IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 3, 2020, accepted July 8, 2020, date of publication July 20, 2020, date of current version August 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3010372

Design of Layered and Heterogeneous Network Routing Algorithm for Field Observation Instruments

JIUYUAN HUO, JIGUANG YANG, AND HAMZAH MURAD MOHAMMED AL-NESHMI
School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Corresponding author: Jiuyuan Huo (huojy@mail.lzjtu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant 61862038, in part by the Lanzhou Talent Innovation and Entrepreneurship Technology Plan Project under Grant 2019-RC-14, and in part by the Foundation of A Hundred Youth Talents Training Program of Lanzhou Jiaotong University.

ABSTRACT Field stations and observation systems in cold and arid areas are mostly distributed in harsh natural environments which lead to problems such as the poor real-time data collection, transmission and processing, lower accuracy of data and failure to form integrated network observation, etc. To solve these problems which severely restrict scientific monitoring and research in these areas, a Layered and Heterogeneous Clustering routing algorithm LHC is proposed for field observation instrument networks based on the classical clustering routing algorithm LEACH. First, the LHC algorithm adopts the mechanism of heterogeneous node energy to divide nodes into advanced nodes and normal nodes. The advanced nodes have more initial energy than normal nodes, which increases their probability to be elected as cluster heads (CHs). Then, a hierarchical structure is used to divide the network into several layers, each layer elects a fixed number of CHs, and the distribution of CHs is improved. Finally, by analyzing the influence of the residual energy of the node and the distance between the node and the base station (BS) of the network, the mechanism based on energy and distance factors is introduced into each round of CH election to improve the CH election method. In Matlab experiments, the improved LHC algorithm was compared with LEACH, SEP, and DEEC through analysis and comparison from the aspects of network life cycle, energy consumption and CH number. The experimental results show that the LHC algorithm has the advantages of uniform CH distribution and balanced node energy consumption which effectively improve the energy efficiency and data transmission capacity, and prolong the life cycle of the observing instrument network. The LHC algorithm provides an important routing protocol for observing instrument network and real-time reliable data transmission.

INDEX TERMS Observation instrument network, wireless sensor network, cluster routing algorithm, layered architecture, energy heterogeneous.

I. INTRODUCTION

The cold and arid areas of China form about two-thirds of the total land area. Although the ecological environment in these areas is fragile, they contain indispensable resources for the national economy and have a prominent strategic position. Observation stations deployed in the field are important bases for obtaining scientific data and carrying out scientific research and scientific research data is mainly obtained through long-term and continuous monitoring of geological elements by field observation instruments [1]. In order to establish the observation network system in cold and arid areas, we made a detailed investigation of the field observation stations in cold and arid areas in China. As shown in Fig. 1, most of these observation stations are located harsh natural environments with high altitude, cold, and pathless locations, which result in a series of problems, such as poor real-time transmission and lower accuracy of data collection and processing. In particular, field stations mainly focus on in-station observation and research to result in insufficient regional comprehensive research. These problems have seriously restricted the development of field monitoring and scientific research in cold and arid areas [2].

A field observation instrument network is a wireless ad-hoc network technology to achieve real-time monitoring of specific areas in the field. It has many characteristics such as

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

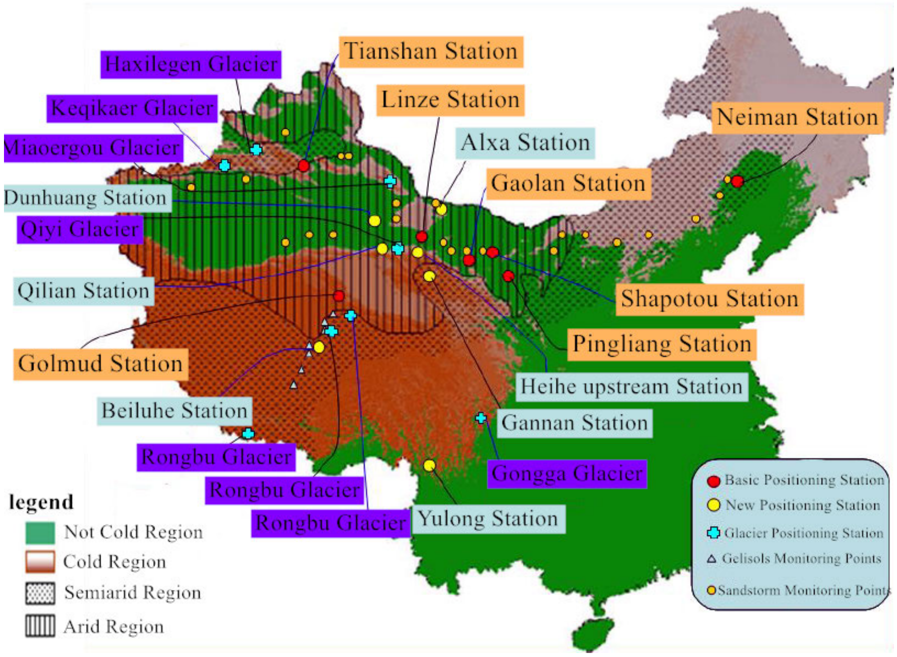

FIGURE 1. Field monitoring network for ecology and environment in cold and arid regions of China [3].

many observation elements, a large amount of transmitted data, complex structure, etc. Furthermore, it requires high real-time and accuracy of data collection, transmission and processing. Although the nodes of the field observation instrument network have a certain amount of solar cell auxiliary power supply, due to the harsh environmental impact of the deployment, energy consumption is still the primary problem that needs to be addressed. Therefore, according to the characteristics of observation instrument networks, it is necessary to build a routing protocol with an efficient energy balance and robust data transmission to solve the problem of network "hot spots" and the short network life cycle caused by uneven energy consumption of nodes. After the investigation and analysis of the field observation environment in cold and arid areas, the observation instrument network and wireless sensor network (WSN) [4] were compared, and it was found that both of them had similar characteristics such as self-organization, wireless transmission and dynamic topology, etc. Thus, according to the unique characteristics of instrument networks, the wireless sensor network can be adapted to meet the requirements of the observation instrument network.

Wireless sensor networks are a new type of network [5]–[7], which can conveniently monitor areas that need real-time monitoring and send data to Base Station (BS) to achieve specific functions. Nodes of WSN have the advantages of low cost and a self-organizing network, but also have the disadvantage of energy limitations [8], [9]. How to balance the energy consumption of each node to extend the life cycle of the network has great significance to the application of the improved WSN to the network of observation instruments mentioned in this paper.

The key point of the routing protocol for field station observation instrument networks is network clustering optimization and the combination of network clustering and inter-cluster routing. In this way, the problem of network death caused by the uneven energy consumption of nodes can be alleviated and avoided. However, current solutions of clustering have not fully solved the problem of uneven distribution of CHs and the unbalanced energy consumption of CHs in the network. Therefore, based on the clustering routing algorithm LEACH (Low Energy Adaptive Clustering Hierarchy), a layered and heterogeneous routing algorithm for field observation instrument network, LHC, is designed to achieve real-time, long-term and stable observations in cold and arid areas and provide high real-time, high quality and stable data for scientific research in these areas.

The main contributions of this study can be summarized as follows: (1) We model the network with a layered heterogeneous structure and the mechanism of energy and distance factors to improve the CH election method. (2) We propose that the LHC algorithm is not only suitable for field observation instrument networks but also to provide efficiency of nodes in such networks. (3) Through comprehensive simulations, we verify the performance of the proposed algorithm.

The rest of this paper is organized as follows. In **Section II**, relevant research work is described. In **Section III**, the network model of field observation instruments is established. The design idea and detailed steps of the LHC algorithm are described in detail in **Section IV**. The simulation experiments are described and analyzed in **Section V**. And in **Section VI**, the conclusions and future work are summarized and described respectively.

II. RELATED WORKS

A. ROUTING ALGORITHM

The primary function of a wireless sensor network routing algorithm is to construct the data transmission mode and path between the nodes and the BS, which is the core technology of the network layer of WSN. The quality of the routing algorithm has an essential impact on the performance of wireless sensor networks [10], [11]. To meet the requirements of various application scenarios, scholars have designed a variety of routing algorithms. Routing algorithms can be divided into different types according to different requirements. For example, based upon the data transmission path, they can be divided into single-path routing protocols and multi-path routing protocols [12], [13]; based upon whether or not the geographical location of nodes is utilized, they can be divided into location-based routing protocols and non-location-based routing protocols [14]; based upon the network topology, they can be divided into plane routing protocol and cluster routing protocol [15].

B. CLUSTERING ROUTING ALGORITHM

In the routing algorithms of WSNs, the clustering routing algorithm has the advantages of low energy consumption, high stability and good expansibility, and so on. In clustering routing algorithms, clustering is grouping, that is, all nodes in the network are divided into different groups according to specific application requirements, and each group is a cluster. Each cluster consists of a CH and multiple member nodes within the cluster. The working principle of the clustering routing algorithm can be summarized as follows: firstly, the member nodes in the cluster send the monitored data to the CH, and then the CH forwards its monitored data and all the received data to the next-hop node or BS after data fusion.

Cluster routing algorithm has the following advantages [16]:

1) Low energy consumption. On the one hand, the cluster head fuses the data of member nodes and then transmits it, which reduces data redundancy and transmission volume, and correspondingly reduces the energy consumption of nodes. On the other hand, a sleeping mechanism is introduced into intra-cluster communication, and member nodes in the cluster can be in a sleeping state when they are not sending data, which also reduces energy consumption to some extent.
2) High stability. The subnetwork formed by clustering is relatively stable, so the influence of the change of network topology structure on the network is reduced.
3) Good Scalability. Clustering can make the network adapt to the change of topological structure dynamically and has high expansibility, so it is suitable for large-scale networks.

The advantages of clustering routing algorithms make it a hot research topic for domestic and foreign scholars. In recent years, researchers have proposed many clustering routing algorithms for WSNs, which have made great breakthroughs in energy consumption and network life cycle [17].

Clustering routing algorithms of WSNs can be divided into two types: homogeneous clustering algorithms and heterogeneous clustering algorithms. At present, most clustering routing algorithms are homogeneous types, such as LEACH [18], LEACH-C [19], PEGASIS [20], HEED [21], etc. LEACH algorithm adopts a distributed random selection method of CHs, in which all nodes are periodically selected as CHs to achieve the goal of uniform energy consumption. LEACH-C is a centralized CH election algorithm, which requires that only nodes with energy higher than the average residual energy of the network can be elected as CHs. LEACH-C requires each node in the whole network to communicate directly with the BS to report its location information and energy information, to estimate the average value of the remaining energy of the network. The BS selects CHs according to the information sent by each node. Although the process consumes a lot of communication energy, it prolongs the stable period of the network. In literature [22], the Genetic Algorithm (GA) is used to replace the simulated annealing algorithm in selecting CHs based on the LEACH-C algorithm, which makes the process of CHs selection better. In Literature [23], fuzzy inference systems have been adopted into the CH selection. The goal of this paper is not only to prolong the lifetime of a mobile sensor network for a certain amount of energy, but also to reduce the packet loss in a mobile sensing environment. In literature [24], fuzzy inference systems have been adopted into the CH selection, too. In literature [25], a power distribution method is used to select CHs, and the network throughput is significantly increased. The LEACH-EPM algorithm proposed in literature [26] optimizes the selection of CHs. It uses the Dijkstra algorithm to establish the inter-cluster routing mechanism, which reduces the communication energy consumption between clusters. The EMESLSC algorithm [27] is a probabilistic clustering algorithm. The probability of nodes being selected as CHs depends on the residual energy, the length of the buffer and the power of received signals. PEGASIS algorithm [20] organizes all nodes into chains, and greatly reduces the number of nodes directly communicating with the BS through the effective chain structure. However, it requires each node to know the location information of other nodes in the network, which costs a large amount of energy. Moreover, if the network size is large, the chain length of the algorithm will increase the data transmission delay. Same as the LEACH algorithm, the HEED algorithm is a completely distributed clustering algorithm [21], which mainly studies and improves the uneven distribution of CHs. The HEED algorithm introduces the residual energy of nodes and the communication cost within the cluster as parameters, so that the selected CHs are well distributed in the network which results in faster clustering speed and a more uniform network energy consumption. However, this algorithm needs to carry out multiple data iterations when generating clusters, and the communication cost is high. Literature [28] found that the HEED algorithm is only applicable to a homogeneous network environment, and proposed a DWEHC algorithm [28] for heterogeneous networks. The DWEHC algorithm further improves the message iteration, makes the CH distribution more balanced, reduces the number of message iterations, and relatively reduces the energy consumption. The combination of WSNs and intelligent algorithms in clustering routing algorithms is also a research hotspot in recent years, such as the STBCP algorithm [29], MOFPL algorithm [30] and so on, and good research results have been obtained. In addition, the RRDLA algorithm [31] studied the QoS routing of WSNs and modeled the reliable routing problem as an MCOP problem. RRDLA algorithm achieved a good balance among multiple QoS constraints such as end-to-end delay and energy consumption. Literature [32] investigated the problem of self-protection in WSN and devised the SPLA algorithm, which is a novel and efficient algorithm. The SPLA algorithm not only preserves the self-protection but also provides efficiency of nodes.

The deployment environment of field observation instrument networks are more suitable for the clustered routing algorithms of the heterogeneous network. On the one hand, some nodes in an observation instrument network use solar cells for auxiliary power supply which forms network nodes with heterogeneous energy. On the other hand, the difference in the power supply capacity of solar cells will also lead to the energy heterogeneity of network nodes. In the related heterogeneous network research, Qing *et al.* improved the heterogeneous energy algorithm SEP [33] and proposed the DEEC algorithm [34]. The SEP algorithm is based on a second-level heterogeneous network. That is, there are advanced nodes and normal nodes in the network. The algorithm allocates different rotation cycles for nodes with different initial energy with the purpose of extending the stable period of the network. However, the CH election of the SEP algorithm is only based on the initial energy of nodes and does not consider the remaining energy of nodes. The DEEC algorithm extends the second-level heterogeneous network to the multi-level heterogeneous network, and the selection of CHs is determined by the residual energy level of nodes and the heterogeneity of the network, which can not only make full use of the heterogeneity of the network but also adapt to the change of energy of nodes. However, the DEEC algorithm uses an estimation scheme of the average residual energy of the network. It assumes that the network energy consumption is uniform, which is inconsistent with reality, thus weakens its practicability. The optimal number of CHs and the formula of average energy per round of the DEEC algorithm are improved in literature [35] to reduce the overall network energy consumption, but the node location is not considered. Literature [36] takes into account the location distribution of nodes, making it easier for nodes close to the BS to be selected as CHs. However, low-energy nodes are not protected, which can easily lead to the rapid death of nodes near the BS due to the repeated selection of CHs. Literature [37], [38] also as modified the SEP algorithm to address multi-level energy heterogeneity and adopted a probabilistic calculation method based on energy ratio in the stage of the CH election, which protects low-energy nodes to a certain extent. Aderohumu *et al.* proposed the E-SEP algorithm [39] which added intermediate nodes and divided the nodes in the network into three categories according to energy. Compared with the SEP algorithm, it further improves the stability of the network, but the distance between the nodes and the BS was not taken into account by the E-SEP algorithm. Faisal *et al* proposed the Z-SEP algorithm [40], in which the heterogeneous energy nodes in the network are deployed in different regions according to different energies, and the CHs are elected only in the regions where the advanced nodes are deployed. This partitioning deployment strategy can save energy consumption to some extent and prolong the network life cycle. Literature [41] proposed an adaptive and efficient dynamic clustering routing algorithm, EDFCM, which is similar to the DEEC algorithm. It also predicts the average energy consumption in the next round of the ideal state by estimating the average residual energy of the network. The EDFCM algorithm takes the predicted value and historical energy consumption as the reference factors for calculating the probability. In literature [42], an improved and heuristic clustering algorithm was proposed to obtain the optimal number and location of CHs. Compared with the SEP algorithm, it extended the network life cycle, but it shortened the stable period of the network.

To sum up, although these algorithms have improved the life cycle of the network to some extent, they still have not fully solved the problems such as the uneven distribution of CHs and the unbalanced energy consumption of CHs in the network. How to solve these problems effectively is still the research hotspot of WSNs, and it is also a key problem that needs to be solved as we introduce WSN into the observation instrument network.

III. NETWORK MODEL OF OBSERVATION INSTRUMENTS NETWORK

A. NETWORK OF OBSERVATION INSTRUMENTS

Based on the investigation and analysis of the field observation environment in cold and arid areas, and the comparative analysis of a field observation instrument network and a WSN, it is concluded that the network routing protocol for an observation instrument network should have the following characteristics:

1) The observation instrument network has fewer nodes than a WSN. However, in cold and arid areas with complex geographical environments, the deployment model of nodes of field observation instruments networks is the same as WSNs, where nodes are usually deployed in a random manner in specific areas.
2) Nodes of WSNs use battery power, and once the battery runs out, the nodes will not work. A field observation instrument network is usually powered by solar cells, and the power supply capacity is stronger than a wireless sensor network. However, due to the influence of harsh weather in the field, the battery may not be able to be charged for a long time and the power supply

efficiency cannot be guaranteed. Therefore, the main design strategy is that observation instrument networks should be efficient and have balanced utilization of energy.

3) WSNs usually have a single observation element and data transmission content is generally simple. However, observation instrument network has many observation elements, large data transmission volume and a complex structure, which require the observation instrument networks to have a large network throughput.
4) Generally, the hardware design of the nodes in observation instrument networks is complex, and the cost of the nodes is higher than that of normal wireless sensor nodes, which requires the node density to have less influence on the network performance.
5) The network of observation instruments should have the character of redundancy. The network can be automatically adjusted or rebuilt when the original cluster is destroyed due to node death or failure.

Solar cells are used as an auxiliary power supply in a part of the nodes of observation instrument networks, which inevitably leads to the energy heterogeneity of network nodes. Secondly, due to the different energy consumption rates of nodes in the process of network operation, energy heterogeneity is ubiquitous in the network. Therefore, the heterogeneous energy model is adopted to design the network routing algorithm for observation instrument networks. In addition, the routing protocol focuses on the network clustering optimization and the combination of network clustering and inter-cluster routing to alleviate and avoid the problem of rapid network death caused by the uneven energy consumption of nodes. In this paper, the main reasons for designing the network routing algorithm for observation instrument networks based on LEACH algorithm are as follows: (1) The LEACH algorithm is a classical clustering routing algorithm, which has the advantages of high stability, simple data fusion and high energy utilization; (2) The LEACH algorithm has a unique logical hierarchy, which can better support other important WSN technologies such as a security mechanism.

B. LEACH ALGORITHM

The Low Energy Adaptive Clustering Hierarchy algorithm (LEACH) is the first WSN cluster routing algorithm, which was proposed by Heinzekman *et al*. of Massachusetts Institute of Technology. It is a typical clustering and data fusion clustering routing algorithm [43]–[45], and most clustering routing algorithms were developed based on its idea of clustering. LEACH's basic idea is described as follows: CHs are selected periodically utilizing equal probability to reduce the number of direct communications between nodes and BS. New CHs will be selected in each round, and the energy load of the whole network will be averaged to each node, to reduce the energy consumption of network communication and extend the whole network life cycle [18]. The network architecture of the wireless ad-hoc network clustering network routing protocol based on LEACH is shown in Fig. 2.

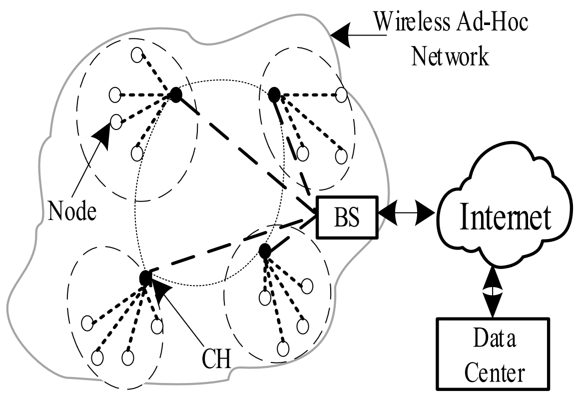

FIGURE 2. Network architecture of wireless ad-hoc network clustering network routing protocol.

The basic process of CHs election by the LEACH algorithm is as follows: each node in the network selects a random number between 0 and 1. If the value selected at random in the current round is less than the set threshold $T(n)$, then this node will be selected as the CH in the round.

The calculation formula of $T(n)$ is shown in Equation (1).

$$T\ (n) = \begin{cases} \dfrac{p}{1 - p[rmod(\frac{1}{p})]} & n \in G \\ 0 & others \end{cases} \tag{1}$$

where p is the probability of each node in the network being elected as CH, r is the number of current rounds, and G is the set of nodes that are not elected as CHs in $1/p$ round. When $T(n)$ is set to 0, all nodes on the network will start to elect CHs again. After the election of CHs, the CH nodes broadcast their CH status to the whole network, and the CSMA MAC protocol is adopted in the broadcast process to avoid conflicts. All non-CH nodes in the network judge which cluster to join according to the strength of the received signal and inform the relevant cluster heads, so the establishment of the cluster is completed.

The LEACH clustering routing algorithm has a unique logical hierarchy which can better support data fusion, security mechanisms and other important WSN technologies. Thus, it is suitable to be adapted for field observation instrument networks in cold and arid areas. However, the LEACH algorithm still has the following disadvantages [46]:

1) The energy consumption of CHs is difficult to balance. In the process of the CH election, the LEACH routing algorithm does not consider the remaining energy of nodes. Since each node has the same probability of becoming the CH, a node with lower energy can be elected as a CH, which will accelerate the death of this node and affect the performance of the whole network.
2) The CHs far away from the BS consume too much energy. From the network energy consumption

mode [18] adopted by the LEACH routing algorithm, it can be seen that there is a threshold value for the distance between the BS and the CH. When the distance is larger than the threshold value, the energy consumption of sending messages of CHs will be proportional to the fourth power of the distance, which will result in excessive energy consumption of the CHs far away from the BS.
3) The distribution of CHs is difficult to optimize. The CH selection mechanism of the LEACH routing algorithm makes it difficult to achieve the optimal number and distribution of CHs.

C. HETEROGENEOUS ENERGY NETWORK

The heterogeneous WSN is a network composed of a variety of different types of sensor nodes [33]. In heterogeneous networks, node energy heterogeneity is ubiquitous. It is because different types of sensor nodes may have different initial energies; even for the same type of nodes, due to the different rate of energy consumption, the remaining energies of nodes are varied in the process of network operation. Secondly, when new nodes are added to the network or some advanced nodes can replenish energy, the energy of network nodes will be heterogeneous [34]. Solar cells are used for auxiliary power supply in part of nodes of field observation instrument networks, and the heterogeneous energy network is more suitable for practical application scenarios.

In this paper, a two-level heterogeneous energy network is adopted, which consists of two types of sensor nodes with different initial energy: advanced nodes and normal nodes. Suppose E_0 is the initial energy of normal nodes, m is the proportion of advanced nodes in the network, and the energy of advanced nodes is a times higher than normal nodes. Of the total number of nodes N, the initial energy of m^*N advanced nodes is $E_0 * (1 + a)$, and the initial energy of $(1 - m) * N$ normal nodes is E_0. Then the total initial energy of the two-level heterogeneous energy network is shown in Equation (2).

$$E_{total} = N * E_0 * (1 + a * m) \tag{2}$$

D. BASIC ASSUMPTIONS

In order to facilitate simulation, we assume that the observation network of the study is distributed in a square area of M^*M, which is composed of N nodes randomly distributed in an observation instrument network. The BS is located at the edge of the area, and the whole network carries out the data collection periodically. The network is assumed as follows:

1) The nodes and the BS will not move after deployment.
2) Each node has the ability to conduct data fusion and each node has a unique ID.
3) Two-level heterogeneous energy of nodes.
4) The node can freely adjust its transmission power according to the distance of the receiver and the link is symmetrical. If the transmission power of the other party is known, the node can calculate the distance from the sender to itself according to the received signal strength indicator (RSSI) 47].
5) The BS is not limited by energy and does not need to consider energy consumption. When the BS transmits signals at specific transmitting power, all sensor nodes in the whole network can receive the signals.

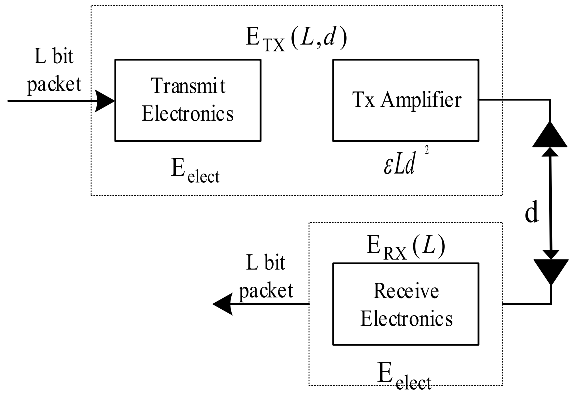

FIGURE 3. Energy consumption model of wireless communication.

E. ENERGY MODEL

The LHC algorithm adopts the same wireless communication energy consumption model as the LEACH algorithm [18], as shown in Fig. 3. The energy consumption model is divided into three components: E_{TX}, E_{RX} and E_{DF}, which represent the energy consumption of sending data, receiving data and data fusion, respectively. The energy consumed by a node sending L bits of data a distance of d consists of two parts: transmitting circuit loss and power amplification loss, as shown in Equation (3).

$$E_{TX}(L, d) = \begin{cases} L * E_{elect} + L * \varepsilon_{fs} d^2 & d < d_0 (Free\ space\ model) \\ L * E_{elect} + L * \varepsilon_{mp} d^4 & d \geq d_0 (Multipath\ decay\ model) \end{cases} \tag{3}$$

where E_{elect} represents the transmitting circuit loss, $L * E_{elect}$ is the energy consumed by the node to receive L bits of data, ε_{fs} and ε_{mp} represent the energy required for power amplification in the Free space model and Multipath decay model, respectively. d_0 is the distance threshold, and its calculation formula is shown in Equation (4).

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{4}$$

F. THE PRACTICAL APPLICATION

Observation instrument networks mainly include sensors, data acquisition instruments, nodes, a BS and sensor networks. In this paper, on the basis of the original sensors and data acquisition instrument on the observation platform of field stations, the key technology of observation instrument network is designed independently.

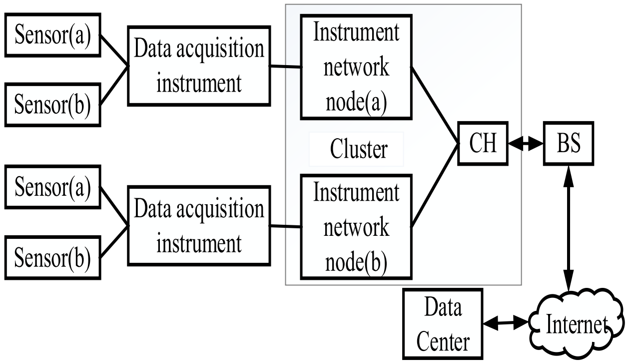

FIGURE 4. The practical application scenario of an observation instrument network.

Data acquisition instruments are connected with a number of sensors and the instrument network nodes deployed in the observation area. Nodes are connected with the deployed data acquisition instrument through RS232 serial port and send control commands to the data acquisition instrument to realize data automation, digital collection and reverse control of the data acquisition instrument.

The instrument network nodes constitute the network with the routing protocol of a wireless ad-hoc network, and transmit the packets to the BS, which returns the data to the data center through network technologies of the network transport layer. The BS is mainly responsible for receiving the observation data transmitted from the instrument network nodes. The data is then fused and transmitted back to the data center through network technologies. The practical application scenario of an observation instrument network is shown in Fig. 4.

IV. DESIGN OF LHC ALGORITHM FOR FIELD OBSERVATION INSTRUMENT NETWORK BASED ON LEACH ALGORITHM

In this paper, based on the characteristics of field observation instrument networks, a Layered and Heterogeneous Clustering routing algorithm LHC is proposed based on LEACH, which considers the residual energy and transmission distance of nodes.

A. NETWORK INITIALIZATION

Working the same as the LEACH algorithm, the LHC algorithm uses rounds as the working cycle of the algorithm, and each round includes the stages of CH election, cluster establishment, schedule establishment, and data acquisition and transmission. In order to ensure the effective working time of the network, the total working time of the first three stages of the algorithm is much less than that of the data acquisition and transmission stage, and the whole working cycle is automatically triggered to start again after a specified time.

The initialization of the observation instrument network is carried out only once in the initial stage of network operation, and there is no initialization stage in each subsequent cycle.

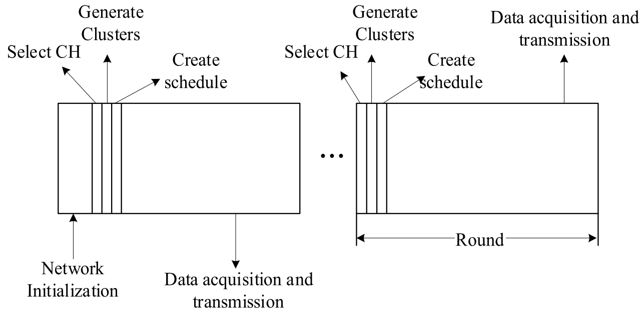

FIGURE 5. Sequence diagram of LHC algorithm.

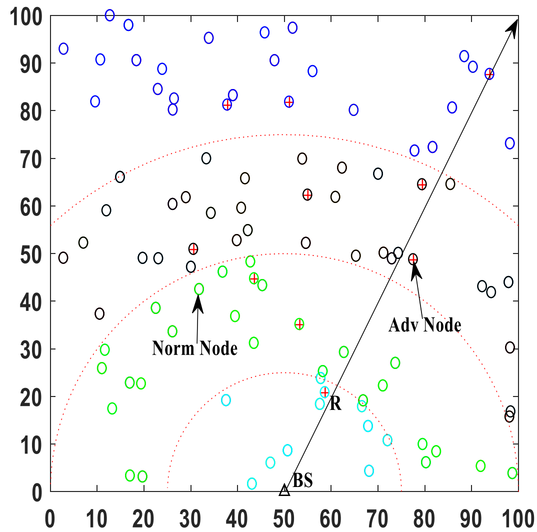

FIGURE 6. Layered structure of LHC algorithm.

The BS initiates network initialization. When the network deployment is completed, the BS will broadcast the message of starting. All nodes will start to work after receiving the message. The working sequence of the network is shown in Fig. 5.

Different from the LEACH algorithm, the LHC algorithm uses a hierarchical structure similar to the Progressive Multi-hop Rotational Clustered (PMRC) algorithm [48] to divide the network into multiple adjacent virtual sub-layers according to the size of the network in the initial stage of the network, and select a fixed number of CHs in each layer.

The structure is a hierarchical structure proposed by Mei *et al.* [48], which is suitable for large-scale WSNs. The structure divides the network into multiple layers, each of which is composed of several clusters. The CHs are responsible for forwarding the data of the cluster members to the inner CHs, and finally to the BS. The PMRC structure is designed for WSNs based on multi-hop clustering. As the CH needs to forward more data, it consumes too much energy, which limits the network life. Thus, only the similar hierarchical structure of PMRC is used to optimize the cluster head distribution in this paper.

Fig. 6. shows the hierarchical architecture adopted in this paper. According to the network model, all nodes are randomly distributed in the monitoring area, and the BS is

deployed in the central position at the edge of the monitoring area. When the network is initialized, the BS based on the network size and user requirements divides the observation area into k layers by the circles $O_i(i = 1, 2, 3 \ldots)$ which iR $(i = 1, 2, 3\ldots)$ are the radius of the circle, respectively. And the center of the circles is the BS. The first layer takes the BS as the CH because nodes located in it are close to the BS, and $k_{opt}/(k-1)$ CHs are selected for each layer of the other layers. k_{opt} denotes the number of optimal CHs of the network, which is an important parameter related to the clustering routing algorithm [37]. In literature [17], the author deduced the calculation formula of k_{opt} as Equation(5).

$$k_{opt} = \sqrt{\frac{N}{2\pi}} \cdot \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \cdot \frac{M}{d_{toBS}^2} \tag{5}$$

where d_{toBS} is the average distance from the CH to the BS, and is shown in Equation (6).

$$d_{toBS} = 0.765\frac{M}{2} \tag{6}$$

In order to evenly distribute the CHs, the relationship between the number of layers k and the optimal number of CHs (k_{opt}) is expressed in Equation (7).

$$k - 1 = \lceil \sqrt{k_{opt}} \rceil \tag{7}$$

Assuming that the side length of the observation area is M and the radius of the first circle (O_1) is R, then, the number of layers (k) should satisfy Equation (8).

$$k = \lfloor \frac{\sqrt{(\frac{M}{2})^2 + M^2}}{R} \rfloor = \lfloor \frac{M\sqrt{5}}{2R} \rfloor \tag{8}$$

Suppose the total number of sensor nodes in the network is N and number of sensor nodes in each layer is $n_i(i = 1, 2, \ldots, k)$, then the relationship between N and n_i satisfies Equation (9).

$$N = \sum\nolimits_{i=1}^{k} n_i \tag{9}$$

Then, the probability that a node in the $k-1$ layers will be elected as CHs is shown in Equation (10).

$$pi = \frac{k_{opt}}{N - n_1} \tag{10}$$

The probability of CH election for nodes in the $k-1$ layers of the LHC algorithm is pi. From Equation (5), it can be seen that the number of optimal CHs is related to the number of nodes and the size of the network. For example, $N = 100$ and $M = 100$, $k_{opt} \approx 9$. Then, when 3 CHs are selected from each layer, the radius R should be set to 25.

This layered structure can make the distribution of the elected CHs more uniform, which solves the problems of uneven distribution of the CHs in LEACH and other improved algorithms to some extent.

B. CLUSTER HEAD ELECTION

After dividing the whole network into k layers, each layer is equivalent to a small network. The mechanism of candidate CHs was introduced in this paper. Firstly, the candidate CHs are elected in each layer except the first layer, and then the optimal CHs are selected from the set of candidate CHs to be the final CHs.

1) ELECTION OF CANDIDATE CLUSTER HEADS

The election of CHs is the core of the whole clustering algorithm, and the elected CHs directly affects the performance of the algorithm. Whether a node is suitable for CH is closely related to the following two points:

a: THE RESIDUAL ENERGY OF THE NODE

The residual energy of node directly affects its viability. If nodes with relatively high residual energy elected to be CHs, the life cycle of the whole network could be extended. However, if nodes with lower energy are elected as CHs, they will die quickly due to the massive energy consumption, which will have a certain impact on the network topology, data transmission, life cycle and other aspects.

b: DISTANCE BETWEEN THE NODE AND THE BS

As can be seen from the energy model Equation (3), when the distance between the node and the BS is less than the given threshold value (d_0), the communication overhead is $O(n^2)$; but when the distance is higher than d_0, the communication overhead becomes $O(n^4)$. Therefore, if a node far away from the BS is elected as a CH, due to the long communication distance, the power consumption will be too high and the death of the node will be accelerated. Thus, to elect a node as the CH the residual energy of the node should be considered, as well as its distance from the BS. And if the residual energy is equivalent for all nodes, it is more advantageous for nodes close to the BS to be elected as CHs.

The residual energy of the node reflects its ability to process a message, which is called the energy factor. The distance between the node and the BS reflects the energy consumption rate of the node, which is called the distance factor. Through the above analysis, a cost function for the node(i) in the cluster(j) based on energy factor and distance factor is shown in Equation (11).

$$cost_{i,j}(r) = E_i(r)^2 + (1 - E_i(r)^2) \cdot D_i(r)^2 \tag{11}$$

where r is the current number of rounds, $E_i(r)$ represents the ratio of the residual energy of node(i) to the initial energy of normal nodes in the network, and $D_i(r)$ represents the ratio of d_0 to the distance between node(i) and the BS.

The expression of $E_i(r)$ is shown in Equation (12).

$$E_i(r) = \frac{E_i}{E_0} \tag{12}$$

The expression of $D_i(r)$ is shown in Equation (13).

$$D_i(r) = \frac{d_0}{d_i} \tag{13}$$

The LHC algorithm considers these two factors comprehensively, and sets the threshold value for nodes to be elected as CHs according to Equation (14).

$$T'(n) = \begin{cases} \dfrac{p_{nrm}}{1-(r\,mod(\frac{1}{p_{nrm}}))} \times cost_{i,j}(r) & n \in G_1 \\ \dfrac{p_{adv}}{1-(r\,mod(\frac{1}{p_{adv}}))} \times cost_{i,j}(r) & n \in G_2 \\ 0 & others \end{cases} \tag{14}$$

where, p_{nrm} is the probability that normal nodes are elected to be the CH, and the calculation formula is shown in Equation (15).

$$p_{nrm} = \frac{pi}{1 + a \cdot m} \tag{15}$$

p_{adv} is the probability that advanced nodes are elected to be the CH, and the calculation formula is shown in Equation (16).

$$p_{adv} = \frac{pi(1+a)}{1 + a \cdot m} \tag{16}$$

where r is the current number of rounds, G_1 is the set of normal nodes that have not been elected to be the CHs in the previous *1/p* rounds, and G_2 is the set of advanced nodes that have not been elected to be the CH in the previous *1/p* rounds.

In the process of CH election, the larger the cost ($cost_{i,j}(r)$) of a node and the larger value of $T'(n)$, then, the higher probability that it will be elected as CH. The probability that a normal node and an advanced node are selected as CHs is proportional to the residual energy of the node, and is inversely proportional to the distance from the node to the BS. After the completion of network initialization, the residual energy of all normal nodes is E_0, and that of advanced nodes is $E_0 * (1 + a)$. Thus, the probability that advanced node to be elected as CH is higher than normal nodes. With the operation of the network, the larger the residual energy and the closer to the base station a node is, the higher probability the nodes will be elected as a CH.

2) DETERMINATION OF THE FINAL CLUSTER HEAD

Although the LHC algorithm divides the network into layers in the initial stage, the distribution of cluster heads elected in each layer is more random. The mechanism to control the distance between cluster heads and set a minimum spacing value of CHs was adopted to make sure that the distance between CHs elected in each round in each layer will not be larger than this threshold value. In this way, data redundancy caused by the close distance of CHs can be avoided and the CHs can be distributed more evenly. If the observation network is evenly divided into K clusters in a square area of M^*M, and the region radius of the cluster is set as d, then the maximum d_{max} and minimum d_{min} of d can be expressed in Equation (17) and (18), respectively.

$$d_{min} = \frac{1}{2} \times \sqrt{\frac{M^2}{K}} = \frac{M}{2\sqrt{K}} \tag{17}$$

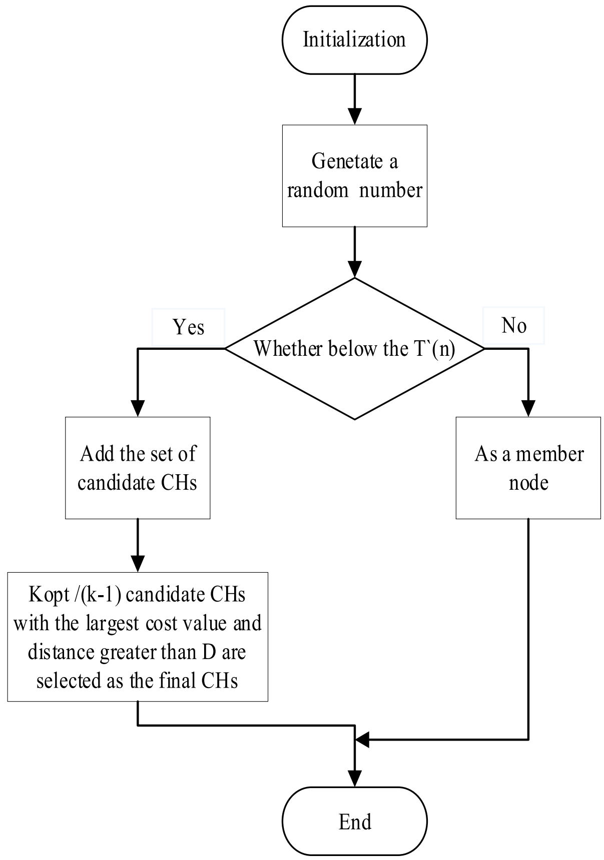

FIGURE 7. Flow chart of cluster head election in LHC algorithm.

$$d_{max} = \sqrt{\frac{1}{\pi} \times \frac{M^2}{K}} = \frac{M}{\sqrt{\pi K}} \tag{18}$$

Then, the region radius of each cluster should be between the interval $[\frac{M}{2\sqrt{K}}, \frac{M}{\sqrt{\pi K}}]$, and the minimum distance D between CHs is defined in Equation (19).

$$D = d_{min} = \frac{M}{2\sqrt{K}} \tag{19}$$

If in a round of CH election, two distances between candidate CHs are less than D, then the candidate with the lower value of the cost function $cost_{i,j}(r)$ will be discarded.

The process of electing CHs for each layer of the LHC algorithm is shown in Fig. 7.

C. CLUSTERING PROCESS

Each node that has elected itself a CH for the current round broadcasts an advertisement message to the rest of the nodes. For this "CH-advertisement" phase, the CHs use a CSMA MAC protocol, and all CHs transmit their advertisement using the same transmit energy. The advertisement message of CH includes node location, residual energy, current number of rounds, etc. The non-CH nodes must keep their receivers on during this phase of set-up to hear the advertisements of all the CH nodes. After other nodes receive the advertisement message, they decide which cluster (the nearest cluster) to join according to the strength of the received signal, and send the join information to the corresponding CH. After each node has decided to which cluster it belongs, it must inform the CH node that it will be a member of the cluster. Each node transmits this information back to the CH again using a CSMA MAC protocol. During this phase, all cluster-head nodes must keep their receivers on.

The cluster head node receives messages from all the nodes joining the cluster. Based on the scale of the cluster, the cluster head node creates a TDMA (schedule) and broadcasts it to the members of the cluster to inform them of the start time to transmit data, so that they can transmit data within the corresponding time, so as to ensure that the members of the cluster will not have conflicts when they transmit data.

D. THE SELF-HEALING MECHANISM OF CLUSTER

In the operation process of the whole network, whenever a node dies, the established cluster is destroyed, and the self-healing mechanism of the cluster should be considered. Node death can be divided into two cases: cluster member node death and cluster head node death.

1) The member node of the cluster dies. Members of the cluster should broadcast the news that they are about to die before running out of energy. Therefore, all normal nodes should have an energy threshold that is larger than the energy required for this broadcasting. Once their energy is smaller than this threshold, they will immediately start broadcasting their dying message. The message includes the ID information, location information, current number of rounds, and so on.
2) Cluster head node dies. The energy threshold of the CH is larger than the energy used to send the dying message to the BS. When the BS receives the news of a CH death, it immediately organizes a new round of CH election.

E. DATA ACQUISITION AND TRANSMISSION

After the establishment of the cluster, the network nodes begin to collect and transmit data. If the node in the cluster wants to send data, it will start collecting and sending data to the CH in a time slot allocated by the CH. If the node does not transmit data, it will turn off the wireless signal and sleep automatically to reduce the energy consumption of the node. When the CH receives the data transmitted by all the nodes in the cluster, the data is fused and forwarded to the BS. The purpose of data fusion is to eliminate redundant data, reduce data volume and energy consumption of transmission.

When the algorithm finishes executing for a period, the network restarts and enters the next round of CH election, clustering and data transmission.

As described above, the operation flow chart of LHC algorithm is shown in Fig. 8.

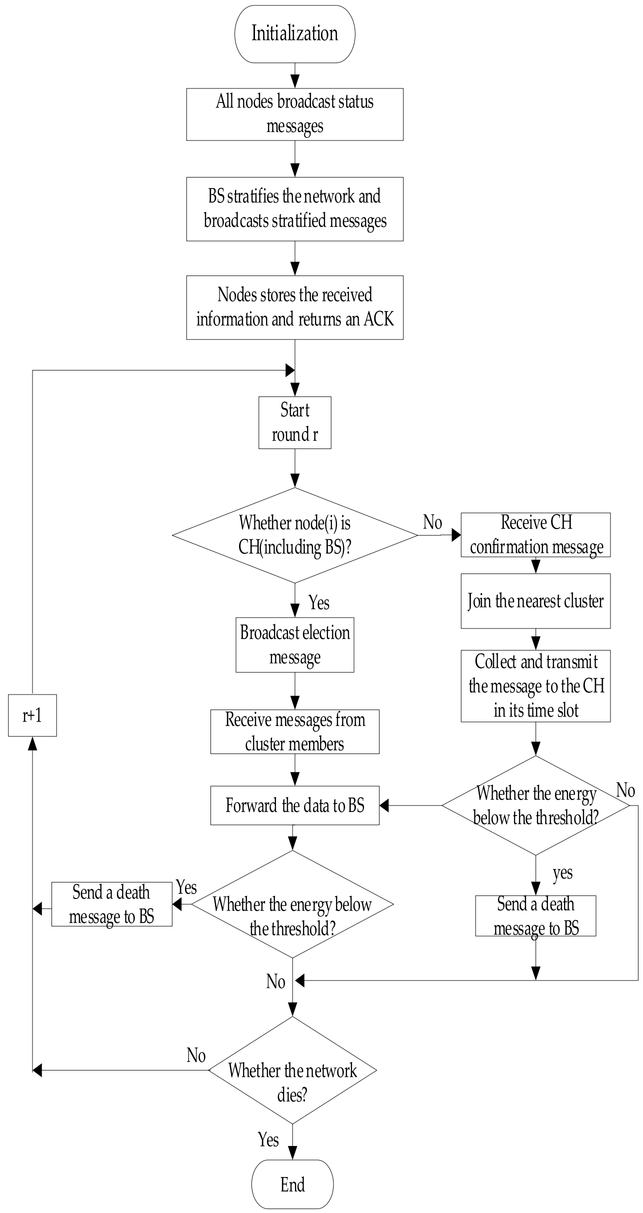

FIGURE 8. Workflow of LHC algorithm.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. ANALYSIS OF ALGORITHM COMPLEXITY

The complexity of the routing algorithm has an essential influence on the communication overhead of the network [49]. The complexity of the field observation instrument network should be kept constant or changed linearly with the number of nodes.

The messages generated by the LHC algorithm in each round are: in the stage of CH election, k_{opt} nodes are elected as CHs, and each cluster head broadcasts an elected message, the total number of messages is k_{opt}; after the BS receives the selected message of all CHs, it broadcasts a CH confirmation message to all nodes, and the total number of messages is 1; each node sends a cluster-added message according to the calculated result of the threshold value $T'(n)$, and the total number of messages is $N - k_{opt}$; Each CH confirms the addition of nodes, and the total number of messages to be sent is $N - k_{opt}$.

TABLE 1. Parameters configuration in simulation experiments.

Parameter	Value
Network area (*m*)	(100,100)
Coordinates of BS (*m*)	(50,0)
Total number of nodes *N*	100
Initial energy of node (*J*)	0.05
Data packet size L (*bits*)	4000
E_{elect} (*nJ/b*)	50
ε_{fs} ($pJ/b/m^2$)	10
ε_{mp} ($pJ/b/m^4$)	0.0013
d_0 (*m*)	87.7
Multiple to initial energy of normal nodes *a*	1
Percentage of advanced nodes number *m*	0.1

In a single round of the network running, the total number of messages generated is: $2N + 1 - k_{opt}$, so the message complexity of this algorithm is $O(N)$. This shows that the LHC algorithm has low message overhead and high energy efficiency. According to the analysis of algorithm complexity, the threshold ($T'(n)$) determines the message overhead of the algorithm. Choosing the appropriate $T'(n)$ can not only make more suitable nodes become CHs, but also avoid excessive messages overhead.

B. NETWORK SIMULATION EXPERIMENT

MATLAB was used to run the simulation experiments of the LHC algorithm in this paper, and the LEACH, SEP and DEEC algorithms were also implemented in the same experimental environment to evaluate and analyze the algorithm's performance. First, because some of the nodes in the field observation instrument network use solar cells for auxiliary power supply, the influence of heterogeneous parameters *a* and *m* on the improved algorithm were simulated and analyzed. Then, the LHC algorithm was compared with the other three algorithms in terms of CH distribution, network life cycle, network energy consumption and packet quantity.

1) SIMULATION PARAMETER SETTING

The parameters of the simulation experiments selected in this paper are shown in Table 1. In literature [33], [34], sufficient simulation experiments have been carried out on the SEP and DEEC algorithms for heterogeneous energy networks. The initial value of the heterogeneous parameters in LHC algorithm is consistent with SEP algorithm and DEEC algorithm.

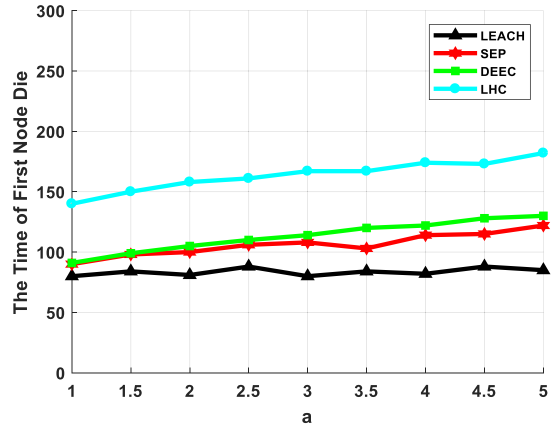

FIGURE 9. Round of the first node dies when *a* is varying from 1 to 5.

2) INFLUENCE OF HETEROGENEOUS PARAMETERS A AND M ON THE LHC ALGORITHM

Field observation instrument networks are different from the general WSNs in terms of the power supply method. Due to the deployment of observation instrument networks in harsh environments, some of the nodes use solar cells as an auxiliary power supply. In this paper, the heterogeneous parameter *a* is a multiplier that describes the initial energy of an advanced node based on the initial energy in a normal node. In the LHC algorithm, it is used to describe the power supply capacity of the node using solar cells as a power supply. We assume $a = 1$, and varying *a* from *1-10* represents the increase in power supply capacity. The parameter *m* is the proportion of advanced nodes. It is used to describe the number of nodes using solar cells as an auxiliary power supply. We assume that $m = 0.1$, and when *m* varies from 0.1 to 0.9, this represents the increase in number of advanced nodes.

The parameters *a* and *m* are mainly used to verify whether the LHC algorithm can make full use of supplementary energy. Therefore, *a* and *m* are essential indexes for evaluating the performance of an observation instrument network.

In the second-level heterogeneous network, the parameter values of advanced nodes to the total number of nodes (*m*) and the multiple (*a*) to initial energy of normal nodes were changed gradually in the simulation experiments, and the performances of four clustering algorithms (LEACH, SEP, DEEC and LHC) were observed.

Fig. 9. describes the time (round) of first node death of the four algorithms in the cases of $m = 0.1$ and as *a* increases from 1 to 5. It can be seen from Figure 9 that, the increase in energy of the advanced nodes in the LEACH algorithm has almost no effect on the death time of the first node. This is because all nodes in the LEACH algorithm can be elected as CHs with the same probability, and the residual energy of nodes is not considered in the CH election. In SEP, DEEC and LHC algorithms, the increase in energy of the advanced nodes leads to a delay in the death time of the first node, which increases the network stability period. Among them, LHC has the best effect on extending the stable period of the network.

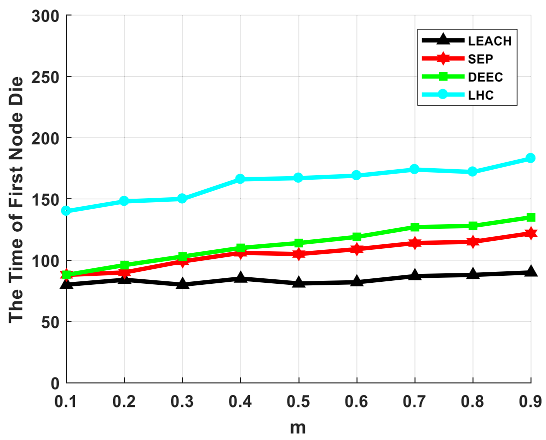

FIGURE 10. Round of the first node death when m is varied from 0.1 to 0.9.

Fig. 10. describes the first node death time (round) of the four algorithms when $a = 1$ and m increases from 0.1 to 0.9. As it can be seen from the figure, the increase in the number of advanced nodes does not affect the death time of the first node of the LEACH algorithm, which also indicates that it is more suitable for isomorphic networks. For the other three algorithms, increasing the number of advanced nodes will prolong the death time of the first node.

As it can be seen from Fig. 9. and Fig. 10, the LEACH algorithm cannot make good use of the increase in energy by the changes to m and a, which result in a shorter stable period for LEACH in the whole process. It suggests that the LEACH algorithm is not a suitable algorithm for heterogeneous network clusters, thus, it cannot be applied to field observation instrument networks directly.

In addition, the LHC algorithm can make good use of the energy increase by the change of a, and its stable period is longer than that of the LEACH, SEP and DEEC algorithms. In other words, as the number of deployed advanced nodes increases or the power supply capacity of solar cells increases, the stable period of an observation instrument network can be significantly extended.

3) DISTRIBUTION ANALYSIS OF CLUSTER HEADS

The distribution of CHs has a significant influence on the performance of the network. Fig. 11 shows the CH distribution when each algorithm runs to 100 rounds in the same simulation experimental environment. In the figure, "o" represents normal nodes, "+" represents advanced nodes, "∗" represents CHs, and "" represents the BS.

It can be drawn from Fig. 11. (A), (B), and (C) that the CHs of the LEACH, SEP and DEEC algorithms are not evenly distributed. At some rounds, the number of nodes between each cluster varies greatly, and there are cases of massive clusters and tiny clusters. In this way, the energy consumption of CHs in clusters with many nodes will be large, and may result in the "hot spot" problem in the network. Also, cluster heads may be dense in some regions and absent in other regions. As a result, nodes far away from the CH have to increase their transmission power in order to maintain communication with the CH which results in unbalanced energy consumption. As can be seen from Fig. 11(D), due to the implementation of hierarchical architecture in the LHC algorithm, a fixed number of CHs are selected in each layer and the minimum distance between CHs is set, the CHs are evenly distributed, and the distance between nodes in the cluster and CHs is relatively short. The LHC algorithm can maintain better network communication and extend the network life cycle.

Cluster heads play a very important role in clustering networks. Whether the number of CHs can be kept stable is an important criterion to measure and evaluate the clustering algorithms. Table 2 and Fig. 12(A) show the number of CHs of four algorithms from 1 to 100 rounds in the simulation process. As can be seen from the chart, the number of CHs in the LEACH, SEP and DEEC algorithms fluctuate severely. In particular, the DEEC algorithm changes the probability of nodes being selected as CHs, resulting in the phenomenon of zero CHs. Therefore, it is difficult for the number of CHs to be stable near the optimal number of CHs. However, in the LHC algorithm, the number of CHs keeps near the optimal number of CHs and remains stable. As shown in Fig. 12(B), this stability will not be broken until the number of dead nodes starts increasing.

TABLE 2. The number of CHs generated during the execution of four algorithms.

Algorithm	Rounds				
	10	20	50	80	100
LEACH	6	5	8	16	5
SEP	15	8	13	8	12
DEEC	10	3	4	12	4
LHC	9	9	9	9	9

4) THE LIFE CYCLE OF NETWORK

Table 3 and Fig. 13 show the network lifecycle of four different algorithms. Due to the improvements of the SEP, DEEC and LHC algorithms over the LEACH algorithm, the performances of these three algorithms are all better than the LEACH algorithm. The improvements of the SEP and DEEC algorithms were carried out under the assumption that the BS was located at the center of the observation area. Thus, in the case of BS locating at the edge of the observation area, the two algorithms perform moderately. In practical application, the BS is generally located outside the observation area. The LHC improves the CH election method and adopts a hierarchical structure to make the CHs distributed evenly. Therefore, when the BS is located at the edge of the observation area, the number of rounds before the first node death and the total rounds before network death of LHC algorithm are all better than the other two algorithms.

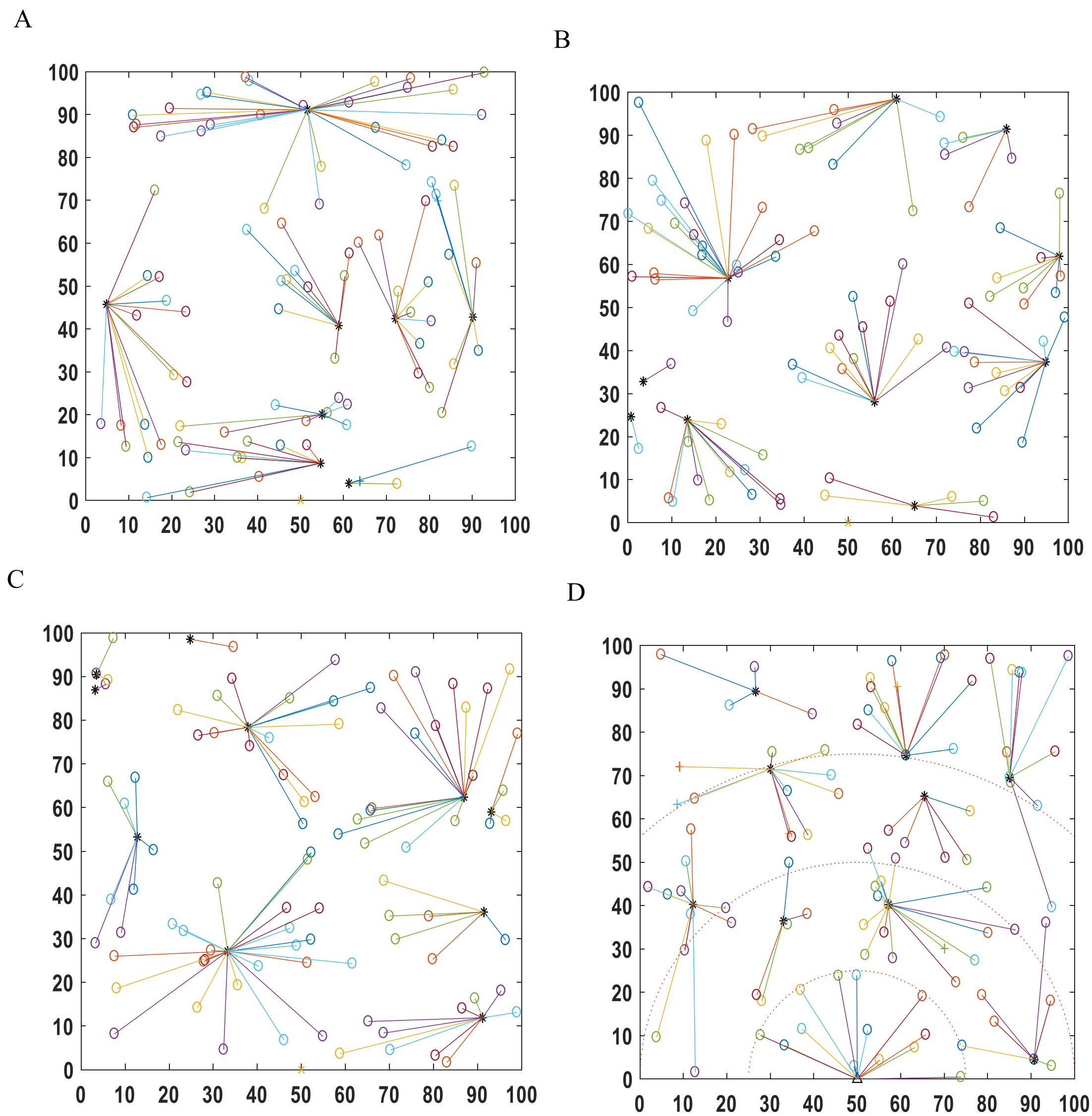

FIGURE 11. Cluster head distribution diagram of four algorithms at 100 rounds, among which (A) LEACH algorithm, (B) SEP algorithm, (C) DEEC algorithm, and (D) LHC algorithm.

TABLE 3. Comparison of network life cycles for the four algorithms.

Algorithm	First Node Dies	70% Node Dies	Last Node Dies
LEACH	80	133	329
SEP	90	136	255
DEEC	91	156	207
LHC	140	297	658

As can be seen from Table 3, the first node death occurs when the LHC algorithm runs to about 140 rounds, and 70% of the nodes die when it runs to about 300 rounds. The death time of the first node of the LHC algorithm is 75% higher than that of the LEACH algorithm, and the death time of 70% nodes is 120% higher than the LEACH algorithm. It is because the nodes with lower residual energy and nodes far away from the BS in the LEACH algorithm may also be selected as CHs, which will cause the premature death of the CHs due to excessive energy consumption. The LHC algorithm first divides the network into k layers according to the optimal number of CHs, and then selects the optimal number of CHs in each layer, which can ensure the CHs are evenly distributed in the network and can balance the energy consumption of the whole network. The LHC algorithm also takes into account the energy factor and distance factor in the CH election to make the elected CH the most suitable node to be the CH, so as to prolong the network life cycle.

Fig. 14 shows the comparison diagram of the remaining energy of the four algorithms in the whole network. It can be

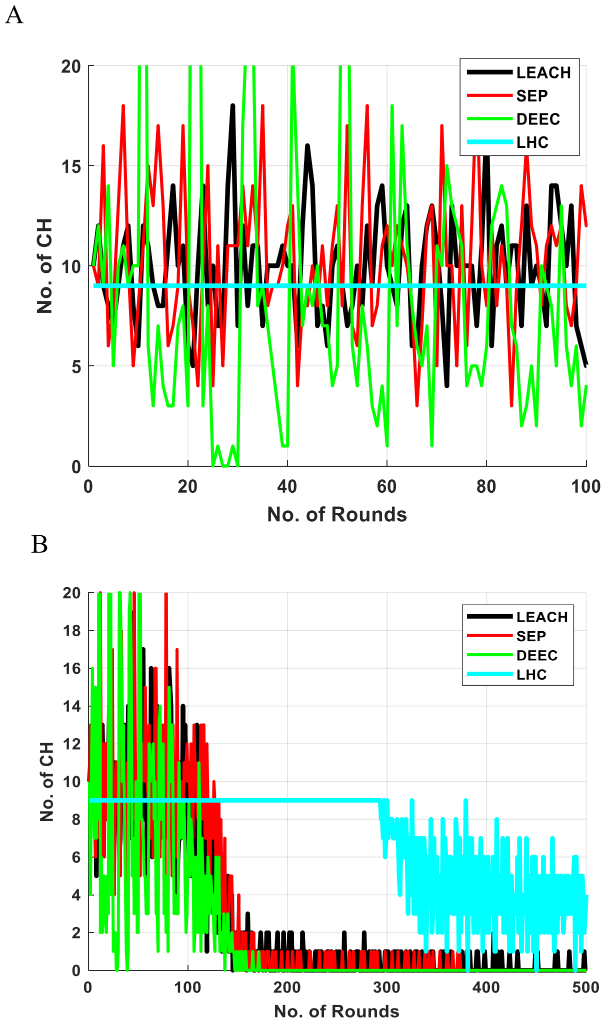

FIGURE 12. Comparison of the number of CHs in a round between LHC and other three algorithms. (A) Comparison of the number of CHs in the first 100 rounds. (B) Comparison of the number of CHs in the whole network life cycle.

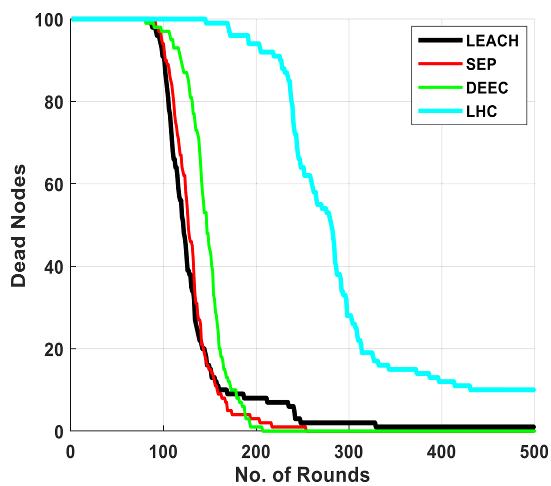

FIGURE 13. Comparison of life cycle of LHC and other three algorithms.

seen from the figure that the remaining energy of the LHC algorithm is obviously higher than the other three algorithms

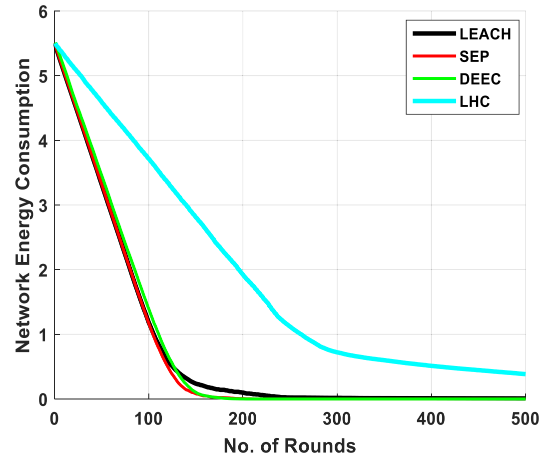

FIGURE 14. Comparison of residual energy of LHC and other three algorithms.

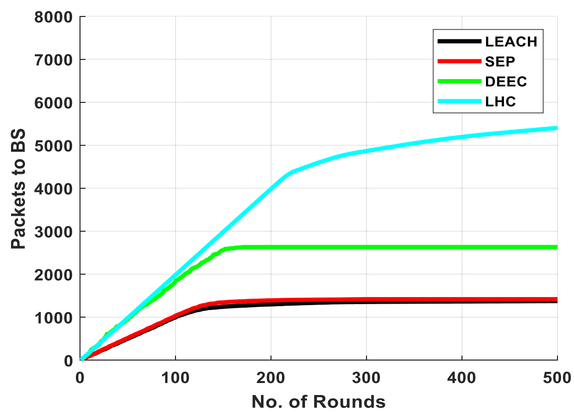

FIGURE 15. Comparison of data packets received by the BS with the number of rounds.

which indicates it has the highest energy efficiency of the algorithms.

5) NUMBER OF PACKETS RECEIVED

Table 4 shows the statistical comparison of packets received by the BS after 50, 100, 150, 200 and 500 rounds of the four algorithms. It can be seen that when running to 500 rounds, the number of packets received by using the LHC algorithm is about 5400, which is much larger than the other three algorithms. This indicates that the LHC algorithm can meet the demand of large data transmission volume in observation instrument networks.

Fig. 15 is a comparison of the number of packets received by the BS in the four algorithms with the number of rounds. As can be seen from the figure, the total number of packets received by LHC algorithm before the death of the first node was two times that of both the SEP algorithm and LEACH algorithm, and 1.1 times that of the DEEC algorithm. The number of packets received by LHC algorithm before 200 rounds increased linearly, and most data was received in the algorithm process, indicating that LHC was stable in operation and higher in energy efficiency.

TABLE 4. Comparison of packets received by the BS.

Algorithm	Rounds				
	50	100	150	200	500
LEACH	507	980	1240	1380	1419
SEP	518	1006	1343	1391	1427
DEEC	986	1926	2559	2632	2632
LHC	1000	1969	2987	3966	5398

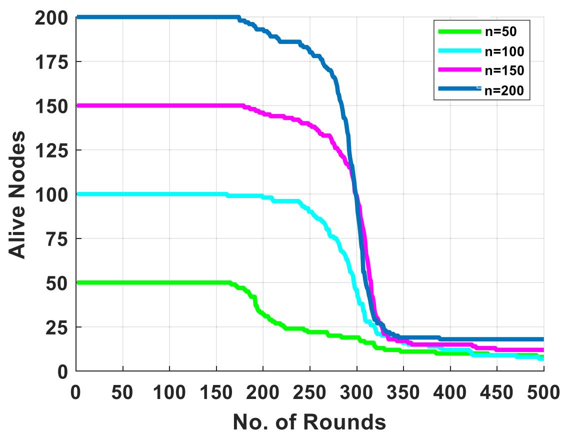

FIGURE 16. Comparison of the life cycles with different density.

The number of packets received by the DEEC algorithm fluctuated significantly, which is mainly due to the large fluctuation in the number of generated CHs leading to the fluctuation in the amount of sent to the BS. The amount of data received by LEACH algorithm and SEP algorithm increased approximately linearly before about 100 rounds and then decreased slowly. The main reason was that the number of dead nodes increased rapidly after 100 rounds, resulting in a decrease in the amount of data received by the BS.

6) INFLUENCE OF NODE DENSITY ON THE LHC ALGORITHM

Generally, there are fewer nodes in a field observation instrument network, and the hardware design of the nodes is sophisticated which results in a higher cost per node compared to WSNs nodes. Therefore, the influence of node density on the network life cycle is also an important indicator to measure the performance of the observation instrument network. Fig. 16 shows the comparison of the LHC algorithm life cycle when the number of nodes is 50, 100, 150 and 200 in the network area of 100m∗100 m. It can be seen clearly that the node density has little influence on the stable period of the LHC algorithm, which is of considerable significance to observation instruments networks due to the high cost of nodes.

VI. CONCLUSION

To address a series of problems that exist in field observation stations, such as poor real-time, poor accuracy of data collection and processing, and failure to realize networked observation, a Layered, Heterogeneous and Clustering network routing algorithm (LHC) was designed based on the LEACH algorithm. This paper mainly does the following: (1) Due to the fact that some nodes in field observation instrument networks use solar cells as an auxiliary power supply, the routing algorithm is designed by using the heterogeneous energy model; (2) To improve the uneven distribution of CHs in clustering routing algorithms, a layered mechanism is adopted to improve the CH distribution effectively; (3) Considering that the selected CHs in clustering routing algorithms do not adequately consider the residual energy and location of nodes, a cost function is proposed to improve the threshold equation of CH election, and a candidate CH mechanism is introduced to make the selected CHs better. (4) To solve the problem of unbalanced energy load caused by an unstable number of CHs in clustering routing algorithms, the optimal number of CHs strategy is adopted in the improved algorithm which selects a stable number of CHs and reduces the energy burden of CHs in each round. The comprehensive and in-depth simulation experiments show that the CH distribution in the LHC algorithm is uniform and the energy consumption of nodes is balanced, which effectively improves the energy efficiency and data transmission capacity of the observation instrument network. It prolongs the life cycle of the network and provides an important routing method for realizing observation and reliable data transmission of field instruments networks. However, the drawbacks of the LHC algorithm can be summarized as follows: (1) The LHC algorithm adopts single-hop data transmission, which will cause the nodes far away from the BS consuming more energy, resulting in an uneven distribution of dead nodes. (2) The candidate CH mechanism of the LHC algorithm will cause the speed of cluster generation to be slower than the LEACH algorithm.

At present, research on the routing protocols of field observation instrument networks is still in the initial stage. This paper conducts exploratory research in theory, proposes a routing algorithm which can be suitable for field observation instrument networks, and carries out theoretical analysis and relevant simulation experiments. In the future, we will try to develop and implement the protocol on a hardware node, and study the performance and adaptability of the LHC algorithm in the actual deployment environment.

REFERENCES

[1] J. Y. Huo, Z. N. Ren, Y. R. Yang, Z. Zhang, Y. N. Zhang, and M. L. A. Ren, "A kind of instrument network node equipment and monitoring system," Chinese Patents, 2 017 210 303 230, May 1, 2018.

[2] J. Huo, Y. Zhang, and L. Liu, "The research and design of network routing protocol for field observation instruments," *Int. J. Sensor Netw.*, vol. 16, no. 2, pp. 87–99, 2014.

[3] Y. J. Ding, *Cultivated in the Ice and Snow Desert, Formed the Strong Foundation of Scientific Research in the Cold and Arid Areas—Written in the Field Observation and Experimental Research in the Cold and Arid Areas for 50 Years*. Science Press, 2009, pp. 572–574.
[4] A. Boukerche, A. Mostefaoui, and M. Melkemi, "Efficient and robust serial query processing approach for large-scale wireless sensor networks," *Ad Hoc Netw.*, vol. 47, pp. 82–98, Sep. 2016.
[5] C. Zhu, L. T. Yang, L. Shu, V. C. M. Leung, T. Hara, and S. Nishio, "Insights of top-*k* query in duty-cycled wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 62, no. 2, pp. 1317–1328, Feb. 2015.
[6] Q. Jiang, J. Ma, X. Lu, and Y. Tian, "An efficient two-factor user authentication scheme with unlinkability for wireless sensor networks," *Peer Peer Netw. Appl.*, vol. 8, no. 6, pp. 1070–1081, Nov. 2015.
[7] F. Akhtar and M. H. Rehmani, "Energy replenishment using renewable and traditional energy resources for sustainable wireless sensor networks: A review," *Renew. Sustain. Energy Rev.*, vol. 45, pp. 769–784, May 2015.
[8] H. Shen and G. Bai, "Routing in wireless multimedia sensor networks: A survey and challenges ahead," *J. Netw. Comput. Appl.*, vol. 71, pp. 30–49, Aug. 2016.
[9] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.
[10] L. Guo and Q. Tang, "An improved routing protocol in WSN with hybrid genetic algorithm," in *Proc. 2nd Int. Conf. Netw. Secur., Wireless Commun. Trusted Comput.*, Apr. 2010, pp. 289–292.
[11] W. Rehan, S. Fischer, M. Rehan, and M. H. Rehmani, "A comprehensive survey on multichannel routing in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 95, pp. 1–25, Oct. 2017.
[12] X. Cao, R.-C. Wang, H.-P. Huang, L.-J. Sun, and F. Xiao, "Multi-path routing algorithm for video stream in wireless multimedia sensor networks," *J. Softw.*, vol. 23, no. 1, pp. 108–121, Mar. 2012.
[13] K.-W. Chin, J. Judge, A. Williams, and R. Kermode, "Implementation experience with MANET routing protocols," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 5, p. 49, Nov. 2002.
[14] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance," *IEEE Trans. Mobile Comput.*, vol. 2, no. 4, pp. 337–348, Oct. 2003.
[15] J. Zhao, A. T. Erdogan, and T. Arslan, "A novel application specific network protocol for wireless sensor networks," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2005, pp. 5894–5897.
[16] C. Gherbi, Z. Aliouat, and M. Benmohammed, "A survey on clustering routing protocols in wireless sensor networks," *Sensor Rev.*, vol. 37, no. 1, pp. 12–25, Jan. 2017.
[17] H. El Alami and A. Najid, "Optimization of energy efficiency in wireless sensor networks and Internet of Things: A review of related works," in *Nature-Inspired Computing Applications in Advanced Communication Networks*. Hershey, PA, USA: IGI Global, 2020, pp. 89–127.
[18] 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.*, Jan. 2000, pp. 10–22.
[19] A. P. Chandrakasan, H. Balakrishnan, and A. C. Smith, *Application-Specific Protocol Architectures for Wireless Networks*. Boston, MA, USA: Massachusetts Institute of Technology, 2000.
[20] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power efficient gathering in sensor information systems," in *Proc. IEEE Aerosp. Conf.*, Mar. 2002, pp. 23–29.
[21] 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. 2004.
[22] A. Rahmanian, H. Omranpour, M. Akbari, and K. Raahemifar, "A novel genetic algorithm in LEACH-C routing protocol for sensor networks," in *Proc. 24th Can. Conf. Electr. Comput. Eng. (CCECE)*, May 2011, pp. 1096–1100.
[23] J.-S. Lee and C.-L. Teng, "An enhanced hierarchical clustering approach for mobile sensor networks using fuzzy inference systems," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 1095–1103, Aug. 2017.
[24] H. El Alami and A. Najid, "A new fuzzy clustering algorithm to enhance lifetime of wireless sensor networks," in *Proc. Int. Afro-Eur. Conf. Ind. Advancement*. Cham, Switzerland: Springer, 2016, pp. 68–76.
[25] M. K. U. Khan and K. S. Ramesh, "A modified LEACH algorithm for WSN: MODLEACH," in *Proc. Int. Conf. E-Bus. Telecommun.*, 2019, pp. 632–642.
[26] Y. Chen, L. Jiang, and Y. Mu, "A LEACH-based WSN energy balance routing algorithm," in *Proc. World Symp. Softw. Eng.*, Sep. 2019, pp. 37–41.
[27] D. N. Kanellopoulos and P. Gite, "A probability-based clustering algorithm with CH election for expanding WSN life span," *Int. J. Electron., Commun., Meas. Eng.*, vol. 9, no. 1, pp. 1–14, Jan. 2020.
[28] P. Ding, J. A. Holliday, and A. Celik, "Distributed energy-efficient hierarchical clustering for wireless sensor networks," in *Proc. 1st IEEE Int. Conf. Distrib. Comput. Sensor Syst.*, Jun. 2005, pp. 322–339.
[29] G. Kia and A. Hassanzadeh, "STBCP: An energy efficient sub-threshold bee colony-based protocol for wireless sensor networks," *Int. J. Sensors, Wireless Commun. Control*, vol. 9, no. 4, pp. 443–453, Sep. 2019.
[30] R. Bhardwaj and D. Kumar, "MOFPL: Multi-objective fractional particle lion algorithm for the energy aware routing in the WSN," *Pervas. Mobile Comput.*, vol. 58, Aug. 2019, Art. no. 101029.
[31] H. Mostafaei, "Energy-efficient algorithm for reliable routing of wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5567–5575, Jul. 2019.
[32] H. Mostafaei and M. S. Obaidat, "Learning automaton-based self-protection algorithm for wireless sensor networks," *IET Netw.*, vol. 7, no. 5, pp. 353–361, Sep. 2018.
[33] S. I. M. G and A. Bestavros, *SEP:A Stable Election Protocol for Clustered Heterogeneous Wireless Sensor Networks*. Boston, MA, USA: Univ. Computer Science Department, 2004.
[34] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Comput. Commun.*, vol. 29, no. 12, pp. 2230–2237, Aug. 2006.
[35] C. Divya, N. Krishnan, and P. Krishnapriya, "Modified distributed energy-efficient cluster for heterogeneous wireless sensor networks," in *Proc. IEEE Int. Conf. Emerg. Trends Comput., Commun. Nanotechnol. (ICECCN)*, Mar. 2013, pp. 611–615.
[36] M. Shaji and S. Ajith, "Distributed energy efficient heterogeneous clustering in wireless sensor network," in *Proc. 5th Int. Conf. Adv. Comput. Commun. (ICACC)*, Sep. 2015, pp. 130–134.
[37] T. N. Qureshi, N. Javaid, A. H. Khan, A. Iqbal, E. Akhtar, and M. Ishfaq, "BEENISH: Balanced energy efficient network integrated super heterogeneous protocol for wireless sensor networks," *Procedia Comput. Sci.*, vol. 19, pp. 920–925, Jan. 2013.
[38] N. Javaid, T. N. Qureshi, A. H. Khan, A. Iqbal, E. Akhtar, and M. Ishfaq, "EDDEEC: Enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks," *Procedia Comput. Sci.*, vol. 19, pp. 914–919, Jan. 2013.
[39] F. Aderohunmu and J. D. Deng, "An enhanced stable election protocol (SEP) for clustered heterogeneous WSN," Dept. Inf. Sci. Univ. Otago, Dunedin, New Zealand, Tech. Rep., 2010, pp. 19–24, vol. 23.
[40] S. Faisal, N. Javaid, A. Javaid, M. A. Khan, S. H. Bouk, and Z. A. Khan, "Z-SEP: Zonal-stable election protocol for wireless sensor networks," 2013, *arXiv:1303.5364*. [Online]. Available: https://arxiv.org/abs/1303.5364
[41] Z. Hong, L. Yu, and G.-J. Zhang, "Efficient and dynamic clustering scheme for heterogeneous multi-level wireless sensor networks," *Acta Automatica Sinica*, vol. 39, no. 4, pp. 454–460, Apr. 2013.
[42] B. A. Attea and E. A. Khalil, "A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks," *Appl. Soft Comput.*, vol. 12, no. 7, pp. 1950–1957, Jul. 2012.
[43] R. K. Kodali, V. S. K. A., S. Bhandari, and L. Boppana, "Energy efficient m- level LEACH protocol," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Aug. 2015, pp. 973–979.
[44] H. Li and J. Liu, "Double cluster based energy efficient routing protocol for wireless sensor network," *Int. J. Wireless Inf. Netw.*, vol. 23, no. 1, pp. 40–48, Mar. 2016.
[45] X. W. Ma and X. Yu, "Improvement on LEACH protocol of wireless sensor network," *Appl. Mech. Mater.*, vols. 347–350, pp. 1738–1742, Aug. 2013.
[46] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados, "Energy-efficient routing protocols in wireless sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 551–591, Qua. 2013.
[47] X. Li, H. Shi, and Y. Shang, "A sorted RSSI quantization based algorithm for sensor network localization," in *Proc. 11th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Jul. 2005, pp. 557–563.
[48] M. Yang, S. Wang, A. Abdelal, Y. Jiang, and Y. Kim, "An improved multi-layered architecture and its rotational scheme for large-scale wireless sensor networks," in *Proc. 4th IEEE Consum. Commun. Netw. Conf.*, Jan. 2007, pp. 855–899.
[49] J. Bilal, F. Haleem, J. Huma, and M. Khan, "Energy efficient hierarchical clustering approaches in wireless sensor networks: A survey," *Wireless Commun. Mobile Comput.*, vol. 17, pp. 1–14, Oct. 2017.

JIUYUAN HUO received the master's degree from Lanzhou University, in 2007, and the Ph.D. degree from the Chinese Academy of Sciences, in 2012. He is currently a Professor with Lanzhou Jiaotong University. His research interests include optimization algorithms, data mining, and computer networks.

JIGUANG YANG received the bachelor's degree in computer science and technology from the Gansu University of Political Science and Law, in 2018. He is currently pursuing the master's degree in technology of computer with Lanzhou Jiaotong University. His research interests include wireless sensor networks and the Internet of Things, specifically in fields of algorithm improvement.

HAMZAH MURAD MOHAMMED AL-NESHMI received the bachelor's degree in computer networks from the University of Science and Technology, Yemen, in June 2017. He is currently pursuing the master's degree in computer application technology with Lanzhou Jiaotong University, China. His research interests include networking and machine learning, specifically in machine learning algorithms optimization.