WILEY | Hindawi

Research Article

Energy Balance Multihop Clustering Routing Protocol for Large-Scale Water Quality Monitoring

Yulin Gong [iD], Jiannan Cao, Chengfei Han, and Yunqing Liu

School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130000, China

Correspondence should be addressed to Yulin Gong; gongyulin@cust.edu.cn

Due to the influence of the environment on the water quality wireless sensor network, it is difficult to replace the node energy at any time, so it must make the most of the little energy available. This work provides a technique that combines cluster head selection, cluster structure creation, and data transmission into one optimal scheme. Firstly, we optimize the cluster head election probability threshold formula on the basis of LEACH and introduce overlap ratio in the competitive mechanism, to avoid excessive overlap between cluster heads. Subsequently, to alleviate the "hot spot" problem caused by multihop, the competitive mechanism of I-LEACH is optimized, which is based on the nonuniform competitive mechanism in the EEUC algorithm when electing the cluster head. Meanwhile, in the structural planning between cluster heads, the optimal path is searched based on the parallelism of genetic algorithm (variable path coding strategy); the combination of I-LEACH and EEUC-IGA, named Energy Balance Multihop Clustering Routing Protocol (BEBMCR), can avoid the emergence of "hot paths" and reduce running time. The simulation results show that BEBMCR still has a longer stable period and higher energy utilization rate under large-scale networks, and node energy consumption is more balanced.

1. Introduction

In recent years, with the rapid development of the Internet of Things (IoT), wireless sensor networks, one of the key technologies to realize the Internet of Things [1], have been widely used. It also brings new ideas to the field of water environment monitoring, which can be applied to harsh and inconvenient applications that are difficult for humans to reach and require long-term large-area monitoring.

The wireless sensor network is mainly composed of low-power sensor nodes and forms a multihop network in a self-organizing manner, with real-time perception, data processing, storage, and transmission capabilities. Since the general water area is widely distributed, it is difficult to maintain the network after deployment, so it can work in unattended mode [2, 3]. Limited by the size, weight, and cost of water quality sensor nodes, batteries are usually used to power nodes with limited energy. At present, the research hotspots of wireless sensor networks are mainly to save energy and extend the life of nodes.

1.1. Routing Algorithms of Homogeneous WSNs. Heinzelman et al. (2000) proposed Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol [4]. This protocol randomly selects some nodes as cluster head nodes; the node is sent to the base station through the cluster head. The random selection of cluster heads balances the energy of nodes to a certain extent and prolongs the working life of the network. However, the protocol also has excessive energy consumption caused by random election of cluster heads and single-hop data transmission. In recent years, scholars have proposed various improvement methods for these problems. Distance-and-Energy-Aware Routing (DEARER) combines energy and distance to make network energy consumption more balanced [5]. The same idea, combining the energy quality function and the distance quality function to determine the cluster head, can make the selection of the cluster head more uniform and reasonable and reduce the power consumption in data transmission [6]. However, the data transmission stage is not considered, and it is easy to form a single intercluster route. In the process of random selection

---

Input: Node i energy $E_0(i)$, node i position $d_0(i)$.
Output: Optimal path for data transmission.
1 **for** $i = 1$ to max **do**
2    $T_{\text{new}} = (1 - p)T(n) + pf(i, j)$
3: Where $f(i, j) = w_1(E(i)/\overline{E_{\text{cur}}}) + w_2(\text{Number}.N_{\text{nbor}}/(1/p) - 1) + w_3(1/d_{\text{toBS}}) + w_4(1/d_{(i,j)})$
4: **if** $k_{\text{or}} \leq 0.03$ **then**
5    $n_i \longrightarrow CH$
6    $R_i = 1.16(1 - c(d_{\max} - d_i/d_{\max} - d_{\min}))r_{\text{opt}}$
7   Calculate $d(CH_i, CH_j)$
8   **end if**
9   **if** $d(CH_i, CH_j) \leq d_0$ **then**
10   min $f(x_i)$
11   Choose the best path $X'' = \{x_1'', x_2'', \cdots, x_i'', \cdots, x_n''\}$
12 **end if**
13 **end for**

ALGORITHM 1: BEBMCR.

of cluster heads, the energy of cluster heads needs to meet certain standards to replace new cluster heads, which reduces excessive energy consumption caused by repeated election of cluster heads by nodes and extends the network life cycle [7]. On the basis of the original cluster head, a secondary cluster head is selected as a guarantee that the cluster head node fails or cannot complete the work alone, which can improve the robustness of network transmission [8]. Elmonser et al. proposed Dynamic Multihop LEACH (DMH-LEACH), which combines dynamic clustering, multihop transmission, and node mobility to improve node capabilities, but ignores the impact of other locations and node overlap on clustering. For homogeneous WSNs, Shahbaz et al. proposed a multi-path routing solution, in which the wireless sensor network is clustered using the firefly algorithm, but the search method relies too much on excellent individuals, which reduces the convergence speed [9, 10].

1.2. Routing Algorithms of Heterogeneous WSNs. Research on the two-level energy heterogeneous model, based on the LEACH algorithm, combined with the difference in node initial energy, proposes the SEP algorithm. This protocol optimizes the number of cluster heads and sets higher cluster head elections for nodes with higher energy. Thereby, node energy consumption is more balanced, the disadvantage is that the routing adopts single-hop transmission, and the transmission energy consumption is high [11]. For the three-level energy heterogeneous model, the weighted cluster head election strategy is dynamically selected through the initial energy of the node. Combined with the remaining energy of the node, the purpose of extending the life cycle of the network is achieved [12]. In the Energy-Efficient Uneven Clustering (EEUC) protocol, in order to reduce the direct energy consumption of the cluster head, the protocol uses a nonuniform cluster competition mechanism. However, the disadvantage is that the residual energy of the cluster head node and the uneven distribution of the cluster head are not considered [13]. The Distribute Energy-Efficient Clustering Algorithm (DEEC) is based on a general multi-level energy heterogeneous model. The rationality of selecting cluster heads needs to consider the current energy of the node and the average remaining energy in the network. However, only the remaining energy is considered, and the actual distance from the node to the sink node is not fully considered [14]. A centralized nonuniform clustering routing protocol (C-EEUC) takes into account the residual energy of nodes and the cost of communication between nodes and cluster heads. However, the optimal number of cluster heads is not considered for clustering [15].

In order to solve the above problems, the proposed scheme improves the basic idea:

(1) According to the remaining energy of the node, the location of the node, and the density of neighbor nodes, optimize the threshold formula for electing cluster heads

(2) Introducing overlap rate into the competition mechanism, combined with the concept of competition radius, solves the overlap of the cluster head node coverage area, stabilizes the number of cluster heads in the network, and helps improve the throughput and energy utilization of cluster heads [7, 8]

(3) The topological structure of the water quality monitoring network is designed using a genetic algorithm, and the algorithm is made suitable for large-scale network architectures by combining single-hop and multihop communication

(4) The BEBMCR Algorithm 1 draws on the EEUC nonuniform competition mechanism in the election of cluster heads [13]. In the selection of the next hop, it searches for the optimal path based on the efficiency and parallelism of the improved genetic algorithm and uses the variable path coding strategy to integrate communication distance, the relative remaining energy, and the number of nodes in the cluster
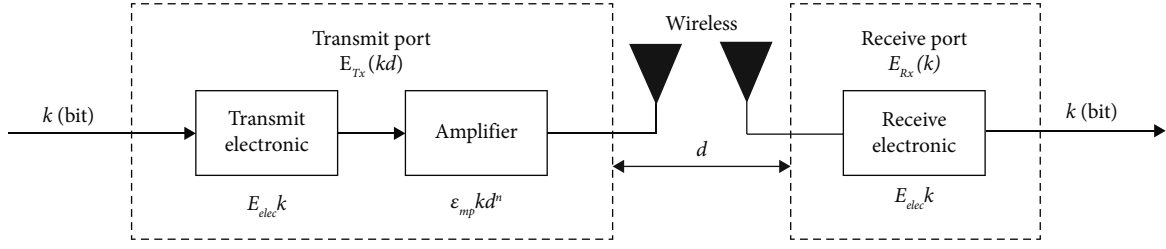
FIGURE 1: Energy consumption model.

2. Establishment of Network Model

The homogeneous WSN routing algorithm [4–10] and the heterogeneous WSN routing algorithm [11–15] use the same first-order radio energy consumption model for analysis. The energy consumption analysis of the network wireless communication model is exactly the same as the LEACH algorithm, so the analysis will not be repeated here.

2.1. Energy Consumption Model. The wireless communication energy consumption model is shown in Figure 1. It needs to consume data to transmit k bit data to the sink node with a distance of d_{toBS} [16, 17].

$$E_{Tx}(k, d) = \begin{cases} E_{\text{elec}}k + \varepsilon_{\text{fs}}kd_{\text{toBS}}^2, & d_{\text{toBS}} < d_0, \\ E_{\text{elec}}k + \varepsilon_{\text{mp}}kd_{\text{toBS}}^4, & d_{\text{toBS}} \geq d_0, \end{cases} \quad (1)$$

where E_{elec} is the energy consumption of sending 1 bit data, ε_{fs} is amplification characteristic constant, ε_{mp} is the multi-path attenuation model amplification characteristic constant, and d_0 is the distance threshold of the two models.

The calculation formula of d_0 is as follows:

$$d_0 = \sqrt{\frac{\varepsilon_{\text{fs}}}{\varepsilon_{\text{mp}}}}. \quad (2)$$

The calculation formula for the energy consumed by the node receiving k bit data is as follows:

$$E_{\text{Rx}}(k) = E_{\text{elec}}k. \quad (3)$$

In addition to the above energy consumption, the cluster head node also needs to perform data fusion after receiving the data. The energy E_d consumed by the k bit data is:

$$E_d = kE_{da}, \quad (4)$$

where E_{da} is the energy of the cluster head node compressing 1bit data.

2.2. Network Model. Random water quality sensor and random point distribution within a certain area, outside the base station area, and outside the area have the following contents:
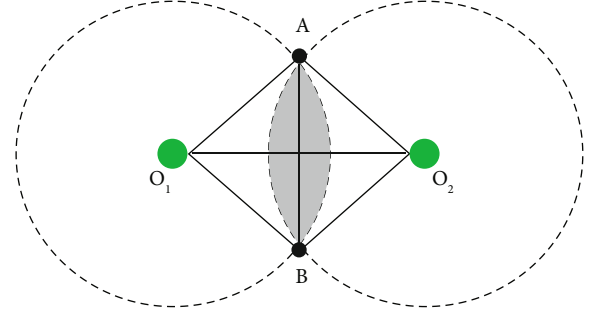

FIGURE 2: Overlap of two CH fields.

(1) The water quality sensor nodes in the network are spread at random, and the positions of the sensor nodes and base station nodes will not alter once the deployment is complete

(2) Except for the base station, each sensor node in the network has a unique ID number, and the computing power, initial energy, and storage space of each node are exactly the same

(3) Each water quality sensor node can judge the location of the sending end node by analyzing the angle and intensity of the received signal

(4) The data sent by the sensor node can adjust the communication power according to the communication distance

(5) The sensor node can automatically adjust the working and sleep modes according to the TMDA time slot [18]

2.3. Overlap Ratio. In the ideal state of network energy consumption balance, the DEEC algorithm calculates the cluster head probability by averaging the remaining energy, and other factors are ignored, such as the distance between the node and the base station and the density of adjacent nodes [14]. It is ensured that nodes with high residual energy, geographical location close to the base station, and high density of adjacent nodes have a greater probability of becoming cluster heads, but there are too many overlapping areas between cluster heads, which are not relatively independent, as shown in Figure 2, CH O_1 and CH O_2 repeatedly cover a larger area, and will increase the energy consumption of the node to a certain extent in this range [19].

Taking into the overlap of the cluster head competition area, the cluster head enters the final determination link combined with the calculation formula of the competition radius R_c. The fields of cluster head O_1 and cluster head O_2 are circles with radius R, the area of which is recorded as S, and the area of quadrilateral O_1AO_2B is recorded as $S_{O_1AO_2B}$. The areas of sector AO_1B and sector AO_2B are recorded as $S_{A\widehat{O_1}B}$ and $S_{A\widehat{O_2}B}$, respectively, and the shadow area is recorded as S_{over}, and get:

$$\begin{cases} S_{A\widehat{O_1}B} = S_{A\widehat{O_2}B} = \dfrac{\theta \pi R^2}{360}, \\ S_{O_1AO_2B} = |O_1O_2| \cdot |AB|, \\ S_{over} = S_{A\widehat{O_1}B} + S_{A\widehat{O_2}B} - S_{O_1AO_2B}. \end{cases} \tag{5}$$

where $|O_1O_2|$ and $|AB|$ can be expressed as:

$$\begin{aligned} |O_1O_2| &= \sqrt{\Delta x^2 + \Delta y^2}, \\ |AB| &= 2R \sin(\theta/2), \end{aligned} \tag{6}$$

where $\theta = \angle AO_1B = AO_2B$, $(\Delta x, \Delta y)$ is the relative position coordinates of O_1 and O_2.

Using the law of cosines, we can get:

$$\cos \theta = \frac{|O_1O_2|}{2R}. \tag{7}$$

The value of θ can be calculated using formula (7):

$$\theta = 2R \cos^{-1} \frac{|O_1O_2|}{2R}. \tag{8}$$

From formulas (5), (6), and (8), the overlap area ratio of the region can be calculated as k_{or} (shaded area):

$$\begin{aligned} k_{or} &= \frac{S_{over}}{S} = \frac{S_{A\widehat{O_1}B} + S_{A\widehat{O_2}B} - S_{O_1AO_2B}}{S} \\ &= \frac{\cos^{-1}\left(\sqrt{(\Delta x)^2 + (\Delta y)^2}\right)}{90} - 4R \cdot \frac{(\Delta x)^2 + (\Delta y)^2}{\pi R}. \end{aligned} \tag{9}$$

When the overlap area ratio is $k_{or} = 30\%$, which can get $|O_1O_2| = 1.16R$, the range of the water quality sensor's cluster head broadcast message is $1.16R_0$. In this way, in order to ensure that adjacent CH nodes have competition radius alternation, broadcasting within the defined range can save node energy consumption to a certain extent. The main purpose is to ensure that the nodes are evenly distributed and to avoid the repeated area of cluster heads being too large.

3. Improved LEACH Protocol for Cluster Head Election

In the cluster head election phase, add node remaining energy or node position or node density to improve the cluster head election formula, or Heinzelman et al. adopted non-

uniformity; the idea of clustering is mainly to reduce the scale of clusters close to the sink node. Although the phenomenon of "hot spots" is alleviated, when the actual water quality sensor detects water quality indicators, due to the influence of geography and weather, adjacent CH nodes have alternate radii of competition. This part of the improvement of LEACH (I-LEACH) is an improved algorithm proposed for the poor scalability of the above algorithms.

3.1. Cluster Head Selection Stage. According to the LEACH algorithm analysis in Section 2.1, it can be seen that the average energy considered by the cluster head election probability is in an ideal state, and regardless of this factor, the distance between the node and the base station, but the energy consumption of the node is affected by the distance. In addition, the random selection of cluster heads will also lead to uneven distribution of cluster heads in the network. In this section, the selection of the temporary cluster head will be based on the LEACH algorithm, combined with the current energy and location conditions of the node to obtain a new threshold formula [20]. Figure 3 summarizes the cluster head selection process.

Step 1. Information exchange between node and neighbor nodes

In order to ensure that all nodes are active, the base station first broadcasts a message and then broadcasts an initialization message to announce the start of the cluster head election.

After receiving the message, the node stores the data in the message and estimates the distance from itself to the base station d_{toBS} and calculates k_E, at the same time, to decide whether to run for cluster heads. The nodes participating in the campaign broadcast messages with the maximum communication radius to obtain neighbor node information [7].

$$k_E = \frac{E_i(0)}{\overline{E_{cur}}}, \tag{10}$$

where $E_i(0)$ is the remaining energy of the node at the beginning and $\overline{E_{cur}}$ is the average remaining energy of the network.

Step 2. Election of temporary CHS by new threshold formula

By comparing and selecting the largest fitness value, if the fitness value is the same, the second round of calculation is performed. The candidate cluster head with a relatively large fitness value is selected as the cluster head during the cluster structure formation stage, and the election message is broadcast. Takele et al. proposed a new threshold calculation method, which mainly considers the method including the current remaining energy of the node, the density of adjacent nodes, the distance between the node and the base station, and the average distance between the node and the adjacent nodes [21].

$$f(i,j) = w_1 \frac{E(i)}{\overline{E_{cur}}} + w_2 \frac{Number.N_{nbor}}{(1/p) - 1} + w_3 \frac{1}{d_{toBS}} + w_4 \frac{1}{d_{(i,j)}}, \tag{11}$$
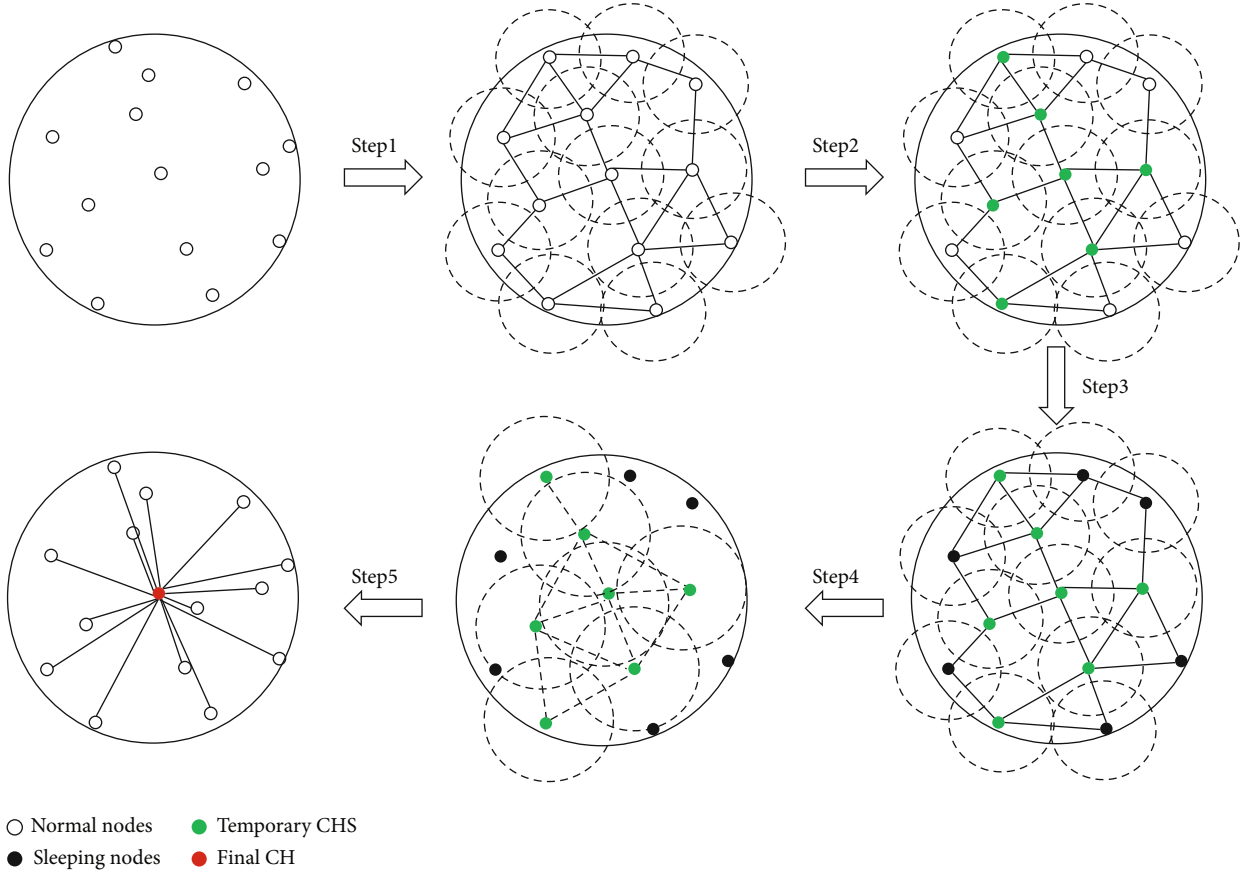
Figure 3: I-LEACH algorithm cluster head selection process diagram.

where $E(i)$ is the current remaining energy of the node; d_{toBS} is the distance between the node and the base station; $d_{(i,j)}$ is the average distance between the node and neighbor nodes, and Number.N_{nbor} is the number of neighbor nodes within the radius, and $(1/p) - 1$ is the number of nearby nodes in the standard cluster.

The maximum coverage communication range of an area with dense water quality sensor nodes is R, that is, the node distance is less than R to be defined as a dense node area.

$$N_{nbor} = \left\{ j d\left(CM_i, CM_j\right) \leqslant R, j \in N \right\}, i \in \{1, 2, \cdots, N\},$$

$$R = \sqrt{\frac{S}{\pi N \rho}},$$

(12)

where $d(CM_i, CM_j)$ is the distance between the ordinary node CM_i and the ordinary node CM_j of the cluster where it is located, ρ is the percentage of the number of surviving nodes in the CH node number station, and $S = M \times M$ is the area of the monitoring area.

The calculation formula of $d_{(i,j)}$ is as follows:
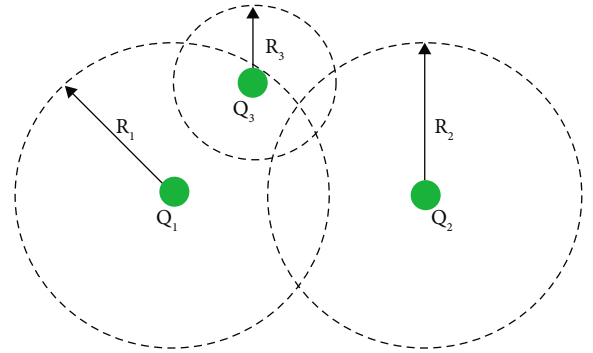
$$d_{(i,j)} = \frac{\sum d(i,j)}{Number.N_{nbor}},$$

(13)

Figure 4: Improve the uneven competitive mechanism.

where $d(i, j)$ is the distance between node i and node j and w_1, w_2, w_3, and w_4 are weighting coefficients, and $w_1 + w_2 + w_3 + w_4 = 1$.

In the improved threshold formula, the probability of a node being selected as a temporary cluster head is related to the current remaining energy of the node, the relative density of the node, the distance from the base station, and the average distance from the node.

Step 3. The candidate cluster heads that have failed the election temporarily go to sleep

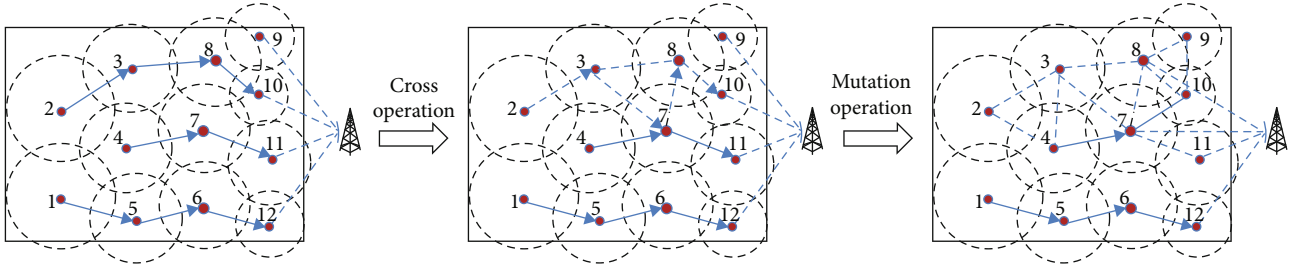*Step 4.* Information exchange between temporary CH and neighbor CHs

FIGURE 5: Searching for optimal path process by IGA.

TABLE 1: Setting of simulation parameters.

Parameters	Value
Field size ($M \times M$)	100m × 100m
Location of BS	(150, 50)
Number of water quality sensor nodes, N	100
Initial energy of water quality sensor nodes, E_0	0.5 J
The energy consumption of transceiver circuit, E_{elec}	50nJ/bit
The energy for aggregation, E_{DA}	5nJ/bit/signal
Free space, ε_{fs}	10pJ/bit/m^2
Multipath fading, ε_{mp}	
The control packet size, k	4000bits
The data packet size, c	32bits
w_1, w_2, w_3, w_4	0.3, 0.2, 0.3, 0.2
α, β	0.5, 0.5
Overlap ratio, k_{or}	30%

For specific content, refer to Section 2.3 and Section 3.2.

Step 5. The final CH node is selected, and the final CH node broadcasts the elected message and necessary information.

3.2. Overlap Ratio Add to Competitive Mechanism. In the final determination of the cluster head node, the overlap rate is introduced in I-LEACH. The temporary cluster head takes itself as the center and broadcasts the message with $1.16R$ as the competition radius. If there are other temporary cluster heads, compare the remaining energy of these temporary cluster heads, specify that the temporary cluster head with the most remaining energy is the final cluster head in the area, and broadcast the elected cluster head message to the network and the remaining temporary cluster heads that fail the broadcast announced its withdrawal from the competition. When the temporary cluster heads in the same competition area receive the news that other cluster heads are elected, they will automatically announce their withdrawal from the competition.

$$R = \frac{M}{\sqrt{\pi k_{opt}}}, \tag{14}$$

where M is the side length of the monitoring area and k_{opt} is the theoretical optimal number of CHS.

I-LEACH optimizes the probability threshold formula when selecting cluster head to ensure high standby energy and high distance to base station; nodes with high density of adjacent nodes have more action opportunities than cluster head. At the same time, the overlap rate is introduced to reduce the cost. Since there is not much overlap between cluster heads, they are relatively independent, which is conducive to improving the throughput and energy utilization of cluster heads.

4. Algorithm for the "Hot Paths" Problem

In order to reduce the complexity of the wireless sensor network without affecting the effectiveness of the EEUC protocol, in the routing algorithm for searching for the optimal path, this paper chooses a variable-length path coding strategy to improve the genetic algorithm. On the one hand, the nonconformity of the EEUC protocol is used, which is the idea of uniform clustering. On the other hand, it improves the efficiency of genetic algorithm to solve the optimal value [13, 22].

4.1. Improved EEUC Protocol. In order to address the issue of "hot paths" in the network, cluster heads near the base station consume too much energy, causing these nodes to perish quickly. The optimization in EEUC's nonuniform competition mechanism is based on the election of cluster
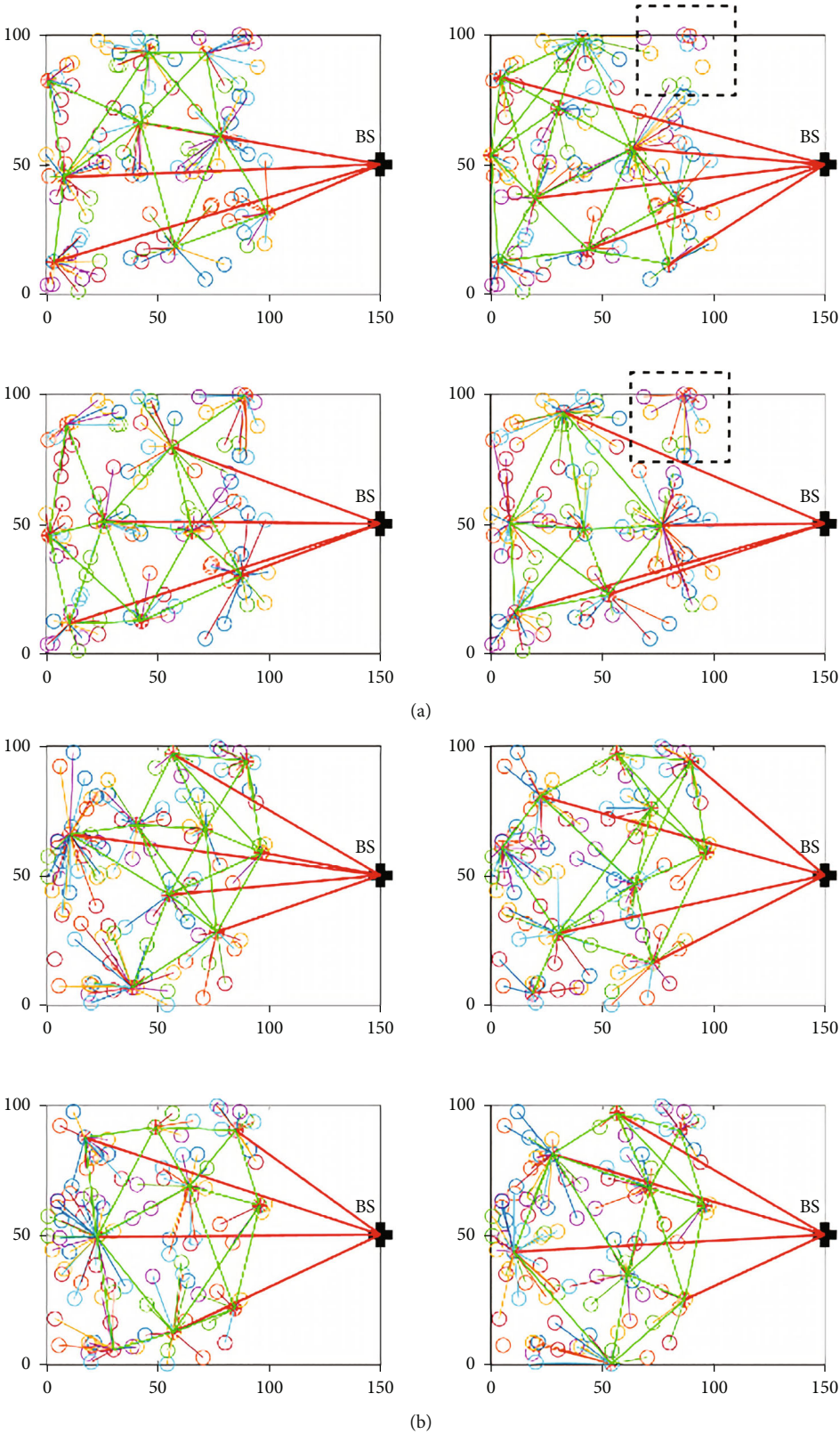
FIGURE 6: Simulation network model: (a) with EEUC routing protocol, (b) with BEBMCR routing protocol.
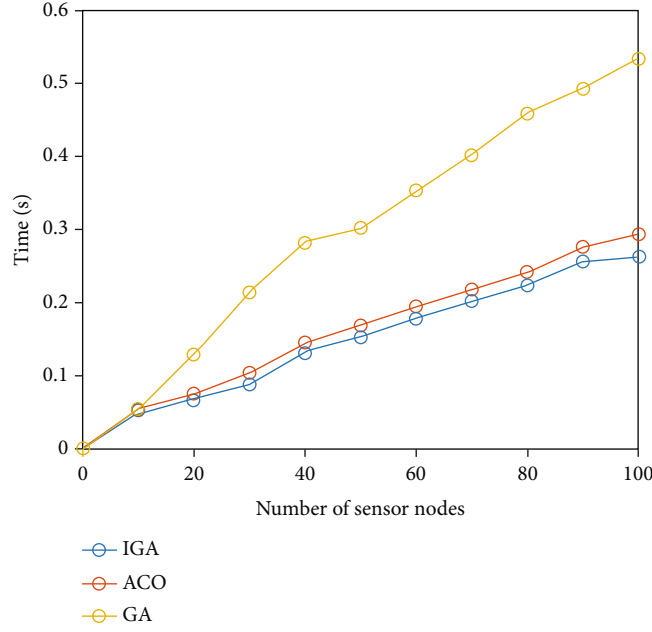
FIGURE 7: Calculation time to find the best path.

head nodes. Competitive radius issue: nodes form clusters of different sizes by calculating their own competition radius, so that the cluster radius will be closer to the base station, reducing energy consumption for receiving member node information and increasing the ability to send and receive information between internal and external.

Figure 4 is a schematic diagram of temporary cluster head competition, because Q_1 and Q_2 do not have a competitive relationship. While Q_3 is within the competition radius of Q_2, it is necessary to compare the current remaining energy of the two.

After the CH is ultimately determined, the remaining half is completed using the I-LEACH clustering stage's remaining energy gradient, which will not be reproduced here.

4.2. Structure between CHS. The encoding method affects the efficiency of the genetic algorithm to solve the optimal value. Usually, binary coding or real number coding is selected, but since the nodes distributed in the network are all set to integers, these integers form chromosomal individuals in the population, so this article chooses the variable-length path coding strategy. For example, the sequence of nodes from node 1 to node n: $1, 2, \cdots n$, and the path code: $1 \longrightarrow 2, \cdots, n$. This coding method makes the path constituted without loops, which obviously improves the search efficiency of the genetic algorithm [23].

Let $X = \{x_1, x_2, \cdots, x_i, \cdots, x_n\}$ represent a path from CH1 to CHn, and the fitness function is shown in the formula.

$$f(x_i) = \frac{1}{\alpha Cost(CH_i, CH_j) + (1 - \alpha)\sum_{i=1}^{n} \tau_{\mathrm{opt}}(i)}, \quad (15)$$

where α is the weight of the energy consumption of the communication link, $\mathrm{Cost}(CH_i, CH_j)$ refers to the path consumption of path X, and $\tau_{\mathrm{opt}}(i)$ is the busy state of node i.

$$\tau_{\mathrm{opt}}(i) = \beta \frac{d_{\mathrm{toBS}}(i)}{d_0} + 1 - \beta \frac{E_{\mathrm{total}}(i)}{E_0(i)}, \quad (16)$$

where $E_{\mathrm{total}}(i)$ is the total energy consumed by node i in the network communication process, $E_0(i)$ is the initial energy of node i, and $d_{\mathrm{toBS}}(i)$ is the distance between node i and BS.

$$\mathrm{Cost}(CH_i, CH_j) = \frac{d(CH_j, BS) \cdot d(CH_i, CH_j)}{E_{\mathrm{curj}}}, \quad (17)$$

where $d(CH_i, CH_j)$ is the distance from CH_i to the next CH_j, $d(CH_j, BS)$ is the distance from the next CH_j to BS, and E_{curj} is the current energy of the next CH_j.

Crossoperation: suppose $X_1 = \{1, 5, 6, 12\}, X_2 = \{4, 7, 11\}, X_3 = \{2, 3, 8, 10\}$, randomly select x_7 as the intersection in path X_1, then select that intersection makes the path after this node and x_7 form a feasible path X_2, that is, the intersection is x_8. After the two paths are exchanged and reorganized, two new paths are formed: $X_2' = \{4, 7, 8, 10\}$ and $X_3' = \{2, 3, 7, 11\}$. When performing the mutation operation, select the node x_7, x_8 with more energy in the path as the mutation point, then the path from the initial node will not change, and energy of the node is less than x_7, x_8, which is a new path: $X_2'' = \{2, 4, 7\}, X_3'' = \{11, 7\}, X_4'' = \{3, 8\}, X_5'' = \{9, 8\}$, and $X_6'' = \{10, 8\}$. The specific process is shown in Figure 5.

4.3. Combination of I-LEACH and EEUC-IGA. The proposed protocol—Energy Balance Multihop Clustering
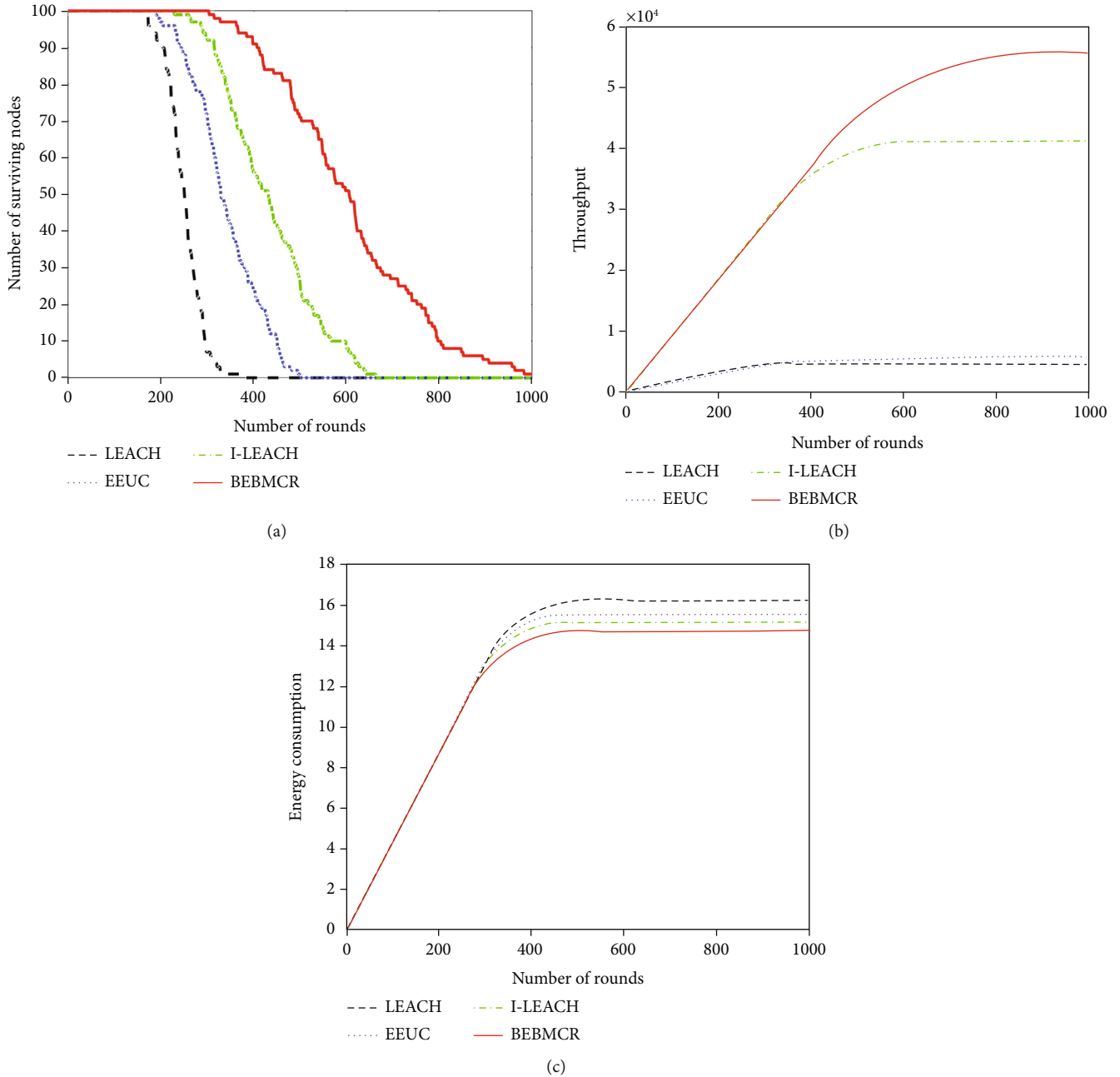
(a)

(b)

(c)

FIGURE 8: Comparison of LEACH, I-LEACH, EEUC, and BEBMCR; (a) number of surviving water quality nodes, (b) data transmission, and (c) energy consumption.

Routing Protocol for Large-Scale Water Quality Monitoring (BEBMCR)—is as follows.

We combine the advantages of the two algorithms. Aiming at the overlapping problem of the cluster head broadcast field, based on the LEACH protocol, the cluster head election probability is improved according to the remaining energy of the node, the density of neighbor nodes, and the location of the node; combined with the concept of the proposed competition radius, the energy consumption of nodes in the repeated coverage area of the network is reduced; secondly, in view of the dead node consumption of the EEUC protocol and the problem of some nodes not joining the cluster, we optimized the genetic algorithm and adopted path coding strategy to improve the efficiency of genetic algorithm to solve the optimal value and reduce the energy consumption of unnecessary network path transmission.

5. Experimental Simulation

5.1. Simulation Setup. In order to facilitate the creation of a GUI to verify each algorithm, this article chooses the MATLAB environment to implement and simulate the routing protocol and compare LEACH, I-LEACH, EEUC, and BEBMCR. Table 1 shows the environmental parameter
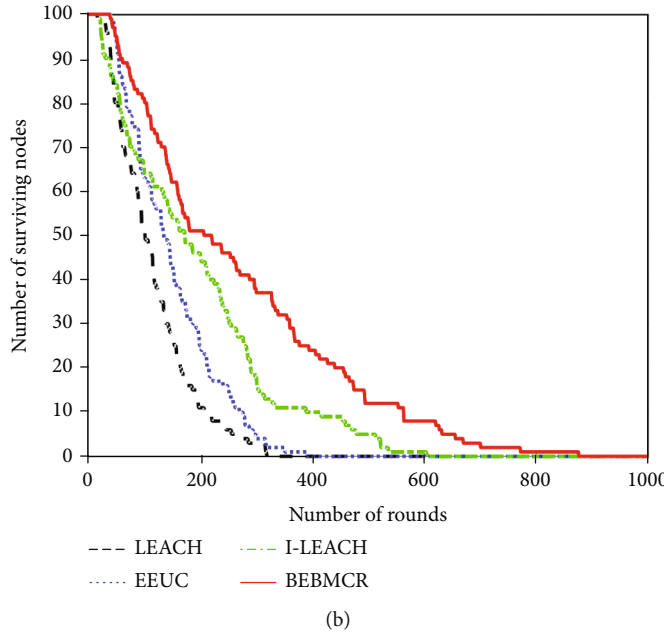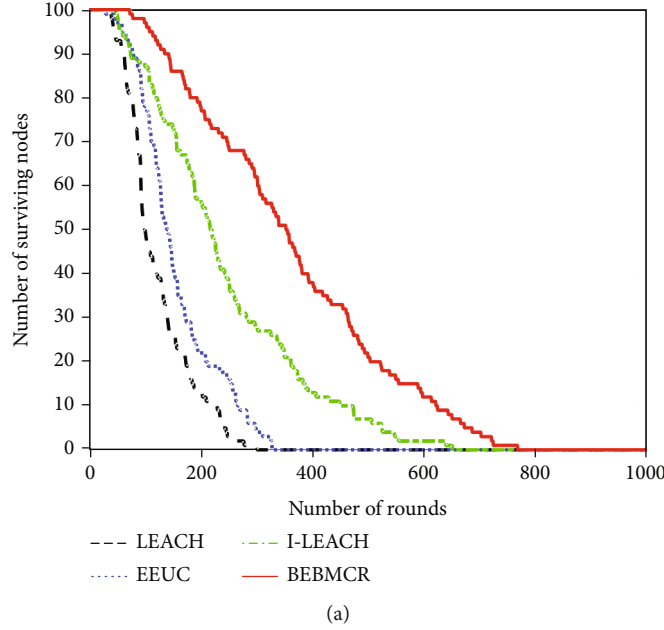
(a)

(b)

FIGURE 9: Number of surviving water quality nodes in different areas; (a) with 200 m × 200 m and 100 water quality nodes, (b) with 500 m × 500 m and 100 water quality nodes.

values used in the simulation. The comparison method is mainly used to simulate and analyze the number of surviving nodes in the entire network, network data transmission volume, and network energy consumption.

The simulation network model is shown in Figure 6. CHS uses multihop data transmission, and its graphical view represents the cluster of each round node and the routing topology of the network, in which the network is divided into formation. There is a cluster head represented by a red node, and the area below is called a cluster. Each cluster in the whole network has a BS, which collects data from all other nodes. The red line indicates the CH node sending data to BS, and the green line indicates the transmission path between CH nodes. The red and green lines represent the multihop data transmission between CHS.

Figure 6(a) shows the simulation iteration process of the EEUC routing protocol. The competition range of nodes decreases with the decrease of the distance from the base station. The result is that the closer the cluster is to the base station, the smaller the scale is, and the less energy is consumed. Within the cluster, the data processing process consumes less energy, which provides more energy for inter-cluster communication. But in EEUC routing in each round of data transmission, if the distance from BS is farthest or
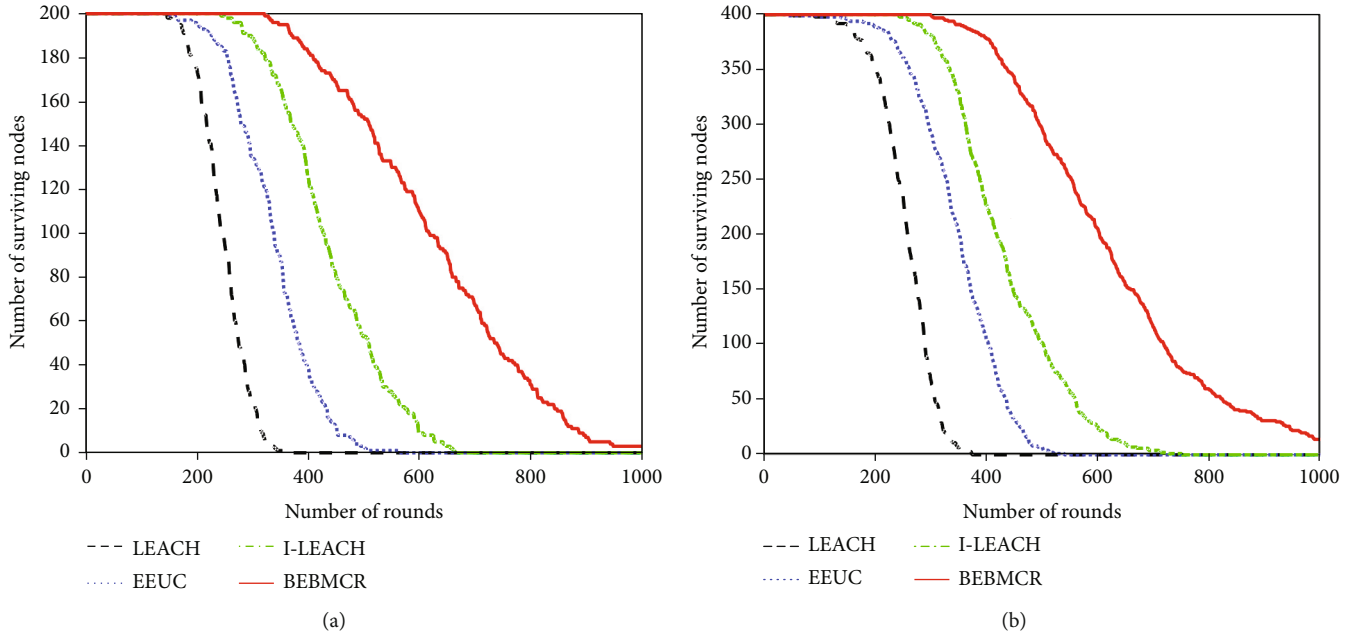
FIGURE 10: Number of surviving water quality nodes in different areas; (a) with 200 m × 200 m and 100 water quality nodes, (b) with 500 m × 500 m and 100 water quality nodes.

closest, all nodes need to recalculate their respective competition radius and inform other nodes in the network. Therefore, it is difficult to avoid that some nodes in the network do not join any clusters or some cluster heads do not communicate. The second and fourth parts of Figure 6(a) show this problem very well.

Figure 6(b) shows the simulation iteration process of the proposed BEBMCR routing protocol. The proposed algorithm combines I-LEACH and EEUC and adopts I-LEACH competition mechanism when clustering, thereby reducing the energy of the cluster head, which is close to the base station. The purpose of IGA is to optimize the multihop mode between EEUC clusters to communicate with the base station, thereby alleviating the "hot spot" problem in the algorithm.

This paper also simulates the classic ACO path search algorithm, the routing protocol (GA) based on the traditional genetic algorithm, and the improved genetic algorithm (improved GA) from the calculation time of the optimal path [23, 24]. As is shown in Figure 7, when the number of surviving water quality nodes is relatively small, these three algorithms will search for the best path in time. As the number of water quality nodes increases, the ACO algorithm needs to search one by one according to the path cost, while the genetic algorithm finds the best path according to its efficiency and parallelism. Therefore, compared with the ACO algorithm and the GA algorithm, the time for the improved GA algorithm to search for the optimal path is greatly reduced, which reduces the delay of the algorithm. Because the path coding strategy proposed in this paper is easier to perform crossover and mutation operations and correspondingly reduces the calculation time, it also shows that the improved GA algorithm improves the efficiency of the algorithm.

5.2. Simulation Results. Figure 8(a) shows that in LEACH, I-LEACH, and EEUC, the first dead node appeared in the first 300 rounds, but in the proposed BEBMCR, it appeared in about 350 rounds. When all water quality sensor nodes fail, the network operation rounds of leaching, I-LEACH, and EEUC are the first 400, 500, and 700, and our algorithm is about 1000. The reason is that our algorithm not only considers the remaining energy of the node, the location of the node, and the density of adjacent nodes but also increases the competition radius. Compared with the above three protocols, the network lifetime of BEBMCR has increased by 30% to 60%.

Figure 8(b) shows that the throughput of BEBMCR is much higher than the other protocols. This is the case of limited energy, according to the proposed algorithm, the network can send more data, so BEBMCR has higher energy efficiency than other energy efficiency. At the same time, since LEACH adopts a competition mechanism of randomly electing cluster heads, LEACH has the smallest throughput. For example, when the round is 400, the throughputs of LEACH, EEUC, and I-LEACH are, respectively, 0.42×10^4, 0.47×10^4, 3.87×10^4, and 4.21×10^4.

Figure 8(c) shows the trend system of the energy consumption of the routing protocol with the number of rounds. It is not difficult to find that the energy consumption of the BEBMCR protocol is always less than the energy consumption of other protocols, because the BEBMCR protocol uses IGA to optimize EEUC's intercluster multihop communication with the base station, alleviating the "hot spot" problem in the algorithm and designing a reasonable communication mode for water quality sensor nodes around the receiver k.

Figure 9 shows that, while keeping the number of water quality sensors in the monitoring area and other conditions

in the network unchanged, the network area is changed to 200 m × 200 m and 500 m × 500 m, respectively. In LEACH, I-LEACH, and EEUC, in the first 100 rounds, the first node died; with the increase of the network area, the network life of BEBMCR has been prolonged compared to the other three algorithms. Overall, the network life of BEBMCR has remained above 50%.

Figure 10 shows that while keeping the field size in the monitoring area and other conditions in the network unchanged, the number of water quality sensor nodes is changed to 200 and 400, respectively. In LEACH and EEUC, the first dead nodes appear in the first 200 rounds. I-LEACH is the first node death in the first 300 rounds. That is, as the number of water quality sensor nodes increases, the first dead node in the BEBMCR algorithm is maintained at about 300 rounds. The network lifetime of BEBMCR is still maintained at an increase of about 50%. When all the water quality sensor nodes are dead, the network operation rounds of LEACH, I-LEACH, and EEUC have basically not changed, and our algorithm is after 1000. Therefore, BEBMCR is reliable when there are water quality sensor nodes on the large network.

6. Conclusion

In order to extend the running time of the monitoring network and realize long-term online monitoring of river water quality parameters, an energy balance clustering routing algorithm (BEBMCR) for large-scale water quality monitoring is proposed. The main innovations of this article are as follows:

(1) The cluster head election probability is increased based on the remaining energy of the node, its location, and the density of neighboring water quality nodes, according to the research of LEACH, the classic cluster routing method

(2) Considering the overlap rate of water quality sensors, combined with the concept of competition radius, the overlap rate is introduced into the competition mechanism to stabilize the number of cluster heads in the network

(3) The combination of single-hop communication and multihop communication makes the algorithm suitable for large-scale water quality sensor network structure

(4) In the selection of the next hop, the optimal path is searched based on the efficiency and parallelism of the genetic algorithm, and the variable path coding strategy is used to integrate the communication distance, the relative remaining energy, and the number of water quality nodes in the cluster, to avoid the appearance of "hot spot"

(5) Combining the characteristics of I-LEACH and EEUC-IGA, four groups of experiments were compared

Experimental results show that the routing algorithm is more suitable for large-scale network deployment and pro-vides a theoretical basis and reference for river water quality monitoring research. In this paper, the distribution of sensor nodes is fixed and static, but in some cases, the nodes need to be mobile. How to ensure the optimal energy efficiency of the network brought by the routing protocol when the nodes are constantly changing is an important issue that needs to be considered in the next research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Authors' Contributions

Yulin Gong (corresponding author) is responsible for the conceptualization, formal analysis, funding acquisition, methodology, and project administration; Jiannan Cao for the data curation, software, visualization, and original draft; Chengfei Han for the investigation, resources, and validation; and Yunqing Liu for the supervision, review, and editing.

Acknowledgments

References

[1] H. Zhang, "Application of wireless sensor network based on ZigBee technology in marine ecological environment monitoring," *Journal of Coastal Research*, vol. 110, no. sp1, pp. 54–56, 2020.

[2] B. A. I. Qiuchan, J. Wu, and J. I. N. Chunxia, "The water quality online monitoring system based on wireless sensor network," in *IEEE International Symposium on Computational Intelligence and Design*, pp. 234–237, Hangzhou, China, Dec. 2020.

[3] S. Kavi Priya, G. Shenbagalakshmi, and T. Revathi, "Applied fuzzy heuristics for automation of hygienic drinking water supply system using wireless sensor networks," *The Journal of Supercomputing*, vol. 76, no. 6, pp. 4349–4375, 2020.

[4] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd annual Hawaii international conference on system sciences*, vol. 2, p. 10, Maui, HI, USA, Jan. 2000.

[5] Y. Dong, J. Wang, B. Shim, and D. I. Kim, "DEARER: a distance-and-energy-aware routing with energy reservation for energy harvesting wireless sensor networks," *Proceedings of Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3798–3813, 2016.

[6] A. Thakkar, "DEAL: distance and energy based advanced LEACH protocol," in *International conference on information*

and communication technology for intelligent systems, vol. 2, pp. 370–376, Springer, Cham, 2017.

[7] D. Singh and S. K. Nayak, "Enhanced modified LEACH (EMODLEACH) protocol for WSN," in *2015 International symposium on advanced computing and communication*, p. 328, Silchar, India, Sept. 2015.

[8] S. D. Sasikala, N. Sangameswaran, and P. Aravindh, "Improving the energy efficiency of LEACH protocol using VCH in wireless sensor network," *International Journal of Development Research*, vol. 3, no. 2, pp. 918–924, 2015.

[9] M. Elmonser, H. B. Chikha, and R. Attia, "Mobile routing algorithm with dynamic clustering for energy large-scale wireless sensor networks," *Wireless Sensor Systems*, vol. 10, no. 5, pp. 208–213, 2020.

[10] A. N. Shahbaz, H. Barati, and A. Barati, "Multipath routing through the firefly algorithm and fuzzy logic in wireless sensor networks," *Peer-to-Peer Networking and Applications*, vol. 14, no. 2, pp. 541–558, 2021.

[11] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: a stable election protocol for clustered heterogeneous wireless sensor networks," *Second International Workshop on Sensor and Actor Network Protocols and Applications*, vol. 3, 2004.

[12] D. Kumar, T. C. Aseri, R. B. Patel, and D. I. Kim, "EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks," *Computer Communications*, vol. 32, no. 4, pp. 662–667, 2009.

[13] C. Li, M. Ye, G. Chen, and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," in *Proceedings of International Conference on Mobile Adhoc and Sensor Systems Conference*, p. 604, Washington, DC, USA, Nov. 2005.

[14] S. Singh, A. Malik, and R. Kumar, "Energy efficient heterogeneous DEEC protocol for enhancing lifetime in WSNs," *Engineering Science and Technology, an International Journal*, vol. 20, no. 1, pp. 345–353, 2017.

[15] W. Chen, B. Zhang, X. Yang, W. Fang, W. Zhang, and X. Jiang, "C-EEUC: a cluster routing protocol for coal mine wireless sensor network based on fog computing and 5G," *Mobile Networks and Applications*, vol. 278, pp. 1–14, 2019.

[16] L. Kong, J. S. Pan, V. Snášel, P. W. Tsai, and T. W. Sung, "An energy-aware routing protocol for wireless sensor network based on genetic algorithm," *Telecommunication Systems*, vol. 67, no. 3, pp. 451–463, 2018.

[17] S. Mottaghi and M. R. Zahabi, "Optimizing LEACH clustering algorithm with mobile sink and rendezvous nodes," *Proceedings of AEU-International Journal of Electronics and Communications*, vol. 69, no. 2, pp. 507–514, 2015.

[18] V. A. A. Shiny and V. Nagarajan, "Energy efficient routing protocol for mobile wireless sensor network," *International Journal of Computer Applications*, vol. 43, no. 21, pp. 1–5, 2012.

[19] Y. Xu, W. Jiao, and M. Tian, "Energy-efficient connected-coverage scheme in wireless sensor networks," *Sensors*, vol. 20, no. 21, p. 6127, 2020.

[20] I. Diané, R. Kacimi, Z. Mammeri, and I. Niang, "Energy optimization based on the redundancy in WSNs," in *6th Joint IFIP Wireless and Mobile Networking Conference*, Dubai, United Arab Emirates, April 2013.

[21] A. K. Takele, T. J. Ali, and K. A. Yetayih, "Improvement of LEACH protocol for wireless sensor networks," in *International Conference on Information and Communication Tech-nology for Development for Africa*, pp. 250–259, Springer, Cham, 2019.

[22] P. Sivakumar and M. Radhika, "Performance analysis of leach-ga over leach and leach-c in wsn," *Procedia Computer Science*, vol. 125, pp. 248–256, 2018.

[23] J. L. Liu and C. V. Ravishankar, "LEACH-GA: genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks," *Proceedings of International Journal of Machine Learning and Computing*, vol. 1, no. 1, pp. 79–85, 2011.

[24] S. Anu, "RLEACH-ACO: revised LEACH-ACO routing protocol for enhancing energy efficiency of WSN," *Journal of Emerging Technologies and Innovative Research*, vol. 5, no. 8, pp. 356–364, 2018.