

Research Article

A Task Assignment Method Based on User-Union Clustering and Individual Preferences in Mobile Crowdsensing

Zihao Shao , Huiqiang Wang , Yifan Zou , Zihan Gao , and Hongwu Lv 

College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Huiqiang Wang; wanghuiqiang@hrbeu.edu.cn

Received 12 January 2022; Revised 19 March 2022; Accepted 15 April 2022; Published 9 May 2022

Academic Editor: Barbara Guidi

Copyright © 2022 Zihao Shao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mobile crowdsensing (MCS) offers a novel paradigm for large-scale sensing with the proliferation of smartphones. Task assignment is a critical problem in mobile crowdsensing (MCS), where service providers attempt to recruit a group of brilliant users to complete the sensing task at a limited cost. However, selecting an appropriate set of users with high quality and low cost is challenging. Existing works of task assignment ignore the data redundancy of large-scale users and the individual preference of service providers, resulting in a significant workload on the sensing platform and inaccurate assignment results. To tackle this issue, we propose a task assignment method based on user-union clustering and individual preferences, which considers the influence of clustering data quality and preference-based sensing cost. Firstly, we design a user-union clustering algorithm (UCA) by defining user similarity and setting user scale, which aims to balance user distribution, reduce data redundancy, and improve the accuracy of high-quality user aggregation. Then, we consider individual preferences of service providers and construct a preference-based task assignment algorithm (PTA) to achieve the diversified sensing cost control needs. To evaluate the performance of the proposed solutions, extensive simulations are conducted. The results demonstrate that our proposed solutions outperform the baseline algorithm, which realizes the individual preference-based task assignment under the premise of ensuring high-quality data.

1. Introduction

The pervasive adoption of mobile smart devices and the rapid development of communication network technologies have accelerated the unprecedented expansion of mobile crowdsensing (MCS) in many aspects of our daily lives. MCS [1] is a compelling paradigm that allows a large group of individuals to collaboratively sense data and extract information about social events and national phenomena with common interest using mobile devices (e.g., smartphones, smart glasses).

Task assignment is a critical problem in MCS, where service providers attempt to recruit a group of brilliant users to complete the sensing task at a limited cost. Thus, the core

goal of task assignment is to make a good balance between data quality and sensing cost [2]. Specifically, suppose the sensing platform always assigns inappropriate tasks to users and keeps users away from daily activities. In that case, users will refuse to perform tasks, leading to revenue loss for service providers and reducing the sensing utility. In addition, service providers may have individual preferences (i.e., maximizing benefits, minimizing costs) when selecting user data, further increasing the complexity and diversity of the task assignment.

By now, various methods to improve data quality have been proposed for task assignment. While service providers can enjoy the convenience provided by user data, large-scale user data will lead to the increase of redundant data.

Data redundancy is a potential threat to data quality, which will increase the workload of the platform and reduce the accuracy of the task assignment. Shahraki et al. [3] pointed out that cluster analysis can solve the problem of data redundancy. The most common data clustering mechanism is the k -means algorithm, despite the effectiveness, applying the k -means algorithm in MCS to realize task assignments still need to tackle complexity and balance challenges. Firstly, the k -means algorithm may require a high calculation time. Unfortunately, most task assignment methods are time-sensitive. That is, the acquisition of high-quality data should not be at the expense of time. Secondly, these algorithms are still a risk of poor data balance due to work neglecting the difference in the scale of clustered users. Thus, leveraging clustering analysis to form balanced, low-redundancy, and high-quality data in sensing regions with a limited time is a crucial problem in the MCS task assignment.

Sensing cost is another important issue to consider in task assignment, which aims to select suitable users to achieve task assignment. Up to now, the research works on sensing cost mainly focused user recruitment cost [4–6], user travel cost [7–9], and data transferring cost [10–13]. Most of the existing works for sensing cost consider a homogeneous preference model, which assumes all service providers have the same preference. Each service provider selects user data independently and randomly according to the same task preference. However, this model is at best an approximation, because different service providers indeed have various tastes and preferences. Such heterogeneity in preferences of service providers has been observed in [14].

The shortcomings of existing works drive us to explore a new task assignment method from data quality and sensing cost for realistic MCS applications. Our research efforts aim to achieve a practical task assignment in different individual preferences for real MCS with varying user data quality, while ensuring high-quality clustering data and preference-based sensing cost. More specifically, we first formulate the problem of task assignment. This formulation carefully considers the quality of clustering data and the individual preference for sensing cost. Afterward, a task assignment method is proposed based on user-union clustering and individual preferences. Different from prior works on task assignment, we first considered data redundancy caused by large-scale user data, leveraged the clustering method to reduce data redundancy and improve the accuracy of high-quality user aggregation. Then, based on this solution, we analyzed the impact of individual preferences and solved the diversified task assignment under the individual preference sensing cost.

In summary, this paper makes the following contributions:

- (i) We formulate the task assignment problem from two perspectives. High-quality clustering data and the individual preference sensing cost are considered in our formulation
- (ii) A UCA-based solution is proposed to balance user data scale, reduce data redundancy, and ultimately improve platform efficiency and data quality

- (iii) A PTA-based solution is proposed to solve the task assignment under the individual preference sensing cost. To the best of our knowledge, this is the first work that validates from different perspectives of the task assignment the benefits of exploiting individual preferences and that gains insights through simulations based on real-world data

The rest of the paper is organized as follows: Section 2 discusses related work. Section 3 introduces our system model and problem formulation. Our UCA and PTA solutions are presented in Section 4. In Section 5, we evaluate our proposed method and present evaluation results. Finally, we conclude this paper in Section 6.

2. Related Work

Data quality and sensing cost become the main criterion to assign the tasks. Much work has been done to support the efficient task assignment in MCS. In the following, we shall introduce existing work in these two criteria.

2.1. Data Quality. Improving the accuracy of data quality is an essential design objective for most task assignments. Several factors have a significant impact on data quality, including data collection times, task duration, and data spatial-temporal coverage.

Data collection times refer to the number of times a target phenomenon is expected to be sensed. On the one hand, multiple measurements can reduce sensor reading errors and make sensing results approach the ground truth. Gong et al. [15] pointed out data quality keeps increasing as the collection times increase, characterized by a non-decreasing sub-modular function. On the other hand, there are tiny fluctuations of sensing data even in short durations and small areas. Xiong et al. [16] proposed that data quality will no longer increase when the collected data exceeds a certain threshold. However, multiple measurements are necessary to improve data quality in most cases.

Task duration is the period from the instant a task is published to the deadline. Wang et al. [17] proposed a two-level heterogeneous pricing mechanism based on the timeliness and location dependence of random arrival in MCS. The proposed greedy task selection algorithm can help users choose the appropriate task to maximize the total revenue and realize task assignment. Zeng et al. [18] took the execution time of workers as the optimization goal, and proposed an adaptive Top- k worker selection algorithm to select the most appropriate workers and achieve efficient task assignment. Huang et al. [19] investigated and formulated the time-dependent task allocation problem, and characterized the cost of performing a sensing task for each mobile user. They proposed an efficient task assignment algorithm called the optimized allocation scheme of time-dependent tasks (OPAT), which can maximize the sensing capacity of each mobile user.

Data spatial-temporal coverage is another important metric to evaluate data quality and has been extensively studied. To evaluate the time coverage provided by a group

of users over a period of time, Alagha et al. [20] considered users' location and mobility mode. They designed a stable coverage recruitment parameter to realize task assignment. To reduce the system cost, Song et al. [21] migrated certain qualified users to less popular tasks to increase data coverage and optimize other performance factors. To satisfy both the service provider's coverage sensing preference and the user's revenue preference, Yucel et al. [22] proposed a coverage-aware stable task assignment method and proved that the user's revenue is proportional to the task coverage scale. Experimental results show that this method achieved accurate task assignment on the premise of ensuring user satisfaction and coverage quality.

The above works are good references for addressing the data quality of task assignment. However, most of these studies have not considered the importance of clustering analysis. Guo et al. [23] analyzed some common problems in task assignment and pointed out that cluster-based task assignment is necessary for future MCS task assignment. In the research of user data clustering, Du et al. [24] combined the data quality of users and proposed a Bayesian co-clustering truth discovery model to capture the fine-grained reliability of users on different clusters. This model enhances the usability of each user under the most appropriate task, which is conducive to observing aggregated tasks. Jin et al. [25] proposed a novel MCS system framework that integrates an incentive, a data aggregation, and a data perturbation mechanism. The data aggregation mechanism incorporated workers' reliability to generate highly accurate aggregated results. So far, the research of user clustering data is still in its infancy. It is crucial to consider clustering data quality evaluation to reduce redundancy and improve platform efficiency.

2.2. Sensing Cost. Sensing cost is the costs paid to perform tasks, including user recruitment cost, user travel cost, and data transferring cost. The first is paid by the sensing platform to recruited users for their involvement; the latter two are paid by users for their movements for data collection and data upload, respectively.

User recruitment cost includes per-user recruitment cost and per-data collection cost. To control the recruitment cost of users, Liu et al. [4] studied the user recruitment problem on both the user's and subarea's sides and proposed a three-step strategy, including user selection, subarea selection, and user-subarea-cross (US-cross) selection. Extensive experiments on two real-world data sets show that user recruitment algorithms can effectively enhance the data inference accuracy under a budget constraint. In practical application, Campioni et al. [5] improved recruitment algorithms for vehicular crowdsensing networks, which aims to select participants within a crowdsensing network such that the most sensing data is obtained for the lowest possible cost. Zhao et al. [6] classified the extrinsic utility into the task payoff shared with other participants and the resource cost incurred by participation. Based on this, they proposed a social-aware incentive mechanism by deep reinforcement learning (DRL-SIM) to control user recruitment cost and derive the optimal long-term sensing strategy for all vehicles.

User travel cost relies on the traveling paths of users, which could be fixed, predetermined, or predictable based on users' historical trajectories [7]. In fixed/predetermined-path-based MCS, each user can perform tasks alone or near their traveling path. In this case, the task assignment problem can often be transformed into a set cover problem or bipartite graph matching problem. Wei et al. [8] considered user moving cost and sensing level. They proposed a greedy task assignment algorithm, GP-BS, to select the most cost-effective participant iteratively. In predictable-path-based MCS, the traveling path of each worker is not predetermined. It is tough to accurately predict the specific locations of users in the future at a fine granularity. Wang et al. [9] proposed an approach that exploits the spatial-temporal causality among travel speeds of road sections by a time-lagged correlation coefficient function, which aims to overcome the uneven spatial-temporal distribution of vehicles and the variation of their data-offering intervals. For the sparse MCS scene, Wang et al. [10] propose a deep learning-enabled industrial sensing and prediction scheme, aiming to achieve high-precision prediction of future moments under the hypothesis of sparse historical data.

Data transferring cost is the cost generated for uploading sensing data. Wang et al. [11] considered that the users' main concern is the cost of data uploading, which affects their willingness to participate in a crowdsensing task. The proposed efficient prediction-based user recruitment for MCS can achieve a lower recruitment payment and the highest delivery efficiency. In [12], a data transfer solution for crowdsensing was proposed to minimize the number of users under the constraints of the quality of sensing data and coverage area of all cell towers. When multiple tasks share a pool of staff with bandwidth constraints, a multi-task allocation strategy is proposed in [13] to ensure platform revenue.

Task assignment algorithms for MCS were designed following the different sensing costs. However, the algorithms proposed in the existing works are usually designed based on a fixed choice. That is, they all neglect the individual preferences for sensing cost. In our previous work [26], we have pointed out the influence of individual preferences on selection. Therefore, it is necessary to consider the individual preferences to ensure the practicality of task assignment.

In summary, despite the variety of the literature on data quality and sensing cost in MCS task assignments, the goal is defined chiefly from the overall system's point of view without considering the individual preferences and the importance of clustering data. Hence, they may not necessarily achieve high accuracy and rationality in the task assignment.

3. System Model and Problem Formulation

In this section, we first give the system model for task assignment in MCS. Then, we formulate the task assignment problem.

3.1. System Model

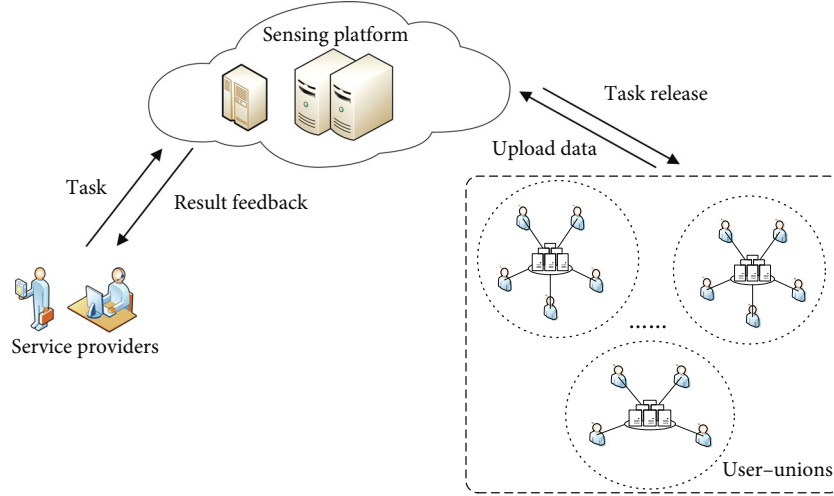


FIGURE 1: Framework for the mobile crowdsensing.

3.1.1. Model Construction. We consider a typical MCS architecture, including a trusted sensing platform, a set of m sensing users, and a set of k service providers, as shown in Figure 1.

For the task assignment, service providers can publish different sensing tasks and task centers to the sensing platform, denoted by $T = \{t_1, t_2, \dots, t_k\}$, $t_{j\text{-center}}$, respectively. The sensing platform assigns tasks to sensing users, denoted by $U = \{u_1, u_2, \dots, u_m\}$. To reduce data redundancy and the burden on the sensing platform, users form the user-union before uploading data. Service providers select appropriate users to realize task assignment according to individual preferences.

In this paper, we make the following assumptions.

- (i) The initial locations of users are uniformly distributed in a specific region
- (ii) The sensing platform is only responsible for the data calculation between users and service providers
- (iii) Service providers have different individual preferences and decide the final choice. Service providers can only select one sensing user to achieve the task assignment

Such assumptions are practical in enterprise or agreement-based cooperation scenarios [27].

3.2. Problem Formulation

3.2.1. Data Quality Problem

(1) *Data Quality Problem Formulation.* In data quality research, considering a large number of sensing users, each user uploading data in an independent way will lead to a decrease in sensing utility. Therefore, we leverage the clustering method to reduce data redundancy and improve the accuracy of high-quality user aggregation. We evaluate the data quality and transform the user clustering problem into

the maximum similarity matching problem, which can be expressed as follows:

$$\text{maximize sim} = \sum_{U, T} \text{sim}_{u_i, t_{j\text{-center}}} = \sum_{U, T} \sum_{u_i \in U, t_{j\text{-center}} \in T} f(u_i, t_{j\text{-center}}) \quad (1)$$

The goal of Equation (1) is to form a union with the highest user similarity from large-scale participating users, so as to reduce data redundancy and improve the efficiency of the sensing platform. $f(u_i, t_{j\text{-center}})$ represents the similarity between u_i and $t_{j\text{-center}}$, which can be expressed as follows:

$$\text{sim}_{u_i, t_{j\text{-center}}} = f(u_i, t_{j\text{-center}}) = \frac{1}{1 + \sqrt{\sum_{a=1}^n \omega_a * (c_{a, u_m} - c_{a, t_{j\text{-center}}})^2}}, \quad (2)$$

where c_{a, u_m} represents the value of u_m under evaluation index a and $c_{a, t_{j\text{-center}}}$ represents the value of $t_{j\text{-center}}$ under evaluation index a .

(2) *Method Construction.* Step 1: Define the user similarity as a two-tuple.

Definition 1. Define the user similarity as $\text{sim}_{u_i, t_{j\text{-center}}} = f(u_i, t_{j\text{-center}})$.

$f(u_i, t_{j\text{-center}})$ is the participants of both clustering data, where u_i denotes the user, and $t_{j\text{-center}}$ denotes the center of task t_j .

Step 2: Calculate the similarity between u_i and $t_{j\text{-center}}$, sort the calculation results, and construct the user-union clustering.

Step 3: Set a maximum user limit τ in each user-union to ensure the balance of the union, i.e., $\|\text{sim}_{u_i, t_{j\text{-center}}}\| \leq \tau$.

3.2.2. Sensing Cost Problem

(1) *Sensing Cost Problem Formulation.* In Section 3.2.1, we use user data to build user-unions, which realize user clustering, reduce data scale and ensure data quality. Based on this solution, we consider the diversity and individual preferences for service providers, and solve the diversified task assignment under the individual preference sensing cost.

For each task assignment problem, each user-union has n sets of user schemes and m sets of data sensing cost evaluation indexes, denoted by $Y = \{Y_1, Y_2, \dots, Y_n\}$ and $G = \{G_1, G_2, \dots, G_m\}$. Each user scheme represents a sensing cost requirement, which can be evaluated by the sensing cost indexes, denoted by $\{h_{11}, h_{12}, \dots, h_{1m}\}$. According to the decision selection sample matrix, service providers select the appropriate user to realize task assignment. The decision selection sample matrix is expressed as follows:

$$H = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1m} \\ h_{21} & h_{22} & \dots & h_{2m} \\ \vdots & \vdots & & \vdots \\ h_{n1} & h_{n2} & \dots & h_{nm} \end{bmatrix}, \quad (3)$$

Based on the above conditions, we normalize the decision information matrix, and use the prospect theory [28] to obtain the positive and negative prospect value matrix. Finally, the acceptability advantage solution is used to sort the schemes and select the most suitable users. Therefore, we transform the preference-based sensing cost problem into the maximum comprehensive prospect value, which can be expressed as follows:

$$\begin{aligned} \max V & \sum_{i=1}^n \sum_{j=1}^m v_{ij}^+ \psi^+(\omega_j) + \sum_{i=1}^n \sum_{j=1}^m v_{ij}^- \psi^-(\omega_j) \\ \text{s.t.} & \omega_j \in [0, 1] \\ & \sum_{j=1}^m \omega_j = 1 \end{aligned} \quad (4)$$

The goal of Equation (4) is to solve the maximum comprehensive prospect value, so as to achieve the preference-based task assignment. The objective function in the first line is to solve the optimal evaluation index weight. The second and third lines define the range of each index, respectively.

(2) *Method Construction.* Step 1: Normalize the decision matrix of user scheme. We define the user sensing costs as the cost index and the benefit index, denoted by h_{ij}^b and h_{ij}^c

, $h_{ij}^b, h_{ij}^c \in H$, $h_{ij}^b \cup h_{ij}^c = H$, and $h_{ij}^b \cap h_{ij}^c = \emptyset$, which can be expressed as:

$$z_j = \frac{1}{n} \sum_{i=1}^n h_{ij}, \quad (5)$$

$$h_{ij}^b = \frac{h_{ij} - z_j}{\max \left\{ \max_j (h_{ij}) - z_j, z_j - \min_j (h_{ij}) \right\}}, \quad (6)$$

$$h_{ij}^c = \frac{z_j - h_{ij}}{\max \left\{ \max_j (h_{ij}) - z_j, z_j - \min_j (h_{ij}) \right\}}. \quad (7)$$

Step 2: Determine the positive and the negative prospect value matrix. The normalized decision matrix is recorded as $\bar{O} = (\bar{h}_{ij})_{n \times m}$. We construct the positive and the negative prospect value matrix, which can be expressed as:

$$\begin{cases} h_j^+ = \max \{h_{ij} \mid 1 \leq i \leq n\} \\ h_j^- = \min \{h_{ij} \mid 1 \leq i \leq n\} \\ j = 1, 2, \dots, m \end{cases}, \quad (8)$$

where $Y^+ = \{h_1^+, h_2^+, \dots, h_m^+\}$ and $Y^- = \{h_1^-, h_2^-, \dots, h_m^-\}$ represent the positive and the negative ideal scheme, respectively.

Step 3: Calculate the correlation coefficient. A proper task assignment usually needs a reference node to measure the prospect value of the scheme, rather than the actual value of the decision result. Therefore, we use the values of positive and negative ideal schemes as reference points, which can be expressed as:

$$\begin{cases} \zeta_{ij}^+ = \frac{\min_{i,j} |h_{ij} - h_j^+| + \varphi \max_{i,j} |h_{ij} - h_j^+|}{|h_{ij} - h_j^+| + \varphi \max_{i,j} |h_{ij} - h_j^+|} \\ \zeta_{ij}^- = \frac{\min_{i,j} |h_{ij} - h_j^-| + \varphi \max_{i,j} |h_{ij} - h_j^-|}{|h_{ij} - h_j^-| + \varphi \max_{i,j} |h_{ij} - h_j^-|} \end{cases}, \quad (9)$$

where $\zeta_{ij}^+, \zeta_{ij}^-$ represent the positive and the negative correlation coefficients, respectively, $0 \leq \zeta_{ij}^+ \leq 1$ and $-1 \leq \zeta_{ij}^- < 0$, φ represents the resolution coefficient, define $\varphi = 0.5$.

Step 4: Construct prospect decision matrix. We construct a prospect value function to represent the subjective feelings of service providers about the user scheme selection, which can be expressed as:

$$v(h_i) = \begin{cases} (1 - \zeta_{ij}^-)^\alpha, & \zeta_{ij}^- \text{ is a reference point} \\ -\lambda \left[-(\zeta_{ij}^+ - 1) \right]^\beta, & \zeta_{ij}^+ \text{ is a reference point} \end{cases}, \quad (10)$$

where α and β represent the concave and the convex degree of the benefit and the cost value functions at the reference point, respectively, $0 < \alpha < 1$, $0 < \beta < 1$. λ represents the degree of loss aversion of the service provider.

According to Equation (10), we achieve the positive and the negative values of Y_i , which expressed as:

$$\begin{cases} v^+(h_{ij}) = (1 - \varsigma_{ij}^-)^1 \\ v^-(h_{ij}) = -\lambda \left[-(\varsigma_{ij}^+ - 1) \right]^\beta \end{cases} \quad (11)$$

Probability weight is the subjective judgment made by the service provider according to the probability ω of the result of the task assignment, which can be expressed as:

$$\begin{cases} \psi(\omega)^+ = \frac{\omega^\eta}{[\omega^\eta + (1 - \omega)^\eta]^{1/\eta}}, \varsigma_{ij}^- \text{ is a reference point} \\ \psi(\omega)^- = \frac{\omega^\gamma}{[\omega^\gamma + (1 - \omega)^\gamma]^{1/\gamma}}, \varsigma_{ij}^+ \text{ is a reference point} \end{cases}, \quad (12)$$

where $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\eta = 0.61$, $\gamma = 0.69$ [28], η and γ represent the fitting parameters of the probability weight function on the left and right of the reference point, respectively.

We calculate the comprehensive prospect value of each user scheme, which can be expressed as:

$$V_i = \sum_{j=1}^m v_{ij}^+ \psi^+(\omega_j) + \sum_{j=1}^m v_{ij}^- \psi^-(\omega_j). \quad (13)$$

Step 5: Weight optimization. The weight of the user scheme should be reasonably assigned, aiming to obtain the maximum comprehensive prospect value, which can be expressed as:

$$V_i^* = \sum_{j=1}^m v_{ij}^+ \psi^+(\omega_j^*) + \sum_{j=1}^m v_{ij}^- \psi^-(\omega_j^*). \quad (14)$$

The multi-attribute hesitant fuzzy evaluation matrix is transformed into the multi-attribute comprehensive prospect matrix.

Step 6: Sort user schemes to determine the preference-based task assignment.

According to the comprehensive prospect matrix, we calculate the positive (i.e., f^+) and the negative (i.e., f^-) ideal solutions of each index, which can be expressed as:

$$\begin{cases} f^+ = \left\{ \max_i v_i(h_{i1}), \max_i v_i(h_{i2}), \dots, \max_i v_i(h_{im}) \right\} \\ f^- = \left\{ \min_i v_i(h_{i1}), \min_i v_i(h_{i2}), \dots, \min_i v_i(h_{im}) \right\} \end{cases} \quad (15)$$

We also need to calculate the group benefit value (i.e., B_i

), individual regret value (i.e., R_i), and comprehensive index value (i.e., BR_i).

$$\begin{cases} B_i = \sum_{j=1}^m \omega_j (f^+ - v(h_{ij})) / (f^+ - f^-) \\ R_i = \max_j \{ \omega_j (f^+ - v(h_{ij})) / (f^+ - f^-) \} \end{cases}, \quad (16)$$

$$BR_i = \kappa \frac{B_i - B_i^-}{B_i^+ - B_i^-} + (1 - \kappa) \frac{R_i - R_i^-}{R_i^+ - R_i^-}, \quad (17)$$

where B_i^+ , B_i^- represent the maximum and minimum group benefit value, R_i^+ , R_i^- represent the maximum and minimum individual regret value, and κ represents the decision preference. When $\kappa > 0.5$, it means that the service provider adopts the maximum group benefit to formulate the task assignment scheme. When $\kappa < 0.5$, it means that the service provider adopts the minimum individual regret to formulate the task assignment scheme. When $\kappa = 0.5$, it means that the service provider adopts the balance principle to formulate the task assignment scheme.

According to the judgment criteria of the VIKOR method [29], the value of B_i , R_i , and BR_i are arranged in descending order. We use BR_i to determine the first (i.e., Y_1) and second (i.e., Y_2) user schemes and realize the preference-based task assignment.

Condition 1 (Acceptability advantage). $BR(Y_2) - BR(Y_1) \geq 1/(n-1)$, where m is the number of options.

Condition 2 (Acceptability stable). Y_1 has the best B_i or R_i .

When Condition 1 and Condition 2 are both satisfied, Y_1 is the optimal user scheme to realize task assignment. When only Condition 1 is satisfied, Y_1 and Y_2 are compromise solutions. When only Condition 2 is satisfied, Y_1, Y_2, \dots, Y_N are approximate ideal schemes.

4. Proposed Task Assignment Solutions

4.1. User-Union Clustering Algorithm. Traditional clustering algorithms are deficient in the efficiency and balance of clustering results. To solve this issue, we propose the user-union clustering algorithm (UCA), as shown in Algorithm 1.

Algorithm 1 realizes the generation of user-union. UCA provides a guarantee for the balance of clustering data by setting an upper limit. The function of ProperCluster (x_j) is to assign u_i to a suitable user-union. CS_j is a two-tuple, which represents the storage of existing user data and the similarity value of the center task in the j th user-union. From 1 to 4, the algorithm is used to calculate the similarity between u_i and $t_{j-\text{center}}$, which aims to quantify the behavioral characteristics of each user. From 5 to 18, the algorithm is used to control the scale of users, which can balance the number of users in the user-union.

Computational complexity. The k -means algorithm is a simple and efficient clustering algorithm, and the computational complexity of the algorithm is $O_2(tkmn)$, where t

```

Input: User data  $U = \{u_1, u_2, \dots, u_m\}$ , task set  $T = \{t_1, t_2, \dots, t_k\}$ , maximum user limit  $\tau$ 
Output: Set of  $K$  task clusters  $tc = \{tc_1, tc_2, \dots, tc_k\}$ 
1: ProperCluster( $x_i$ )
2: Determine the center of the initial task sets and user data evaluation indexes ( $C_{u_i} = (c_{1,u_i}, c_{2,u_i}, \dots, c_{n,u_i})$ )
3: Calculate the user similarity by Equation (1), and sort data in descending order  $STC = \{\text{sim}_{u_i, t_{j-\text{center}}} \mid j = 1, 2, \dots, K\}$ 
4:   for  $j \leftarrow 1$  to  $K$  DO
5:     if  $\|\text{sim}_{u_i, t_{j-\text{center}}}\| < \tau$ 
6:        $u_i$  enter  $tc_j$ 
7:        $u_i$  and  $\text{sim}_{u_i, t_{j-\text{center}}}$  are saved in  $CS_j = \{\{u_i, \text{sim}_{u_i, t_{j-\text{center}}}\} \dots\}$ 
8:       break
9:     else
10:      if  $\text{sim}_{u_i, t_{j-\text{center}}}$  is less than the minimum similarity value in  $CS_j$ 
11:        continue
12:      else
13:         $u_i$  joins the  $j$ th union and deletes edge user ( $u_e$ )
14:        ProperCluster( $u_e$ )
15: repeat
16:   for  $i \leftarrow 1$  to  $N$  DO
17:     ProperCluster( $x_i$ )
18: until saturate task requirements or reach the maximum number of iterations
19: End

```

ALGORITHM 1: User-union clustering algorithm (UCA).

represents the number of iterations, m represents the user scale, n represents the type of user data evaluation index, and k represents the number of clustering tasks. UCA is an improvement of the k -means algorithm, which uses user similarity to realize user clustering and improves the balance of user scale. First, each user needs to calculate the similarity with k , and the complexity is kmn . Next, the value of user similarity is compared with the edge point, when the number of users in the user-union reaches saturation, and the complexity is 1. In the worst case, UCA spends k times for comparison. Therefore, the computational complexity of UCA is $O_1(k^2mn)$. In practical application scenarios, to ensure the clustering accuracy, the number of algorithm iterations (i.e., t) is usually greater than clustering tasks (i.e., k); therefore, $O_2 - O_1 = (t - k)kmn > 0$.

Space complexity. The k -means algorithm needs to store user data and the clustering tasks data, and the space complexity of is $(k + m)n$. Like the k -means algorithm, UCA also needs to store user data and clustering task data, and the space complexity is $(k + m)n$.

4.2. Preference-Based Task Assignment Algorithm. Algorithm 1 provides high-quality data. Then, we propose the preference-based task assignment algorithm (PTA) to solve the diversified task assignment under the individual preference sensing cost, as shown in Algorithm 2.

Algorithm 2 realizes the reasonable and diverse task assignment by calculating the value of group benefit, individual regret, and a comprehensive index. This is a mode of task assignment selection from individual preference, which guides the decision of service providers. From 1 to 2, the algorithm is used to normalize the decision matrix. From 3 to 4, the algorithm is used to determine the positive and negative ideal solutions and calculates the correlation

coefficient. From 5 to 6, the algorithm mainly constructs the prospect decision matrix through optimized weights. From 7 to 19, the algorithm is used to sort user schemes, and achieve preference-based task assignment based on the VIKOR method.

5. Performance Evaluation and Discussion

5.1. Basic Simulation Setup. In our experiments, the data we used came from the real Dartmouth College Wi-Fi campus trace data set [30], which was an experiment on the open-source middleware NSense. This data takes sound collection as an example, including timestamps, the distance between test points and sensing nodes, data collection methods, and data collection environments. We define data collection methods and environments as benefit indexes. Other metrics are defined as cost indexes. We consider two user distribution spaces [31] (i.e., sparse and dense regions) and employ different metrics to measure the performance in UCA and PTA.

In the research of task assignments, high-quality user data can improve the accuracy of assignments. User data clustering can reduce data redundancy and improve the overall quality of user data. Therefore, we first verify the performance of UCA. We compare the performance with three common clustering algorithms [14] (i.e., K -means, K -means improve, and fuzzy C -means clustering algorithm) by calculating the accuracy (ACC), normalized mutual information (NMI), and running time. The K -means improve algorithm limits the number of users of the K -means, aiming to control the balance of user distribution scale.

ACC is used to measure the accuracy of the users' classification after clustering, and compared to the actual

Input: Decision sample matrix H
Output: Optimal task assignment

1: **Initialization**
2: Normalize the sample matrix by Equations (5)–(7)
3: Determine $Y^+ = \{h_1^+, h_2^+, \dots, h_m^+\}$ and $Y^- = \{h_1^-, h_2^-, \dots, h_m^-\}$ by Equation (8)
4: Calculate the correlation coefficient by Equation (9)
5: Build a prospective decision matrix and calculate the prospective value
6: Optimize the index weights to obtain the best comprehensive prospect value by Equations (10)–(14)
7: Calculate B_i , R_i , and BR_i by Equations (15)–(17), confirm the first and second value of BR_i (i.e., Y_1 and Y_2)
8: for Y_1 and Y_2 do
9: if only meet Condition 1 then
10: Y_1 and Y_2 are compromise solutions
11: end if
12: else if only meet Condition 2 then
13: Calculate the largest N by $BR(Y_2) - BR(Y_1) < 1/(N - 1)$, and Y_1, Y_2, \dots, Y_N are approximate ideal schemes
14: end if
15: if both meet Conditions 1 and 2 then
16: Y_1 is the optimal solution
17: end if
18: **end for**

ALGORITHM 2: Preference-based task assignment algorithm (PTA).

classification in the prior knowledge, which can be expressed as:

$$ACC = \frac{\sum_{i=1}^N \nu(s_i, \text{map}(r_i))}{N}, \quad (18)$$

where N is the number of users, map is a mapping function that maps the classification of the clustering results to the original data set, s_i is the original classification of user data in prior knowledge. When $s_i = \text{map}(r_i)$, the value of ν is 1. Otherwise, the value of ν is 0.

NMI is used to evaluate the similarity between the clustering results and the distribution of the original dataset, which can be expressed as:

$$NMI(X, Y) = \frac{I(X, Y)}{\sqrt{H(X)H(Y)}}, \quad (19)$$

where $I(X, Y)$ represents mutual information between X and Y . $H(X)$ and $H(Y)$ represent the information entropy of distributions X and Y , respectively.

Next, we verify the performance of PTA on the premise of obtaining high-quality user data, which aims to realize preference-based task assignment under the individual preference sensing cost. We compare the performance with two methods (i.e., VIKOR [29] and TOPSIS [32]) by calculating the compatibility degree and execution time. The VIKOR method determines the optimal task assignment scheme without the prospect value. The TOPSIS method is a common method to solve the ideal point.

Compatibility degree [33] is used to verify the rationality of task assignment, which can be expressed as:

$$\begin{cases} \text{compd}_{\text{met}i} = \frac{1}{l-1} \sum_{\text{met}j=2}^l p_{\text{met}i, \text{met}j} \\ p_{\text{met}i, \text{met}j} = 1 - \frac{6}{m(m^2-1)} \sum_{d=1}^m f_d^2 \end{cases}, \quad (20)$$

where $\text{compd}_{\text{met}i}$ represents the compatibility of the i th method, $p_{\text{met}i, \text{met}j}$ represents the degree of correlation between i and j , m represents the number of schemes, and f_d represents the sorting difference of the d th scheme in i and j .

5.2. Experiment Results of UCA. It is meaningless to use UCA to reduce data redundancy for the small scale of users in remote regions. Therefore, for the experiment of UCA, we analyze the clustering effect of large-scale users, when the number of users varies from 100 to 1000, respectively. Figures 2–4 show the performance in terms of ACC, NMI, and running time, achieved by the four algorithms.

Apparently, UCA outperforms the three baselines (i.e., higher ACC, higher NMI, and lower running time), no matter how the number of users varies. In Figure 2, the augmentation of user data decreases the accuracy of all clustering algorithms. The reason is that the increase of user scale leads to the rise in low-similarity users, which reduces the clustering accuracy. The accuracy of UCA is better than these three algorithms, and the ACC is basically above 0.82. Compared with the best performance K -means algorithm, the accuracy is improved by about 10%. The reason is that UCA calculates user similarity and sets boundary user replacement

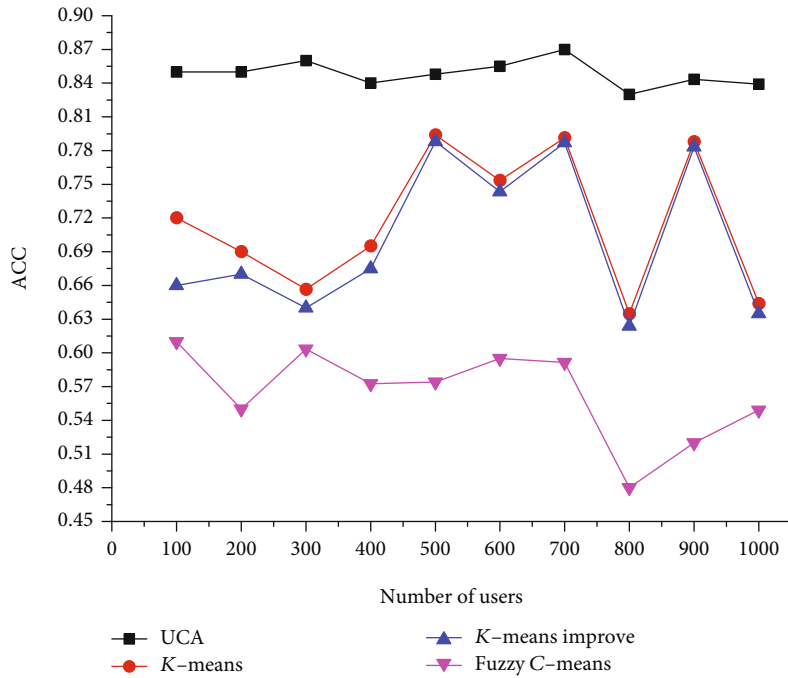


FIGURE 2: ACC vs. number of users.

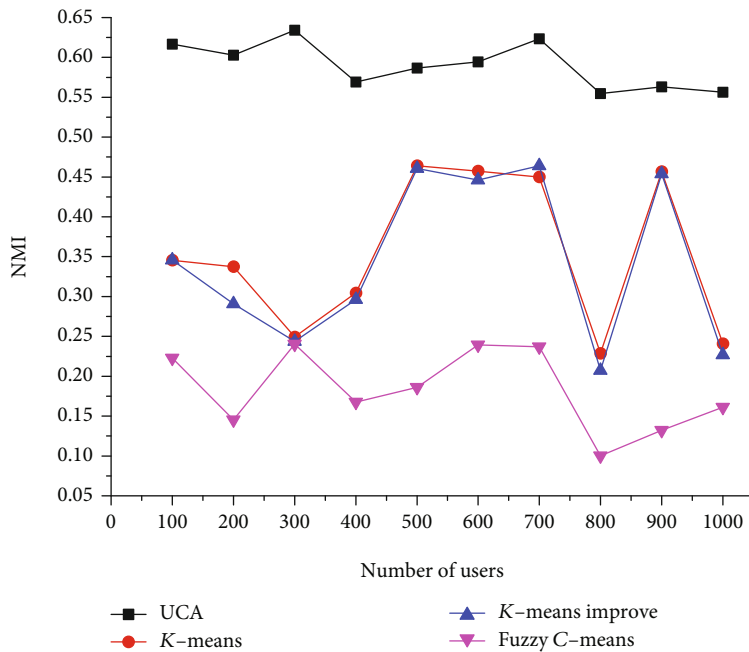


FIGURE 3: NMI vs. number of users.

rules, which can balance the number of users in different unions and ensure that users with high similarity are clustered together as much as possible. In addition, the accuracy of the *K-means improve* is lower than the *K-means* algorithm. This means that a single restriction on the size of users is not conducive to the formation of high-quality user clustering.

We also perform extensive simulations to validate the reduction of running time achieved by UCA under various user scales, as shown in Figure 5. As seen, the results of the four algorithms show an upward trend, and UCA has the lowest running time. The reasons are as follows: Firstly, the *K-means* algorithm uses random clustering centers to achieve user clustering through multiple iterations. The

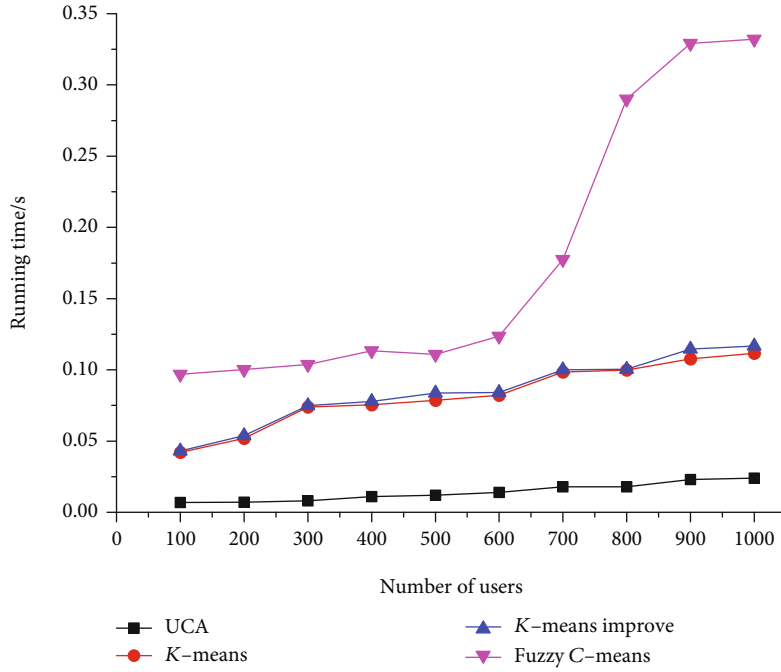


FIGURE 4: Running time vs. number of users.

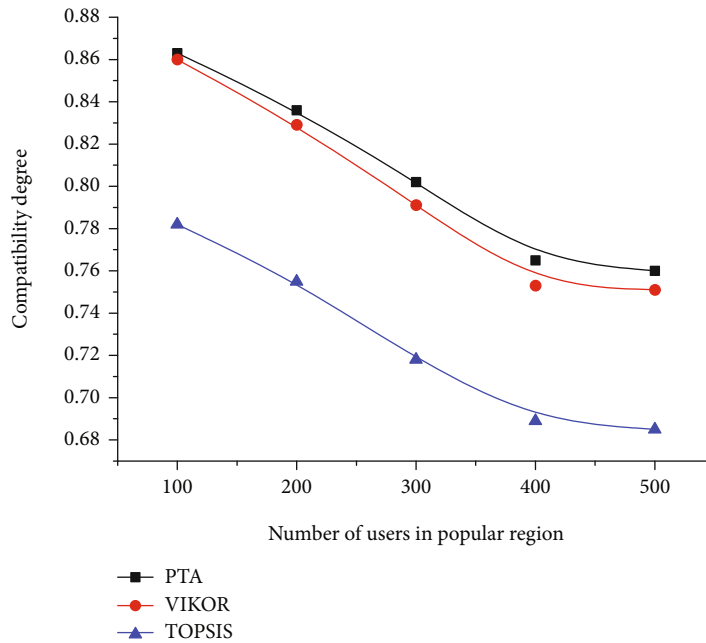


FIGURE 5: Compatibility degree vs. popular region.

growth of the user scale leads to more iterations and time overhead. In addition, setting user boundaries in this algorithm may cause more time costs. Unlike the *K-means* algorithm, *UCA* only needs to calculate the similarity between users and task centers, and compare boundary users to realize the user-union, which can reduce running time. Secondly, the *Fuzzy C-means* algorithm provides more flexible clustering results, but is more sensitive to boundary users.

With the growth of user scale, the existence of enormous boundary users will require a longer time overhead for this algorithm.

In general, for the large-scale user clustering scenario, the proposed user-union clustering algorithm has the characteristics of high classification accuracy and fast calculation speed. It can provide high-quality user data for the preference-based task assignment.

TABLE 1: Decision matrix of task assignment scheme.

Scheme	G_1	G_2	G_3	G_4
Y_1	35	1521	0.69	0.68
Y_2	29	1487	0.78	0.67
Y_3	30	1495	0.77	0.72
Y_4	25	1503	0.72	0.74
Y_5	32	1508	0.74	0.69

TABLE 2: The value of B_i , R_i , and BR_i .

Scheme	B_i	R_i	BR_i
Y_1	0.984	0.3	1
Y_2	0.261	0.161	0
Y_3	0.413	0.183	0.184
Y_4	0.393	0.216	0.289
Y_5	0.673	0.222	0.504

5.3. Experiment Results of PTA. On the premise of ensuring high-quality clustering users, we perform the performance of PTA to realize the diversified task assignment under the individual preference sensing cost. In addition, different user scales may have various user characteristics, which affect the performance of execution time and compatibility. As a result, we first use an example to demonstrate the feasibility of small-scale data. Then, we consider PTA performance in two scenarios by calculating execution time and compatibility degree.

5.3.1. Example. According to UCA, we achieve five alternative task assignment schemes (i.e., the sensing cost for five users), as shown in Table 1.

Step 1. Normalize the sample matrix by Eqs. (5)–(7), G_1 and G_2 are the cost index, G_3 and G_4 are the benefit index. The decision selection sample matrix is

$$H = \begin{bmatrix} -1 & -1 & -1 & -0.5 \\ 0.2 & 0.619 & 0.5 & -0.75 \\ 0 & 0.238 & 0.333 & 0.5 \\ 1 & -0.143 & -0.5 & 1 \\ -0.4 & -0.381 & -0.167 & -0.25 \end{bmatrix}. \quad (21)$$

Then, we achieve the positive ideal assignment scheme (i.e., $Y^+ = \{1, 0.619, 0.5, 1\}$), and the negative ideal assignment scheme (i.e., $Y^- = \{-1, -1, -1, -0.75\}$).

Step 2. According to Equation (8), the correlation coefficient of positive and negative ideal scheme is

$$\zeta^+ = \begin{bmatrix} 0.333 & 0.333 & 0.333 & 0.368 \\ 0.556 & 1 & 1 & 0.333 \\ 0.500 & 0.680 & 0.818 & 0.636 \\ 1 & 0.515 & 0.429 & 1 \\ 0.417 & 0.447 & 0.529 & 0.412 \end{bmatrix}, \quad (22)$$

$$\zeta^- = \begin{bmatrix} 1 & 1 & 1 & 0.778 \\ 0.455 & 0.333 & 0.333 & 1 \\ 0.500 & 0.395 & 0.360 & 0.412 \\ 0.333 & 0.486 & 0.600 & 0.333 \\ 0.625 & 0.567 & 0.474 & 0.636 \end{bmatrix},$$

respectively.

Step 3. According to Equation (11), the positive and negative prospect value matrix of each scheme is

$$v^+ = \begin{bmatrix} 0 & 0 & 0 & 0.266 \\ 0.586 & 0.700 & 0.700 & 0 \\ 0.543 & 0.643 & 0.675 & 0.627 \\ 0.700 & 0.557 & 0.446 & 0.700 \\ 0.422 & 0.479 & 0.568 & 0.411 \end{bmatrix}, \quad (23)$$

$$v^- = \begin{bmatrix} -1.575 & -1.575 & -1.575 & -1.503 \\ -1.101 & 0 & 0 & -1.575 \\ -1.223 & -0.826 & -0.502 & -0.925 \\ 0 & -1.190 & -1.374 & 0 \\ -1.399 & -1.336 & -1.159 & -1.410 \end{bmatrix},$$

respectively.

Step 4. Optimize the index weights to obtain the best comprehensive prospect value by Equation (4), where $\omega_1, \omega_2, \omega_3, \omega_4 \in [0.1, 0.3]$. We achieve the optimal solution (i.e., $\omega^* = \{0.3, 0.3, 0.3, 0.1\}$) and the comprehensive prospect matrix is

$$v^* = \begin{bmatrix} -0.516 & -0.516 & -0.516 & -0.206 \\ -0.174 & 0.223 & 0.223 & -0.268 \\ -0.228 & -0.066 & 0.049 & -0.041 \\ 0.223 & -0.212 & -0.308 & 0.130 \\ -0.324 & -0.285 & -0.199 & -0.163 \end{bmatrix}. \quad (24)$$

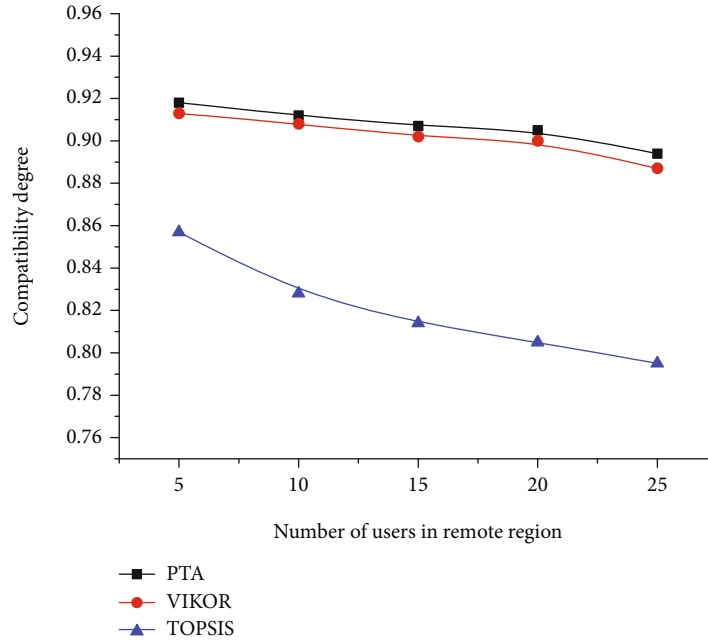


FIGURE 6: Compatibility degree vs. remote region.

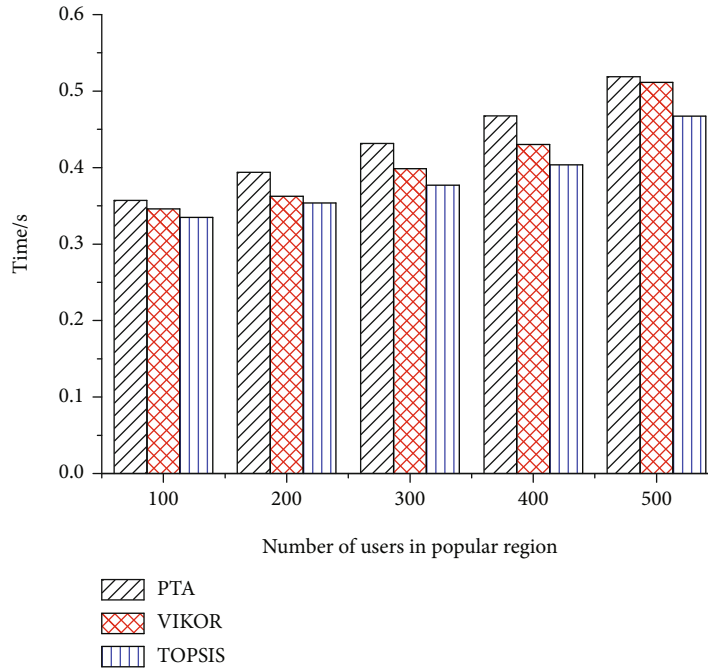


FIGURE 7: Execution time vs. popular region.

Step 5. Use the VIKOR method to sort the schemes. Calculate the value of B_i , R_i , and BR_i , as shown in Table 2.

Table 2 presents the value of B_i , R_i , and BR_i for the five users. As seen, Y_2 has the optimal value of BR_i and B_i , which satisfies Condition 2. Y_3 has the sub-optimal value of BR_i , and $BR_3 - BR_2 = 0.184 < 1/(5 - 1)$, which does not satisfy Condition 1. Obviously, Y_2 and Y_3 are both acceptable and ideal solutions. The service provider can choose Y_2 or Y_3 according to individual preference, and PTA implements

the preference-based task assignment. In addition, we also found an interesting phenomenon that PTA usually chooses low-cost and high-quality schemes. The reason is as follows. Firstly, PTA solves the diversified task assignment under the individual preference sensing cost. That is, service providers play a decisive role in the task assignment. Considering the profit orientation of service providers, low-cost and high-quality schemes are more competitive in selection. Secondly, sensing users are competitive and work hard. Users try to

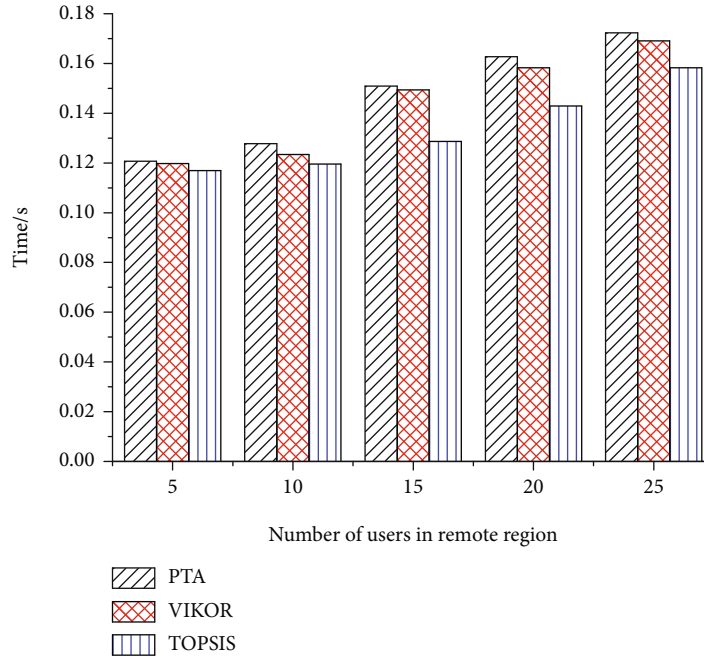


FIGURE 8: Execution time vs. remote region.

improve the quality of uploaded data to win in a task as much as possible. Thirdly, in the calculation results of the best comprehensive prospect value, the weight of the cost index is much greater than the benefit index, which further promotes PTA to choose low-cost and high-quality user solutions.

5.3.2. Performance Comparison. Next, we provide simulation results by three methods in various scenarios.

(1) Compatibility Degree. According to our definition of location regions [31], we conduct simulations to observe the effect of compatibility degree on different solutions when users are in different regions (i.e., popular region and remote region), as shown in Figures 4 and 6.

Generally, high compatibility degree means that the user data is representative and reliable, which means the higher accuracy of task assignment. Figures 4 and 6 show the performance of compatibility degrees under different numbers of users and regions. It is seen that as the number of users increases, the compatibility degree of these methods decreases. The compatibility degree of PTA is better than these two methods, and the compatibility degree in the remote region is better than in the popular region. The reason is as follows. First, as the number of users increases, more similar users participate in sensing tasks, especially in a popular region, which reduces the differences between users. As a result, the sensing platform is challenging to select suitable users, which leads to a decrease in the compatibility of these solutions. Second, compared with the VIKOR method, PTA adds prospect theory to reflect that decision-makers are more sensitive to losses than revenues. Poor indexes are more difficult to compensate by superior

indexes, and the selected user data is more balanced to ensure the accuracy of task assignment. Third, compared with the TOPSIS method, PTA does not need to satisfy both the optimal positive ideal solution and the worst negative ideal solution. The final selection meets the individual preferences of the service provider. Furthermore, we also found that the performance of the three methods in remote regions is better than in popular regions. The reason is that large-scale users in popular regions lead to the high similarity between data, which makes it difficult to assign tasks accurately. On the contrary, the small scale of users in remote regions is conducive to accurate task assignment.

(2) Execution Time. We perform extensive simulations to validate the execution efficiency of PTA under various regions, compared with two solutions, as shown in Figures 7 and 8.

Figures 7 and 8 both show the execution time of PTA under various regions. We find that the execution time of the three solutions increases stably when the number of users enlarges. More alternative users in the sensing platform lead to more computational overhead. In addition, the execution time in the popular region is generally higher than that in the remote region. The reason is that more similar users are contained in the popular region, and more calculations are needed to find suitable candidate users. PTA is slightly worse than VIKOR and TOPSIS methods in execution time, because PTA makes user selection from multiple perspectives, which increases the execution time.

In general, the performance of PTA is acceptable in the preference-based task assignment. The reason is as follows. Firstly, PTA has a more significant advantage in the accuracy of user selection (i.e., the highest compatibility degree),

which can ensure the accuracy of task assignment. Besides, as another part of the task assignment method, UCA has the characteristics of high classification accuracy and fast calculation speed, which can make up for the lack of execution time of PTA.

6. Conclusions

In this paper, we addressed a task assignment problem in MCS. We proposed a task assignment method based on user-union clustering and individual preferences. Specifically, we analyzed and formulated the task assignment problem from two perspectives, respectively. We first define the user similarity and propose a user-union clustering algorithm (UCA) to reduce data redundancy and achieve high-quality clustering data. Based on this solution, we further consider individual preferences of service providers and propose a preference-based task assignment algorithm (PTA) to meet the needs of diversified sensing cost and achieve the task assignment with individual preference. To evaluate the performance of the proposed solutions, we conducted extensive simulations. The results show that our method realizes the individual preference-based task assignment under the premise of ensuring high-quality clustering data. However, our method usually chooses low-cost and high-quality user data, which may suppress the revenues of users. At the same time, for the user-union, using exact values to evaluate data may reduce the accuracy of evaluation. In future works, we will balance the revenues between users and service providers, improve the accuracy of clustering data quality evaluation, and develop a task assignment method with lower complexities.

Data Availability

The authors declare that all the data and materials in this manuscript are available.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was funded by the Nature Science Foundation of China, grant number 61872104, and Fundamental Research Fund for the Central Universities in China, grant number 3072020CF0603.

References

- [1] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.
- [2] W. Gong, B. Zhang, and C. Li, "Task assignment in mobile crowdsensing: present and future directions," *IEEE Network*, vol. 32, no. 4, pp. 100–107, 2018.
- [3] A. Shahraki, A. Taherkordi, Ø. Haugen, and F. Eliassen, "Clustering objectives in wireless sensor networks: a survey and research direction analysis," *Computer Networks*, vol. 180, article 107376, 2020.
- [4] W. Liu, Y. Yang, E. Wang, and J. Wu, "User recruitment for enhancing data inference accuracy in sparse mobile crowdsensing," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1802–1814, 2020.
- [5] F. Campioni, S. Choudhury, K. Salomaa, and S. G. Akl, "Improved recruitment algorithms for vehicular crowdsensing networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1198–1207, 2019.
- [6] Y. Zhao and C. H. Liu, "Social-aware incentive mechanism for vehicular crowdsensing by deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2314–2325, 2021.
- [7] W. Gong, B. Zhang, C. Li, and Z. Yao, "Task allocation in semi-opportunistic mobile crowdsensing: paradigm and algorithms," *Mobile Networks and Applications*, vol. 25, no. 2, pp. 772–782, 2020.
- [8] X. H. Wei, Z. J. Li, Y. Y. Liu, S. Gao, and H. S. Yue, "SDLSC-TA: subarea division learning based task allocation in sparse mobile crowdsensing," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1344–1358, 2021.
- [9] C. Wang, Z. Xie, L. Shao, Z. Zhang, and M. Zhou, "Estimating travel speed of a road section through sparse crowdsensing data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 9, pp. 3486–3495, 2019.
- [10] E. Wang, M. Zhang, X. Cheng et al., "Deep learning-enabled sparse industrial crowdsensing and prediction," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 6170–6181, 2021.
- [11] E. Wang, Y. Yang, J. Wu, W. Liu, and X. Wang, "An efficient prediction-based user recruitment for mobile crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 16–28, 2018.
- [12] H. Xiong, D. Zhang, L. Wang, J. P. Gibson, and J. Zhu, "EEMC," *ACM transactions on intelligent systems and technology*, vol. 6, no. 3, pp. 1–26, 2015.
- [13] J. T. Wang, Y. S. Wang, D. Q. Zhang, F. Wang, Y. D. He, and L. T. Ma, "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*, pp. 1139–1151, Portland, Oregon, USA, 2017.
- [14] A. Pugazhenthil and L. S. Kumar, "Selection of Optimal Number of Clusters and Centroids for K-means and Fuzzy C-means Clustering: A Review," in *2020 5th International Conference on Computing, Communication and Security (ICCCS)*, pp. 1–4, Patna, India, 2020.
- [15] W. Gong, B. Zhang, and C. Li, "Location-based online task scheduling in mobile crowdsensing," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, Singapore, 2017.
- [16] H. Xiong, D. Zhang, G. Chen, L. Wang, V. Gauthier, and L. E. Barnes, "iCrowd: near-optimal task allocation for piggyback crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 15, no. 8, pp. 2010–2022, 2016.
- [17] Z. B. Wang, R. Tan, J. H. Hu et al., "Heterogeneous incentive mechanism for time-sensitive and location-dependent crowdsensing networks with random arrivals," *Computer Networks*, vol. 131, pp. 96–109, 2018.

- [18] B. Zeng, X. Yan, X. Zhang, and B. Zhao, "BRAKE: bilateral privacy-preserving and accurate task assignment in fog-assisted mobile crowdsensing," *IEEE Systems Journal*, vol. 15, no. 3, pp. 4480–4491, 2021.
- [19] Y. Huang, H. Chen, G. Ma et al., "OPAT: optimized allocation of time-dependent tasks for mobile crowdsensing," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2476–2485, 2022.
- [20] A. Alagha, R. Mizouni, S. Singh, H. Otrok, and A. Ouali, "SDRS: a stable data-based recruitment system in IoT crowdsensing for localization tasks," *Journal of Network and Computer Applications*, vol. 177, article 102968, 2021.
- [21] S. Song, Z. Liu, Z. Li, T. Xing, and D. Fang, "Coverage-oriented task assignment for mobile crowdsensing," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7407–7418, 2020.
- [22] F. Yucel, M. Yuksel, and E. Bulut, "Coverage-aware stable task assignment in opportunistic mobile crowdsensing," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3831–3845, 2021.
- [23] W. Z. Guo, W. P. Zhu, Z. Y. Yu, J. T. Wang, and B. Guo, "A survey of task allocation: contrastive perspectives from wireless sensor networks and mobile crowdsensing," *IEEE Access*, vol. 7, pp. 78406–78420, 2019.
- [24] Y. Du, Y. E. Sun, H. Huang et al., "Bayesian co-clustering truth discovery for mobile crowd sensing systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1045–1057, 2020.
- [25] H. Jin, L. Su, H. Xiao, and K. Nahrstedt, "Incentive mechanism for privacy-aware data aggregation in mobile crowd sensing systems," *IEEE/ACM Transactions on Networking*, vol. 26, no. 5, pp. 2019–2032, 2018.
- [26] Z. H. Shao, H. Q. Wang, and G. S. Feng, "PUEGM: a method of user revenue selection based on a publisher-user evolutionary game model for mobile crowdsensing," *Sensors*, vol. 19, no. 13, p. 2927, 2019.
- [27] J. Peng, Y. Zhu, Q. Zhao et al., "Fair energy-efficient sensing task allocation in participatory sensing with smartphones," *The Computer Journal*, vol. 60, no. 6, pp. 850–865, 2017.
- [28] A. Tversky and D. Kahneman, "Advances in prospect theory: cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.
- [29] M. Shumaiza, A. N. A.-k. Akram, and J. C. R. Alcantud, "Group decision-making based on the VIKOR method with trapezoidal bipolar fuzzy information," *Symmetry*, vol. 11, no. 10, p. 1313, 2019.
- [30] S. Firdose, W. Lopes, R. Moreira, and P. Sofia, "CRAWDAD dataset copelabs/usense (v. 2017-01-27)," <http://crawdad.org/copelabs/usense/20170127>.
- [31] Z. H. Shao, H. Q. Wang, Y. F. Zou, Z. H. Gao, and H. W. Lv, "From centralized protection to distributed edge collaboration: a location difference-based privacy-preserving framework for mobile crowdsensing," *Security and Communication Networks*, vol. 2021, Article ID 5855745, 8 pages, 2021.
- [32] K. Palczewski and W. Salabun, "The fuzzy TOPSIS applications in the last decade," *Procedia Computer Science*, vol. 159, pp. 2294–2303, 2019.
- [33] C. Ceballos and V. Pilaud, "Denominator vectors and compatibility degrees in cluster algebras of finite type," *Transactions of the American Mathematical Society*, vol. 367, no. 2, pp. 1421–1439, 2015.