by the above mechanism, we calculate the reward for selected the user for the corresponding assigned tasks based on the Algorithm 1. Here the reward is calculated based on the data quality of the selected user (defined in line 1). The factor  $k_i$  is defined based on the received data quality by the user  $i$  for the task  $j$  (in line 2). Later, the reward is calculated in two ways, based on the conditions of whether the user completed the task or not (in line 3). Finally, the algorithm returns the reward for the selected user  $i$  for the task  $j$ .

**Algorithm 1:** Reward calculation for selected user  $i$  for task  $j$

**Input :** Selected user  $i \in U^*$ , Required data quality  $d_i^j \quad j \in S_i$ , cost  $c^j$   
**Output:** Reward  $\mathcal{R}_i$  for user  $i$  for the task  $j$

- 1  $d_r \leftarrow d_i^j$ /\* data quality provided by user  $i$  for the task  $j$  required for task  $j$  \*/
- 2  $k_i = 1 - e^{-d_r}$
- 3  $\mathcal{R}_i = \begin{cases} (1 - k_i) \times \alpha + c^j & \text{if task } j \text{ completed by } i \\ -k_i \times \alpha + c^j & \text{if task } j \text{ is not completed by } i \end{cases}$
- 4 return  $\mathcal{R}_i$ ;

**5.3 Time complexity analysis**

For the selection of efficient users we use the Krill Herd Optimization (KHO) algorithm. For  $n$  individuals, the algorithm requires computing the distance between each pair of krill particles which requires  $O(n^2)$  time. Additionally, the algorithm requires a search for the best solution or the global optimum which takes  $O(n)$  time. Therefore total complexity of the KHO algorithm is  $O(n^2 + n) = O(n^2)$ .

The time complexity of Algorithm 1 is  $O(1)$  as an assignment (here data quality assigned to variable  $d_r$ ) takes  $O(1)$  or constant time. If there are  $U$  selected users therefore complexity for  $U$  selected is  $O(1 * U) = O(U)$ .

**6 Performance evaluation**

In this section, we present the simulation result and evaluate the performance of the proposed reward mechanism. We compare our proposed methods with state-of-the-art QAIM (Jiang et al. 2017).

**6.1 Simulation environment**

For this simulation we use a T-drive trajectory data sample of Shanghai, China (Zheng 2011). The data set contains approx 10K taxi trajectories. Each entry of the dataset record contains Taxi-ID, Time-stamp, pickup point, and

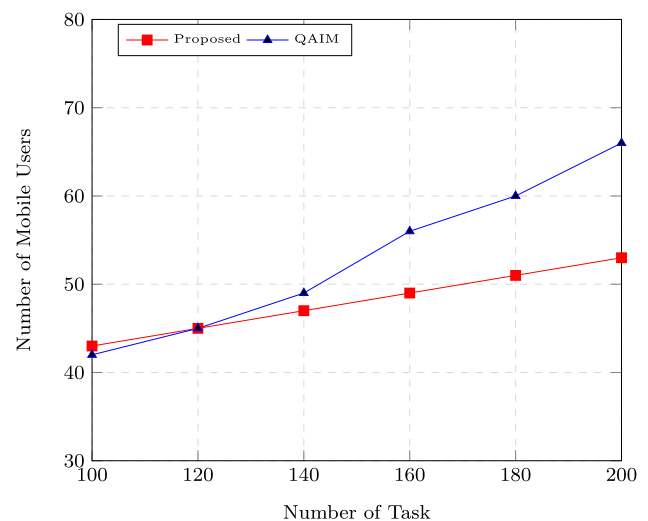
**Table 2** Simulation parameters

Parameter	Values
Number of mobile users	100–300
Number of tasks	100–200
Minimum data quality for each task	0.6
Minimum PoS requirement for each task	0.75
Average cost	20

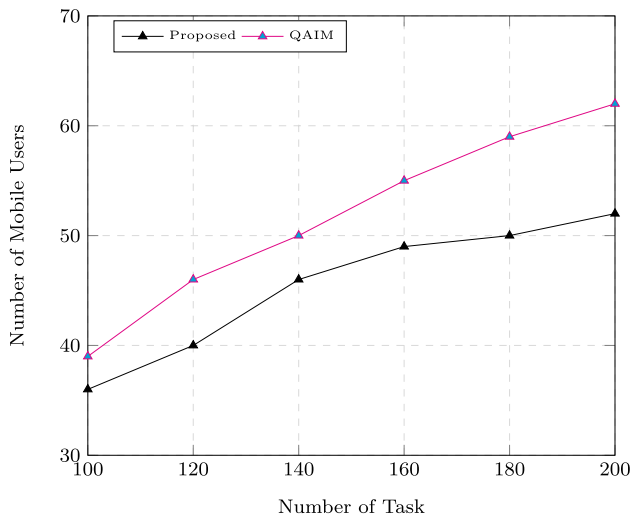
dropping points of the passengers. To define the probability of success (PoS), we calculate the probability that a taxi pickup and drops a passenger. To generate the mobility pattern of the taxi Markov Chain is used. The initial location of the taxi is generated randomly and the task set is generated randomly with uniform distribution. We consider a  $5000m \times 5000m$  region of interest (RoI) where mobile users (here taxi with a mobile device) and tasks are distributed randomly throughout the RoI. For simplicity, we consider homogeneous mobile devices with a sensing range 45 m/second. The communication range for a mobile device is set to 70 m. Two mobile users are considered to be connected if they belong to the communication range of each other. However, the sensed data can be transmitted to the service provider or BS in a single or multi-hop manner. Simulation parameters and their default values used in this paper are given in Table 2.

**6.2 Result analysis**

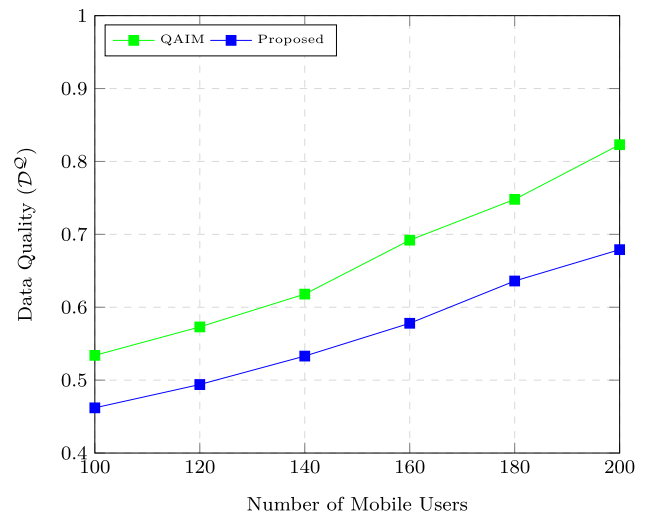
To evaluate the performance of the proposed reward mechanism, the following metrics are used:



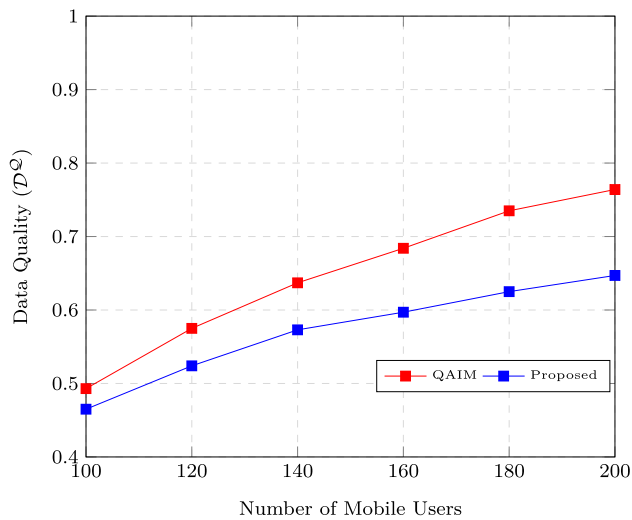
**Fig. 3** Number of task vs number of mobile users (total available mobile users 200, total Task 200)



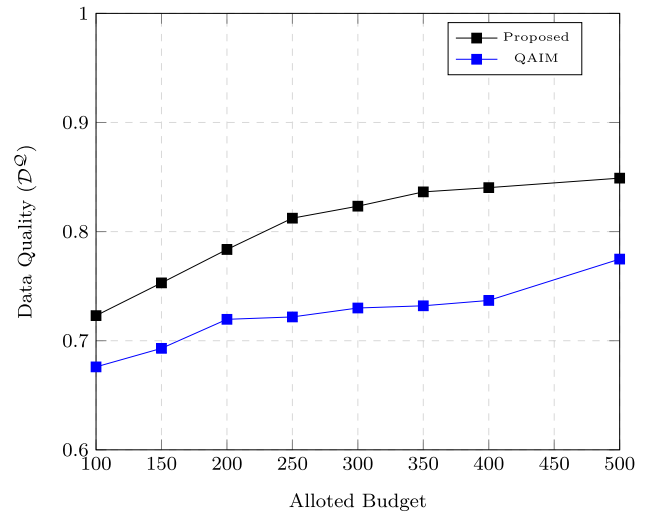
**Fig. 4** Number of task vs number of mobile users (total available mobile users 300, total task 200)



**Fig. 6** Number of mobile users vs data quality when total number of targets 200



**Fig. 5** Number of mobile users vs data quality when total number of targets 200



**Fig. 7** Total number of participants 300, number of task 200

- Impact of budget on data quality
- Impact of budget on number of mobile users
- Impact of data quality on number of mobile users
- Effect of number of mobile users on the number of tasks

The primary concern that we need to investigate is the impact of the number of mobile users on the number of tasks distributed/provided by the service provider. It is observed from Fig. 3 that, as the number of tasks increases, the number of mobile users increases. It's very true that to complete more tasks, a more mobile device is required, however, the ratio of tasks to mobile users is comparatively less in the proposed method compared to QAIM.

Another important concern is the number of available mobile users. From Figs. 3 and 4, we can see that for the

same number of task, a less number of mobile users are required when the availability of mobile users are more. It is obvious because if the availability of mobile users is more, a mobile user can be selected for more than one task.

Similarly, the availability of mobile users has a great impact on the data quality of the system. It can be observed from the Figs. 5 and 6 that, for both the methods, when the number of mobile users increases, data quality also increases. It happens because the data quality depends on the distance of the mobile users from the location of targets. So if we increase the availability of mobile users, the data quality is increased. However, our proposed methods give better data quality with the same number of mobile users.

If we concentrate on the impact of data quality budget, We can observe from Fig. 7 that with the increase in



budget, data quality also increased for both methods. However, our proposed methods provide better data quality compared to the existing one on the same budget.

## 7 Conclusion and future scope

This paper investigates the data quality issue in the mobile crowdsensing network with a constrained budget. We studied how to choose efficient mobile users among the available mobile users so that data quality can be maximized. For the selection of an efficient user, we have considered a realistic scenario where a user may fail to finish the assigned tasks. By increasing the data quality, service costs may be created, thus, to optimize the platform cost with data quality CKHO optimization technique is used. The simulation results are validated by our proposed methods and observed from the simulation results that, our proposed methods outperform the state-of-the-art.

Future research will focus on ways to make better use of the temporal-geographic characteristics of mobile users as well as POIs to increase the coverage of mobile crowdsensing networks.

**Data availability** The website for downloading the dataset <https://www.kaggle.com/datasets/arashnic/tdriver>.

## References

- Ding S, He X, Wang J (2017) Multiobjective optimization model for service node selection based on a tradeoff between quality of service and resource consumption in mobile crowd sensing. *IEEE Internet Things J* 4(1):258–268
- El Khatib RF, Zorba N, Hassanein HS (2018) Cost-efficient multi-tasking in coverage-aware mobile crowd sensing. In 2018 14th International Wireless Communications Mobile Computing Conference (IWCMC), p. 594–599
- Ganti RK, Ye F, Lei H (2011) Mobile crowdsensing: current state and future challenges. *IEEE Commun Mag* 49(11):32–39
- Gao H, Harold LC, Jian T, Dejun Y, Pan H, Wendong W (2019) Online quality-aware incentive mechanism for mobile crowd sensing with extra bonus. *IEEE Trans Mobile Comput* 18(11):2589–2603
- Guo B, Chen H, Han Q, Zhiwen Y, Zhang D, Wang Y (2017) Worker-contributed data utility measurement for visual crowdsensing systems. *IEEE Trans Mob Comput* 16(8):2379–2391
- Jiang L-Y, Fan He Y, Wang L-JS, Huang H-p (2017) Quality-aware incentive mechanism for mobile crowd sensing. *J Sens* 2017:5757125
- Jiaoyan C, Jingsen Y (2019) Maximizing coverage quality with budget constrained in mobile crowd-sensing network for environmental monitoring applications. *Sensors* 19(10):2399
- Jin H, Su L, Chen D, Nahrstedt K, Xu J (2015) Quality of information aware incentive mechanisms for mobile crowd sensing systems. In Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 15. New York, NY, USA., p. 167–176, Association for Computing Machinery
- Jin H, Su L, Nahrstedt K (2017) Theseus: incentivizing truth discovery in mobile crowd sensing systems. In Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 17. New York, NY, USA. Association for Computing Machinery
- Ko H, Sangheon P, Leung Victor CM (2019) Coverage-guaranteed and energy-efficient participant selection strategy in mobile crowdsensing. *IEEE Internet Things J* 6(2):3202–3211
- Li T, Liu Y, Gao L, Liu A (2017) A cooperative-based model for smart-sensing tasks in fog computing. *IEEE Access* 5:21296–21311
- Li M, Gao Y, Wang M, Guo C, Tan X (2019) Multi-objective optimization for multi-task allocation in mobile crowd sensing. *Procedia Computer Science*, 155:360–368. The 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2019), The 14th International Conference on Future Networks and Communications (FNC-2019), The 9th International Conference on Sustainable Energy Information Technology
- Luo T, Kanhere SS, Huang J, Das SK, Fan W (2017) Sustainable incentives for mobile crowdsensing: auctions, lotteries, and trust and reputation systems. *IEEE Commun Mag* 55(3):68–74
- Mondal S, Ghosh S, Khatua S, Das R, Biswas U (2018) Cost effective algorithms for participant selection problem in mobile crowd sensing environment. In 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), p. 453–458
- Mondal S, Mitra S, Mukherjee A, Ghosh S, Khatua S, Das A, Das RK (2022) Participant selection algorithms for large-scale mobile crowd sensing environment. *Microsystem Technol* 28(12):2641–2657
- Ogie RI (2016) Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework. *HCIS* 6(1):24
- Peng D, Fan W, Chen G (2018) Data quality guided incentive mechanism design for crowdsensing. *IEEE Trans Mob Comput* 17(2):307–319
- Pouryazdan M, Kantarci B, Soyata T, Foschini L, Song H (2017) Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing. *IEEE Access* 5:1382–1397
- Saremi S, Mirjalili SM, Mirjalili S (2014) Chaotic krill herd optimization algorithm. *Procedia Technology*, 12:180–185. The 7th International Conference Interdisciplinarity in Engineering, INTER-ENG 2013, 10–11 October 2013. Petru Maior University of Tirgu Mures, Romania
- Song S, Liu Z, Li Z, Xing T, Fang D (2020) Coverage-oriented task assignment for mobile crowdsensing. *IEEE Internet Things J* 7(8):7407–7418
- Wang J, Tang J, Yang D, Wang , Xue G (2016) Quality-aware and fine-grained incentive mechanisms for mobile crowdsensing. In 2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS), p. 354–363
- Wang J, Wang Y, Zhang D, Wang F, He Y, Ma L (2017) Psallocator: multi-task allocation for participatory sensing with sensing capability constraints. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 17, New York, NY, USA, p. 1139–1151 Association for Computing Machinery
- Wang C, Li C, Qin C, Wang W, Li X (2018a) Maximizing spatial-temporal coverage in mobile crowd-sensing based on public transports with predictable trajectory. *Int J Distrib Sens Netw* 14(8):1550147718795351

- Wang Liang Y, Zhiwen GB, Fei Y, Fei X (2018b) Mobile crowd sensing task optimal allocation: a mobility pattern matching perspective. *Front Comput Sci* 12:231–244
- Wang J, Wang L, Wang Y, Zhang D, Kong L (2018c) Task allocation in mobile crowd sensing: state-of-the-art and future opportunities. *IEEE Internet Things J* 5(5):3747–3757
- Wang J, Wang Y, Zhang D, Wang F, Xiong H, Chen C, Lv Q, Qiu Z (2018d) Multi-task allocation in mobile crowd sensing with individual task quality assurance. *IEEE Trans Mob Comput* 17(9):2101–2113
- Wang G-G, Gandomi AH, Alavi AH, Gong D (2019) A comprehensive review of krill herd algorithm: variants, hybrids and applications. *Artif Intell Rev* 51(1):119–148
- Xia L, Tianhua C, Caixia X, Huaiyu D, Vincent Poor H (2018) Mobile crowdsensing games in vehicular networks. *IEEE Trans Veh Technol* 67(2):1535–1545
- Yang S, Tang S, Gao X, Yang B, Chen G (2017) On designing data quality-aware truth estimation and surplus sharing method for mobile crowdsensing. *IEEE J Sel Areas Commun* 35(4):832–847
- Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. *Comput Netw* 52(12):2292–2330
- Yu J, Xiao M, Gao G, Hu C (2016) Minimum cost spatial-temporal task allocation in mobile crowdsensing. In: International Conference on Wireless Algorithms, Systems, and Applications, volume 9798, p. 262–271, 08
- Zhang M, Yang P, Tian C, Tang S, Gao X, Wang B, Xiao F (2016) Quality-aware sensing coverage in budget-constrained mobile crowdsensing networks. *IEEE Trans Veh Technol* 65(9):7698–7707
- Yuan J, Zheng Y, Zhang C, Xie W, Xie X, Sun G, Huang Y (2010) T-drive: driving directions based on taxi trajectories. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10. New York, NY, USA, ACM, pp 99–108

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.