An Incentive Mechanism in Expert-decision-based Crowdsensing Networks

Bing Jia[a], Hao Gong[a,*], Zhaopeng Zong[a], Tao Zhou[a], Thar Baker[b], Ahmed Al-Shamma'a[c] and Yan Jia[d,*]

[a]College of Computer Science, Inner Mongolia University, Inner Mongolia 010021, China
[b]College of Computing and Informatics, University of Sharjah, Sharjah, UAE
[c]College of Engineering, University of Sharjah, Sharjah, UAE
[d]College of Cyber Science, NanKai University, Tianjing 300071, China

ARTICLE INFO

Keywords:
crowdsensing
expert-decision
statistical analytics
MRAI-UCDQ
sensing cost data
collected sensing data

ABSTRACT

With the rapid development of wireless communication and smart devices, crowdsensing applications became popular due to their flexibility to deploy and low cost use. Incentive mechanism is one of the most important research contents in crowdsensing, about crowdsensing incentive mechanism, most existing data quality evaluation methods measure the contributions of users only in terms of data quality, and ignore to measure the sensing cost of users. This leads to the problems of different quality evaluation standards, difficult to measure the data quality and difficult to give a reasonable and effective evaluation to complex problems. However, Expert-decision can effectively solve these problems and give high-quality evaluation decision for complex and numerous data results. In this paper, aiming at the shortcomings of existing research, we propose an expert-decision-based crowdsensing framework and gives the multidimensional rating for incentive mechanism based on user cost and data quality(MRAI-UCDQ), which consists of user cost evaluation model, data quality evaluation model, contribution quantification and reward distribution, by analysing user sensing cost data and collected sensing data (comprehensive evaluation with quantitative and qualitative analysis). Finally, through nearly 30 days of real experiments, 159 volunteers were recruited and 7000 pieces of sensory data were collected. The result shows the MRAI-UCDQ improves the evaluation performance of data quality and stimulates the user's perceived participation.

1. Introduction

Crowdsensing is the frontier research direction of the combination of Internet of Things (IoT) and Artificial Intelligence (AI) [1, 2], which is a new sensing mode that combines crowdsourcing ideas and mobile device perception capabilities by users [3]. In crowdsensing, users carry mobile devices instead of traditional sensors, because of the convenience of carrying mobile devices, the collection of sensing data becomes more convenient. In this paper, we develop a mobile app based on sensors to collect sensing data. It consists of sensing platform, users (the collectors of sensing data) and requesters (the initiators of sensing task), compared with traditional sensing methods [4], it has the characteristics of flexible deployment and low management cost. Due to the resource consumption and privacy security problems arising from the participation in the sensing task, users are prevented from actively participating in the task, resulting in issues such as the insufficient number of users in the sensing task [5] and the poor quality of data collected [6]. The incentive mechanism of crowdsensing has played a pivotal role, which is mainly divided into five modes: bonus incentive [7, 8], game incentive [9, 10], social relationship incentive [11, 12], virtual integral incentive [13, 14, 15] and mixed incentive [16, 17, 18]. The main goal of the research on the incentive mechanism of crowdsensing is to encourage the holders (sensing users) of mobile devices to join the sensing tasks under the management of the server platform, actively participate in sensing tasks, and submit high-quality and reliable sensing data [19, 20]. The execution of crowdsensing tasks depend on the participation of a large number of users, which need an effective and accurate evaluation mechanism for users. Without a good evaluation mechanism, users' trust in the platform may decrease. So it is more important to design an intelligent and efficient evaluation method in incentive mechanism [21, 22].

Expert-decision system can effectively improve the quality of decision and make professional and intelligent decision for complex question evaluation. In crowdsensing, an expert-decision-based crowdsensing system can encourage

*Corresponding author

more user participation, analyse sensing data by inference engine and give reasonable results. To be specific, the server publishes the sensing task, and the user uploads the sensing data. The expert-decision system is primarily involved in transaction verifiction, data encapsulation and user contribution quantification. Expert-decision system puts sensing data into inference engine, analyses the data and regards the result as the evaluation of user contribution quantification. Meanwhile, platform will pay users according to the quantification results. However in most existing research [23, 24, 25, 26] about the incentive mechanism of crowdsensing, only the quality of sensing data is considered as the evaluation standard of user contribution, ignore to measure sensing cost of user. Therefore, there is a need for an incentive mechanism of crowdsensing that can efficiently and correctly evaluate user contributions from multiple perspectives. In this paper, we propose an incentive mechanism of crowdsensing based on expert-decision, which can not only evaluate collected data but also users' sensing cost data, elevate the accuracy and efficiency of data evaluation. Besides, in order to deal with the comprehensive evaluation of different types of data, in terms of estimating, we give a comprehensive assessment to collected data quality and users' sensing cost based on fuzzy comprehensive evaluation method. Specifically, we add user sensing cost to the evaluation standards of user contribution, and we evaluate users' space cost, users' time cost, users' preferences cost and users' privacy cost as the user sensing cost evaluation standards. Specific data quality evaluation indicators and evaluation methods (quantitative and qualitative) in process of sensing data analytics are given based on the data quality evaluation method and indicator system written by the data application environment construction and service project team of Chinese academy of sciences, so that the accuracy of multi-angle evaluation results can be guaranteed, and more users can be encouraged to participate in the sensing tasks.

The rest of the paper is organized as follows. Section II reviews the literature. Section III shows the framework of expert-decision-based crowdsensing system and introduces the multidimensional rating for incentive mechanism based on user cost and data quality(MRAI-UCDQ). In Section IV, experimential results and performance evaluation are reported. We conclude this paper in Section V.

2. Related work

The applications of crowdsensing rely on a large number of users participation. However, due to the insufficient number of users in the task and the low quality of data provided, the development of swarm intelligence perception has been seriously affected. The expert-decision-based crowdsensing system can give reasonable quantification resluts of user contribution and have an impact on user engagement. A reasonable contribution quantification mechanism can be regarded as an effective incentive mechanism, which can motivate more users to accept the sensing task and upload the data. Therefore, an effective incentive mechanism should be considered.

In the traditional structure of crowdsensing, the incentive mechanism is a kind of fixed reward scheme with platform, which can be divided into monetary incentive mechanism and non-monetary incentive mechanism [27]. Either way, it is to compensate for the cost of power consumption, computing, storage, communication and other resources, as well as the cost of time and space. For the former way, based on the online auction model, Zhao D [28] investigate the problem that users submit their private types to the crowdsourcer when arriving, and the crowdsourcer aims at selecting a subset of users before a specified deadline for maximizing the value of services provided by selected users under a budget constraint. Feng Z [29] introduce a reverse auction framework to model the interactions between the platform and the smartphones, this way guarantees that submitted bids of smartphones reflect their real costs of performing sensing tasks. For the latter way, an incentive mechanism based on the social ties among users is presented to recruit sufficient users with high quality [30]. Jordan K [31] use the relations between movement and landmarks in order to reason about the structural significance of locations in a city park, based on the movement behavior exhibited by the players of the location-based game called Ostereiersuche.

In addition to the cost paid by users in the process of sensing tasks, users also want an effective and accurate evaluation mechanism to guarantee their profits. Therefore, it is necessary to design an intelligent and efficient decision method in incentive mechanism to ensure users' benefits and complete data collection tasks as much as possible. Some researchers focus on data analysis process in incentive mechanism, the evaluation platform can accurately measure the data quality of users, remove abnormal data and provide relatively accurate and reliable information for service damanders by using advanced machine learning and data mining technology [32, 33]. But the analysis results lack intelligence and specificity.

In recent years, more and more researches focus on the use of expert-decision system to ensure the accuracy of evaluation results in the IoT (Internet of Things). Guo, et al [34] propose an expert-decision system of coal mine safety production based on IoT, the use of expert-decision system improves the accuracy of prediction, reduces the possibility of safety accidents, and makes monitoring more and more intelligent. Zhang, et al [35] improve an agricultural irrigation expert-decision system based on IoT by integrating multiple influence parameters, the expert-decision system shows good performances in quantification and evaluation. For the evaluation of data in expert-decision sysytem, Hamedan Farahnaz, et al [36] develop a fuzzy logic-based expert system for diagnosis and prediction of chronic kidney disease and evaluate its robustness against noisy data. Zhao, et al [37] study the inference methods in expert-decision system, add fuzzy set theory to the evaluation mechanism of data, which improves the effectiveness of evaluation results. Above all, in this paper, we propose an expert-decision crowdsensing system to evaluate the contribution of users.

In most existing research, only the quality of sensing data is considered in the quantification of the user's contribution, because the quality of sensing data is regarded as the only index of the embodiment of the user's sensing level. We consider this to be inadequate, because users are bound to consume some resources in the process of performing sensing tasks, i.e. The cost of power consumption, computing, storage, communication and other resources, as well as the cost of time, space, privacy and preference. These must be compensated.

The existing work mentioned above is summarized in Table 1. Therefore, in view of the problems of the existing work, we propose a comprehensive evaluation method in expert-decision system to evaluate user contributions, which not only takes the quality of sensing data as the evaluation standard, but also takes the costs of users, such as their space cost, their time cost, their preferences cost and the privacy cost as the evaluation standards. Besides, most literature propose that the evaluation of data quality is estimated by falling in different discrete intervals of data, which is lack of scientific evaluation. In this paper, multidimensional data quality evaluation indicators and evaluation methods are given from quantitative analysis and qualitative analysis respectively using data quality index which consists of quality of form, quality of content, and quality of utility. These three data quality indicators reflect data quality from three levels of grammar, semantics and pragmatics. This data quality evaluation standard is referred to the data quality evaluation method and index system written by the data application environment construction and service project team of Chinese academy of sciences [38], so that the accuracy of the results can be guaranteed by quantifying from multiple perspectives, and more users can be encouraged to participate in the sensing task.

Table 1
Classification of incentive mechanism for mobile crowdsensing.

existing work	incentive method	evaluation standard
Peng et al [23]	bonus payment	assess participants' data quality issues
Gao et al [24]	bonus payment	maximize social welfare and the level of particaipation
Sun et al [25]	bonus payment	maximize the expected profit of the platform, data quality problems, and participation level issues
Nan et al [26]	bonus payment	user participation level and data quality issues
Zhao et al [28]	bonus payment	maximize platform utility and budget balance
Feng et al [29]	bonus payment	minimize social costs
Jordan et al [31]	entertainment leisure	the game player's movement data has the significance of improving the landmark system

3. MRAI-UCDQ incentive mechanism

The first part of the section will introduce the structure model of the expert-decision-based crowdsensing system, the rest part will introduce the incentive mechanism MRAI-UCDQ in detail. MRAI-UCDQ is composed of user cost evaluation model, data quality evaluation model, contribution quantification and reward distribution. And each part of the cost is quantified in expert-decision crowdsensing system. Due to the parameters limitations in models and the requirements of the sensing tasks, the data processing based on MRAI-UCDQ can only be done offline. The extension to models in online state, which is left as our future work.

3.1. The structure model of the expert-decision-based crowdsensing network

The framework of expert-decision-based crowdsensing networks is shown in Fig. 1, which consists of six steps, the details are as follows:

(1) The requester initiates a task request to the platform server.
(2) The platform server reviews and publishes the task to users.
(3) Users accept the task and task can be redistributed among users.
(4) Users accomplish the task and upload the sensing data to expert-decision system.
(5) Expert-decision system verifies the sensing data to evaluate the user contributions, and give the results to the platform server.
(6) The server pays for users according to their contribution quantified results.

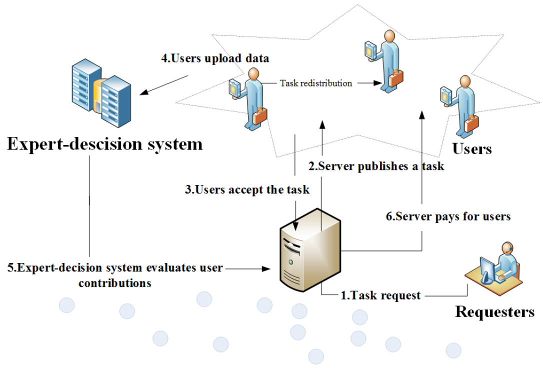

Fig. 1: The framework of expert-decision-based crowdsensing network.

3.2. User cost assessment

User cost assessment model in expert-decision system evaluates users' cost from four aspects: user space cost, user time cost, preference cost and privacy protection cost. The first is used to measure the activity range of users, the second is used to measure the time when users perform tasks, the third is used to measure the relationship between user preferences and task attributes and the last is used to measure users' requirements for location privacy protection.

3.2.1. User space cost

In crowdsensing, most researchers directly use the GPS location information of users, which leads to the disclosure of users' location information and reduces users' participation enthusiasm. In this paper, the geographic location is divided into regions, and the regional coordinate information transformed by GPS position coordinates of users is used to protect the location privacy information of users. Region division not only protects the location information of users to a certain extent, but also adjusts the size of region division according to the task demand to obtain the location information that satisfies the task requirements in (1).

Firstly, determine the maximum value of the longitude in the information collection area $longitude_{max}$ and the minimum value $longitude_{min}$, according to the longitude partition adjustment parameter α calculate the maximum number of longitude partition $longitude_{zone}$.

Secondly, determine the maximum $latitude_{max}$ and minimum $latitude_{min}$ values of the latitude in the information collection area, and calculate the maximum number of latitude partitions $latitude_{zone}$ according to the latitude partition adjustment parameter β.

$$\begin{cases} longitude_{zone} = \frac{longitude_{max} - longitude_{min}}{\alpha} \\ latitude_{zone} = \frac{latitude_{max} - latitude_{min}}{\beta} \end{cases} \tag{1}$$

According to the above latitude and longitude division formula can obtain partition $latitude_{zone} \times longitude_{zone}$. Define it as $Zone_{i,j} \quad i \in [0, latitude_{zone}],\ j \in [0, longitude_{zone}]$. When the user uploads data in the divided information collection area, the model calculates the position of the user in the custom coordinate system by the regional coordinate formula, and the calculation is as follows:

The calculation formula of the longitude area value $Zone_j$ and $Zone_i$ in the custom coordinate system is shown in (2). lo and la are the GPS longitude value and GPS latitude value of the user's location, $longitude_{min}$ and $latitude_{min}$

are the minimum value of the longitude and latitude of the measured area, α and β are the adjustment parameters of the longitude partition and latitude partition.

$$\begin{cases} Zone_j = \frac{lo-longitude_{min}}{\alpha} \\ Zone_i = \frac{la-latitude_{min}}{\beta} \end{cases} \tag{2}$$

According to the longitude area value $Zone_j$ and the latitude area value $Zone_i$ in the custom coordinate system, the area $Zone_{i,j}$ in which the user is in the custom coordinate system when the data is uploaded is determined.

A crowdsensing task requires users to collect perceptual information at various locations. The collection of perceptual information in densely populated areas can result in redundant information and reduced information value. Therefore, users are encouraged to collect data in areas where people are sparse by defining position activity. Hot spot area is an important factor in measuring position activity. Hot spot area $ZoneHot_{i,j}$ is the area in which the users have the most frequent activities, that is the number of users and the area where the users upload the most data. The further the user's area is from the hot spot area, the stronger the user's ability to move and the higher the positional activity. This encourages more users to collect data away from hot spot area and collect as much data as possible from the entire collection area. According to the situation of users uploading data, the partition type can be divided into four types: information-intensive, more information, less information, and minimal information.

Hot spot areas often appear in the first type information-intensive. Before defining a hot spot area, we need to define the user's participation in each area, so the area busy value $Zone(i,j)$ is defined in (3).

$$Zone(i,j) = \frac{N\_Zone_{(i,j)}}{N\_all} + \frac{NU\_Zone_{(i,j)}}{NU\_all} \tag{3}$$

N_all is the number of data uploaded by all users of a task, and $N\_Zone_{(i,j)}$ is the number of data collected by all users in the area numbered $Zone_{i,j}$. NU_all is the number of users participating in a task, and $NU\_Zone_{(i,j)}$ is the number of users participating in the task in the area numbered $Zone_{i,j}$. Calculate the zone busy value $Zone(i,j)$ for each zone and define the maximum value of the zone busy value as $max(i,j)$. Usually there is only one hot spot area, the maximum value in $Zone(i,j)$ (the area where $max(i,j)$ is located) is defined as the hot spot area $ZoneHot_{i,j}$, and the value $Zone(i,j)$ of the hot spot area is defined as $Zone_{hot}(i,j)$. But there may be a situation that the maximum value and the second largest value of $Zone(i,j)$ are very close. In this case, it is inappropriate to define a hot spot area, so we can define the areas where the maximum and the second largest value are located as the hot pot areas. And so on, multiple hot spot areas can be set. According to the size of the area, the number of hot spot areas cannot keep increasing all the time, and there is a certain amount limit. This peper quantifies the hot spot area in expert-decision system as shown in (4).

$$Zone_{hot}(i,j)\{Zone(i,j)|max(i,j) - Zone(i,j) \le d_\alpha\} \tag{4}$$

d_α is the difference between $max(i,j)$ and $Zone(i,j)$. All areas in which the zone busy value difference is within this range are defined as hot spot areas, as long as one area staisfies $Zone(i,j) = Zone_{hot}(i,j)$, it is the hot spot area $Zone_{hot}(i,j)$.

The user's location activity is defined based on the distance between the user and the hot spot area. In this paper, the acquisition area is divided into differerent areas of equal size, which is closer to the definition of Chebyshev distance [40]. When there are multiple hot spot areas, the distance equals the average distance between the area where the point is located and the hot spot area is taken. The distance between the user's area and the hot spot area is quantified in expert-decision system as shown in (5). $Zone_{i,j}$ is the area where the user is located, and $(i,j)$is the zone coordinate value. $ZoneHot_{i,j}$ is the hot spot area, and (i_{hot}, j_{hot}) is the coordinate value of the hot spot area.

$$|Zone_{i,j}ZoneHot_{i,j}| = max(|i - i_{hot}, |j - j_{hot}|) \tag{5}$$

The user space cost $PZone$ is quantified in expert-decision system as shown in (6). C is the capability parameter, n_{hot} is the number of hot spot areas, $|Zone_{i,j}ZoneHot_{i,j}|$ is the distance between the user's area and the hot spot area.

$$PZone = \frac{\sum_{m=1}^{n_{hot}} |Zone_{i,j}ZoneHot_{i,j}|}{n_{hot}} \times C \tag{6}$$

3.2.2. Privacy protection cost

Based on the above description of user's location (longitude latitude mapping) information, we add the user's level selection for privacy protection of their location information data in expert-decision system. The higher the level of privacy protection is, the lower the cost of privacy protection will be, and the lower the quantification of users' contribution to their privacy protection in expert-decision system. Next, we will give the definition of privacy protection level.

In this paper, we simply set three levels for the cost of privacy protection. Three levels can be described as: no need for privacy protection, need low level of privacy protection, need high level of privacy protection. Here we define the user's privacy protection cost parameter as p, and the three levels of parameter P can be defined as 1, 0.5, 0.

Three levels of privacy protection can be expressed as: for users who do not need privacy protection, the published location information is their real location; for users who need lower level of privacy protection, the published location information is any value in the circle area with their real location as the center and radius as r; for users who need higher level of privacy protection, the published location information is any value in the circle area with their real location as the center and radius as R ($R > r > 0$). And the arrows stand for the r and R. R and r will be set according to the area coordinate range of the collected data.

3.2.3. User time cost

In the user's time participation, to protect the user's time privacy information, the day is divided into 24 time zones, 0 time zone is defined from 00:00 to 01:00, and 1 time zone is defined from 01:00 to 02:00 and so on. When users upload data, they do not directly upload the exact time of the collected data but the time zone number. The time in which users perform tasks is also an important indicator of the ability of users. The longer the participation time, the stronger the users' willingness to participate. And the actual participation time of the users $AcTask$ is shown in (7). at stands for the number of the time zones, and $Time_t$ is users' participation time in time zone t.

$$AcTask = \sum_{t=1}^{at} Time_t \tag{7}$$

Due to the dispersion of uploading data time, we define the content about time hot spot here. Because in normal working hours, users have high enthusiasms to upload sensing data, and in non-working hours (rest time), users have low enthusiasms to upload sensing data, which will result in less sensing data received in non-working hours and data collected exists local defect. Therefore, we affirm the working time (we define the time as 6:00-18:00) as the time hot spot, and Th is given to evaluate the time hot spot of users upload data. In this paper, we define Th as 0 for users who submit data in the time hot spot and 0.1 for users who are not in the time hot spot. If the data is uploaded multiple times, the average value can be taken as the final Th.

User Time Cost TC is defined as shown in (8). $ATask$ indicates the total time of the task. $AcTask$ is the actual participation time of the users, Th indicates the hot time of the user upload data and o is the adjustment parameter.

$$TC = \frac{AcTask}{ATask} \times o + Th \tag{8}$$

3.2.4. User preference cost

In this paper, for each user, before accepting the sensing task, we ask users to make their own preference for the task type, and then we add labels to the published task. If the users' preferences are consistent with the task labels they participate in, then we set the parameter Pr in (9) for them. The Pr is the cost of user performance. Here, $Pr = 1$ when the defined preferences match, otherwise $Pr = 0$.

3.2.5. User cost quantification

This paper evaluates the user's cost to participate in the four aspects of the user's participation in the task's position activity and the user's participation in the task's time engagement which is quantified in expert-decision system as shown in (9).

$$Cost = \theta \times \frac{\sum_{i=1}^{n} PZone_i}{n} + \mu \times TC + \iota \times P + \lambda \times Pr \tag{9}$$

$$\theta + \mu + \iota + \lambda = 1$$

In each sensing task, each time users collect the sensing data, they will record the users' current positional activities. TC is the user's time participation, $PZone_i$ is the user's location activity, n is the number of times the user collects data in the task, P is the cost of user privacy protection and Pr is the cost of user preference, θ, μ, ι, λ are the weight.

The quantification algorithm of user cost is Algorithm 1.

Algorithm 1: User Cost Quantification Algorithm

Input: M users, each user has n crowdsensing data

Output: Participation ability of each user

Initialize the Hot spot area and mark the hot spot area and the number of hot spot is n_{hot} and the privacy protection parameter is p and the preference parameter is Pr and the number of time hot spot is Th;

$i = 1; j = 1; k = 0; Pos = 0;$

while $j \leq m$ **do**

    while $i \leq n$ **do**

        Calculate coordinate points

        $Zone_{j'} = \frac{lo-longitude_{min}}{\alpha}$

        $Zone_{i'} = \frac{la-latitude_{min}}{\beta}$

        Calculate the distance of users from the hotspot area for every data

        while $k \leq n_{hot}$ **do**

            $Pos = Pos + |Zone_{i,j} ZoneHot_{i,j}|;$

            $k = k + 1;$

        $PZone_i = \frac{Pos}{n_{hot}} \times \tau;$

        $i = i + 1;$

    while $i \leq n$ **do**

        $TC = \frac{Actask}{Atask} \times o + Th;$

    $j = j + 1;$

while $j \leq m$ **do**

    while $i \leq n$ **do**

        $PZ = PZ + PZone_i;$

        $i = i + 1;$

    $Cost = \theta \times \frac{PT}{n} + \mu \times TC + \iota \times P + \lambda \times Pr;$

    $j = j + 1;$

In the part of user cost in expert-decision system, firstly, quantify the space cost of the user, partition the measured area and mark the position information of the user in the partition (ie, initialize the data information), obtain the hot spot area and calculate the distance between each piece of data and the hot spot area, define the user's position activity according to the distance between the user and the hot spot area. The farther the user collects data from the hot spot area, the stronger the ability of the user to participate in the task, and the more valuable the data the user collects. Secondly, quantify the users' time participation ability and statistic the time users participate in the task. The longer the user participates in a task, the higher the enthusiasm of the user to participate in the task. Thirdly, get the user privacy protection cost. Forthly, get the user preference cost. Finally, based on the characteristics of user position activity, time engagement, privacy protection cost and prefernece cost, expert-decision system defines the user's cost quantification function to reflect the user's overall cost.

3.3. Data quality

Due to the single and limited problems of the fixed data assessment methodology, it can not truly reflect the data quality. In this paper, we use expert-decision system to evaluate the data collected by users based on the given evaluation indicators, which consists of quantitative indicators (integrity, uniqueness, validity, timeliness, reliability) and qualitative indicators (comprehensibility, accessibility, objectivity, relevance). These indicators have been selected based on the data quality evaluation method and indicator system written by the data application environment construction and service project team of Chinese academy of sciences.

3.3.1. Quantitative analysis for data quality

The main workflow of the assessment methodology in expert-decision system is:

a Obtaining data information collected by a certain user in a certain task from the data set;
b According to the task's demand for data, choose one or more or even all of the quantitative and qualitative evaluation indicators;
c The corresponding judging formula is given for each indicator to calculate the data quality based only on quantitative analysis. And the corresponding grade division is given for each qualitative index;
d Firstly, calculate the data quality based only on quantitative analysis. Secondly, according to the former results and the qualitative grade division, calculate the comprehensive data quality based both on the quantitative analysis and qualitative analysis.

In the quantitative evaluation algorithm for the data quality, different data evaluation indicators and the corresponding judging formulas are used to form a set of data quality assessment template, which makes the data quality assessment more authentic and reasonable.

Algorithm 2: Quantitative Evaluation Algorithm For Data Quality

```
Input: IArray contains I_n data evaluation indicators, DArray represents the D_n crowd-sensing data of the
       user in one task.
Output: Quantitative evaluation result for data quality of the user
select m evaluation indicators that meet needs and m ∈ I_n;
E_j = 0; F_j = 0; QTDQ = 0;
∑_{j=1}^{m} b_j = 1;
i = 1; j = 1;
while j ≤ m do
    while i ≤ D_n do
        if data i passes the evaluation rule then
            E_j = E_j + 1;
        end if
        i = i + 1;
    F_j = E_j / D_n;
    QTDQ = QTDQ + b_j × F_j;
    j = j + 1;
return QTDQ;
```

The alogrithm is shown in Alogrithm 2. I_n represents the number of evaluation indicators; D_n represents the number of crowd-sensing data; E_j represents the number of qualified data for each evaluation index; F_j represents the degree of completion of the indicator; $QTDQ$ represents the quantitative data quality;

3.3.2. Comprehensive analysis for data quality based on fuzzy comprehensive evaluation method

In this paper, expert-decision system achieves fuzzy comprehensive evaluation based on the eclectic fuzzy multi-attribute comprehensive evaluation method which is described: given a user set $A = |A_1, A_2, \cdots, A_n|$ and the data quality attribute set corresponding to each user $C = |C_1, C_2, C_3, C_4, C_5|$, we define 5 data quality attributes, including $QTDQ$ which we get above, comprehensibility, availability, objectivity, relevance. The weight set that describes the relative importance of each data quality attribute $W = W_1, W_2, \cdots, W_m$. Becsuse the representation of attribute index and weight size can be digital or natural language, and the data structure involved can be accurate or imprecise. All natural language or imprecise attribute indexes, weight size and data structure are expressed as fuzzy subsets or fuzzy numbers in decision space [38], so in this paper, we describe the four qualitative indexs (including comprehensibility, availability, objectivity, relevance) with degree A, B, C, D, and we define the degrees with different score regions. The following describes the main steps of the method in expert-decision system.

The first step, express data quality attribute with triangle fuzzy number, the method is shown in (10). M which represents user is a triangle fuzzy number and a subset of all fuzzy sets on real number set R, and μ_M is membership function of M, M also can be expressed as $M = (k, m, p)$. After all data quality attribute indexes are transformed into triangular fuzzy numbers, the obtained fuzzy index matrix is $F = (f_{ij})_{m*n}$.

$$\mu_M(x) = \begin{cases} \frac{x-k}{m-k}, & x \in [k, m] \\ \frac{x-p}{m-p}, & x \in [m, p] \\ 0, & x < k \quad or \quad x > p \end{cases} \tag{10}$$

The second step, normalize matrix F, and the triangular fuzzy number M above can be described as $M = (k, m, p)$. Assume there are N evaluation objects, for a fuzzy index value $x_i = (a_i, b_i, c_i)$, $(i = 1, 2, \cdots, N)$ (N is the numbr of users), its normalization formula of cost index is shown in (11), and the income index is shown in (12). Finally, we can get the normalized fuzzy index matrix $R = (y_{ij})_{m*n}$, j represents the data quality evaluation index number.

$$y_i = \left(\frac{\min\left(a_i\right)}{c_i}, \frac{\min\left(b_i\right)}{b_i}, \frac{\min\left(c_i\right)}{a_i} \wedge 1 \right) \tag{11}$$

$$y_i = \left(\frac{a_i}{\max\left(c_i\right)}, \frac{b_i}{\max\left(b_i\right)}, \frac{c_i}{\max\left(a_i\right)} \wedge 1 \right) \tag{12}$$

The third step, construct fuzzy decision matrix, the fuzzy decision matrix $D = (r_{ij})_{m*n}$ can be obtained by weighting matrix R. Define $w = \left(w^{(1)}, w^{(2)}, w^{(3)}\right)$, $y_{ij} = \left(y_{ij}^{(1)}, y_{ij}^{(2)}, y_{ij}^{(3)}\right)$, and the defination of r_{ij} is shown in (13).

$$r_{ij} = w \Theta y_{ij} = \left(w^{(1)} y_{ij}^{(1)}, w^{(2)} y_{ij}^{(2)}, w^{(3)} y_{ij}^{(3)}\right) \tag{13}$$

Finally calculate fuzzy positive ideal and fuzzy negative ideal, the fuzzy positive ideal and negative ideal are composed of the maximum value and the minimum value of the fuzzy index value in each index. Calculate the weighted Euclidean distance between each candidate object and the fuzzy positive ideal and the fuzzy negative ideal, and then calculate the membership degree of each candidate object belonging to the fuzzy positive ideal. The larger the membership degree is, the higher the evaluation of the user is. And we define the selected user's corresponding $r_i = \sum_j r_{ij}$ with CDQ_i.

3.4. Contribution quantification and reward distribution

According to the sensing data uploaded by users, expert-decision system quantifies the contribution of user cost and data quality.

3.4.1. Contribution quantification

In addition to collecting sensing data, users will also give pricing to the data. This paper selects the average pricing given by the users as the final pricing of the users' current task. User pricing is defined as shown in (14).

$$price = \frac{\sum_{j=1}^{N} (p_j)}{\varphi * N} \tag{14}$$

price is the price of the user's task. N is the total number of data collected by the user. p_j is the price of each piece of data collected by user. φ is the adjustment parameter. The difficulty and bonus of the task determine the pricing interval, and the task publisher gives the appropriate pricing range to provide a reference for the users. Users' participation performance definitions are shown in (15).

$$Chip_i = \omega 1 \times Cost_i + \omega 2 \times \frac{CDQ_i}{\epsilon} + \omega 3 \times \frac{\delta}{price_i} \tag{15}$$

$$\omega 1 + \omega 2 + \omega 3 = 1$$

Among them, $Chip_i$ stands for the user's participation performance, CDQ_i is the user's comprehensive data quality, $price_i$ is the user's pricing, $\omega 1, \omega 2, \omega 3$ are the weight, δ, ϵ are the adjustment parameters which will be adjusted according to the experimental data. It can be observed from (15) that $Chip_i$ is related to $Cost_i$, CDQ_i and $price_i$. The larger the scope of user activities, the longer the participation time, the higher the user's participation performance, so $Chip_i$ is positively correlated with the $Cost_i$; the higher the quality of data uploaded by the user, the higher the participation performance of the user, so $Chip_i$ is positively correlated with CDQ_i; when the total task bonus is fixed, the lower the user's bid, the more task cost can be saved, so $Chip_i$ is inversely related to $price_i$. Since $Cost_i$, CDQ_i and $price_i$ are all obtained according to above algorithms, $chip_i$ can be easily calculated in (15). The expert-decision system further sends the quantified contribution to the Server, and then the Server sends the corresponding reward to the user.

3.4.2. Reward distribution

User's bonus distribution principle can be defined as follows: only if the user's participation performance value above threshold $aver\_chip$, the bonus can be distributed. The user's bonus distribution threshold $aver\_chip$ is defined in (16). n_{user} is the number of users. $chip_j$ is the participating performance value for each user. γ is the adjustment parameter, $j \in n_{user}$.

$$aver\_chip = \gamma \times \frac{\sum_{j=1}^{n_{user}} Chip_j}{n_{user}} \tag{16}$$

After the collection of task data, the algorithm statistical analyses sensing data of each user and calculates the participation performance chip of each user. The bonus distribution algorithm pseudo-code for the incentive mechanism is shown in Algorithm 3.

Algorithm 3: Reward Distribution

```
Input: Chip_i, aver_chip, n_user, Tbonus, Price_i
Output: Users reward
i = 1, m = 1, k = 1, W_k = 0;
while i ≤ n_user do
    if Chip_i ≥ aver_chip then
        Mark the user i as the winner_m ;
        m = m + 1;
while i ≤ n_user do
    if Tbonus > 0||m ≠ 0 then
        if User i is the winner then
            Reward_i = Price_i ;
            Tbonus = Tbonus − Price_i;
            m = m − 1;
        else
            Reward_i = 0;
    else
        break;
```

The condition for bonus distribution is satisfied when the user's participation performance value is higher than the threshold $aver\_chip$. Users with participation performance values are higher than the threshold are sorted by the size of the participating performance values from largest to smallest. Bonuses are distirbuted from large to low by server according to the users' participation performance values. The number of bonuses distributed is the number of users' pricing. It ends when the total prize $Tbonus$ is finished distributed by server or the number of users who meet the conditions is zero.

In the last, we summarize notations in formulas in Table 2. The adjustment parameters involved in the above formulas are α, β, o, φ. The purpose of setting them is to perform unified calculations on parameters of different orders of magnitude without affecting the experimental results, which also can be understood as weight parameters that balances parameters of different orders of magnitude.

Table 2

Notation definitions in formulas.

notation	definition
$ZoneHot$	hot spot area
N_all	the total number of data uploaded by users
NU_all	the total number of users participating in the task
N_Zone	the total number of data uploaded by users in the area $Zone$
NU_Zone	the total number of users participating in the task in the area $Zone$
$PZone$	user space cost
$AcTask$	the actual participation time of the users
$ATask$	the total time of the task
$Time_t$	users' participation time in time zone t
Th	the hot time of the user uploaded data
TC	user time cost
Pr	user preference cost
$Cost$	user cost quantification
$QTDQ$	quantitative data quality
CDQ	user comprehensive data quality
$Chip$	user participation performance
$aver\_chip$	user bonus distribution threshold

4. Experiments and results

4.1. Experimental setup

In the experiments, three tasks with the same mission objectives but using different incentive mechanisms (recommended incentives, MAA [41] and RADP-VPC [42], MAA is a multi-attribute auction incentive mechanism, which also can be thought as an auction method in which the seller and the buyer conduct multiple negotiations on price and other attributes. RADP-VPC is a dynamic price reverse auction incentive mechanism based on virtual participation in credit, which selects the participant with the lowest bid as the winner and pay. They are all aution incentive mechanisms, so it's better to compare with them.) are published through the server, and the different incentive mechanisms' bonus distribution methods are described in details in the task description. The users select the tasks under the corresponding incentive mechanism to participate in the collection of the sensing data according to their own needs and wishes. The experimental environment is the campus of Inner Mongolia University, specifically including canteens, teaching buildings, dormitory areas, etc. Participants are students of Inner Mongolia University. And the information collected are noise information, light intensity information and location information. In the experiment of multi-task data collection, eleven groups of tasks with the same target are published through the server but only eight groups of tasks receive data. Each group of tasks includes three tasks which use three different incentive mechanisms, and we use these data to form a data set for analysis, the total number of tasks is 24.

The users in this experiment are students of Inner Mongolia University, in which the ratio of male to female is 33 to 17, and their respective user participation degree in three incentive mechanisms is shown in Fig. 2. The ratio of bachelor, postgraduate and doctor is about 28 to 17 to 5, and their respective data user participation degree in three incentive mechanisms is shown in Fig. 3. And the number of tasks in the fiture stand for cumulative tasks. The total users in this experiment is 159, and the average number of data in each round is about 1500, but there are some duplicate data, so after technical screening, we can use about 900 data in each round on average. Although there are not many users participating in the data collection tasks, the proportion of users' participation in the total number of users fluctuates to a certain extent, and the overall trend is the same, so the change in the number of users has little impact on the results of following experiments.

Before comparing the incentive effects of the three incentive mechanisms, the parameters of the MRAI-UCDQ incentive mechanism need to be calculated and calibrated according to the experimental environment and the collected data. The parameters are adjusted for satisfying calculations between different orders of magnitude and getting better experimental results. The parameters are shown in Table 3.

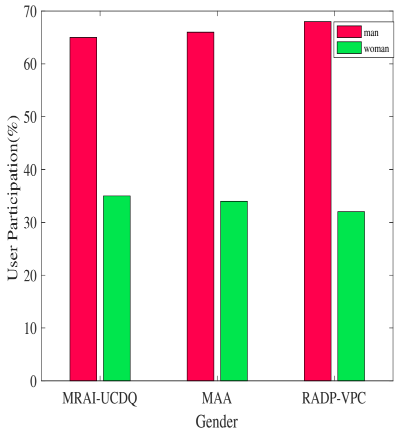

Fig. 2: User participation of different genders.

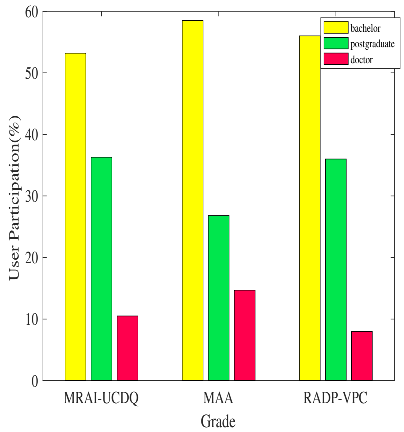

Fig. 3: User participation with different educational backgrounds.

Table 3
Parameter representation.

Parameter name	Symbol	Value
Longitude partition parameter	α	0.001
Latitude partition parameter	β	0.00039
Position activity parameter	C	0.1
Time cost adjustment parameter	o	3000
User cost weight	$\mu \setminus \theta$	$0.5 \setminus 0.5$
Data quality weight	$b1 \setminus b2 \setminus b3 \setminus b4 \setminus b5$	$0.3 \setminus 0.3 \setminus 0.2 \setminus 0.1 \setminus 0.1$
Participate in performance tuning parameters	ϑ	30
Participate in performance tuning parameters	σ	100
Participate in performance weights	$\omega1 \setminus \omega2 \setminus \omega3$	$0.3 \setminus 0.4 \setminus 0.3$
Threshold adjustment parameter	γ	0.65
Total bonuses	$Tbonus$	2000
Total time zone	at	24

4.2. Analysis of experimental results

User participation is shown in Fig. 4. The user participation of the MRAI-UCDQ incentive mechanism is higher than that of MAA and RADP-VPC. The changing trend of the MAA incentive mechanism's user participation is rising first and then stable; RADP-VPC incentive mechanism's price winner is only one user, so its user participation is the lowest and its changing trend is the same as MAA incentive mechanism.

Use the data quality assessment model in expert-decision system to evaluate the data quality of the three incentive mechanisms. The data scoring range is from 0 to 100. The higher the data score, the better the quality. The quantitative evaluation indicators of the data quality are set as uniqueness, completeness, validity, reliability and timeliness according to the attributes of the sensing data received by the task. The average data quality of the users is shown in Fig. 5. In this three incentive mechanisms, MRAI-UCDQ incentive mechanism's data quality is the highest; MAA incentive mechanism's average data quality changing trend is descending with the number of tasks; RADP-VPC incentive mechanism's data quality score is the lowest and declines gradually. And the membership grade results of comprehensive evaluation of quantitative and qualitative indexes (including comprehensibility, objectivity, et al) based on fuzzy comprehensive evaluation is shown in Fig. 6. In this three incentive mechanisms, MRAI-UCDQ incentive mechanism also shows the best performance in membership grade; MAA incentive mechanism's membership grade changing trend is also declining with the number of tasks increases; RADP-VPC incentive mechanism's membership grade changing trend is the same as the MAA and MRAI-UCDQ, but shows the worst performance.

The winner's average bonus is shown in Fig. 7. The number of winners selected in each task is different. Therefore, the winner's average bonus is chosen to compare the difference in bonuses among the three incentives. The total bonuses won by all the winners are shown in Fig. 8. The MAA is the most expensive incentive mechanism for the three incentive mechanisms. The RADP-VPC incentive mechanism only selects one winner, so the RADP-VPC average bonus is the least consumed. And the MRAI-UCDQ bonus consumption is in between. MRAI-UCDQ collects the data with the best quality but not spends the most, so MRAI-UCDQ is better than MAA and RADP-VPC.

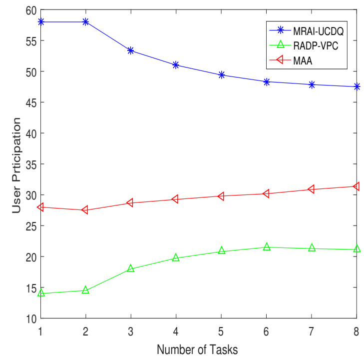

Fig. 4: User prticipation in different incentive mechanisms.

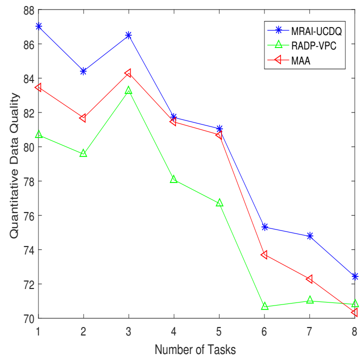

Fig. 5: Quantitative analysis result of data quality.

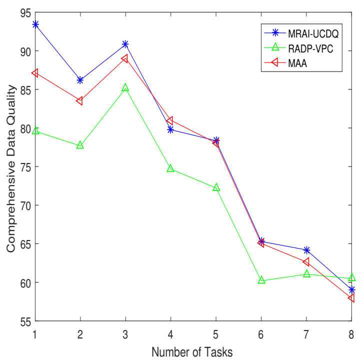

Fig. 6: Quantitative analysis and qualitative analysis result of data quality.

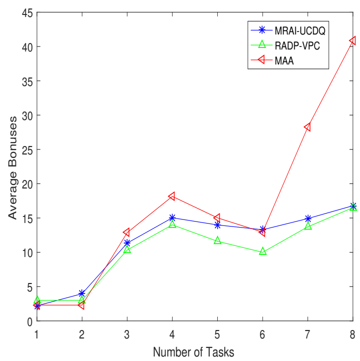

Fig. 7: Winner average bonuses.

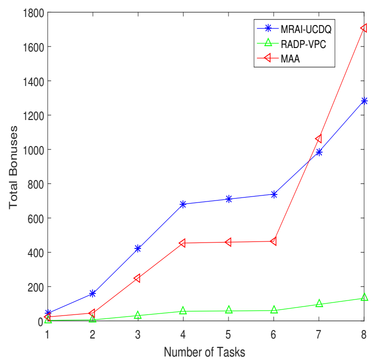

Fig. 8: Winner total bonuses.

5. Conclusion

In summary, most existing crowdsensing incentive mechanisms do not have effective user evaluation methods and ignore the quantification to the sensing cost of users. In this paper, we propose MRAI-UCDQ, which is a multi-attribute reverse auction incentive mechanism based on expert-decision. MRAI-UCDQ quantifies user contribution which includes user cost and data quality in the expert-decision system. In specific, MRAI-UCDQ analyses the user sensing cost data and collected sensing data by the inference engine, and gives reasonable evaluation results. Experiment results prove that MRAI-UCDQ is superior to MAA and RADP-VPC, and MRAI-UCDQ can increase the participation enthusiasm of users and ensure the quality of data collected by sensing tasks. For future work, we will add user economic condition to user sensing cost quantification parameters, elevate the number of experimental users, increase the complexity of the data collection environment, improve the existing functions of the experimental app and add more practical functions.

References

[1] M. Wieczorek, J. Sika, M. Wozniak, S. Garg and M. Hassan, "Lightweight CNN model for human face detection in risk situations," in IEEE Transactions on Industrial Informatics, doi: 10.1109/TII.2021.3129629.

[2] Marcin Woźniak, Jakub Siłka, and Michal Wieczorek. 2021. Deep learning based crowd counting model for drone assisted systems. In Proceedings of the 4th ACM MobiCom Workshop on Drone Assisted Wireless Communications for 5G and Beyond (DroneCom '21). Association for Computing Machinery, New York, NY, USA, 31–36. DOI:https://doi.org/10.1145/3477090.3481054.

[3] Wu Y, Zeng JR, Peng H, Chen H, Li CP. Survey on incentive mechanisms for crowd sensing. Ruan Jian Xue Bao/Journal of Software, 2016,27(8):20252047 (in Chinese).

[4] Lv Z, Halawani A, Feng S, et al. Touch-less interactive augmented reality game on vision-based wearable device[J]. Personal and Ubiquitous Computing, 2015, 19(3-4): 551-567.

[5] Deterding S, Dixon D, Khaled R, et al. From game design elements to gamefulness: defining gamification[C]// International Academic Mindtrek Conference: Envisioning Future Media Environments. 2011.

[6] Wen Y, Shi J, Zhang Q, et al. Quality-Driven Auction-Based Incentive Mechanism for Mobile Crowd Sensing[J]. IEEE Transactions on Vehicular Technology, 2015, 64(9):4203-4214.

[7] Lee J S, Hoh B. Sell your experiences: a market mechanism based incentive for participatory sensing[C]// IEEE International Conference on Pervasive Computing & Communications. 2010.

[8] Gao L, Hou F, Huang J. Providing long-term participation incentive in participatory sensing. arXiv preprint arXiv:1501.02480,2015. [doi: 10.1109/INFOCOM.2015.7218673].

[9] Kawajiri R, Shimosaka M, Kashima H. Steered crowdsensing: incentive design towards quality-oriented place-centric crowdsensing[C]// Acm International Joint Conference on Pervasive & Ubiquitous Computing. 2014.

[10] Han K, Graham E A, Vassallo D, et al. Enhancing Motivation in a Mobile Participatory Sensing Project through Gaming[C]// IEEE Third International Conference on Privacy. 2011.

[11] Luo T, Kanhere SS, Tan HP. SEW-Ing a simple endorsement Web to incentivize trustworthy participatory sensing. In: Proc. Of 2014 the 11th Annual IEEE Int'l Conf. on Sensing, Communication, and Networking (SECON). IEEE, 2014. 636644. [doi:10.1109/ SAHCN.2014.6990404]

[12] Bigwood G, Henderson T. IRONMAN: Using Social Networks to Add Incentives and Reputation to Opportunistic Networks[C]// IEEE Third International Conference on Privacy. 2012.

[13] Yu T, Zhou Z, Da Z, et al. INDAPSON: An incentive data plan sharing system based on self-organizing network[C]// IEEE Infocom -ieee Conference on Computer Communications. 2016.

[14] Chou C M, Lan K C, Yang C F. Using virtual credits to provide incentives for vehicle communication[C]// International Conference on Its Telecommunications. 2013.

[15] Lan K, Chou C M, Wang H Y. An Incentive-Based Framework for Vehicle-Based Mobile Sensing[J]. Procedia Computer Science, 2012, 10:1152-1157.

[16] Wang Y, Jia X, Jin Q, Ma J. QuaCentive: A quality-aware incentive mechanism in mobile crowdsourced sensing (MCS). The Journal of Supercomputing, 2015. [doi: 10.1007/s11227-015-1395-y]

[17] Lee J S, Hoh B. Dynamic pricing incentive for participatory sensing[J]. Pervasive & Mobile Computing, 2010, 6(6):693-708.

[18] Deterding S, Dixon D, Khaled R, et al. From game design elements to gamefulness: defining gamification[C]// International Academic Mindtrek Conference: Envisioning Future Media Environments. 2011.

[19] LiuZhao, LiKenli, ZhouXu, et al. Incentive Mechanisms for Crowdsensing[J]. ACM Transactions on Sensor Networks (TOSN), 2020.

[20] Lin J , Li M , Yang D , et al. Sybil-Proof Online Incentive Mechanisms for Crowdsensing[C]// IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. IEEE, 2018.

[21] X. Liu, W. Wang, B. Yang and Y. Fan, "Incentive Mechanism for Data Uploading in Mobile Crowdsensing," 2019 International Conference on Networking and Network Applications (NaNA), Daegu, Korea (South), 2019, pp. 391-395, doi: 10.1109/NaNA.2019.00074.

[22] J. Liu, Y. Yang, D. Li, x. deng, S. Huang and H. Liu, "An Incentive Mechanism Based on Behavioural Economics in Location-based Crowdsensing Considering an Uneven Distribution of Participants," in IEEE Transactions on Mobile Computing, doi: 10.1109/TMC.2020.3002586.

[23] Wang L, Zhang D, Wang Y, et al. Sparse mobile crowdsensing: challenges and opportunities[J]. IEEE Communications Magazine, 2016, 54(7):161-167.

[24] Dong Z, Huadong M A, Liang L. Quality Measuring and Assurance for Mobile Crowd Sensing[J]. ZTE Technology Journal, 2015.

[25] Wang, Jiangtao, et al. "Multi-Task Allocation in Mobile Crowd Sensing with Individual Task Quality Assurance." IEEE Transactions on Mobile Computing 17.9(2018):2101-2113.

[26] Bokhari A , Buhari S M , Alzahrani A A . Optimal Quality Assurance Task Assignment for Freelance Crowdsourcing Applications[C]// 2019 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). IEEE, 2019.

[27] Füller J, Hutter K, Fries M. Crowdsourcing for Goodness Sake: Impact of Incentive Preference on Contribution Behavior for Social Innovation[J]. 2012, 23(2012):137-159.

[28] Zhao D, Li X Y, Ma H. Budget-feasible online incentive mechanisms for crowdsourcing tasks truthfully[J]. IEEE/ACM Transactions on Networking, 2014, 24(2): 647-661.

[29] Feng Z, Zhu Y, Zhang Q, et al. TRAC: Truthful auction for location-aware collaborative sensing in mobile crowdsourcing[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE, 2014: 1231-1239.

[30] Yang G , He S , Shi Z , et al. Promoting Cooperation by the Social Incentive Mechanism in Mobile Crowdsensing[J]. IEEE Communications Magazine, 2017, 55(3):86-92.

[31] Jordan K, Sheptykin I, Grüter B, et al. Identification of structural landmarks in a park using movement data collected in a location-based game[C]//COMP@ SIGSPATIAL. 2013: 1-8.

[32] Xiaoguang Niu, Meng Li, Qianyuan Chen, et al. EPPI: An E-cent-based privacy-preserving incentive mechanism for participatory sensing systems[C]// 2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC). IEEE, 2015.

[33] Xuewen Dong, Tao Zhang, Guangxia Li, et al. A Privacy-Preserving Incentive Mechanism on Multi-Bid Crowdsourced Spectrum Sensing[C]// 2018 International Conference on Networking and Network Applications (NaNA). IEEE Computer Society, 2018.

[34] Guo Yongli, Wang Dan, Guo Hongzhuan. Research on coal mine safety production expert system based on Internet of things [J]. Coal technology, 2018,37 (08): 317-319.

[35] Zhang Ting, Lu Jun, Shen Jingjing. Intelligent irrigation expert decision system based on Internet of things [J]. Modern agricultural science and technology, 2017 (21): 176-177.

[36] Sadoughi Farahnaz, Kazemy Zahra, Hamedan Farahnaz, et al. Artificial intelligence methods for the diagnosis of breast cancer by image processing: a review.. 2018, 10:219-230.

[37] Zhang Guodong, Liu Jiayin, MI Tiezhu, Wan Jili, Wang Kexiang, Xu Chunying, GE xujuan. An agricultural Internet of things system based on big data [P]. Shandong Province: cn109460098a, March 12, 2019.

[38] Data quality evaluation method and index system. https://wenku.baidu.com/view/539d20adf605cc1755270722192e453610 665bcf.html, 2009-09.

[39] Zhang M, Yang P, Chang T, et al. Quality-Aware Sensing Coverage in Budget-Constrained Mobile Crowdsensing Networks[J]. IEEE Transactions on Vehicular Technology, 2016, 65(9):7698-7707.

[40] Karjanto N. Properties of Chebyshev polynomials[J]. 2020.

[41] Krontiris I, Albers A. Monetary incentives in participatory sensing using multi-attributive auctions. Int'l Journal of Parallel,Emergent and Distributed Systems, 2012,27(4):317336.

[42] Lan JB, Xu Y, Liu JZ. The compromise method of fuzzy multi-attribute decision [J]. System engineering and electronic technology, 2004 (01): 40-43.