IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 7, 2019, accepted October 27, 2019, date of publication November 5, 2019, date of current version December 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2951587

Data Collection Through Mobile Vehicles in Edge Network of Smart City

YUEYI LUO[1,3], **XIAOYU ZHU**[2], **AND JUN LONG**[1,4]
[1]School of Computer Science and Engineering, Central South University, Changsha 410083, China
[2]School of Automation, Central South University, Changsha 410083, China
[3]School of Mathematics and Statics, Central South University, Changsha 410083, China
[4]Network Resources Management and Trust Evaluation Key Laboratory of Hunan Province, Central South University, Changsha 410083, China

Corresponding author: Xiaoyu Zhu (zhuxiaoyu@csu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472450, in part by the Key Technology Research and Development Program of Hunan Province under Grant 2018GK2052, in part by the Hunan Provincial Education Department of China under Grant 18B200, and in part by the Hunan Provincial Natural Science Foundation under Grant 2017JJ2333.

ABSTRACT With increasing technological facilities, billions of sensor devices are deployed in the smart city, which allows to sense and collect the status and information of all kinds of infrastructures in the city. Those large number of sensing nodes form many self-organized wireless sensor networks (WSNs). Although the single WSN has been widely studied by researchers. However, few studies have focused on how to effectively connect to Internet of Things (IoT) and collect data in many decentralized WSNs of smart city with a low cost. In this paper, a cluster head rotation joint mobile vehicle data collection (CHR) scheme is proposed to effectively collect data from many decentralized WSNs in smart city with a low cost. In CHR scheme, each WSN selects one or multiple nodes which can connect to mobile vehicles as the cluster heads. Then all the in-network sensor nodes send their data to cluster head through multi-hop communication, due to the cluster head can connect to mobile vehicles, so when mobile vehicles pass by cluster head, the data of cluster head can be sent to mobile vehicles and brought to the data center to process. We first propose a single cluster head rotation joint mobile vehicle data collection (SCHR) scheme to collect data by only using a single CH in a WSN in which the CH rotation and clustering algorithms are carefully designed to balance in-network energy consumption. Then multiple cluster head rotation joint mobile vehicle data collection (MCHR) scheme is proposed to further balance the energy consumption and prolong the network lifetime in which multiple cluster heads are selected to be jointly responsible for the data collection task. The extensive experiments show that the CHR scheme has good performance in the network energy, network lifetime and network transmit capacity.

INDEX TERMS Smart city, wireless sensor network, data collection, mobile vehicle, network lifetime.

I. INTRODUCTION

With the rapid development of Internet of Things (IoT) [1], [2], cloud computing and 5G technology, it is possible to connect different kinds of things into the Internet. More and more devices are connected to the IoT. It's estimated that more than 20 billion devices will join the IoT in 2020, billions of devices that didn't connect into the Internet produce EB-level data every day. The smart city is the trend of future urban development, IoT plays an indispensable role in the construction of smart city [3]–[5]. Sensors are widely used in the process of building smart city, which can be embedded into the infrastructures. IoT uses sensors to connect different kinds of infrastructures into internet, which makes the dream of smart city become true gradually. Through extensive network interconnection, a large amount of data is quickly collected and processed. The processed data provides scientific decision-making support for municipal departments, which promotes the intelligent management and development of cities.

With the development of micro-processing technology, the sensing devices have stronger computation power and more functions, thus they can adapt to more fields and have more applications. A large number of sensing nodes are flexibly deployed in many sensing areas according to the needs of applications. There are lots of sensing devices deployed in

The associate editor coordinating the review of this manuscript and approving it for publication was Xuxun Liu.

the smart city, the smart city can manage the infrastructures through the data collected by the sensors [6]. For example, the grassland is embedded with humidity sensors, if the grassland is not humid enough, then the sensors will send the watering request; The urban buildings are equipped with sensors for measuring carbon emissions, if the carbon emission of the building exceeds the standard, then the sensor will send an alert message. The temperature sensors, fire sensors, traffic cameras and other kinds of sensors are embedded into the infrastructures in smart city. The sensors collect the environmental monitoring data, surveillance data, traffic data and other kinds of data. In the smart city, the data collected by the sensors are transferred to the relevant municipal apartments, then the municipal apartments can take measures to manage the infrastructures. The data collection in smart city can help solve the problems of urban pollution, traffic management, and environment protection.

Although the sensing devices can be quickly and flexibly deployed in the city, which will help to realize the beautiful prospects of the smart city. However, how to collect the data sensed by those sensing devices is a challenging issue, and the data collection of this sensing-based network is different from the traditional widely researched data collection solutions in wireless sensor network (WSN) [7]–[9]. The previous WSN schemes mostly studied a connected network which is composed of many sensor nodes and a sink. The data sensed by each node is routed to the sink through multi-hop communication, and the sink is connected to the Internet, thus the data collection can be realized. However, the sensing network in smart city is different from the traditional WSN [10]–[14]. In the sensing network of smart city, large number of sensing nodes are deployed by the needs of the applications, there is not even a basic communication device to connect to the Internet. Thus, such kind of network can be abstracted as several isolated WSNs composed of many sensors in smart city [15]. For these isolated WSNs, although its interior can be organized into a connected network, these connected networks only contain ordinary nodes, and they don't have sinks which can connect to the Internet. For example, the network deployed with the advance of road construction does not have the basic condition to connect to the Internet. In addition, to deploy a dedicated network to connect Internet also needs to consider the construction cost. To deploy the dedicated networks for many temporary networks is neither timely nor economical. These traditional data collection schemes can only solve the data collection problem in a single WSN, while they cannot be applied to solve the data collection problem in many isolated networks.

Bonola *et al.* [4] proposed a data collection scheme through mobile vehicles using opportunistic communication, they observe that there is a large number of mobile vehicles in the city [16]–[18]. The transceivers are integrated into the chips of the vehicles, which can exchange data with the nearby sensors and data center. In this situation, the sensor nodes near the road can be acted as sink nodes. Thus, we can use opportunistic communication of mobile vehicles to collect data in smart city. When the vehicle travels by the intelligent devices, the sensor nodes embedded in the intelligent devices will send the sensory data of the infrastructures through their communication module. The corresponding applications further analyze and process the data extracted from the data center, which forms intelligent decision-making and realizes the dream of smart city. The solution proposed by Bonola *et al.* [4] is an attempt to collect data through mobile vehicles, they observe that the vehicles can finish the data collection after a period of time since there are many mobile vehicles in smart city. However, the research of [4] is a preliminary attempt, there are many other issues that are not addressed. Their scheme only involves data collection problem of single sensing device along the road, they do not involve that the sensing devices in the smart city that can connect into many networks in a self-organized way.

In this paper, we study the data collection problem in the sensing network of smart city. The practical significance, complexity, and challenge of this kind of data collection studies are far greater than the previous researches. Although many data collection schemes have been proposed, none of the existing schemes can be directly applied to solve the data collection problem with an effective and low-cost way in a large number of isolated WSNs deployed in smart city. We propose a data collection scheme through mobile vehicles to collect data in edge network of smart city. The in-network data can route to the sensor nodes along the road through multi-hop communication, the sensor nodes along the road can collect data through mobile vehicles, which can realize the data collection network cover the whole city in a low-cost way. Compared with the previous researches, the main contributions of this paper are as follows:

(1) A cluster head rotation joint mobile vehicle data collection (CHR) scheme is proposed to effectively collect data from many isolated WSNs in smart city with a low-cost style. Our study in this paper extends the shortcomings of the previous studies which only considers collecting data for the sensing devices near the roads. The CHR scheme not only can collect data for the sensing devices near the roads, but also can effectively collect data in those self-organized WSNs if only one sensor node can connect with the mobile vehicles.

(2) We propose two in-network data collection schemes. The one is the single cluster head rotation joint mobile vehicle data collection (SCHR) scheme, the other is multiple cluster head rotation joint mobile vehicle data collection (MCHR) scheme. The SCHR scheme collects data by only using a single CH in a WSN in which the CH rotation and clustering algorithms are carefully designed to balance the in-network energy consumption. The MCHR scheme is proposed to further balance the energy consumption and prolong the network lifetime in which multiple cluster heads are selected to be jointly responsible for the data collection task, so the energy consumption in the network collection, delay, lifetime and other performance metrics can be furtherly improved. The cluster head rotation algorithm and clustering algorithms based on the cost function are proposed to balance the energy

consumption of nodes and prolong the network lifetime for those two data collection schemes.

(3) The extensive experimental results show that our proposed CHR schemes have good performance in the metrics of energy consumption, network lifetime and network transmission capacity. The MCHR scheme has a better performance than SCHR scheme.

The rest of this paper is organized as follows: In Section II, we introduce the related work. The network model and problem statement are presented in Section III. Then, the CHR schemes are introduced in Section IV. The experiment results of CHR schemes are presented in Section V. Finally, Section VI gives the conclusions.

II. RELATED WORK

The data collection framework in the smart city contains two tiers: the data collection tier and the cloud tier. In the data collection tier, the sensor nodes are embedded into the infrastructures in smart city, which are used to detect the status of the infrastructures. The sensor nodes transferred the status data to the vehicles when the sensor nodes are within the communication range of the vehicles. Then the vehicles send the status data to the data center when the vehicles move near the data center. In the cloud tier, the data centers send data to the cloud, the cloud notifies the municipal departments to maintain the infrastructures.

In this paper, we focus on the data collection tier, which is closely related with cloud computing [19], wireless sensor networks (WSNs) [20]–[23] and vehicular network [24], [25]. The data collection research is a hot topic in WSNs. However, WSNs has some unique shortcomings. The sensor nodes contain limited energy, and the communication energy consumption is large, thus the lifetime of the sensor nodes is short.

A lot of clustering algorithms had been proposed by the researchers, for example, LEACH [26], EEHC [27], HEED [28]. These data collection solutions are based on the base station, the sensor nodes transfer data to the base station through single-hop or multi-hop communication. Although these solutions can extend the lifetime of the WSNs to a certain extent, the nodes around the base station still bear large overhead and have shorter lifetime than the other nodes far away from the base station. When the nodes around the base station consume all the energy, the nodes far away from the base station cannot communicate with the base station although their residual energy is still very high.

In order to solve the problem of unbalanced energy consumption, the mobile sinks (MSs) are introduced into WSNs, MS may visit each sensor node and collect data through single-hop communication [29], or visit some locations of WSNs and the sensor nodes transmit the data to the vehicles through multi-hop communication [30]. The single-hop communication solution can reduce energy consumption, but it has a high delay. The multi-hop communication solution has low delay and high communication overhead. In addition, the vehicles update the location constantly, which results in large routing overhead for the sensor nodes. Some approaches [31], [32] have been proposed to transfer data to some buffer nodes, the buffer nodes cache the received data and transfer data to the vehicles when the vehicles are within the communication range.

In recent years, there are some surveys [33], [34] of MS applied on WSNs. The MSs are divided into three types based on the mobility type. The three types are random mobility, controllable mobility, and constrained mobility, respectively.

In the random mobility approaches [35], [36], the MS is randomly walking in WSNs, the MS receives the data when it travels near the sensor node. The random mobility approaches can balance the energy consumption in some extent, but due to the randomness and uncertainty of MS's moving trajectory, some sensor nodes cannot be accessed for a long time, which result in data overflow and high delay.

In the controllable mobility approaches [37], [38], the MS's trajectory is designed according to the actual deployments of the sensor nodes. The MS can adjust the movement based on the real-time network situation and nodes' feedback information. According to different performance metrics, such as the data collection rate and real-time, the controllable mobility approaches can design different trajectories to achieve the best performance of one performance metric. However, the MS is mainly traveling along with the road segments, the control movements cannot all be implemented, because there are many limitations in the real situation.

In the constrained mobility approaches [24], [39]–[42], the MS moves along the predefined trajectory to collect data. The authors of [24] focus on the maximization problem of the data collection rate, a single-hop communication solution is proposed to solve this problem, all the sensor nodes are within the communication range of MS. In [39], the authors solve the scheduling problem from node to MS, and the proposed solution balanced the data collection rate and energy consumption. The authors of [40] propose a multi-hop communication solution to collect data. Each node knows the shortest path to the nodes within the communication range of MS. In [41], the authors focus on improving the throughput, the proposed solution balances the received data for each node within the communication range of MS and improves the throughput. The authors of [42] focus on the nonuniform node distribution in WSNs and propose an unequal clustering algorithm to balance the energy consumption.

The authors of [4] propose a data collection scheme through mobile vehicles using opportunistic communication. However, their scheme only involves data collection problem of single sensing device along the road, they do not involve that the sensing devices in the smart city that can connect into many networks in a self-organized way. Few studies have focused on how to effectively connect to the Internet of Things (IoT) and collect data in many decentralized WSNs of smart city with a low cost.

In this paper, the sensors are deployed to sense the environment parameters. The core issue is to design an energy

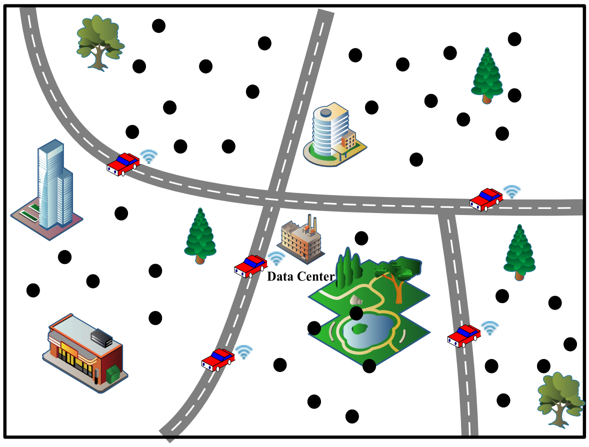

FIGURE 1. **The network model in smart city.**

balancing routing algorithm. In smart city, the sensor nodes are embedded into the infrastructures in smart city, which are used to detect the status of the infrastructures. The collected data are transferred to the vehicles when the sensor nodes are within the communication range of the vehicles. Then the vehicles send the status data to the data center when the vehicles move near the data center. Finally, the data center send data to the cloud to notify the corresponding department to maintain the infrastructures.

III. SYSTEM MODEL AND PROBLEM STATEMENT

A. NETWORK MODEL

The network model contains a lot of sensors, and the network is divided into many isolated WSNs. Figure 1 is the network model in smart city, the network contains:

(1) Sensor nodes: The black point denotes the sensor node. The sensor nodes are embedded in the urban infrastructures, which are used to detect the status of the infrastructure. Once certain conditions are satisfied, the sensor nodes will send the collected data to the vehicle. Some nodes are selected as cluster heads (CHs) from the nodes that can directly communicate with the vehicles. Some other nodes transfer the data to the CHs through single or multi-hop communication, then the CHs transfer the data to the vehicles when the vehicles passing by. We assume that all the sensors have the same energy.

(2) Mobile vehicles: The mobile vehicles travel along the roads. The vehicles are responsible for receiving the data collected by the sensors and transmit the data to the data center. Vehicles can exchange data with sensor nodes and data center deployed in the smart city. The vehicles use opportunistic communication to collect data from the sensor nodes and transmit the data to the data center via short-range communication. In smart city, a large number of vehicles move around the city every day, for example, the taxi cabs work 24 hours a day and their routes cover almost the whole city. We do not force vehicles to drive according to the predefined route, the vehicles follow their own driving habits, for example, the taxi cabs drive along their own planned routes to pick up passengers, the sensor nodes send data to the vehicles through opportunistic communication. When the vehicles move along the city, they happen to pass by the sensor nodes and communicates with the nodes with a low cost. In this way, we can collect data in a low-cost and effective way through vehicles. We assume that the vehicles use batteries as energy resources, and they have no energy limitation. The vehicles integrate with transceiver chips, and they are equipped with wireless communication technology and storage capacity. The vehicles broadcast signals during the movement.

(3) Data center: The data center is located at a fixed position in the smart city. The data center can receive the data transferred by the vehicles through wireless communication, and the data center can be regarded as the endpoint of the data transmission.

B. ENERGY CONSUMPTION MODEL

The energy consumption model mainly considers the following factors: the amount of transferred data, the consumed energy of sending and receiving data. In order to compute the consumed energy, this paper uses a common energy consumption model. The consumed energy for sensor node s to send l bits data from hop h to hop $(h+1)$ is:

$$E_{send}(s,h) = \begin{cases} l \times E_{elec} + l \times \varepsilon_{fs} \times r^2 & r < r_0 \\ l \times E_{elec} + l \times \varepsilon_{amp} \times r^4 & r \geq r_0 \end{cases} \quad (1)$$

E_{elec} represents the consumed energy in the transmit circuit, r is the data transmission distance. If the transmission distance is less than the threshold r_0, the power amplification loss adopts the free space model, and the multi-path decay model is used when the transmission distance is greater than or equal to the threshold r_0. ε_{fs} and ε_{amp} represent the required energy for power amplification in the free space model and the multipath decay model, respectively. The consumed energy for sensor s to receive l bits data is:

$$E_{receive}(s,h) = l \times E_{elec} \quad (2)$$

IV. PROPOSED SOLUTION

A. RESEARCH MOTIVATION

In smart city, a large amount of sensors are deployed according to the needs of the application. The smart city can be divided into many separate regions. In the interior of the region, the sensor nodes can form a connected network in a self-organized way, the interior network cannot connect to the Internet without sink node. The regions cannot communicate with each other due to the limitation of the communication range. Thus, the sensing network of smart city can be abstracted into many disconnected WSNs. The traditional WSN data collection schemes can collect data for a single connected network, but these schemes cannot solve the data collection problem in smart city for many disconnected WSNs.

There are many vehicles in the smart city, which are integrated with the transceiver chips and can communicate with sensors and data center when they travel along the roads. However, the existing data collection schemes based on vehicles only involve the single sensing device along the road, they do not involve the sensing devices which are far away from the roads and cannot communicate with the vehicles. There does not exist a data collection scheme which can collect data for many decentralized and disconnected WSNs in smart city. Therefore, it is necessary to provide a low-cost and effective data collection scheme for the sensing network in smart city.

In this paper, we propose a cluster head rotation joint mobile vehicle data collection (CHR) scheme to collect data in smart city. In CHR scheme, each WSN selects one or multiple nodes which can connect to mobile vehicles as the cluster heads. Then all the in-network sensor nodes send their data to cluster head through multi-hop communication, due to the cluster head can connect to mobile vehicles, so when mobile vehicles pass by cluster head, the data of cluster head can be sent to mobile vehicles and brought to data center to process. Firstly, we propose a single cluster head rotation joint mobile vehicle data collection (SCHR) scheme. Then, we propose a multiple cluster head rotation joint mobile vehicle data collection (MCHR) scheme.

B. THE DESIGN OF SCHR

In this subsection, we propose a cluster head rotation joint mobile vehicle data collection (SCHR). In the sensing network of smart city, the sensor nodes can perceive the status of the infrastructures and send data to the data center through mobile vehicles. The mobile vehicle has a limited communication range, the nodes within the communication range can directly communicate with the vehicle. However, there are some other nodes far away from the vehicles, which need to find a route path to the vehicle.

In SCHR scheme, each WSN selects one single CH to communicate with the vehicle, all the other nodes send the data to the CH, the CH receives the data and aggregates the data, then the CH transfers the aggregated data to the vehicle. The CH needs to consume more energy than other nodes, if the CH is fixed in the WSN, the CH will consume all the energy. If the CH is dead, all the other sensor nodes cannot communicate with the vehicles, the network lifetime will decrease largely. Thus, in order to avoid that the CH exhausts all the energy in the network, we rotate the CH based on the predefined cost function.

The SCHR scheme has several characteristics. Firstly, the sensing network can be abstracted into many isolated WSNs, each WSN selects one single node to be the CH. Secondly, the node closer to the vehicle and contains enough residual energy has a higher probability to be selected as a CH. Thirdly, the CH rotation strategy chooses different CHs in each round, which can balance the workload of the network.

Figure 2 shows the network model of SCHR, where the black points are the sensor nodes, which are responsible for sensing the status of the infrastructures in the smart city, collecting data and routing data to the CH. The red points are the CHs, we can find that each WSN has only one CH. The CH is responsible for collecting the data, aggregating the data and sending data to the vehicles when the vehicles passing by. The CH rotation strategy is based on the cost function, which is related to the location of the vehicles and residual energy.

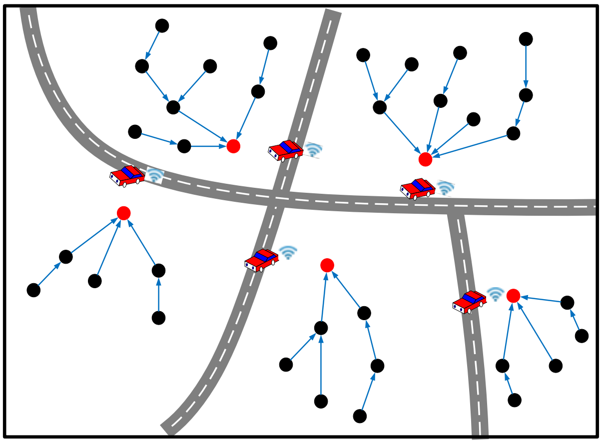

FIGURE 2. **The network model of SCHR.**

1) STAGE 1: OBTAIN THE HOP DISTANCE

Firstly, the vehicle travels along the road once in order to obtain the hop distances from the sensors to the vehicle. When the vehicle moves, the vehicle broadcasts the signal, the nodes within the communication range receive packets (including the hop distances, the initial hop distance is 0). The node increases the hop distance in the packet by 1, and regard it as its own hop distance to the vehicle. Then, the node broadcasts the modified message to the node within its communication range. Repeat this strategy until no new broadcast occurs. One node may receive several messages, if the hop distances are different, then select the smallest hop distance. In the end, each node knows its hop distance to the vehicle.

As shown in Figure 3, the WSN contains 10 sensors, represented as $n_1, n_2, \ldots, n_{10}$. In the initialization phase, each node obtains its hop distance to the vehicle, n_4 and n_8 are within the communication range, the hop distance to the vehicle is 1, n_4 and n_8 increase the hop distance by 1 and broadcast the packet. n_2, n_5, n_7 and n_9 receive the packet, increase the hop distance by 1, the modified hop distance is 2. Each node computes the hop distance until no broadcast signal occurs. Finally, the hop distances of n_4 and n_8 are 1, the hop distances of n_2, n_5, n_7 and n_9 are 2, the hop distances of n_1, n_3, n_6 and n_{10} are 3.

2) STAGE 2: CH ROTATION

After the vehicle receives the signal, the nodes whose hop distance are 1 will be selected as the CH candidate nodes. Then, calculate the transferring costs from the CH candidate nodes to the vehicle according to the defined cost function. The node with the minimum cost will be selected as the CH.

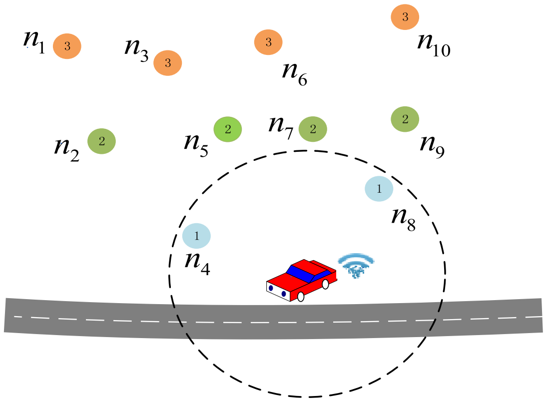

FIGURE 3. Get the hop distance.

In the cost function, the distance from the node to the vehicle has a large impact on the transferring cost, because the consumed energy of the node is proportional to the transferring distance. If the CH's distance to the vehicle is very far, the CH will consume large energy in the data transmission process. Thus, the distance from the node to the vehicle is selected as an impact factor.

In addition, the residual energy of the nodes has a large impact on the cost, if the residual energy of the CH is very small, then it may become a dead node in the very early process, the other cluster members cannot communicate with vehicle if the CH is dead. Thus, we introduce residual energy into the cost function.

In the cost function, the distance from the node to the vehicle and the residual energy of the nodes are selected as the impact factors in the CH rotation process.

The distance factor is defined as:

$$w_D = \frac{D(n_i, MS)}{\max(D(n_i, MS))} \tag{3}$$

The energy factor is defined as:

$$w_E = 1 - \frac{E(n_i)}{\max(E(n_i))} \tag{4}$$

The cost function is defined as:

$$\text{cost} = \alpha w_D + (1-\alpha) w_E \tag{5}$$

$D(n_i, MS)$ represents the distance from the candidate node n_i to the vehicle, $\max(D(n_i, MS))$ represents the maximum value of $D(n_i, MS)$, $E(n_i)$ denotes the residual energy of n_i, $\max(E(n_i))$ denotes the maximum value of $E(n_i)$, α denotes the weight of the factors.

w_D is the weight of distance, which is the ratio of $D(n_i, MS)$ to $\max(D(n_i, MS))$, the smaller the $D(n_i, MS)$, the smaller the cost. w_E is the weight of the energy, which is the ratio of $\max(E(n_i)) - E(n_i)$ to $\max(E(n_i))$, the larger the $E(n_i)$, the smaller the cost. Finally, select the CH with the minimum cost based on equation (5).

We assume that the GPSs of nodes and vehicle are known, and calculate the distances based on the GPSs. Algorithm 1 shows the equation of calculating distance. The GPSs of two points a and b are (lan_a, lon_a) and (lan_b, lon_b), lan_a and lon_a are point a's latitude and longitude, lan_b and lon_b are point b's latitude and longitude. The distance between the two points is calculated according to the GPS.

Algorithm 1 Calculate Distance

```
Input: Two points $a$ and $b$, $R$ is the radius of the earth
Output: The distance between $a$ and $b$
$\Delta\psi = lan_a - lan_b$
$\Delta\gamma = lon_a - lon_b$
$\Delta\eta = \arcsin\left(\sqrt{sin^2\left(\frac{\Delta\psi}{2}\right) + \cos(lan_a)\cos(lan_b)\, sin^2\left(\left(\frac{\Delta\gamma}{2}\right)\right)}\right)$
$D(a, b) = 2 \times \Delta\eta \times R$
```

Algorithm 2 The CH Rotation of SCHR

```
$\max(D(n_i, MS)) = 0$, $\max(E(n_i)) = 0$,
  $\min(\text{cost}) = \infty$
MS broadcasts the CH-Compete-Msg
While $h_i = 1$ Do
    $n_i$ broadcasts the CH-Candidate-Msg
End While
For each CH-Candidate-Msg MS received from $n_i$ Do
    Extract $n_i$ and $E(n_i)$
    Calculate $D(n_i, MS)$ using Algorithm 1
    If $\max(D(n_i, MS)) < D(n_i, MS)$ then
        $\max(D(n_i, MS)) = D(n_i, MS)$
    End if
    If $\max(E(n_i)) < E(n_i)$ then
        $\max(E(n_i)) = E(n_i)$
    End if
End for
For each CH-Candidate-Msg MS received from $n_i$ Do
    $w_D = \frac{D(n_i,MS)}{\max(D(n_i,MS))}$
    $w_E = 1 - \frac{E(n_i)}{\max(E(n_i))}$
    $\text{cost} = \alpha w_D + (1-\alpha) w_E$
    If $\min(\text{cost}) > cost$ then
        $\min(\text{cost}) = cost$
    End if
End for
MS chooses the node $n_i$ with the minimum cost
    $\min(\text{cost})$ as CH
```

In the SCHR scheme, we choose one CH with minimum cost in each WSN. The nodes consume energy in each round, the consumed energy is related to the amount of transferred data and transferring distance. The CH consumes more energy than the other nodes as the CH transfers more data. The CH rotation is based on the cost function, which is related to the distance to the vehicle and the residual energy. The CH rotates as the residual energy of nodes change in each round.

Algorithm 2 shows the CH rotation process of SCHR. The vehicle travels along the road and broadcasts the CH-Compete-Msg, which contains vehicle ID. If the residual energy of n_i is larger than the threshold value of T, and the hop distance of n_i is equal to 1, then n_i becomes a candidate CH node. Next, each candidate node broadcasts the CH-Candidate-Msg, which contains the node ID, node GPS (lan_{n_i}, lon_{n_i}), and the residual energy $E(n_i)$. The vehicle selects the node with the minimum cost as the CH. The network selects the node closer to the vehicle and larger residual energy as the CH, which will enhance the lifetime of the CH and balance the energy consumption.

3) STAGE 3: THE CLUSTERING ALGORITHM

In the clustering algorithm, the CH (hop distance is 1) broadcasts the message that it can be the parent node, the nodes within CH's communication range start to select the parent node, these nodes select CH as the parent node because there is only one CH in one WSN. After nodes (hop distance is 2) select the parent node, they broadcast the message that they can be the parent nodes. This process repeats until all the nodes select the parent nodes. When the clustering process finishes, all the nodes in the WSN can find a path to the CH and transfer their data to the vehicle.

Assume that a node n_i (hop distance is h) is the source node, it needs to select a parent node (hop distance is $h-1$). The source node n_i may receive many messages from the candidate parent nodes, it selects the node with the minimum cost as its parent node.

After receiving the broadcast messages from the candidate parent nodes, the source node calculates the costs to several candidate parent nodes according to the cost function. In the parent node selection process, select the node with minimum cost as the parent node, the impact factors are the transmitting distance and residual energy. Select the node with large residual energy and small transmitting distance as the parent node, which can balance the energy, avoid the premature death of the nodes in the transmitting path, and enhance the network lifetime.

In cost function, the distance from the source node to the vehicle has a large impact on the data transferring cost, because the consumed energy is proportional to the data transferring distance. If the distance from the source node to the vehicle is very far, it will consume large energy. Thus, we introduce the distance from the source node to the vehicle into one impact factor in the cost function.

In addition, in the path from the source node to the vehicle, the minimum energy in the path is the bottleneck of this path. In order to avoid the premature death of the node with minimum energy, we need to assure that the node with the minimum energy in all the paths is the maximum. Thus, the residual energy is introduced in the cost function.

In the process of selecting parent node n_k for the source node n_i, we let the distance from the source node to the vehicle and the residual energy as the impact factors in the cost function.

The distance factor is:

$$w_D(i,k) = \frac{D(n_i, n_k, MS)}{\max(D(n_i, n_k, MS))} \tag{6}$$

The energy factor is:

$$w_E(i,k) = 1 - \frac{E_{min}(n_i, n_k, MS)}{\max(E_{min}(n_i, n_k, MS))} \tag{7}$$

The cost function is:

$$\text{cost}(i,k) = \alpha w_D(i,k) + (1-\alpha) w_E(i,k) \tag{8}$$

where $D(n_i, n_k, MS)$ is the distance from n_i to vehicle MS, $\max((D(n_i, n_k, MS))$ is the maximum value in all the values of $D(n_i, n_k, MS)$, $E_{min}(n_i, n_k, MS)$ is the minimum energy in all the nodes of the path from n_i to n_k, $\max(E(n_i, n_k, MS))$ is the maximum value in all the values of $E_{min}(n_i, n_k, MS)$, α is the weight to adjust the factors.

$w_D(i,k)$ is the weight of the distance, the smaller the distance from the source node to the vehicle, the smaller the consuming energy of transferring data from the source node to the vehicle. Thus, the smaller the $D(n_i, n_k, MS)$, the smaller the $w_D(i,k)$, the smaller the $\text{cost}(i,k)$.

$w_E(i,k)$ is the weight of the residual energy, the larger the minimum energy in the path, the smaller the probability that the node with the minimum energy becomes the bottleneck of the path, the smaller the route cost. Thus, the larger the $E_{min}(n_i, n_k, MS)$, the smaller the $w_E(i,k)$, the smaller the $\text{cost}(i,k)$.

When the source node selects the parent node, it may receive several messages from the candidate parent nodes, the source node can select the parent node according to the calculation results based on the cost function. For example, if n_i receives the messages from n_p and n_q, it needs to select one node from n_p and n_q as the parent node. n_i computes the costs of transferring data from n_i to n_p and n_q. The costs of transferring data from n_i to n_p and n_q are denoted as $\text{cost}(i,p)$ and $\text{cost}(i,q)$, respectively. If $\text{cost}(i,p) \leq \text{cost}(i,q)$, then n_i selects n_p as the parent node. If $\text{cost}(i,p) > \text{cost}(i,q)$, then n_i selects n_q as the parent node.

The parent node selection algorithm is given in Algorithm 3, which describes the process of source node selecting a parent node from several candidate parent nodes. First, initiate $\max(D(n_i, n_k, MS))$ and $\max(E(n_i, n_k, MS))$ as 0, set $\min(\text{cost}(i,k))$ as ∞. For each message received from candidate parent nodes, n_i extracts (x_k, y_k), $D(n_k, MS)$, $E_{min}(n_k, MS)$, and $\max(E(n_i, n_k, MS))$. Next, compute $w_D(i,k)$ and $w_E(i,k)$, then compute the cost $\text{cost}(i,k)$. Finally, n_i selects n_k with minimum cost $\min(\text{cost}(i,k))$ as its parent node.

In the clustering algorithm, each source node needs to select a parent node, which minimizes the data transferring cost from the source node to the parent node. Each node whose hop distance is $h+1$ selects one node with the minimum cost from the candidate nodes whose hop distances are h. In each round of parent node selection process, the candidate parent node and the source node hear and send messages interactively.

The SCHR clustering algorithm is shown in algorithm 4. In SCHR, each WSN has only one CH. Firstly, the source nodes within the communication range of CH select the CH

Algorithm 3 Select Parent Node

```
$\max(D(n_i, n_k, MS)) = 0$, $\max(E\ (n_i, n_k, MS)) = 0$, $\min(\text{cost}(i, k)) = \infty$
For each message $n_i$ received from the candidate node $n_k$ Do
    Extract the GPS of $n_k$, $D(n_k, MS)$, $E_{min}(n_k, MS)$ from the message
    Calculate $D(n_i, n_k)$ using Algorithm 1
    $D(n_i, n_k, MS) = D(n_i, n_k) + D(n_k, MS)$
    If $\max(D(n_i, n_k, MS)) < D(n_i, n_k, MS)$ Then
        $\max(D(n_i, n_k, MS)) = D(n_i, n_k, MS)$
    $E\ (n_i, n_k, MS) = \min(E_{n_i}, E\ (n_k, MS))$
    End if
    If $\max(E_{min}(n_i, n_k, MS)) < E_{min}(n_i, n_k, MS)$ Then
        $\max(E_{min}(n_i, n_k, MS)) = E_{min}(n_i, n_k, MS)$
    End if
End for
For each message $n_i$ received from the candidate node $n_k$ Do
    $w_D(i, k) = \frac{D(n_i,n_k,MS)}{\max(D(n_i,n_k,MS))}$
    $w_E(i, k) = 1 - \frac{E_{min}(n_i,n_k,MS)}{\max(E_{min}(n_i,n_k,MS))}$
    $\text{cost}(i, k) = \alpha w_D(i, k) + (1 - \alpha) w_E(i, k)$
    If $\min(\text{cost}(i, k)) > \text{cost}(i, k)$ Then
        $\min(\text{cost}(i, k)) = \text{cost}(i, k)$
    End if
End for
Node $n_i$ chooses node $n_k$ with the minimum cost $\min(\text{cost}(i, k))$ as its parent node
```

as the parent node. Then start the next round of parent node selection. In each round of parent node selection, for each candidate parent node n_k, first send Offer-Msg to claim that it can be a candidate parent node, then wait for the request message Request-Msg in timer T_{asso}, then send Confirm-Msg to the source node. For each source node n_i, first listen for the claim message Offer-Msg in timer T_{offer}, then calculate the cost from n_i to n_k, and select the node with the minimum cost as the parent node, and send the Request-Msg to this node. Then listen for the Confirm-Msg in timer $T_{confirm}$. Start the next round of selection after one round of selection until the timer $T_{cluster}$.

As shown in Figure 4, the CH n_4 first broadcasts the message that it's the CH. n_2, n_5, n_7 and n_8 select n_4 as the parent node after receiving the message. Then n_2, n_5, n_7 and n_8 broadcast the message that they can be the parent nodes, and the other nodes receiving the messages select the node with minimum cost as the parent node. Finally, the clustering process finishes.

The final goal is to transfer the data to the vehicle, but as the vehicle is continuously moving, thus the strategy of transferring data to the vehicle is different from the traditional strategy. In order to let the CH knows the timing of starting and finishing transferring data, the vehicle will send the data request message periodically. The CH will send the data to the vehicle after receiving the message from the vehicle.

Algorithm 4 SCHR clustering alorithm

```
If $h_i == 1$Then
    Select the CH by Algorithm 2
End If
If $h_i == 2$ Then
    Node $n_i$ chooses CH as its parent node
End If
$h = 2$
While timer $T_{cluster}$ not expires Do
    While $h_k \leq h + 1$ Do
        If $h_k == h$ Then
            Node $n_k$ is a candidate parent node
        End If
        If $h_i == h + 1$ Then
            Node $n_i$ is a child node
        End If
        If node $n_k$ is a candidate parent node Then
            Broadcast the Offer-Msg
            Listen for the Request-Msg until timer $T_{asso}$ expires
            Send Confirm-Msg to the child node $n_i$
        End If
        If node $n_i$ is a child node Then
            Listen for the Offer-Msg until timer $T_{offer}$ expires
            For each message $n_i$ received from the candidate parent node $n_k$ Do
                Calculate the cost $\text{cost}(i, k)$ by Algorithm 3
            End For
            Node $n_i$ chooses node $n_k$ with the minimum cost $\min(\text{cost}(i, k))$ as its parent node
            Send Request-Msg to the parent node $n_k$
            Listen for the Confirm-Msg until the timer $T_{confirm}$ expires
        End If
    End While
    $h = h + 1$
End While
```

C. THE DESIGN OF MCHR

In SCHR, each WSN can only have one CH, but this scheme has some limitations. The nodes transfer the data to the CH when the vehicles passing by, then CH transfers the data to the vehicles. The CH needs to transfer all the data generated from the sensor nodes in the WSN, the energy consumption of the CH is the highest. The CH is easy to become a dead node as it needs to transfer a large amount of data. Thus, we propose to select multiple CHs in one WSN, and we propose a multiple cluster head rotation joint mobile vehicle data collection (MCHR).

The MCHR has three characteristics: first, the network is divided into several disconnected WSNs, multiple nodes are selected as the CHs from the nodes that can directly connect with the vehicles. Second, the sensor nodes in the WSN are partitioned into different clusters based on the CHs.

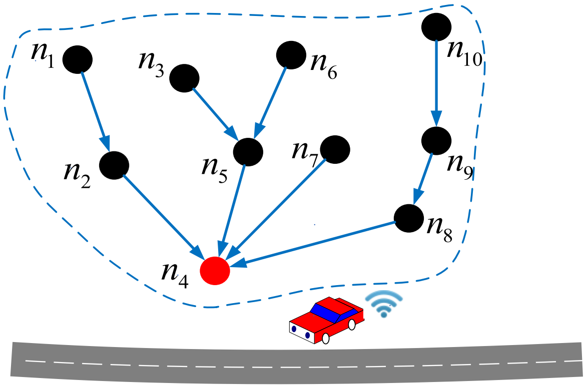

FIGURE 4. The clustering algorithm of SCHR.

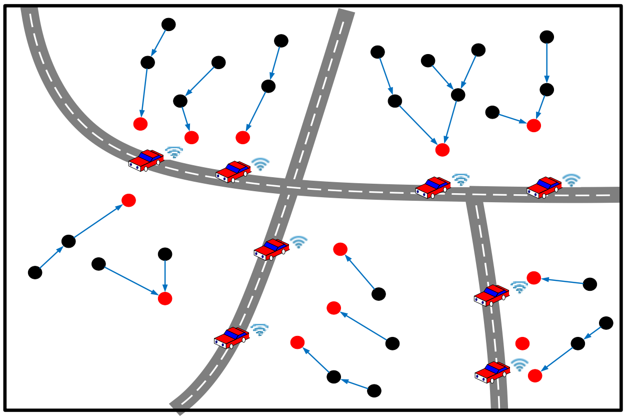

FIGURE 5. The network model of MCHR.

Third, the number of clusters and the clustering will be changed as different paths are selected by the vehicles.

Figure 5 shows the network model of MCHR, the network is partitioned into five WSNs. In each WSN, the sensor nodes are partitioned into several clusters.

The MCHR solution contains three phases: (1) CH rotation. The CHs in each WSN are selected in this phase. (2) The clustering algorithm. In the routing algorithm, each node needs to find the route path to the CH. (3) The sensor node update process. In the update phase, if insert or delete a sensor node, it needs to update the route path.

1) PHASE 1: CH ROTATION

In MCHR, the vehicle first obtains the hop distances of all the sensor nodes as SCHR. The sensor nodes whose hop distances are 1 become the CH candidate nodes. We only allow the sensors whose residual energy exceed the threshold value to participate in the CH rotation process. The CHs are selected according to the transferring costs. In order to control the number of CHs, we propose to choose the top k nodes as CHs according to the ascending order of the costs. The nodes' costs are recalculated and reordered based on the cost function in each round, thus the CHs rotate in different rounds. In Figure 6, in the first round, if k equals 2, then n_4 and n_8 are selected as the CH nodes.

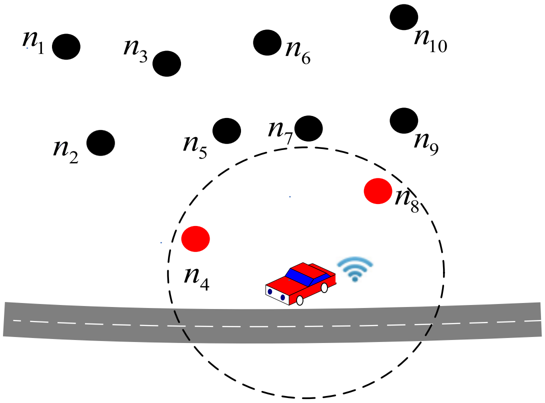

FIGURE 6. The CH rotation process.

Algorithm 5 The CH Rotation of MCHR

```
While $E(n_i) > T$ and $h_i = 1$ Do
    $n_i$ becomes a candidate CH node
End While
Calculating the cost for each candidate CH node $n_i$ by Algorithm 2
MS chooses the top $k$ nodes as CHs according to the ascending order of the costs
```

Algorithm 5 shows the CH rotation process of MCHR.

2) STAGE 2: THE CLUSTERING ALGORITHM

After the CH rotation process, we need to design a clustering algorithm which partitions the nodes in the WSN into several clusters. Each cluster corresponds to a CH, the nodes choose a cluster and route the data to the CH. In the data transferring process, each node needs to find an appropriate cluster and route path in order to balance the network energy consumption. Each node may receive several messages from multiple candidate nodes, the main point of the clustering algorithm is to select one parent node from multiple candidate nodes. We continue to use the cost function in SCHR's clustering algorithm to choose the parent node, the objective of the clustering algorithm is to balance the energy consumption and maximize the network lifetime.

In MCHR, each WSN contains several CHs. Firstly, each CH n_k broadcasts the Offer-Msg that it can be the candidate parent node. The message contains the node ID k, the GPS GPS_{n_k}, the distance from the node to the vehicle $D(n_k, MS)$, and the minimum energy of the nodes in the path from n_k to $MSE_{min}(n_k, MS)$.

Each node n_i within the communication range receives multiple messages from several CHs. n_i begins to calculate the costs to the CHs using the cost function, n_i chooses the CH with the minimum cost and send the Request-Msg to join the CH's cluster. The Request-Msg contains n_i's node ID i. The CH receives Request-Msg and sends the Confirm-Msg

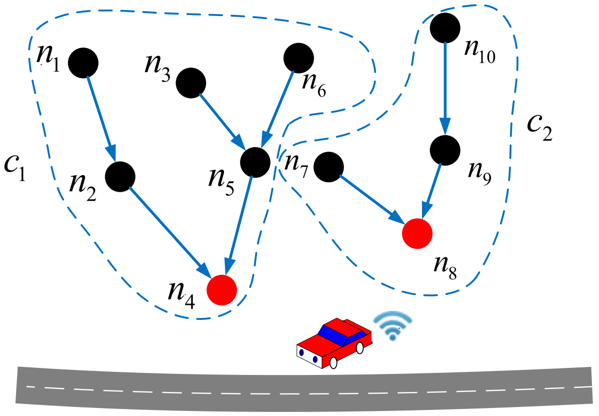

FIGURE 7. **The clustering algorithm of MCHR.**

to invite n_k to join the cluster. The Confirm-Msg contains: n_i's node ID i, n_k's node ID k. When n_i receives the Confirm-Msg, n_i chooses n_k to be its parent node and joins its cluster.

There are some other nodes which are far away from the CH, which cannot directly communicate with CH and need to send data to the CH through multi-hop communication. Thus, after the first round of parent node selection, the nodes whose hop distances equal 2 have chosen the CH and join the cluster. Next, each node whose hop distance equals 2 broadcasts the message that it can be the parent node n_k. Each node n_i whose hop distance equals 3 chooses the parent node based on the cost function. We continue to repeat the parent node selection process until each node has chosen their parent node.

The MCHR clustering algorithm is shown as Algorithm 6. Firstly, initialize the hop distance parameter $h = 1$, and cluster all the nodes in the WSN. In each round, the nodes whose hop distances are $h + 1$ select the parent node whose hop distance is h. In each round of parent node selection, for each candidate parent node n_k, firstly, n_k broadcasts the Offer-Msg to claim that it can be the candidate parent node, then n_k waits for the Request-Msg in timer T_{asso} and sends Confirm-Msg to the source node n_i. For each source node n_i, firstly, n_i listens for the Offer-Msg in timer T_{offer}, then n_i calculates the routing cost with candidate parent node n_k and selects the node with minimum cost. After finishing one round of parent node selection, continue to finish the next round of selection until timer $T_{cluster}$ ends.

As shown in Figure 7, the CHs n_4 and n_8 broadcast the messages. n_2, n_5, n_7 and n_9 receive the messages and compute their costs to send data to n_4 and n_8, then they choose the CH with minimum cost and join its cluster. Next, n_2, n_5, n_7 and n_9 broadcast the messages, n_1, n_3, n_6 and n_{10} find the corresponding cluster and join the cluster. Finally, there are two clusters c_1 and c_2.

3) STAGE 3: THE SENSOR NODE UPDATE PROCESS

In the network of smart city, in order to perceive the status of the infrastructures promptly, the sensor nodes need to update frequently. The WSN may insert or delete a sensor node,

Algorithm 6 MCHR Clustering Algorithm

```
h = 1
While timer $T_{cluster}$ not expires Do
    While $h_k \leq h + 1$ Do
        If $h_k == h$ Then
            Node $n_k$ is a candidate parent node
        End If
        If $h_i == h + 1$ Then
            Node $n_i$ is a child node
        End If
        If node $n_k$ is a candidate parent node Then
            Broadcast the Offer-Msg
            Listen for the Request-Msg until timer $T_{asso}$ expires
            Send Confirm-Msg to the child node $n_i$
        End If
        If node $n_i$ is a child node Then
            Listen for the Offer-Msg until timer $T_{offer}$ expires
            For each message $n_i$ received from the candidate parent node $n_k$ Do
                Calculate the cost cost$(i, k)$ by Algorithm 3
            End For
            Node $n_i$ choose node $n_k$ with the minimum cost min(cost$(i, k)$) as its parent node
            Send Request-Msg to the parent node $n_k$
            Listen for the Confirm-Msg until timer $T_{confirm}$ expires
        End If
    End While
    $h = h + 1$
End While
```

a good routing algorithm should dynamically generate the route path when the sensor node changes.

a: INSERT A SENSOR NODE

If the inserted node is closer to the road and within the communication range of the vehicle, then it becomes a CH, the nodes in the network need to cluster. If the inserted node is not within the communication range of the vehicle, then it needs to join the existing cluster.

As shown in Figure 8, the network inserts two sensor nodes n_{11} and n_{12}, as n_{11} is within the communication range of the vehicle, thus n_{11} becomes a new CH, the sensor nodes are partitioned into three clusters. n_{12} is not within the communication range of the vehicle, first choose the cluster c_1 according to the distance, then select n_2 as its parent node based on the calculated costs.

b: DELETE A SENSOR NODE

If the deleted node is a CH, then the network needs to cluster, the nodes belong to the dead CH will join the other clusters.

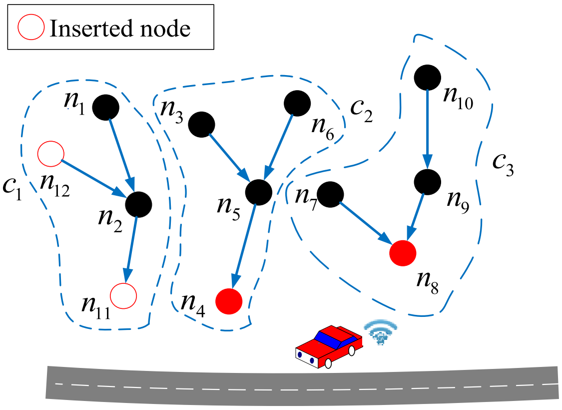

FIGURE 8. **Insert a new sensor node.**

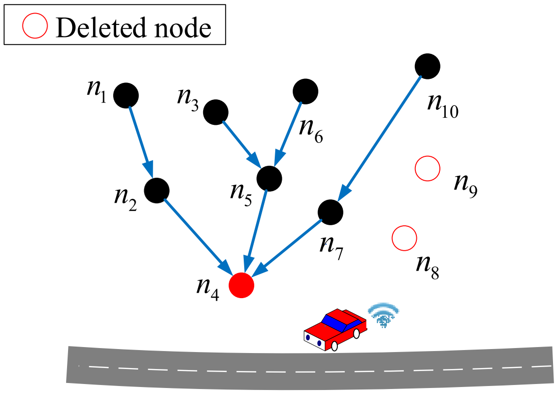

FIGURE 9. **Delete a sensor node.**

If the deleted node is not a CH, then the dead node needs to regenerate the route path.

As shown in Figure 9, the network deletes two sensor nodes n_8 and n_9. n_8 is a CH, the nodes n_7 and n_{10} need to join the cluster of n_4. n_7's parent node is n_4, n_{10}'s parent node is n_7.

V. ANALYSIS OF EXPERIMENTAL RESULTS

A. SIMULATION PARAMETERS

We evaluate the performance of our proposed SCHR and MCHR schemes. We use Matlab to simulate the network. As shown in Figure 10, the network size is $500 \times 500m^2$, 100 nodes are randomly distributed in the network. The network contains four WSNs, each road has two vehicles, the four vehicles are V1, V2, V3, and V4, respectively.

Table 1 shows the parameters in the proposed SCHR and MCHR schemes.

B. NETWORK ENERGY AND LIFETIME

To evaluate the network energy and lifetime, we tested SCHR and MCHR under different values in various aspects.

1) AVERAGE ENERGY OF EACH NODE

Figure 11 shows the average energy of each node for SCHR and MCHR schemes. The network is evaluated in various values of α (0.3, 0.5, 0.7). For the MCHR scheme, we select the top 5 nodes as CHs in each WSN according to the cost function. As the round number increases, the average energy of each node decreases. MCHR's average energy is larger than SCHR, the nodes in SCHR consume more energy than MCHR. The larger the α, the larger the average energy of each node. According to the cost function, α influences the CH rotation, parent node selection and clustering algorithms. As α increases, the weight of distance increases, the node nearby the vehicle has a higher probability to be selected as the CH and parent node. The nodes nearby the vehicle consume more energy and become dead earlier. As the number of nodes which can communicate with the vehicle becomes smaller, each node consumes smaller energy, thus the average energy of each node increases as α increases.

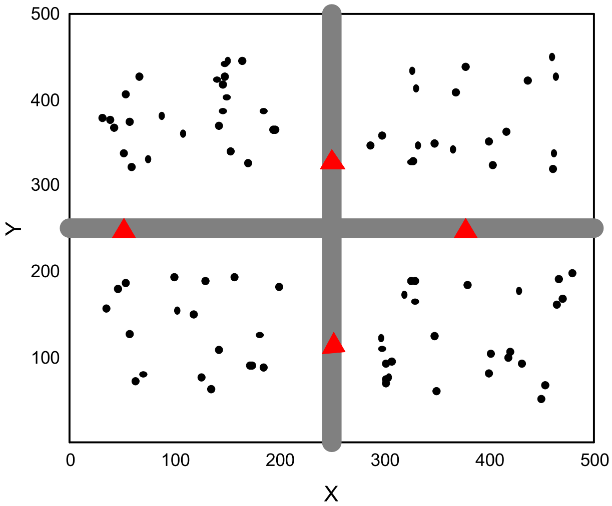

FIGURE 10. **The simulation network.**

TABLE 1. **Parameters of simulations.**

Parameter	Value
Area of deployment	$500 \times 500m^2$
Number of nodes	100
Threshold distance (r_0) (m)	87
Sensing range (r) (m)	≤80
Initial energy(E_{init})(J)	0.5
E_{elec}(nJ/bit)	50
ε_{fs}(pJ/bit/m^2)	10
ε_{amp}(pJ/bit/m^4)	0.0013
Data packet size（bit）	100
Send packets size (bit)	1000
Data acquisition rate (kbps)	200
The vehicle speed (km/h)	30

In order to observe the nodes' residual energy of the whole WSN more intuitively, Figure 12 and 13 show the node energy distribution for SCHR and MCHR, respectively. α is

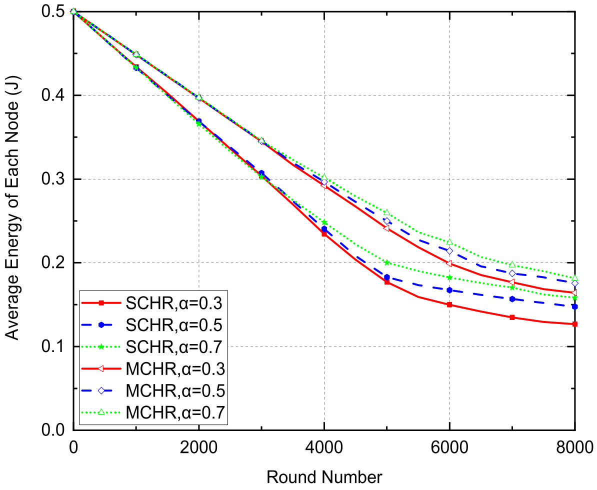

FIGURE 11. **The average energy of each node.**

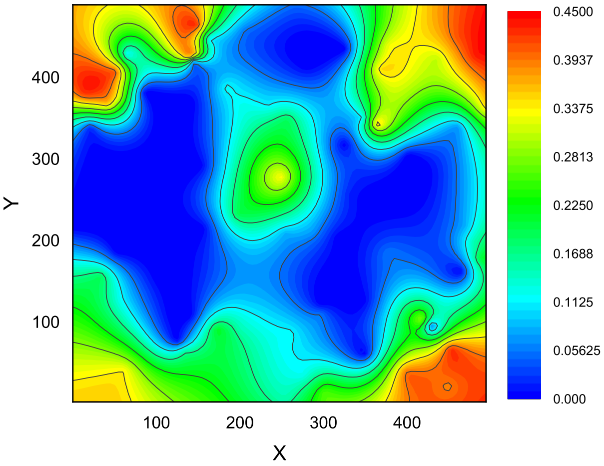

FIGURE 12. **Energy distribution for SCHR.**

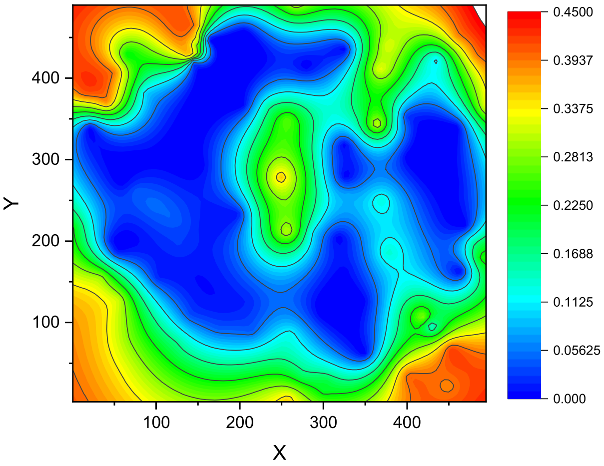

FIGURE 13. **Energy distribution for MCHR.**

fixed as 0.5, the round number is 8000, the blue node represents small residual energy, the green node represents almost half residual energy, the red node represents large residual energy.

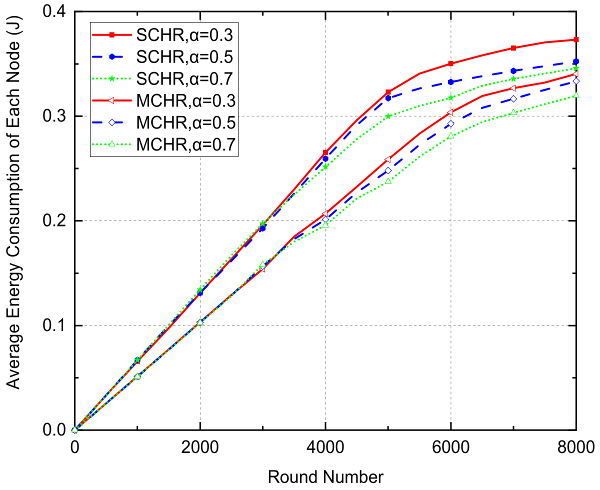

FIGURE 14. **Average energy consumption of each node.**

From the observation, the boundary nodes have more residual energy, the middle nodes have smaller residual energy. Because the middle nodes are closer to the vehicle, which are more overloaded than the boundary nodes. When the middle nodes died after consuming all the energy, the boundary nodes cannot communicate with the vehicle, thus the boundary nodes have more residual energy as they cannot transfer packets. MCHR has more blue nodes and smaller green nodes compared with SCHR, which indicates that MCHR's energy consumption is more balancing than SCHR. Because MCHR has more CHs in each WSN, thus more boundary nodes can transfer packets from CHs to vehicle.

2) AVERAGE ENERGY CONSUMPTION

Figure 14 shows the average energy consumption of each node for SCHR and MCHR schemes. The network is evaluated in various values of $\alpha(0.3, 0.5, 0.7)$. The average energy consumption increases as the round number increases. MCHR's average energy consumption is smaller than SCHR, as MCHR has more CHs than SCHR. The average energy consumption of each node decreases as α increases. As α increases, the weight of distance increases, it avoids the long-hop relaying and decreases the energy consumption, thus the average energy consumption of each node decreases as α increases.

3) MAXIMUM ENERGY OF NODES

Figure 15 shows the maximum energy of nodes for SCHR and MCHR schemes. Each node's initial energy is 0.5 J. As the round number increases, the maximum energy of nodes decreases. The nodes with maximum energy usually represent the boundary nodes in each WSN. SCHR's maximum energy is larger than MCHR. In MCHR, more boundary nodes can transfer their packets through CHs to the vehicle, thus MCHR's maximum energy is smaller than SCHR. MCHR's maximum energy changes slightly with α, SCHR increases as α increases in the first 3800 rounds.

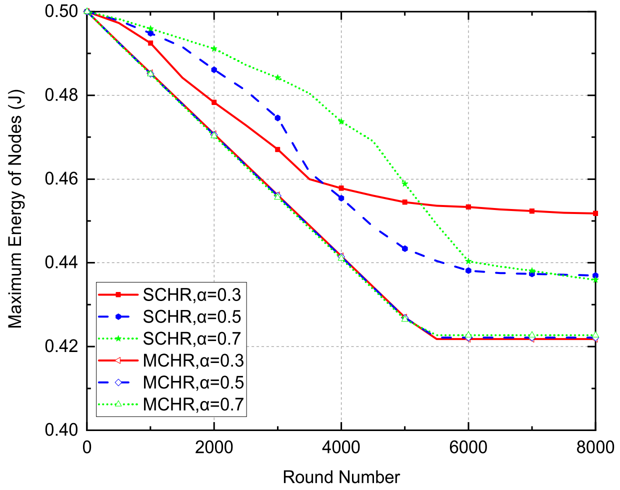

FIGURE 15. **Maximum energy of nodes.**

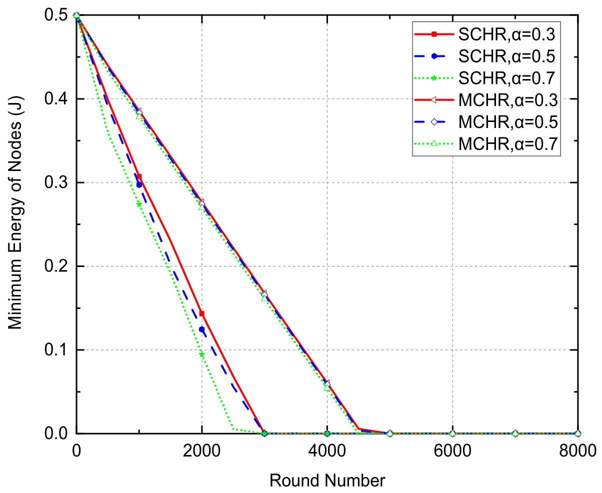

FIGURE 16. **Minimum energy of nodes.**

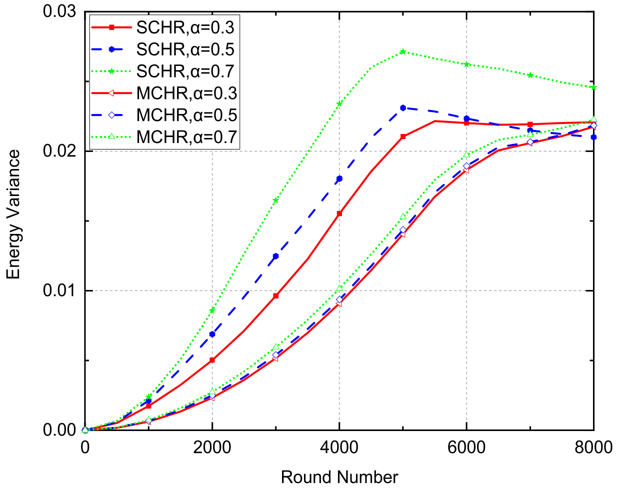

FIGURE 17. **Energy variance.**

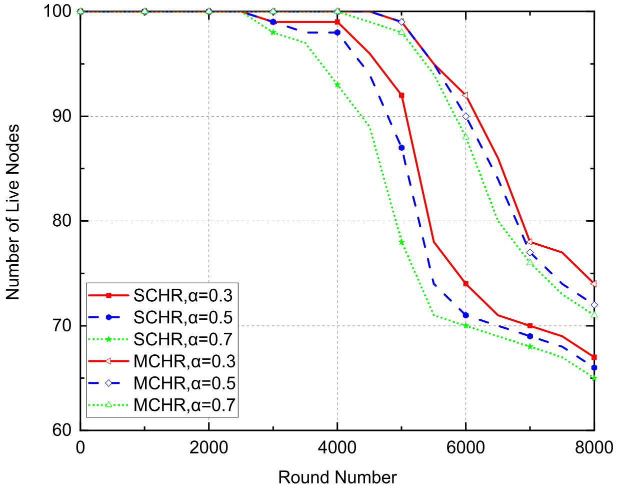

FIGURE 18. **Number of live nodes.**

4) MINIMUM ENERGY OF NODES

Figure 16 shows the minimum energy of nodes for SCHR and MCHR schemes. The nodes with minimum energy usually represent the nodes nearby the vehicle. SCHR's minimum energy is smaller than MCHR, because SCHR's nodes nearby the vehicle are heavily overloaded and consuming energy faster than MCHR. MCHR's minimum energy changes slightly with α, SCHR decreases as α increases.

5) ENERGY VARIANCE

Energy variance is the variance between the maximum energy and the minimum energy. The energy variance reflects whether the energy consumption is balancing or not. Figure 17 shows the energy variance for SCHR and MCHR schemes. SCHR's energy variance is smaller than MCHR, which shows that SCHR is more energy balancing than MCHR. The energy variance of SCHR and MCHR increases as α increases. The larger the α, the larger the weight of the residual energy in the cost function, SCHR and MCHR consider energy balancing more when selecting nodes to transfer packets. Thus, the energy consumption is more balancing when α is smaller.

6) NUMBER OF LIVE NODES

Figure 18 shows the number of live nodes for SCHR and MCHR schemes. SCHR's number of live nodes is smaller than MCHR, and SCHR's first node dies before MCHR, which shows that MCHR achieves longer network lifetime. The smaller the α, the larger the number of live nodes. Because when α is smaller, the weight of residual energy is larger, thus the selected nodes contain more residual energy, and the nodes have a longer lifetime.

Table 2 shows the comparison results of SCHR, MCHR and LEACH [26]. We assume that there is a base station in the center of the simulation network for LEACH. The performance metrics include the round numbers of the first node dies (FRND), 5% of nodes die (FVND), 10% of nodes die (TEND), 20% of nodes die (TWND). The experimental results show that SCHR and MCHR are better than LEACH in FRND and FVND. but LEACH is better in TWND. SCHR and MCHR collect data through vehicles, LEACH collect

TABLE 2. **The comparison results.**

Scheme	FRND	FVND	TEND	TWND
SCHR	2897	4486	5215	5789
MCHR	4613	5571	5921	6875
LEACH	1758	3489	5497	7235

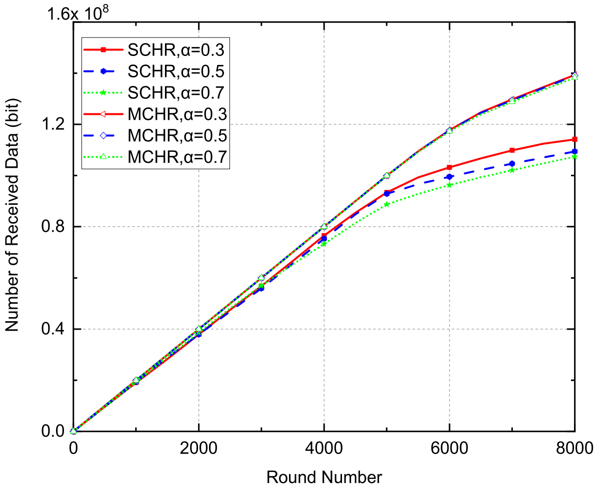

FIGURE 19. **Number of received data.**

data through base stations. The experimental results show that the proposed SCHR and MCHR can select appropriate CHs to transfer data, thus the round number of first node dies is larger than LEACH. In LEACH, after the nodes near the base station die, the other nodes are still living but they cannot transfer data. SCHR and MCHR are more balancing than LEACH, because SCHR and MCHR can rotate the CHs according to the distance to the vehicles and residual energy.

C. NETWORK TRANSMISSION CAPACITY

1) NUMBER OF RECEIVED DATA

Figure 19 shows the number of received data for SCHR and MCHR schemes. The four vehicles receive data from the CHs and transfer the data to the data center. The received data reflects the efficiency of the data collection scheme, the data collection scheme with a larger amount of received data has better efficiency. SCHR's received data is smaller than MCHR, because there are more CHs in MCHR which can transfer data to the vehicle. We can find that the number of received data increases as α decreases. The smaller the α, the larger the number of live nodes, thus the vehicles can receive more data from the nodes.

Figure 20 shows the number of received data for SCHR's four vehicles V1, V2, V3, and V4. The vehicles' received data are from 2.3×10^7 to 3.19×10^7. V3's received data is the largest, and V2's received data is the smallest. We can find that the number of received data for SCHR's four vehicles increases as α decreases. The vehicles' received data is relevant with the nodes in the WSN.

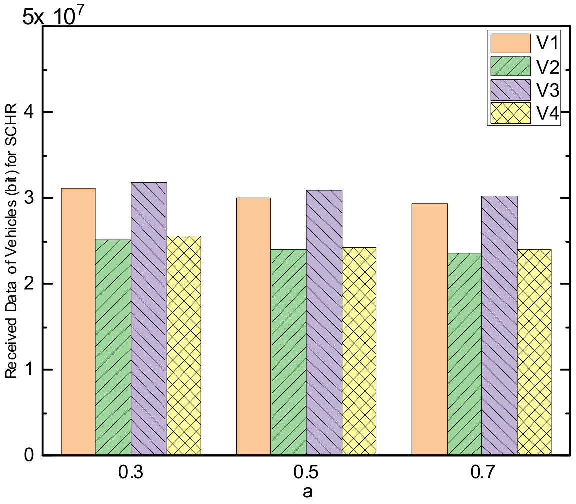

FIGURE 20. **Number of received data for SCHR.**

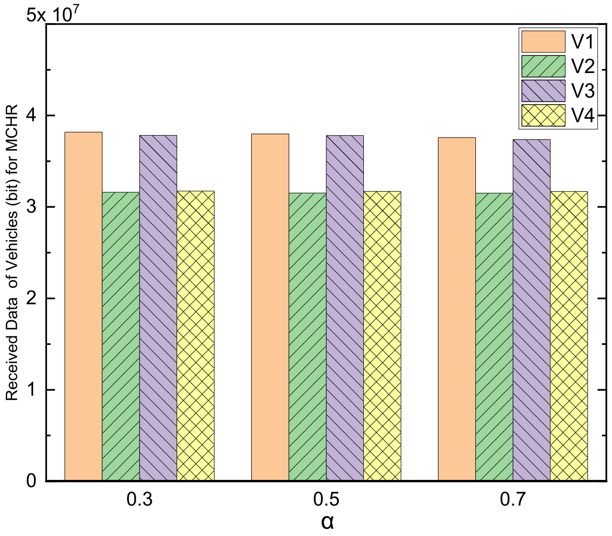

FIGURE 21. **Number of received data for MCHR.**

Figure 21 shows the number of received data for MCHR's four vehicles V1, V2, V3, and V4. The vehicles' received data are from 3.16×10^7 to 3.81×10^7. It's observed that the number of received data for MCHR's four vehicles has small impact with α.

Figure 22 shows the variance of received data for SCHR and MCHR. We can find that SCHR's variance is larger than MCHR, and MCHR's variance of received data is smallest when α is 0.7.

2) NUMBER OF CACHING DATA

Caching data means the number of data cached by the nodes. Figure 23 shows the number of caching data for SCHR and MCHR schemes. SCHR's caching data is larger than MCHR, because in SCHR scheme, when the nodes near the vehicle die after consuming all the energy, the nodes far away from the vehicle cannot communicate with the vehicle, thus these

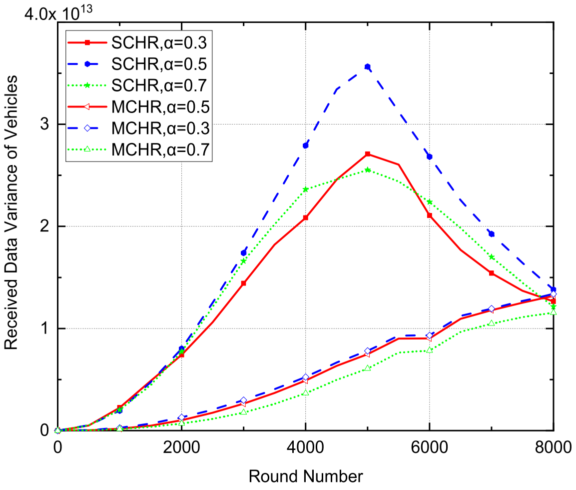

FIGURE 22. Variance of received data.

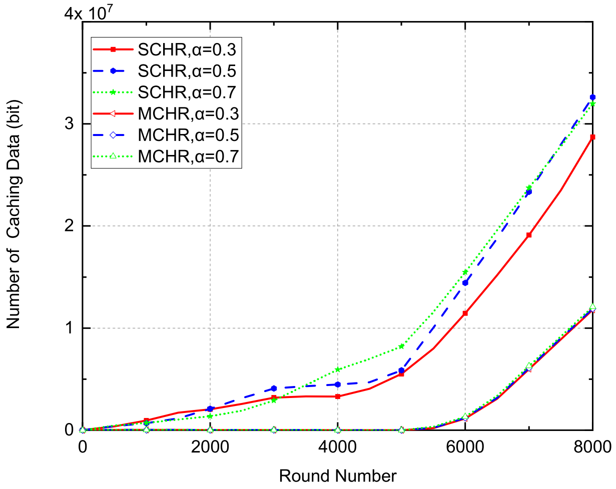

FIGURE 23. Number of caching data.

nodes cache more data in local. We can find that the caching data of SCHR and MCHR increases as α increases. The larger the α, the more probability that the nodes far from the vehicle cannot communicate with the vehicle, thus more data cached by the nodes.

Figure 24 shows the number of caching data for SCHR under $\alpha = 0.5$ in 8000 rounds. We can find that some nodes' caching data is very large, because these nodes cannot transfer their data to the vehicle through multi-hop communication once the relaying node dies.

Figure 25 shows the number of caching data for MCHR under $\alpha = 0.5$ in 8000 rounds. We can find that MCHR's caching data is not very large. The number of caching data in MCHR is much smaller than SCHR.

3) DATA TRANSMISSION RATE

The data transmission rate is equal to the number of received data by the vehicles divides the number of data sensed by the sensors. Figure 26 shows the data transmission rate for SCHR and MCHR schemes. The data transmission rate decreases as the round number increases. SCHR's data transmission rate is smaller than MCHR, because the number of received data in MCHR is more than SCHR. MCHR's data transmission rate is 100% in the first 5000 rounds, the data transmission rate of MCHR changes slightly with α.

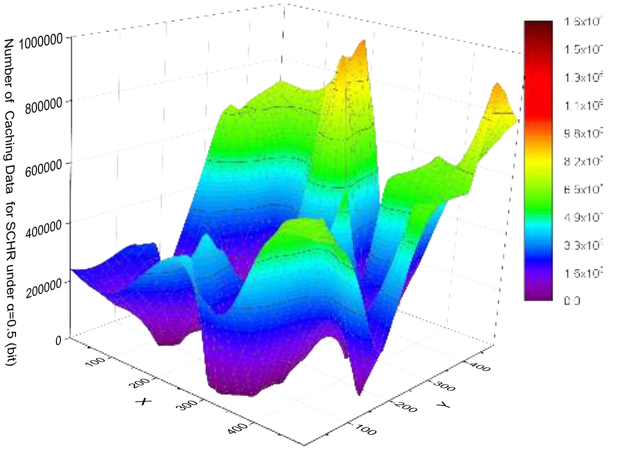

FIGURE 24. Number of caching data for SCHR.

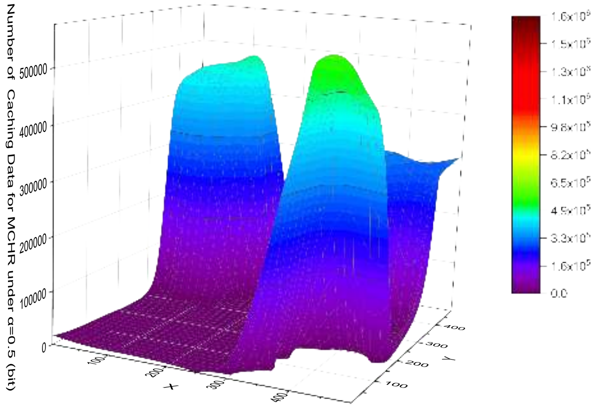

FIGURE 25. Number of caching data for MCHR.

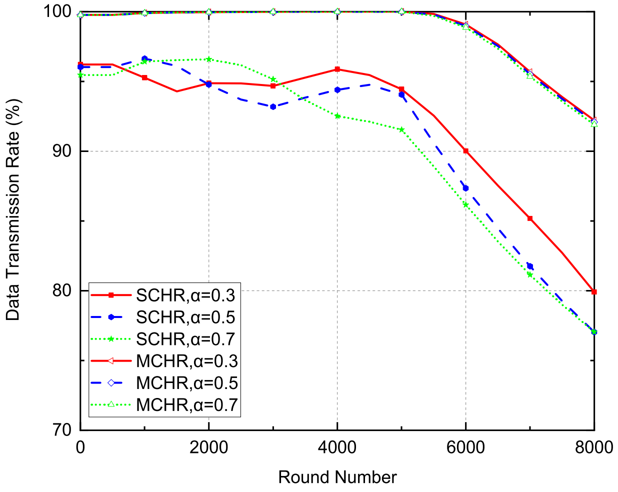

FIGURE 26. Data transmission rate.

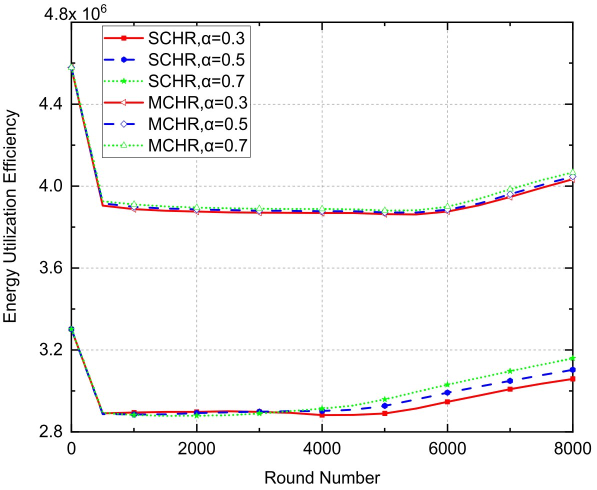

FIGURE 27. **Energy utilization efficiency.**

4) ENERGY UTILIZATION EFFICIENCY

The energy utilization efficiency is equal to the number of received data by the vehicle divides the energy consumption of the nodes. Figure 27 shows the energy utilization efficiency for SCHR and MCHR schemes. MCHR's energy utilization efficiency is larger than SCHR, which shows that MCHR is more energy balancing than SCHR. We can find that the data transmission rate of SCHR and MCHR increases slightly as α increases.

VI. CONCLUSION

In smart city, the sensor nodes are equipped with different components for different applications, which allow to sense and collect the status and information of all kinds of infrastructures in the city. In this paper, we propose two data collection schemes SCHR and MCHR through mobile vehicles in edge network of smart city. The sensing network in smart city forms several decentralized and isolated WSNs. In SCHR, each WSN selects one CH, the CH rotation and clustering algorithm use the cost function based on the distance to the vehicle and residual energy. In MCHR, the energy consumption is more balancing as each WSN contains multiple CHs, the clustering algorithm has good scalability when the network inserts or deletes a sensor node. The SCHR and MCHR schemes are novel data collection schemes that can connect large number of sensing devices into IoT and effectively collect data based on the existing city infrastructures in a flexible and low-cost way. The performance of SCHR and MCHR have been validated through extensive experiments, which show that SCHR and MCHR have good performance in energy consumption and network lifetime.

REFERENCES

[1] X. Liu, A. Liu, T. Wang, K. Ota, M. Dong, Y. Liu, and Z. Cai, "Adaptive data and verified message disjoint security routing for gathering big data in energy harvesting networks," *J. Parallel Distrib. Comput.*, vol. 135, pp. 140–155, Jan. 2020.

[2] Y. Liu, A. Liu, T. Wang, X. Liu, and N. N. Xiong, "An intelligent incentive mechanism for coverage of data collection in cognitive Internet of Things," *Future Gener. Comput. Syst.*, vol. 100, pp. 701–714, Nov. 2019.

[3] T. Wang, H. Luo, W. Jia, A. Liu, and M. Xie, "MTES: An intelligent trust evaluation scheme in sensor-cloud enabled industrial Internet of Things," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2019.2930286.

[4] M. Bonola, L. Bracciale, P. Loreti, P. Amici, A. Rabuffi, and G. Bianchi, "Opportunistic communication in smart city: Experimental insight with small-scale taxi fleets as data carriers," *Ad Hoc Netw.*, vol. 43, pp. 43–55, Jun. 2016, doi: 10.1016/j.adhoc.2016.02.002.

[5] L. Hu, A. Liu, and M. Xie, "UAVs joint vehicles as data mules for fast codes dissemination for edge networking in smart city," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 6, pp. 1550–1574, Nov. 2019, doi: 10.1007/s12083-019-00752-0.

[6] Z. Tang, A. Liu, and C. Huang, "Social-aware data collection scheme through opportunistic communication in vehicular mobile networks," *IEEE Access*, vol. 4, pp. 6480–6502, 2016, doi: 10.1109/ACCESS.2016.2611863.

[7] X. Xiang, W. Liu, T. Wang, M. Xie, X. Li, H. Song, A. Liu, and G. Zhang, "Delay and energy-efficient data collection scheme-based matrix filling theory for dynamic traffic IoT," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, p. 168, Jun. 2019, doi: 10.1186/s13638-019-1490-5.

[8] Q. Li, A. Liu, and T. Wang, "Pipeline slot based fast rerouting scheme for delay optimization in duty cycle based M2M communications," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 6, pp.1673–1704, Nov. 2019, doi: 10.1007/s12083-019-00753-z.

[9] H. Dai, X. Wang, and A. X. Liu, "Wireless charger placement for directional charging," *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1865–1878, Aug. 2018.

[10] X. Liu and P. Zhang, "Data drainage: A novel load balancing strategy for wireless sensor networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 125–128, Jan. 2018, doi: 10.1109/LCOMM.2017.2751601.

[11] X. Sun, W. Liu, T. Wang, Q. Deng, A. Liu, N. Xiong, and S. Zhang, "Two-hop neighborhood information joint double broadcast radius for effective code dissemination in WSNs," *IEEE Access*, vol. 7, pp. 88547–88569, 2019, doi: 10.1109/ACCESS.2019.2925829.

[12] T. Shu, W. Liu, T. Wang, Q. Deng, M. Zhao, N. N. Xiong, X. Li, and A. Liu, "Broadcast based code dissemination scheme for duty cycle based wireless sensor networks," *IEEE Access*, vol. 7, pp. 105258–105286, 2019, doi: 10.1109/ACCESS.2019.2932006.

[13] X. Liu, "A novel transmission range adjustment strategy for energy hole avoiding in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 67, pp. 43–52, May 2016, doi: 10.1016/j.jnca.2016.02.018.

[14] X. Liu, M. Zhao, A. Liu, and K. K. L. Wong, "Adjusting forwarder nodes and duty cycle using packet aggregation routing for body sensor networks," *Inf. Fusion*, vol. 53, pp. 183–195, Jan. 2020, doi: 10.1016/j.inffus.2019.06.020.

[15] X. Sun and N. Ansari, "EdgeIoT: Mobile edge computing for the Internet of Things," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 22–29, Dec. 2016, doi: 10.1109/MCOM.2016.1600492CM.

[16] A. Mehrabi and K. Kim, "General framework for network throughput maximization in sink-based energy harvesting wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1881–1896, Jul. 2017, doi: 10.1109/TMC.2016.2607716.

[17] H. Salarian, K.-W. Chin, and F. Naghdy, "An energy-efficient mobile-sink path selection strategy for wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2407–2419, Jun. 2014, doi: 10.1109/TVT.2013.2291811.

[18] Y. Liu, K.-Y. Lam, S. Han, and Q. Chen, "Mobile data gathering and energy harvesting in rechargeable wireless sensor networks," *Inf. Sci.*, vol. 482, pp. 189–209, May 2019, doi: 10.1016/j.ins.2019.01.014.

[19] T. Peng, Q. Liu, B. Hu, J. Liu, and J. Zhu, "Dynamic keyword search with hierarchical attributes in cloud computing," *IEEE Access*, vol. 6, pp. 68948–68960, 2018, doi: 10.1109/ACCESS.2018.2878268.

[20] N. Wang, J. Wu, and H. Dai, "Bundle charging: Wireless charging energy minimization in dense wireless sensor networks," in *Proc. 39th IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2019, pp. 1–11.

[21] Y. Liu, A. Liu, X. Liu, and M. Ma, "A trust-based active detection for cyber-physical security in industrial environments," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2019.2931394.

[22] N. Wang and J. Wu, "Mutually exclusive data dissemination in the mobile publish/subscribe system," in *Proc. 13th IEEE Int. Conf. Mobile Ad-Hoc Smart Syst. (MASS)*, Oct. 2016, pp. 75–83.

[23] R. Deng, S. He, and J. Chen, "An online algorithm for data collection by multiple sinks in wireless-sensor networks," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 1, pp. 93–104, Jun. 2018, doi: 10.1109/TCNS.2016.2578460.

[24] A. Mehrabi and K. Kim, ''Maximizing data collection throughput on a path in energy harvesting sensor networks using a mobile sink,'' *IEEE Trans. Mobile Comput.*, vol. 15, no. 3, pp. 690–704, Mar. 2016, doi: 10.1109/TMC.2015.2424430.

[25] X. Liu, L. Fu, and J. Wang, ''Multicast scaling of capacity and energy efficiency in heterogeneous wireless sensor networks,'' *ACM Trans. Sensor Netw. (TOSN)*, vol. 15, no. 3, pp. 1–32, 2019, doi: 10.1145/3322497.

[26] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, ''Energy-efficient communication protocol for wireless microsensor networks,'' in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, Jan. 2000, p. 10.

[27] B. Seema and E. J. Coyle, ''An energy efficient hierarchical clustering algorithm for wireless sensor networks,'' in *Proc. Joint Conf. IEEE Comput. Commun. Societies*, San Francisco, CA, USA, vol. 3, Apr. 2003, pp. 1713–1723.

[28] O. Younis and S. Fahmy, ''HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,'' *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct./Dec. 2004, doi: 10.1109/TMC.2004.41.

[29] R. Sugihara and R. K. Gupta, ''Optimal speed control of mobile node for data collection in sensor networks,'' *IEEE Trans. Mobile Comput.*, vol. 9, no. 1, pp. 127–139, Jan. 2010, doi: 10.1109/TMC.2009.113.

[30] S. Basagni, A. Carosi, E. Melachrinoudis, C. Petrioli, and Z. M. Wang, ''Controlled sink mobility for prolonging wireless sensor networks lifetime,'' *Wireless Netw.*, vol. 14, no. 6, pp. 831–858, 2008, doi: 10.1007/s11276-007-0017-x.

[31] M. Ma and Y. Yang, ''SenCar: An energy efficient data gathering mechanism for large scale multihop sensor networks,'' *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 10, pp. 1476–1488, Oct. 2007, doi: 10.1109/TPDS.2007.1070.

[32] G. Xing, T. Wang, Z. Xie, and W. Jia, ''Rendezvous planning in wireless sensor networks with mobile elements,'' *IEEE Trans. Mobile Comput.*, vol. 7, no. 12, pp. 1430–1443, Dec. 2008, doi: 10.1109/TMC.2008.58.

[33] M. Di Francesco, S. K. Das, and G. Anastasi, ''Data collection in wireless sensor networks with mobile elements: A survey,'' *ACM Trans. Sensor Netw.*, vol. 8, no. 1, pp. 1–31, 2011, doi: 10.1145/1993042.1993049.

[34] Y. Gu, F. Ren, Y. Ji, and J. Li, ''The evolution of sink mobility management in wireless sensor networks: A survey,'' *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 507–524, 1st Quart., 2016, doi: 10.1109/COMST.2015.2388779.

[35] R. Shah, S. Roy, S. Jain, and W. Brunette, ''Data mules: Modeling and analysis of a three-tier architecture for sparse sensor networks,'' *Ad Hoc Netw.*, vol. 1, nos. 2–3, pp. 215–233, 2003, doi: 10.1016/S1570-8705(03)00003-9.

[36] S. Jain, R. C. Shah, R. Brunette, G. Borriello, and S. Roy, ''Exploiting mobility for energy efficient data collection in wireless sensor networks,'' *Mobile Netw. Appl.*, vol. 11, no. 3, pp. 327–339, Jun. 2006, doi: 10.1007/s11036-006-5186-9.

[37] M. Zhao and Y. Yang, ''Optimization-based distributed algorithms for mobile data gathering in wireless sensor networks,'' *IEEE Trans. Mobile Comput.*, vol. 11, no. 10, pp. 1464–1477, Oct. 2012, doi: 10.1109/TMC.2011.178.

[38] S. Mottaghi and M. R. Zahabi, ''Optimizing LEACH clustering algorithm with mobile sink and rendezvous nodes,'' *AEU-Int. J. Electron. Commun.*, vol. 69, pp. 507–514, Feb. 2015, doi: 10.1016/j.aeue.2014.10.021.

[39] L. Song and D. Hatzinakos, ''Architecture of wireless sensor networks with mobile sinks: Sparsely deployed sensors,'' *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1826–1836, Jul. 2007, doi: 10.1109/TVT.2007.897213.

[40] A. A. Somasundara, A. Kansal, D. D. Jea, D. Estrin, and M. B. Srivastav a, ''Controllably mobile infrastructure for low energy embedded networks,'' *IEEE Trans. Mobile Comput.*, vol. 5, no. 8, pp. 958–973, Aug. 2006, doi: 10.1109/TMC.2006.109.

[41] S. Gao, H. Zhang, and S. Das, ''Efficient data collection in wireless sensor networks with path-constrained mobile sinks,'' *IEEE Trans. Mobile Comput.*, vol. 10, no. 4, pp. 592–608, Apr. 2011, doi: 10.1109/TMC.2010.193.

[42] H. Huang and A. V. Savkin, ''An energy efficient approach for data collection in wireless sensor networks using public transportation vehicles,'' *AEU Int. J. Electron. Commun.*, vol. 75, pp. 108–118, May 2017, doi: 10.1016/j.aeue.2017.03.012.

YUEYI LUO received the B.Sc. and M.Sc. degrees from Central South University, China, where he is currently pursuing the Ph.D. degree with the Department of Computer science and Technology, School of Computer Science and Engineering. He is also a Lecturer with the School of Mathematics and Statics, Central South University. His research interests include edge networks, vehicular networks, and WSNs.

XIAOYU ZHU received the B.Sc. degree in electronic information engineering, the M.Sc. degree in computer science, and the Ph.D. degree in computer science from Central South University, Changsha, China, in 2010, 2013, and 2019, respectively. She is currently a Postdoctoral Fellow with the School of Automation, Central South University. She has been a Visiting Student with Temple University, Philadelphia, PA, USA. Her research interests include vehicular networks, sensor-cloud computing, and network security.

JUN LONG is currently a Professor with the Department of Computer Science and Technology and also the Director of the Network Resources Management and Trust Evaluation Key Laboratory of Hunan Province, Central South University, China. His major research interests include machine learning and big data analysis.

• • •