Research Article

Task Grid-Based Urban Environmental Information Release Mechanism for Mobile Crowd Sensing

Zhenwei Chen,1 Yang Lu,2 Zusong Li,1 and Xiancun Zhou1

1Faculty of Electronic and Information Engineering, West Anhui University, Lu’an, China
2Institute of Distributed Intelligence and Internet of Things, Hefei University of Technology, Hefei, China

Correspondence should be addressed to Zhenwei Chen; 03000059@wxc.edu.cn

Received 19 April 2022; Revised 3 August 2022; Accepted 23 August 2022; Published 21 September 2022

Copyright © 2022 Zhenwei Chen 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.

With the increased awareness of environmental protection, people have higher requirements for the accuracy of environmental information of surrounding life. The current monitoring of urban environmental information mainly comes from local environmental weather stations. Although the monitoring equipment of environmental weather stations is better than personal monitoring equipment, the monitoring equipment of weather monitoring stations is too expensive and only suitable for large-scale coarse-grained monitoring. Because the environmental information of a city is affected by factors such as landforms, buildings, rivers, factories, population density, and traffic flow, there are great differences in the environmental information of different areas in a city. Therefore, this study proposes a method that can be used for small-scale and fine-grained environmental information monitoring: the task grid-based urban environmental information release mechanism for mobile crowd sensing (MCS). Through this mechanism, the monitoring area is divided into different task grids according to the characteristics of the area, and the environmental information is sensed by mobile crowd sensing. For the sensing data, through an efficient data fusion algorithm designed in this study, the sensing information is fused to obtain the fine-grained environmental information of different task grids in the area. Through the use of this mechanism, differentiated environmental information can be provided to users in different areas of the city. In a simulation, this mechanism showed higher information accuracy than traditional information release methods. Thus, the mechanism is scientific and has good application value.

1. Introduction

With the increased awareness of environmental protection, more people are paying attention to environmental quality. Owing to the lack of diversified data acquisition methods and channels, traditional urban environment research has often been too one-sided and rough owing to objective reasons such as inaccurate data and the lack of data. Traditional urban environmental information is collected and released by weather stations. Environment monitoring equipment is expensive and not suitable for large-scale deployment. Owing to the large scope of an administrative area, the information released is not accurate enough. Therefore, how to improve the accuracy of environmental information release has become the research goal of many scholars.

1.1. Related Work.

In [1], the authors focused on the long-term observations of monitoring stations, which cannot provide customized data services and in-time emergency response under urgent situations (gas leakage incidents). The authors designed a conceptual framework, multiagents for sensing, monitoring, estimating, and determining (MASMED), to provide fine-grained concentration maps, customized data services, and on-demand emergency management. In the framework, they leverage the hybrid design of wireless sensor networks (WSNs) and mobile crowd sensing (MCS) to sense urban air quality and relevant data (e.g., traffic data and meteorological data). Using the sensed data, they can create a fine-grained air quality map for the authorities and relevant stakeholders. The air quality map can provide people an overall understanding of the distribution of air quality in an area and can help the
authorities and relevant stakeholders quickly analyze pollution sources. However, the main users of this framework are the authorities and relevant stakeholders, whose needs are different from those of ordinary users.

In [2], the authors proposed a novel model by integrating the CNN with BRBES to monitor air quality from satellite images with improved accuracy. The model mainly monitors PM2.5 in the air by analyzing satellite images and has a good monitoring effect. The model can monitor PM2.5 but cannot monitor other parameters of air quality or make an overall evaluation of air quality.

In [3], the aim of the authors was to develop an unmanned aerial vehicle (UAV) equipped with low-cost sensors to monitor air quality both on the move and in a stationary position that is difficult to access. It is equipped with sensors for carbon monoxide (CO), ozone (O3), nitrogen monoxide (NO), and nitrogen dioxide (NO2) measurements. The proposed UAV has good application value. However, for large-scale urban air quality monitoring, a large number of screening nodes need to be deployed, and the cost would therefore be relatively high.

In [4], the authors proposed a smart long-range (LoRa) sensing node to collect the air quality information in a timely manner and update it on the cloud. They developed a long-range wide area network (LoRaWAN)-based Internet of Things (IoTs) air quality monitoring system (AQMS). The system is composed of multiple sensors (NO2, SO2, CO2, CO, PM2.5, temperature, and humidity) and can continuously monitor air quality and is equipped with a rechargeable battery with a photovoltaic solar panel via a solar charger shield for sustainable operation. However, for large-scale urban air quality monitoring, a large number of sensing nodes need to be deployed, and the cost would therefore be relatively high.

1.2. Research Background. The popularity of smart terminals, such as smartphones, with a large number of sensors, has enabled ordinary users to obtain sensing data. A large number of sensors also provide the possibility of accurate release of urban environmental information. Ordinary users use their smart devices to obtain data, which are then submitted to the sink node for fusion; the sink node submits the final data to the sensing platform for publication. This method of obtaining information data by ordinary users using their own smart terminal devices is called MCS [5–14]. Currently, MCS is widely used in many fields, including environmental monitoring [15], traffic conditions [16], and medical health [17].

The applications in these fields have shown that MCS is a useful solution for large-scale data acquisition applications. Generally, the data collection process of MCS consists of four steps [18]:

1. The sensing task is assigned to the mobile terminal by the sensing platform
2. The sensing tasks are performed on the mobile terminal
3. Ordinary users collect data, preprocess the data, and submit the preprocessed data to the sensing platform
4. The sensing platform processes the sensing results submitted by the users

For the current application in urban environmental monitoring, owing to the wide range of data monitoring, the granularity of monitoring data is relatively large. To improve the accuracy of monitoring data, the main task of this study was to improve the way MCS collects and processes data and also improve the accuracy of monitoring data in multitask areas.

1.3. Research Method of This Paper. A multidimensional urban environment data-releasing mechanism based on a task grid for MCS is proposed in this paper. The mechanism works as follows:

1. First, one of the task areas of the geographic area is divided into some task grids. Using the clustering idea in WSN, we regard each task grid as an independent cluster.
2. In each independent cluster, a cluster head terminal is elected, which is responsible for preprocessing the sensing data submitted by other participants. In the preprocessing process, an algorithm designed for removing low-quality data can quickly remove low-quality data and form an effective data matrix for submission to the sensing platform.
3. A data fusion algorithm fuses the data matrix submitted by independent clusters and obtains the optimal estimation of the actual truth value in the task grid.
4. The sensing platform pushes the optimal estimation of the actual truth value to the intelligent terminal in the task grid.

The rest of this paper is organized as follows: Section 2 describes the structure of the system model, and the problems to be solved are studied. Section 3 describes the design of the preprocessing algorithm and data fusion algorithm. The performance of the algorithm is evaluated and analyzed in Section 4. In Section 5, the work of this paper is summarized, and future work is proposed.

2. System Model and Problem Description

2.1. System Model. In MCS, the data are provided by ordinary users participating in the sensing task. Although a large variety of information can be collected in this way, the quality of the collected data can be greatly affected by factors, such as participant uncertainty, device performance, and device heterogeneity [7].

A mechanism for releasing urban environmental information using MCS technology is designed in this paper. The information release area is divided into several task grids. Each task grid is regarded as an independent cluster, and each task grid is identified by geographic coordinates. Each participant collects data in the task grid and sends the sensing data to the cluster head node in this task grid for removing low-quality data. Then, the cluster head node
2.2. Problem Description. Suppose there are \( n \) task grids, \( k \) in the task grid, given by

\[
\begin{pmatrix}
\text{Task grid 1} & \text{Task grid 2} & \text{Task grid 3} & \ldots & \text{Task grid } n
\end{pmatrix}
\]

given by

\[
\begin{pmatrix}
S_{11} & S_{12} & \cdots & S_{1s_m}
\end{pmatrix}
\]

\[
\begin{pmatrix}
S_{21} & S_{22} & \cdots & S_{2s_m}
\end{pmatrix}
\]

\[
\begin{pmatrix}
\vdots & \vdots & \ddots & \vdots
\end{pmatrix}
\]

\[
\begin{pmatrix}
S_{n1} & S_{n2} & \cdots & S_{ns_m}
\end{pmatrix}
\]

...)

Our task is to optimize the matrix \( T \), remove the low-quality data in the matrix \( T \), and fuse the other data to obtain the best real data in each task grid.

2.3. Selection of Participants. Participants with higher reliability are selected to participate in the sensing of data based on past sensing records. This approach ensures the reliability of the sensing data. However, due to inappropriate device operation by the participants and differences in sensing equipment, low-quality data still appear. The sensing platform sends tasks to the participants, the users participating in the sensing task send the location coordinates to the sensing platform, and the sensing platform divides the participants into a certain task grid area according to their location to participate in the sensing task.

For air quality data, timeliness is important; so, the sensing platform releases the time requirements of the sensing task when releasing the task. Generally, the data sensing task must be completed within 30 min, and the task participants must complete the sensing of the data within the corresponding time. If the collected data are not within the set time range, the sensing platform will not receive the data.

For the security of sensing data, the cluster head node of each task grid and the sensing platform share a pair of keys, which are used to encrypt the preprocessed data and upload the data to the sensing platform for data fusion. This approach ensures the security of the data. As the main focus of this paper is to explore the release mechanism of air quality, the safety issues are not stated here in detail.

3. Data Fusion Process

3.1. Preprocessing of the Task Grid Cluster Head

**Definition 1 (Similar data):** Assuming that two data points, \( d_i \) and \( d_j \), are given, if \( r_{ij} = \left| d_i - d_j \right| \leq \beta \), \( \beta \) is a preset threshold, and it is considered that \( d_i \) and \( d_j \) are similar. \( r_{ij} \) is the similarity between \( d_i \) and \( d_j \), and \( d_i \) and \( d_j \) are mutually similar data points.

For any vector \( T_i (1 \leq i \leq s_m) \) in the matrix \( T = \begin{pmatrix} T_1 & T_2 & T_3 & \ldots & T_{s_m} \end{pmatrix} \) and vector \( G_i (1 \leq i \leq s_n) \) in the matrix \( G = \begin{pmatrix} G_1 & G_2 & G_3 & \ldots & G_m \end{pmatrix} \), assuming that the number of elements is 1, \( x (1 \leq x \leq s_n) \), then the number of valid data points is \( x \). The detailed algorithm for removing low-quality data in the matrix \( T \) is presented in Algorithm 1.
Sort the elements of each vector in the matrix \( T = [T_1, T_2, T_3, \ldots, T_{sm}] \) from largest to smallest to obtain a matrix \( T' = [T'_1, T'_2, T'_3, \ldots, T'_{sm}] \).

**Input:** the matrix \( T = [T_1, T_2, T_3, \ldots, T_{sm}] \)

**Output:** the resulting matrix is \( M = [M_1, M_2, \ldots, M_{sm}] \)

1. For \( \forall \) vector \( T'_i (1 \leq i \leq sm) \)
2. Begin
3. For the vector \( T'_i (1 \leq i \leq sm) \)
4. \( \bar{T} = 1/\sum_{i=1}^{x} T'_i \)
5. For \( (y = 1, 1 \leq y \leq x, y + x) \)
6. If \( |t'_y - \bar{T}| \leq \beta \)
7. \( t'_y \) save
8. else
9. \( t'_y = 0 \)
10. End if
11. End for
12. End begin

**Algorithm 1:** Algorithm for removing low-quality data in task grid.

**Figure 3:** Schematic diagram of weighted fusion for reclustering.

To all elements in the vector \( T'_i (1 \leq i \leq sm) \)

\[ T'_i = [T'_1, T'_2, T'_3, \ldots, T'_{sm}] \]

To all elements in the vector

\[ T'_i (1 \leq i \leq sm) \]

\[ t'_y - \bar{T} \leq \beta \]

\[ 1 \leq y \leq x \]

Yes

\[ t'_y \] save

No

\[ t'_y = 0 \]

The resulting matrix

\[ M = [M_1, M_2, \ldots, M_{sm}] \]

3.2. Platform Data Fusion. For each vector \( M_i \) in the matrix \( M \), as the elements in the vector \( M_i \) are similar data, the similar data are reclustering according to accuracy and divided into several more accurate data groups. Then, the data in each group are classified as elements in each layer, as shown in Figure 3.

The sensing platform uses the weighted fusion method to fuse the data in each vector. First, calculate the average of the data for each layer. For example, the average value of layer 1 is \( W_1 = 1/m \sum_{i}^{} w_i \), the average value of layer 2 is \( W_2 = 1/\sum_{i}^{} w_i \), and the average value of layer \( p \) is \( W_p = 1/r \sum_{i}^{} w_i \). After fusion, the data are \( W_1 = m/kW_1 + n/kW_2 + \ldots + r/kW_p \), where \( m + n + r < k \).

The above reclustering weighted fusion process is performed for each vector sequence \( M_1, M_2, \ldots, M_{sm} \) in the data matrix \( M = [M_1, M_2, \ldots, M_{sm}] \), and the result vector \( M' = [W'_1, W'_2, \ldots, W'_{sm}] \) is obtained after the data matrix is fused by the sensing platform. The sensing platform
Figure 4: Continued.
pushes the fusion result $M_{\text{thuth}} = [W'_1, W'_2, \ldots, W'_m]$ to all the smart terminals in the corresponding task grid.

For air quality data, in the process of data fusion, the sensing platform can divide the layers into 10 layers to meet the needs of air quality data fusion accuracy. With more layers, the accuracy of the fusion result is higher. However, with more layers and a larger data matrix, the computational cost increases accordingly. Therefore, the number of layers of the fusion algorithm depends on the specific application environment.

3.3. Algorithm Analysis. The traditional urban environmental information release adopts the method based on the administrative area. Because the administrative area generally has a relatively large geographic scope, this information release method can only roughly reflect the environmental conditions of the area. Environmental information is easily affected by factors such as traffic flow, population density, rivers, factories, and green plant coverage. Therefore, the traditional environmental information release methods are less precise.

The information release mechanism of a crowd sensing urban environment based on the task grid proposed in this paper can effectively improve the release accuracy of information. In this mechanism, the administrative area is divided into several task grids with similar geographical characteristics. The environmental information of each task grid is obtained by crowd sensing, the sensing data are fused in the sensing platform, and the fusion result is used as the best estimate of the environmental information of the task grid. As the environmental characteristics of each task grid are similar and the area is smaller, this mechanism has higher accuracy than the traditional one in releasing environmental information.

4. Simulation Experiments

In this section, we evaluate the proposed task grid-based crowd sensing information release mechanism through simulation. The simulation environment was Windows 10, and the simulation platform was Matlab2016b. The data were from an air quality online monitoring and analysis platform (https://www.aqistudy.cn), and five different locations in Lu’an City were used as the task collection area. The air quality comprehensive index is mainly determined by the following six factors: particulate matter (PM10 and PM2.5), ozone ($O_3$), sulfur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), and carbon monoxide (CO) [19]. Thus, the data in the simulation were also six-dimensional air quality data.

4.1. Effect of Algorithm Fusion. As the monitoring instruments of the environmental monitoring station were in task grid 4, we verified the fusion accuracy of the data in task grid 4. There are many types of data fusion algorithms. Air quality data are numerical data. At present, the fusion algorithms widely used for numerical data include weighted average [20] and BP neural network [21]. We used the weighted average algorithm, BP neural network, and the algorithm designed in this paper to fuse the sensing data in task grid 4 and compared the fusion result with the data monitored by the environmental monitoring station to judge the accuracy of the data fusion. The simulation results are shown in Figure 4.
Table 1: Average error of three algorithms in data fusion.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>PM2.5 (%)</th>
<th>PM10 (%)</th>
<th>SO2</th>
<th>NO2 (%)</th>
<th>CO (%)</th>
<th>O3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted average</td>
<td>8.08</td>
<td>7.75</td>
<td>8.96%</td>
<td>19.6</td>
<td>23.56</td>
<td>7.79</td>
</tr>
<tr>
<td>BP neural networks</td>
<td>3.77</td>
<td>3.66</td>
<td>6.31</td>
<td>10.9</td>
<td>16.66</td>
<td>4.37</td>
</tr>
<tr>
<td>The algorithm designed in this paper</td>
<td>3.83</td>
<td>3.03</td>
<td>3.26%</td>
<td>10.3</td>
<td>16.78</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Figure 5: Curve of PM2.5 change in different task grid areas.

Figure 6: Curve of PM10 change in different task grid areas.

Figure 7: Curve of SO2 change in different task grid areas.

Figure 8: Curve of CO change in different task grid areas.
As shown in the experimental data statistics in Table 1, the algorithm designed in this paper and the BP algorithm had the highest accuracy. It can be seen from the statistics in Table 1 that the fusion results of the six parameters of air quality by the fusion algorithm designed in this paper meet the needs of air quality monitoring.

4.2. Simulation of the Release Mechanism. Next, we used the fusion algorithm designed in this paper to simulate the six indicators of air quality in the five task grids. The fusion results in the five task grids and the data released by the weather station were compared to prove the advantages of the air quality information release mechanism designed in this paper. The experimental results are shown in Figures 5–10.

As shown in the experimental data statistics in Table 2, by comparing the average value of the sensing data in the five task grids with the average value of the published data, we found that the data in the five task grids are significantly different from the published data. For example, in task grid area 5, the differences in the six indicator data points and the
The data collection and simulation results show differences between areas, with high information accuracy and releasing air quality information can clearly reflect the diffusion algorithm proposed in this paper. This way of releasing the data to the terminals in the task grid according to the platform through the network. The sensing platform pushes cluster head submits the preprocessed data to the sensing through crowd sensing for data preprocessing, and the sends the collected data to the cluster head of the task grid.

The system mechanism based on a task grid is proposed. The system enables users to obtain information with higher accuracy than the traditional urban environmental information release method.

The data released by the weather station can only represent the air quality around the collection point of the air quality monitoring station and cannot represent the air quality of the entire area well. Moreover, there are large differences in the air quality between different locations. Therefore, the task grid-based crowd sensing urban environmental information release mechanism proposed in this paper has great advantages over the traditional environmental quality release mechanism. This mechanism can accurately release the air quality information of each task grid, so that the people entering the area can understand the air quality of the area more accurately. Moreover, the release mechanism collects and releases information through crowd sensing. Compared with traditional air quality information monitoring, the proposed method does not require installing a large number of air monitoring equipment pieces and has the advantages of low cost, wide coverage, information differentiation, and high information accuracy.

5. Conclusion

Because traditional air quality information is collected and released by a monitoring station through relevant equipment, this method cannot reflect the impact of the environment in each area; so, the traditional air quality information release method is inaccurate. To reflect the difference in air quality information in each task grid and the impact of the surrounding environment and other factors on air quality, we propose the crowd sensing approach to collect multidimensional data, whereby an information release mechanism based on a task grid is proposed. The system sends the collected data to the cluster head of the task grid through crowd sensing for data preprocessing, and the cluster head submits the preprocessed data to the sensing platform through the network. The sensing platform pushes the data to the terminals in the task grid according to the fusion algorithm proposed in this paper. This way of releasing air quality information can clearly reflect the differences between areas, with high information accuracy and low cost. The data collection and simulation results show that the information release mechanism proposed in this paper is scientific and has a good application value.

Data Availability

The data used in the simulation are from the air quality online monitoring and analysis platform (https://www.aqistudy.cn).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank everyone who helped during the research and preparation of the article. This research was supported in part by the 2020 Key Project of Natural Science Research in Colleges and Universities in Anhui Province (KJ2020A0623), the 2021 Visiting Research Program of West Anhui University (wxxygnfx2021002), the 2021 Major Natural Science Research Project of Colleges and Universities in Anhui Province (KJ2021ZD0116), and the 2021 Key Natural Science Research Projects of Colleges and Universities in Anhui Province Project (KJ2021A0937).

References


