

Research Article

A Distributed Unequal Clustering Routing Protocol Based on the Improved Sine Cosine Algorithm for WSN

Fang Zhu  and Wenhao Wang 

School of Computer and Communication Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China

Correspondence should be addressed to Wenhao Wang; 2072030@stu.neu.edu.cn

Received 22 December 2021; Accepted 26 January 2022; Published 23 February 2022

Academic Editor: Jesús Lozano

Copyright © 2022 Fang Zhu and Wenhao Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to balance the overall energy consumption and improve the energy efficiency of wireless sensor network (WSN), a distributed energy-balanced unequal clustering routing protocol based on the improved sine cosine algorithm (DUCISCA) is proposed. Firstly, DUCISCA adopts a time-based cluster head competition algorithm. In this algorithm, the broadcast time depends on the residual energy of the candidate cluster head, the distance to the base station, and the number of neighbour nodes. Secondly, a competition radius considering the distance from node to base station and the residual energy of node is proposed. It can balance energy consumption of nodes in different locations to avoid the “hot spot” problem. At the same time, it adopts a time-based broadcast mechanism. The waiting time depends on the residual energy of CCHs, the distance to the BS, and the number of neighbour nodes, which can effectively reduce the overhead of nodes. Thirdly, the energy of cluster head, the number of neighbour nodes, and the distance from the ordinary node to the cluster heads need to be taken into account to get a better clustering result. Finally, in order to speed up convergence and improve the ability to jump out of local optimum, the improved sine-cosine algorithm (ISCA) based on Latin hypercube sampling and adaptive mutation is proposed. The improvement strategies adopted by ISCA are expressed as follows: Firstly, the diversity of the population is enhanced through LHS population initialization. Secondly, the adaptive weight strategy is introduced to accelerate the convergence speed of the algorithm. Finally, the population is disturbed by Gaussian mutation or Levy flight to jump out of the local optimum. The standard deviation of cluster heads’ residual energy in intercluster communication is taken as the objective function to search the energy-balanced intercluster data forwarding path based on ISCA. Compared with EEUC, DEBUC, I-EEUC, and M-DEBUC, the simulation results prove that DUCISCA can effectively balance the overall network energy consumption and prolong the network lifetime.

1. Introduction

With the rapid development of MEMS and wireless communication technology, wireless sensor network (WSN) has been widely used in many fields, such as environmental monitoring, emergency rescue, medical care, intelligent agriculture, various commercial applications, and military fields [1]. WSN is a self-organizing wireless communication network system which consists of a large number of homogeneous sensor nodes with limited energy deployed in relatively poor environments, base stations (BSs) (or sink nodes), and terminal systems [2]. The sensor nodes communicate through radio within the network, and they are ran-

domly distributed in the monitoring region to collect the information and transmit the information to the BS (or sink node) by single hop or multihop. Users can manage the operation of the WSN through the terminal. Due to the intensive network deployment and unattended character of WSN, the nodes’ energy cannot be replenished in time [3]. Therefore, how to balance the energy consumption of the network and prolong the network lifetime has become the key problem of WSN.

Clustering technology for WSN has many advantages, such as high energy efficiency, collision avoidance, and low delay. However, clustering protocols have also some problems. In a large scale of network, cluster-based protocols

mostly adopt multihop routing in order to save energy. In this situation, the CHs near BS will undertake more data forwarding tasks, which is easy to cause the CHs to run out of energy earlier than other nodes. It leads to network coverage vulnerabilities and data loss. That is “hot spot” problem.

In order to improve these problems mentioned above, a distributed unequal clustering routing protocol based on the improved sine cosine algorithm (DUCISCA) is proposed in this paper, which self-organizes networks through unequal clustering and intercluster multihop routing. DUCISCA is a distributed competition algorithm. CHs are selected by competition. The competition range of nodes varies unevenly with the distance from BS and the residual energy of nodes, which can effectively balance the energy consumption of nodes in different locations. An improved sine-cosine algorithm (ISCA) based on Latin hypercube sampling and adaptive mutation is proposed and applied to intercluster multihop routing to obtain the near optimal multihop path, so as to balance the energy consumption of network.

According to above discussion, our contributions are concluded as follows:

- (1) We have adopted a time-based broadcast technology to select CHs, and the waiting time of CHs is related to the residual energy of CHs, the distance to the BS, and the number of neighbour nodes, which can effectively reduce the overhead of nodes
- (2) We have presented a cluster formation scheme which considers energy, distance, and the number of neighbour nodes to obtain better clustering effect
- (3) We have proposed an improved sine-cosine algorithm (ISCA) based on Latin hypercube sampling and adaptive mutation, which achieves good results on 6 standard test functions. And we have applied ISCA to intercluster multihop routing to obtain the near optimal multihop path, so as to balance the energy consumption of network

The rest of this paper is organized as follows. Some well-known routing protocols proposed for WSN are summarized in “Related Works”. “System Model” describes the network model and energy model in this work. “DUCISCA Protocol” describes proposed unequal clustering routing protocol DUCISCA in detail. “Simulation Experiments and Analysis” presents the simulations and performance evaluation of DUCISCA. At last, “Conclusion” summarizes this paper.

2. Related Works

Earlier clustering routing protocols usually adopted the method of uniform clustering [4] to divide the whole network into clusters with equal size. The number of members and the cluster radius is approximately equal in each cluster. Heinzelman et al. [5] proposed a low-energy adaptive clustering hierarchy (LEACH), a clustering-based protocol that utilizes randomized rotation of CHs to evenly distribute the energy load among the sensors in the network. Younis

and Fahmy [6] presented a hybrid energy-efficient distributed clustering approach (HEED) that periodically selects CHs according to a hybrid of the node residual energy and a secondary parameter (node proximity to its neighbours or node degree). Thakkar and Kotecha [7] proposed a coverage routing protocol based on LEACH (CVLEACH) to make uniform distribution of CHs by creating nonoverlapped cluster regions using overhearing properties of the sensor nodes, which prolongs lifetime of WSN. In uniform clustering routing protocols, CH selection is an important challenge that directly affects the network performance [8]. A protocol for analyzing energy-delay trade-off (EDIT) [9] is proposed, which can better balance nodes’ energy consumption and reduce transmission delay. Thakkar and Kotecha [10] presented a new CH election protocol inspired from Bollinger Bands. Simulation result shows significant improvement in the network lifetime in comparison with other decentralized and ant-based algorithms. Daniel et al. [11] proposed an effective method named tunicate swarm butterfly optimization algorithm (TSBOA) is developed for selecting CH to accomplish effective data transmission between the sensor nodes. Although these protocols can balance the energy consumption, the overall energy consumption of the network is still very high due to the single-hop routing method. Therefore, in order to reduce the energy consumption of CHs, researchers propose many multihop routing protocols, such as EM-LEACH [12], IEE-LEACH [13], and DMEERP [14]. Multihop routing can reduce the energy consumption of CHs far away from the BS, but more data forwarding tasks fall on CHs closer to the BS, resulting in excessive energy consumption and premature failure. It will cause “hot spot” problem and premature death of network lifetime.

From the above analysis, it can be seen that the energy consumption among clusters is unbalanced whether the uniform clustering protocols use single-hop or multihop routing. So many researchers propose unequal clustering routing protocols. In unequal clustering, the cluster size varies proportionally to the distance to BS. Energy conservation and eliminating hot spot problem are the most common objectives of unequal clustering [15]. UCS [16] proposed the idea of unequal clustering for the first time to balance energy consumption of CHs. EEUC [17] partitions the nodes into clusters of unequal size, and clusters closer to the BS have smaller sizes than those farther away from the BS. For intercluster communication, EEUC adopts an energy-aware multihop routing system considering residual energy of CHs to balance the energy consumption of network. I-EEUC [18] improves the CH selection algorithm based on EEUC. CH elections adopt a timed broadcast mechanism to select CHs reasonably, which prolongs the network lifetime compared with EEUC. Jiang et al. [19] proposed a distributed energy-balanced unequal clustering routing protocol (DEBUC). DEBUC adopts a time-based competitive algorithm to select CHs. For intercluster routing, it takes into account the residual energy of nodes, the cost of intracluster, and intercluster communication. The simulation results show that it has better performance than LEACH, EEUC, and I-EEUC. Based on DEBUC, M-

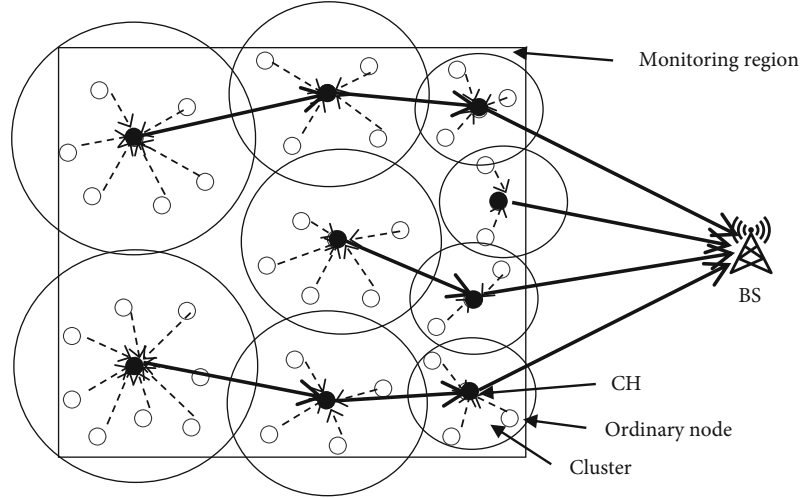


FIGURE 1: Schematic diagram of DUCISCA.

DEBUC [20] has improved its candidate CH election method, competition radius, and wait time. The simulation results show that it has better performance than DEBUC.

3. System Model

3.1. Network Model. Let us consider a sensor network consisting of N sensor nodes uniformly deployed over a two-dimensional (2D) square area. We denote the i -th sensor by S_i and the corresponding sensor node set $S = \{S_1, S_2, \dots, S_N\}$, where $|S| = N$. We make some assumptions about the sensor network model [21]:

- (1) There is a base station with unlimited energy located far away from the square sensing field. Sensors and the base station are all stationary after deployment
- (2) All nodes are homogeneous and have the same capabilities. Each node is assigned a unique identifier (ID)
- (3) All nodes have similar capabilities and equal status and can act as cluster head nodes or ordinary nodes
- (4) Nodes can use power control to vary the amount of transmission power which depends on the distance to the receiver
- (5) Data fusion technology is adopted to reduce the amount of data transmission

3.2. Energy Model. The radio model in this work is the same as described by [14]. With reasonable signal-to-noise ratio guaranteed, the energy consumption for node sending data is shown in (1).

$$E_{Tx}(n, d) = \begin{cases} nE_{elec} + n\epsilon_{fs}d^2, & d \leq d_0, \\ nE_{elec} + n\epsilon_{mp}d^4, & d > d_0, \end{cases} \quad (1)$$

where n is the number of bits transmitted, d is the transmission distance, E_{elec} is the energy consumption for sending or

receiving 1 bit data, ϵ_{fs} is the coefficient of energy consumption for amplifying radio at free-space mode, ϵ_{mp} is the coefficient of energy consumption for amplifying radio at multifading mode, and $d_0 = \sqrt{\epsilon_{fs}/\epsilon_{mp}}$ is the threshold value of distance.

The energy consumption of data received by the node is

$$E_{Rx}(n) = nE_{elec}. \quad (2)$$

Due to the large difference of data between clusters, data fusion between clusters is not considered in this simulation. The assumptions of the data fusion model in the cluster are that the CH receives n bit data sent by each member node and compresses it into n bit data regardless of the number of nodes in the cluster. The energy consumption for CH fusion of n bit data is

$$E_{Fx}(n, d) = nE_{DA}, \quad (3)$$

where E_{DA} (nJ/bit) is the energy consumption for fusing 1 bit data.

In our simulation, communication energy consumption model parameters set to $E_{elec} = 50$ nJ/bit, $\epsilon_{fs} = 10$ pJ/bit/m², $\epsilon_{mp} = 0.0013$ pJ/bit/m⁴, and $d_0 = 87$ m.

4. DUCISCA Protocol

DUCISCA protocol adopts “round” cycle mechanism, and each round includes three stages: cluster formation, multihop routing between clusters, and data transmission. Firstly, clusters with different sizes are formed. Secondly, according to the network clustering results, multihop routing between clusters is established. Finally, the network enters stable stage to complete data transmission. Figure 1 is the basic schematic diagram of DUCISCA.

4.1. CH Selection Algorithm. In the CH selection stage, the broadcast radius of control message is R_{comp}^0 to ensure that

candidate cluster heads can receive message from their neighbour nodes.

4.1.1. Candidate Cluster Head (CCH) Election. In practical applications of WSN, a large number of nodes are usually deployed with a high node density [22]. Therefore, it is not necessary for all nodes to be CCHs. Given an appropriate threshold T to control the proportion of CHs, in this paper, T is set to 0.4 according to [17]. Each node S_i calculates the value of μ . If $\mu < T$, node S_i becomes a CCH; otherwise, node S_i becomes an ordinary node and goes into a dormant state until the final cluster head election is completed. The value of μ is calculated by (4).

$$\mu = \mu_0 \frac{E_{\text{avg}}}{\text{RE}_i}, \quad (4)$$

where μ_0 is a random number evenly distributed between (0,1) and E_{avg} is the average residual energy of all alive nodes in the network. Its calculation process is as follows: The data packets transmitted by the nodes include their own residual energy RE_i . After the BS receives them, the average residual energy E_{avg} of all alive nodes is calculated; RE_i denotes the residual energy of node S_i . Obviously, the larger of RE_i , the smaller of μ , so the greater probability that S_i will become a CCH.

4.1.2. Calculation of Unequal Competition Radius. In DEBUC, the calculation of the competitive radius only considers the distance from the CCH to BS and ignores the energy. This may result in low-energy nodes to be the CHs, and after forming a cluster, the members in the cluster will cause an increase in the energy consumption of CHs. Therefore, a new competitive radius considering both distance and energy is proposed, and the improved competitive radius for CH v_i is calculated in (5).

$$v_i \cdot R_{\text{comp}} = \left[1 - w_1 \frac{d_{\text{max}} - d(v_i, \text{BS})}{d_{\text{max}} - d_{\text{min}}} - w_2 \left(1 - \frac{\text{RE}_{v_i}}{E_0} \right) \right] \times R_{\text{comp}}^0, \quad (5)$$

where d_{max} and d_{min} are the maximum and minimum distances from the alive nodes to the BS, respectively, $d(v_i, \text{BS})$ denotes the distance from v_i to the BS, RE_{v_i} is the residual energy of v_i , w_1 and w_2 are both constants between (0,1), and $w_1 + w_2 = 1$.

4.1.3. Final Cluster Head (FCH) Election

Definition 1. In DUCISCA, given maximum communication radius R_{comp}^0 , for S_i , its neighbour node set Nbr_i is computed as (6).

$$\text{Nbr}_i = \left\{ j \mid d(i, j) \leq R_{\text{comp}}^0, j \neq i, j \in N \right\}. \quad (6)$$

Definition 2. The degree of S_i is calculated as (7).

$$\text{degree}_i = |\text{Nbr}_i|. \quad (7)$$

CCH v_i broadcasts competition message with R_{comp}^0 radius. The message content includes ID and the competition radius $v_i \cdot R_{\text{comp}}$ and the remaining energy RE_i . CCH v_i establishes its neighbour CCH set NT_i based on the received election message.

$$\text{NT}_i = \left\{ v_j \mid d(v_i, v_j) < \max(v_i \cdot R_{\text{comp}}, v_j \cdot R_{\text{comp}}) \right\}. \quad (8)$$

In the CHs' competition stage, DEBUC adopts a time-based broadcast mechanism to calculate the waiting time. It only considers the energy factor of the nodes, which easily causes the "hot spot" problem. Therefore, this paper presents an improved waiting time that takes into account energy, distance, and node degree. The improved waiting time is calculated for CCH v_i which is shown in (9).

$$t_i = k \times T_{\text{CH}} \times \left[c_1 \left(1 - \frac{\text{RE}_i}{E_0} \right) + c_2 \frac{d(v_i, \text{BS})}{d_{\text{max}}} + c_3 \frac{\text{degree}_i}{\max(\text{degree})} \right], \quad (9)$$

where k is a random number evenly distributed between (0.9,1) to reduce the probability of broadcast message conflicts; T_{CH} is the maximum duration of predefined CHs' competition; E_0 is the initial energy of v_i ; c_1 , c_2 , and c_3 are weight factors of energy, distance, and degree, respectively, and $c_1 + c_2 + c_3 = 1$. Using a time-based broadcast mechanism can reduce the number of control messages broadcast and received when CCHs compete for FCHs.

BS broadcasts the CHs' selection message CH_SEL_MSG to synchronize the time of each node. After receiving the CH_SEL_MSG message, the node starts its own clock to start timing and listen to the messages sent by other nodes.

If $\text{RE}_i < \bar{E}_{\text{NT}_i}$ (\bar{E}_{NT_i} is the average residual energy of the neighbour nodes of v_i), v_i gives up the competition of FCH; otherwise, its waiting time is calculated according to (9). It is to be noted from (9) that the CCHs with larger residual energy, closer to the BS, and a small number of neighbour nodes have higher probability for FCH selection.

The pseudo code of CH selection algorithm is shown as Algorithm 1.

4.2. Cluster Formation. After the CH selection is completed, the ordinary nodes quit the sleep statue and the CHs broadcast CH_ADV_MSG message in the network area. In DEBUC, ordinary nodes join the closest cluster according to the strength of the received signal, which is easy to cause uneven energy consumption. In this case, how to select an optimal cluster to join is a key issue of unequal clustering [23]. To solve this problem, a fitness function is set to each CH, which comprehensively considers the energy, degree of CH, and the distance between ordinary

Input: The locations of the BS and each node in the network, the residual energy and the maximum cluster radius of each node.

Output: FCHs.

1. Calculate the value of μ according to Eq. (4)
2. For every node in the network
3. if $\mu < T$ then
4. become a tentative CH
5. end if
6. if become a tentative CH, then
7. broadcast COMPETE_MSG(ID, R_{comp} , RE_i)
8. else
9. sleep
10. end if
11. For every tentative node v_i
12. On receiving a COMPETE_MSG from tentative CH v_j
13. if $d(v_i, v_j) < \max(v_i \cdot R_{comp}, v_j \cdot R_{comp})$.
14. add v_j to v_i neighbour set NT_i
15. end if
16. For every tentative node v_i
17. if $RE_i \geq \bar{E}_{NT_i}$
18. calculate t_i according to Eq. (9)
19. else
20. give up the competition and become an ordinary node
21. end if
22. BS broadcasts CH_SEL_MSG to synchronize the time of each node
23. For every tentative node v_i
24. while (the timer T_{CH} is not expired)
25. if $CurrentTime < t_i$
26. if (heard SUCCESS_MSG from a neighbour NT_i)
27. give up the competition and stop the timer t_i
28. end if
29. else if $CurrentTime == t_i$
30. v_i become FCH
31. SUCCESS_MSG(ID)
32. end if
33. end while

ALGORITHM 1: CH selection algorithm.

nodes and CH. The fitness function is calculated as shown in (10).

$$\text{fitness}(i) = d_1 \frac{E_{\text{res}_i}}{E_0} + d_2 \left(1 - \frac{\text{degree}_i}{N}\right) + d_3 \left(1 - \frac{d_{\text{to}_i}}{d_{\text{to_BS}}}\right), \quad (10)$$

where E_{res_i} is the residual energy of CH_i , E_0 is the initial energy, degree_i is the degree of CH_i , N is the total number of all sensor nodes, d_{to_i} is the distance from ordinary node to CH_i , $d_{\text{to_BS}}$ is the distance from ordinary node to BS, d_1 , d_2 , and d_3 are the weight factors, and $d_1 + d_2 + d_3 = 1$.

The process of cluster formation is shown as Algorithm 2.

Figure 2 is an example of unequal clustering of DUCISCA, in which the BS is located at (100,250). As can be seen from this figure, the size of each cluster is different, which achieves the effect of unequal clustering.

4.3. Intercluster Multihop Routing. In order to balance the energy consumption between clusters, ISCA is used to implement multihop routing between clusters. Each CH adopts ISCA to select its relay node in the list of neighbour CHs. The CH transmits the data collected from the members of the cluster to the relay node. The intercluster multihop routing is established, and the CH sends the data to the BS along the optimal path.

4.3.1. Sine-Cosine Algorithm (SCA). The principle of SCA [24] is simple and easy to implement. It only adopts the properties of sine and cosine functions to achieve global exploration and local development of search space and continuously optimizes the solution set of objective functions through iterative evolution. Assume that there are N search agents in the D -dimensional search space, where the location of the i -th search agent can be expressed as $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, $i \in \{1, 2, \dots, N\}$. The process of optimizing the objective function with SCA is as follows.


```

Input: FCHs and the information of each node and BS.
Output: Cluster structure.
1. All cluster heads broadcast CH_ADV_MSG in the network area
2. For every ordinary node  $O_i$ 
3.    $flag\_i = 0$ 
4.   For every final cluster head  $CH_i$ 
5.     calculate distance  $d_{toi}$  between  $O_i$  and  $CH_i$ 
6.     if  $d_{toi} < R_{comp}^0$ 
7.       calculate  $fitness(i)$  according to Eq. (10)
8.     else.
9.        $flag\_i = 1$ 
10.    end if
11.  end for
12.  if  $flag\_i == 0$ 
13.     $O_i$  join in the cluster head with the maximum fitness
14.     $O_i$  send JOIN_CLUSTER_MSG to the CH
15.  else
16.     $O_i$  join in the cluster head with the minimum distance
17.     $O_i$  send JOIN_CLUSTER_MSG to the CH
18.  end if
19. end for
20. unequal clustering complete
    
```

ALGORITHM 2: Cluster formation algorithm.

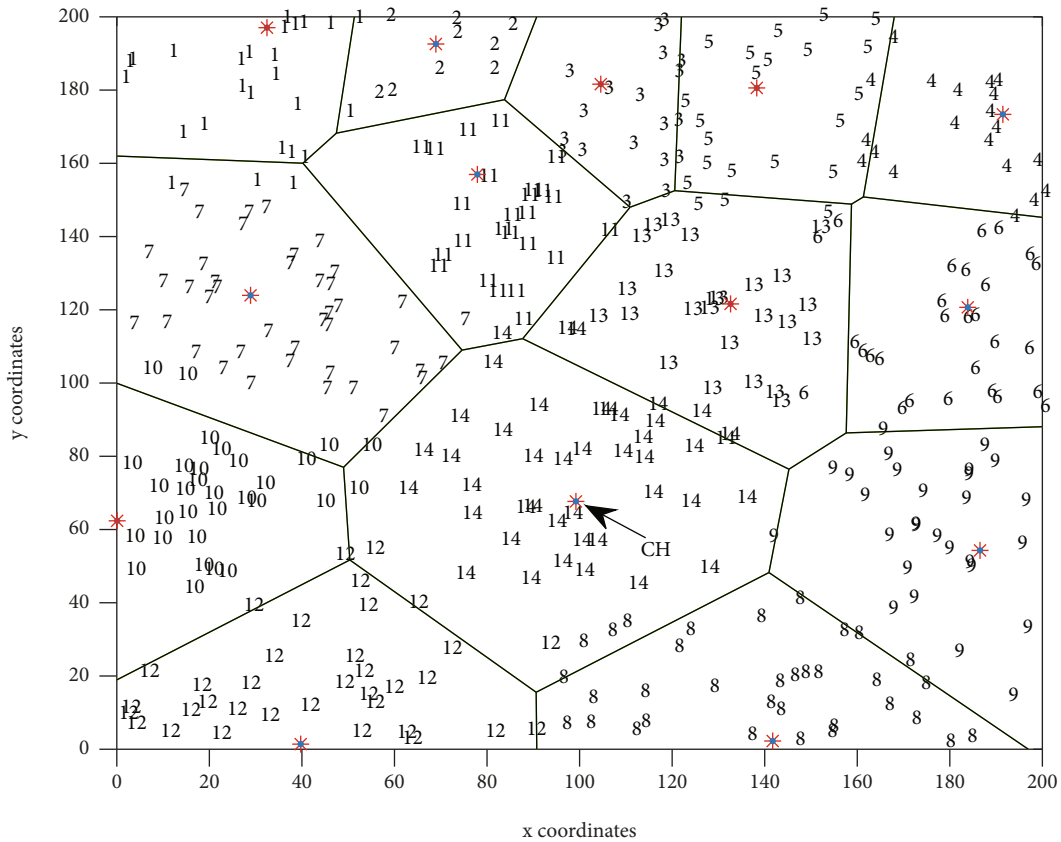


FIGURE 2: An example of unequal clustering for DUCISCA.

TABLE 1: Comparison of optimization performance of 5 algorithms under fixed iteration times.

Algorithms	Functions	Optimum	Worst	Mean	Std	Functions	Optimum	Worst	Mean	Std
BOA		1.34E-11	1.75E-11	1.58E-11	9.45E-13		4.57E-12	1.36E-11	7.81E-12	2.30E-12
WOA		4.89E-80	3.21E-64	2.12E-65	7.75E-65		0	1.46E-01	4.87E-02	2.67E-02
SSA	f_1	0	2.48E-68	8.29E-07	4.53E-69	f_4	0	0	0	0
SCA		1.48E+07	4.80E+10	1.01E+10	1.08E+10		5.46E-01	2.06E+00	1.15E+00	3.19E-01
ISCA		0	0	0	0		0	0	0	0
BOA		8.07E+76	2.95E+90	1.05E+89	5.37E+89		0	3.81E+02	3.61E+01	1.10E+02
WOA		6.20E-129	3.25E-78	1.09E-79	5.94E-79		0	1.14E-13	3.79E-15	2.08E-14
SSA	f_2	6.92E-190	1.69E-45	6.29E-47	3.10E-46	f_5	0	0	0	0
SCA		2.29E+21	2.64E+47	1.19E+46	5.04E+46		1.35E+02	5.38E+03	1.82E+03	1.43E+03
ISCA		0	0	0	0		0	0	0	0
BOA		1.09E-11	1.46E-11	1.27E-11	9.29E-13		0	1.80E+01	2.68E+00	6.09E+00
WOA		5.81E+04	1.20E+05	9.29E+04	1.69E+04		0	1.43E+01	4.78E-01	2.62E+00
SSA	f_3	0	6.77E-56	2.26E-57	1.24E-56	f_6	0	0	0	0
SCA		3.40E+03	2.73E+04	1.11E+04	4.76E+03		6.42E+00	1.89E+02	1.35E+01	2.84E+00
ISCA		0	0	0	0		0	0	0	0

Input: CHs and List of neighbouring CHs.

Output: Routing structure.

1. Random initialization population in [0,1] interval
2. The multi-hop paths represented by each individual are derived from the individual location and the list of neighbouring CHs, and calculate each individual's fitness value. Calculate initial global optimal location g_{best}
3. Update the individual location, fitness values and update g_{best} for t -th iteration according to the strategies of Algorithm 2
4. Skip to step 2 and continue until the maximum number of iterations is reached
5. Routing structure corresponding to output g_{best}

ALGORITHM 3: Intercluster multihop routing based on ISCA.

Firstly, the locations of N search agents are randomly initialized in the search space. Secondly, the individual fitness values are calculated based on the objective function. Finally, the current optimal individual locations are selected and saved. In each iteration of the algorithm, the individual updates the position according to (11).

$$x_{id}^{t+1} = \begin{cases} x_{id}^t + r_1 \times \sin r_2 \times |r_3 P_d^t - x_{id}^t|, & r_4 < 0.5, \\ x_{id}^t + r_1 \times \cos r_2 \times |r_3 P_d^t - x_{id}^t|, & r_4 \geq 0.5, \end{cases} \quad (11)$$

where t is the current iteration, x_{id}^t is the position of the i -th solution in d -th dimension at t -th iteration, and P_d^t is the position of the global optimal solution in the d -th dimension at t -th iteration. There four main parameters in (11), where $r_1 = 2(1 - t/T)$ (T is the maximum number of iterations) is the sine-cosine amplitude adjustment factor, and r_1 determines the direction of the next iteration of the i -th individual; $r_2 \in (0, 2\pi)$, $r_3 \in (0, 2)$, and $r_4 \in (0, 1)$ are random numbers, where r_2 determines the distance for the next iteration of the i -th individual, r_3 is the weight factor of the global optimal individual, and r_4 is the discriminant coefficient.

4.3.2. The Improved SCA Based on Latin Hypercube Sampling and Adaptive Mutation (ISCA)

- (1) Population initialization strategy based on Latin hypercube sampling (LHS)

LHS [25] method is a multidimensional stratified sampling technology proposed by McKay et al. It can efficiently sample in the distribution interval of variables. Its essence is to divide the interval [0,1] into N equally spaced nonoverlapping subintervals and conduct independent equal probability sampling for each subinterval, so as to ensure that the sampling points are evenly distributed in the whole distribution interval.

- (2) Adaptive weight strategy

Affected by the particle swarm optimization (PSO) algorithm [26], this paper adds a weight factor to the original position update formula to speed up the convergence of the algorithm. The new position update formula is shown in (12).

$$x_{id}^{t+1} = \begin{cases} w(t) \cdot x_{id}^t + r_1 \times \sin r_2 \times |r_3 P_d^t - x_{id}^t|, & r_4 < 0.5, \\ w(t) \cdot x_{id}^t + r_1 \times \cos r_2 \times |r_3 P_d^t - x_{id}^t|, & r_4 \geq 0.5, \end{cases} \quad (12)$$

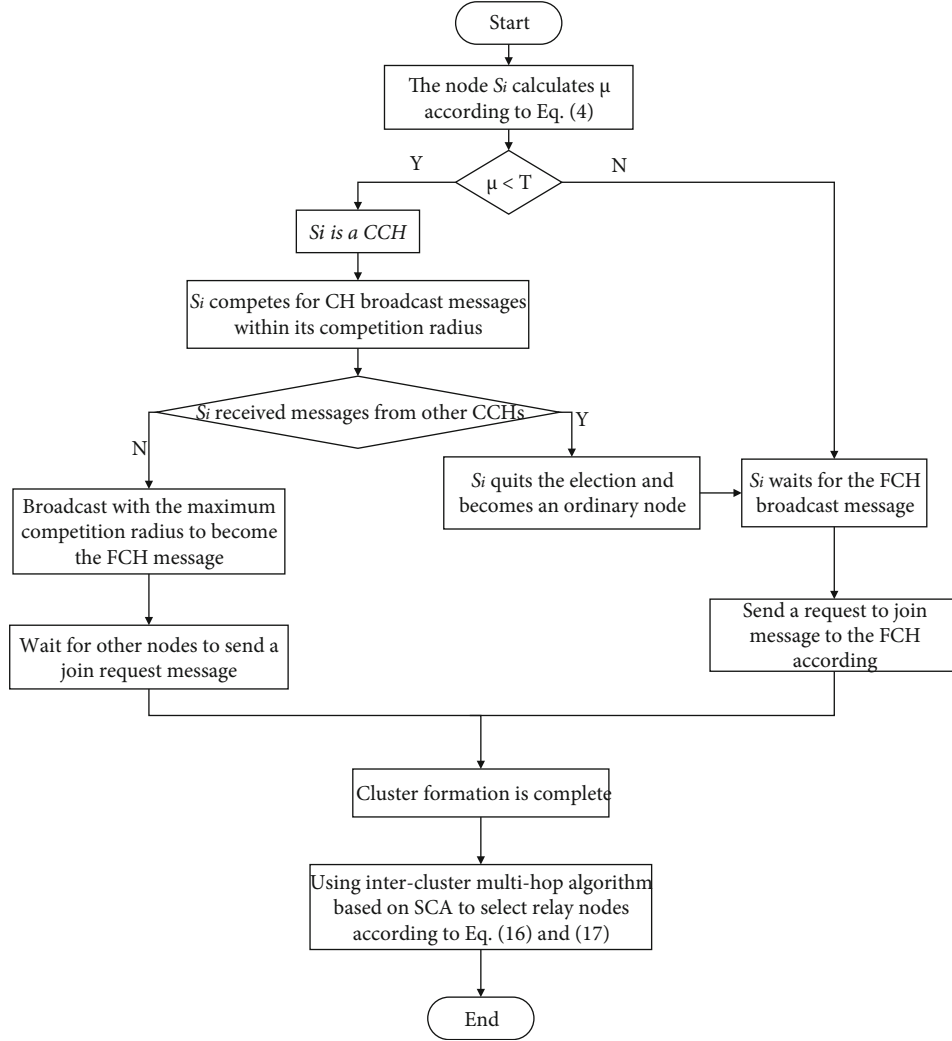


FIGURE 3: The flow chart of DUCISCA protocol.

where $w(t)$ is the weight factor. Its calculation is shown in (13).

$$w(t) = 0.2 \cdot \cos \left[\frac{\pi}{2} \cdot \left(1 - \frac{t}{\text{Max_iter}} \right) \right], \quad (13)$$

where t is the current iteration and Max_iter is the maximum iteration.

(3) Disturbance strategy

In order to expand the local search ability, Gaussian variation and Levy flight strategy are introduced to make the individuals trapped in local extreme points jump out of the limit and continue the search.

Gaussian variation [27] comes from the normal distribution of continuous probability distribution, which has good local development ability. The variation formula is

$$\text{mutation}(x) = x \cdot (1 + \text{randn}(1)). \quad (14)$$

TABLE 2: Network parameters.

Parameters	Values	Parameters	Values
Network coverage	(0,0)~(200,200)m	R_{comp}^0	60 m
Base station location	(100,250)m	c_1, c_2, c_3	0.4,0.3,0.3
Node number	400	d_1, d_2, d_3	0.4,0.3,0.3
Initial energy	0.5 J	w_1, w_2	0.5,0.5
Date packet size	4000 bits	δ	3
Date packet header	100 bits	T_{CH}	30 s
T	0.4		

Levy flight [28] obeys the Levy distribution. Please refer to [28] for specific mathematical model. The variation formula is

$$\text{mutation}(x) = x + (x - g\text{best}) \otimes \text{Levy}(\lambda), \quad (15)$$

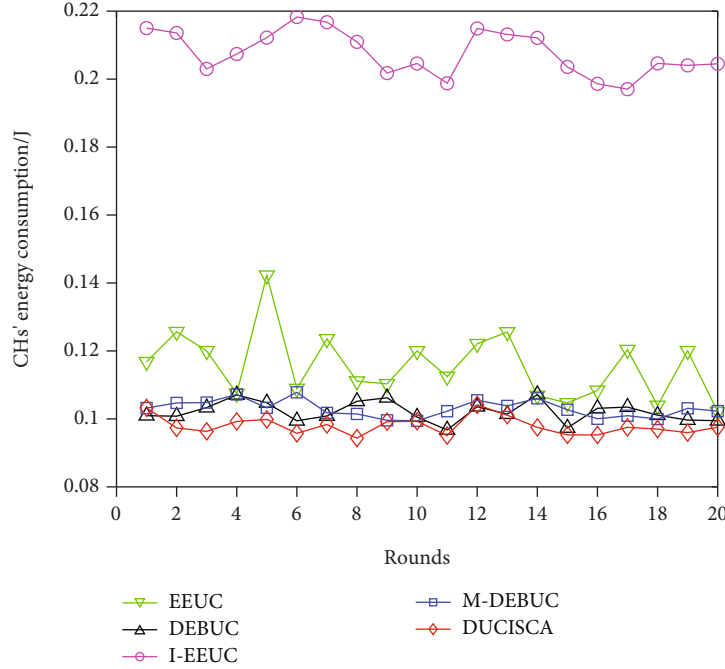


FIGURE 4: CHs' energy consumption.

where x is the original position, $\text{mutation}(x)$ is the position after Gaussian variation or Levy flight, \otimes denotes point to point multiplication, $\text{randn}(1)$ is a random number obeying standard normal distribution, and $\text{Levy}(\lambda)$ is the random vector obeying Levy flight.

4.3.3. Performance of ISCA. In order to verify the performance of ISCA, this paper compares ISCA with BOA [29], WOA [30], SSA [31], and SCA. The population size $N = 30$ and the maximum iteration $\text{Max_iter} = 500$ are set, respectively. Each algorithm operates 30 times independently, and the worst, optimal, average, and standard deviation of these 30 experiments are taken as evaluation indexes. This paper takes the first six 50-dimensional test functions in [32] as examples. The theoretical optimal function values are all 0. Table 1 shows the optimal value, the worst value, the mean value, and the standard deviation of the results obtained from each algorithm running 30 times.

Through the function optimization simulation experiments of five representative comparison algorithms on 6 different feature benchmark functions, the test results prove that the optimization accuracy and stability of ISCA are significantly improved.

4.3.4. Routing Based on ISCA. Suppose there are eight CHs in the network, and the CHs' set is $\{\text{CH}_1, \text{CH}_2, \dots, \text{CH}_8\}$. N individuals are randomly generated in the range of $(0,1)$, in which the dimension of the individual is the number of CHs, and each individual represents a multihop path.

CH_i takes δ times of the competition radius as the broadcast range. Other prior CHs and the BS that receive this message will become CH_i 's neighbour CHs. CH_d selects

TABLE 3: Number of dead node rounds for 5 protocols.

Protocol name	Rounds of FDN	Rounds of HDN
EEUC	356	535
DEBUC	594	633
I-EEUC	558	603
M-DEBUC	596	637
DUCISCA	655	668

the n_d -th CH in its neighbouring CHs' set as the final relay node. The value of n_d is calculated as shown in (16).

$$n_d = \lceil x_{id} \times \text{NeiNum}(\text{CH}_d) \rceil, \quad (16)$$

where x_{id} is the d -th dimensional value of the i -th individual, which denotes the probability that CH_d selects the relay node and $\text{NeiNum}(\text{CH}_d)$ indicates the number of neighbouring CHs of CH_d .

In order to reduce and balance the energy consumption of CHs and prolong the network lifetime, this paper takes the standard deviation of CHs' residual energy as the objective function of ISCA. The specific expression is shown in (17).

$$f(x_i) = \sqrt{\frac{1}{m} \sum_{i=1}^m [E_{\text{res}}(i) - \bar{E}]^2}, \quad (17)$$

where m denotes the number of CHs, $E_{\text{res}}(i)$ denotes the residual energy of CH_i , and \bar{E} denotes the mean value of CHs' residual energy. The residual energy of CH

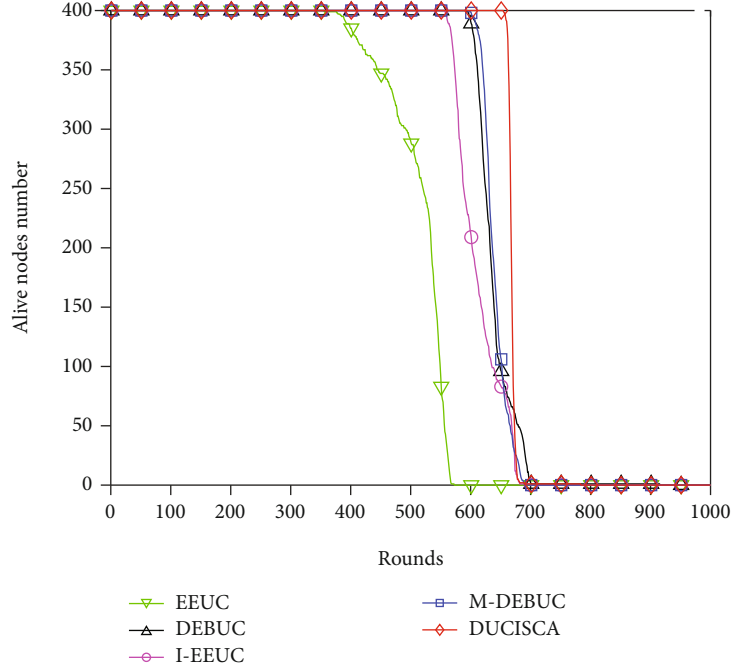


FIGURE 5: Network lifetime.

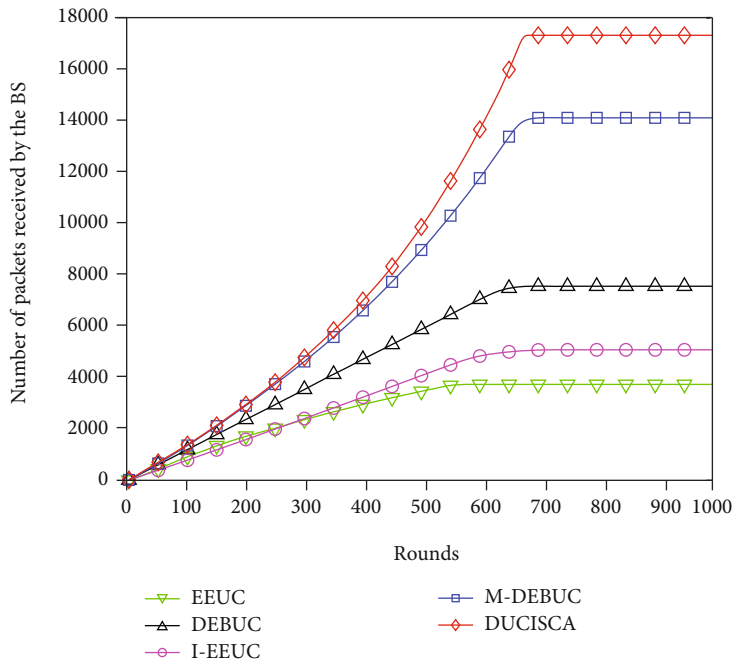


FIGURE 6: Number of packets received by the BS.

needs to subtract the energy consumed in intercluster communication, which is calculated as (18).

$$E_{\text{res}}(i) = E_{\text{bef}}(i) - E_{R_x}(l \times \text{PackNum}(i)) - E_{T_x}(l \times \text{PackNum}(i), d_{\text{next}}(i)), \quad (18)$$

where $E_{\text{bef}}(i)$ denotes the residual energy of CH_i before intercluster communication, $\text{PackNum}(i)$ denotes the number of

TABLE 4: Number of packets received by BS for 5 protocols.

Protocol name	Number of packets received by BS
EEUC	3730
DEBUC	7549
I-EEUC	5075
M-DEBUC	14101
DUCISCA	17312

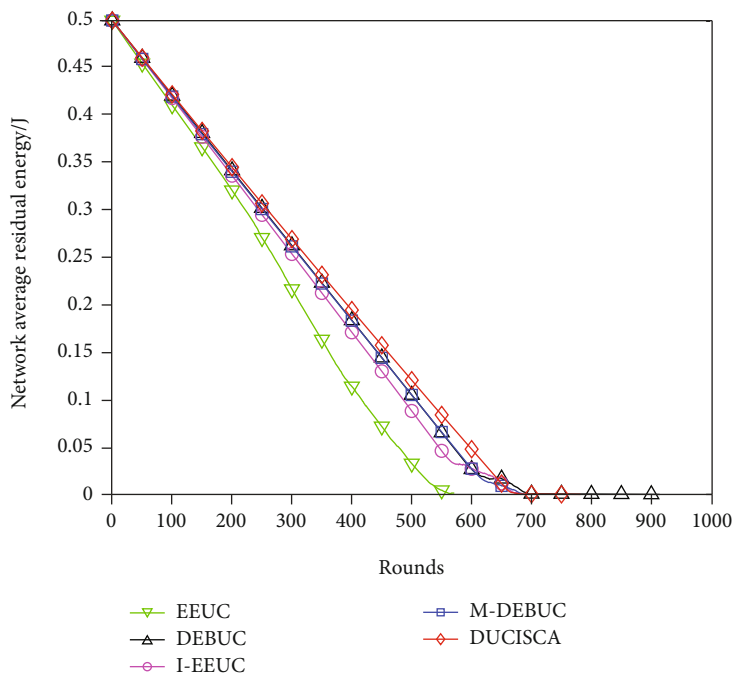


FIGURE 7: Network average residual energy.

packets received by CH_i , and $d_{\text{next}}(i)$ denotes the distance between CH_i and the relay node.

The pseudo code of intercluster multihop routing based on ISCA is shown in Algorithm 3.

The flow chart of the implementation of DUCISCA protocol is shown in Figure 3.

5. Experimental Results and Analysis

5.1. Simulation Parameters. To prove the performance of DUCISCA, the EEUC, DEBUC, I-EEUC, M-DEBUC, and DUCISCA are simulated under the same conditions using MATLAB 2018a, and many performances are compared. We run for 1000 rounds, and each node transmits a data packet each round. The network parameters are shown in Table 2.

5.2. Analysis of Simulation Results

5.2.1. CHs' Energy Consumption. A comparative analysis of CHs' energy consumption after 20 rounds of network clustering is made, as shown in Figure 4. The average CHs' energy consumption of I-EEUC, EEUC, DEBUC, M-DEBUC, and DUCISCA are 0.2077 J, 0.1156 J, 0.1021 J, 0.1030 J, and 0.0980 J, respectively. Obviously, CHs' energy consumption of DUCISCA is lower than other protocols, and the energy consumption of each CH is balanced in each round. The energy consumption of CHs is related to the number of packets forwarded and the forwarding distance. DUCISCA optimizes the selection strategy of CHs and selects high-quality nodes as CHs, and the number of CHs does not change much. In addition, when establishing intercluster forwarding paths, ideal multihop paths are searched through ISCA, and data packets are forwarded accordingly;

the energy consumption of intercluster communication is reduced.

5.2.2. Network Lifetime. Network lifetime is one of the important indicators to evaluate network performance [33]. Table 3 shows the first node death (FDN) and half node death (HDN) times for the five algorithms in Figure 5. If HDN is used as the criterion of network lifetime, the five algorithms have 535, 633, 603, 637, and 668 network lifetimes, respectively. Compared with EEUC, I-EEUC optimizes the CH selection algorithm for more balanced energy consumption. Compared with I-EEUC, DEBUC not only optimizes the rules of CH selection but also takes into account the residual energy of nodes and energy consumption within clusters and energy consumption between clusters in multihop routing, which further balances the network energy consumption. Compared with DEBUC, M-DEBUC improves the competitive radius and waiting time of DEBUC, so it consumes less energy than DEBUC. Compared with M-DEBUC, the proposed DUCISCA considers factors such as node energy, distance to the BS, and node degree when CCHs compete for FCHs, which is more reasonable than M-DEBUC CH selection and the FDN time is later. It can be seen from Figure 5 that DUCISCA can better balance the energy consumption between clusters, so the later death speed is slightly faster than other algorithms.

5.2.3. Number of Packets Received by the BS. Figure 6 is a comparison diagram of the number of packets received by the BS for the five protocols. Table 4 shows the final number of packets received by the BS for the five protocols. Based on this table, the number of packets received by the BS in DUCISCA is about 4.7 times that of EEUC, 3.4 times that of I-EEUC, 2.3 times that of DEBUC, and 1.2 times that of

M-DEBUC. This is because the improved CH selection and multihop routing mode of DUCISCA extends the lifetime of network, especially the improved multihop routing, so that more CHs indirectly send data to the BS through multihop mode, which increases the amount of data received by the BS. Before the arrival of DUCISCA's HDN, most nodes of the compared protocols have died, and there are coverage vulnerabilities in the monitoring area, so it is difficult for nodes to send data to the BS.

5.2.4. Network Average Residual Energy. Figure 7 compares the variation of the network average residual energy for the five protocols with the number of running rounds. Because EEUC only considers energy in routing, the intercluster routing is unreasonable, resulting in less network average residual energy. DEBUC uses a time-based broadcast mechanism to select CH that reduces control overhead effectively. Multihop routing takes into account such factors as CH energy, distance, and node degree. The average residual energy of the network is more than that of EEUC. DUCISCA improves CH selection algorithm and competition radius calculation algorithm, which makes the distribution of clusters more reasonable. Moreover, in order to save and balance energy, DUCISCA adopts intercluster multihop routing based on ISCA strategy that reduces the energy consumption of data transmission effectively. Hence, the energy consumption of DUCISCA is less than other four protocols.

6. Conclusions

The major contribution of this article is to propose a distributed unequal clustering routing protocol based on ISCA (DUCISCA) to enhance WSN performance in terms of energy efficiency and lifetime. In the CH competitive stage, this protocol improves the competitive radius which considers the residual energy of node and the distance to the BS. It can balance energy consumption of nodes in different locations to avoid the "hot spot" problem. This is a precondition for unequal clustering. At the same time, it adopts a time-based broadcast mechanism. The waiting time depends on the residual energy of CCHs, the distance to the BS, and the number of neighbour nodes, which can effectively reduce the overhead of nodes. After the cluster head selection is completed, it enters the cluster formation stage. In order to produce optimal clusters to balance network energy consumption, a rule for joining clusters is proposed, which takes into account the energy of CHs, the degree of nodes, and the distance between ordinary nodes and CHs. Finally, the standard deviation of cluster CHs' residual energy in intercluster communication is taken as the objective function to search the energy-balanced intercluster data forwarding path based on ISCA. Compared with EEUC, DEBUC, I-EEUC, and M-DEBUC, DUCISCA has better effectiveness compared to above protocols in terms of nodes' energy consumption, network lifetime, throughput, and network average residual energy.

In future work, we will further optimize the DUCISCA protocol according to the experimental results to improve the energy saving effect of the protocol.

Data Availability

Please contact with the corresponding author to acquire the underlying data if necessary.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. Prabhu, E. Gajendran, and N. Balakumar, "Prospective applications of wireless sensor node and wireless sensor network," *International Journal of Advanced Computing and Electronics Technology*, vol. 4, no. 1, pp. 18–26, 2017.
- [2] J. Sumathi and R. L. Velusamy, "A review on distributed cluster based routing approaches in mobile wireless sensor networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 835–849, 2021.
- [3] X. He, S. Liu, G. Yang, and N. Xiong, "Achieving efficient data collection in heterogeneous sensing WSNs," *IEEE Access*, vol. 6, pp. 63187–63199, 2018.
- [4] S. Dhiviya, A. Sariga, and P. Sujatha, "Survey on WSN using clustering," in *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, pp. 121–125, Tindivanam, India, 2017.
- [5] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless micro-sensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, vol. 2, p. 10, Maui, HI, USA, 2000.
- [6] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.
- [7] A. Thakkar and K. Kotecha, "CVLEACH: coverage based energy efficient LEACH algorithm," *International Journal of Computer Science and Network*, vol. 1, no. 3, 2012.
- [8] B. Jan, H. Farman, H. Javed, B. Montrucchio, M. Khan, and S. Ali, "Energy efficient hierarchical clustering approaches in wireless sensor networks: a survey," *Wireless Communications and Mobile Computing*, vol. 2017, Article ID 6457942, 14 pages, 2017.
- [9] A. Thakkar and K. Kotecha, "Cluster head election for energy and delay constraint applications of wireless sensor network," *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2658–2664, 2014.
- [10] A. Thakkar and K. Kotecha, "A new Bollinger Band based energy efficient routing for clustered wireless sensor network," *Applied Soft Computing*, vol. 32, pp. 144–153, 2015.
- [11] J. Daniel, S. F. V. Francis, and S. Velliangiri, "Cluster head selection in wireless sensor network using tunicate swarm butterfly optimization algorithm," *Wireless Networks*, vol. 27, no. 8, pp. 5245–5262, 2021.
- [12] S. Al-Sodairi and R. Ouni, "Reliable and energy-efficient multihop LEACH-based clustering protocol for wireless sensor networks," *Sustainable Computing: Informatics and Systems*, vol. 20, pp. 1–13, 2018.
- [13] Y. Liu, Q. Wu, T. Zhao, Y. Tie, F. Bai, and M. Jin, "An improved energy-efficient routing protocol for wireless sensor networks," *Sensors*, vol. 19, no. 20, 2019.

- [14] V. Nivedhitha, A. Gopi Saminathan, and P. Thirumurugan, "DMERP: A dynamic multi-hop energy efficient routing protocol for WSN," *Microprocessors and Microsystems*, vol. 79, 2020.
- [15] S. Arjunan and S. Pothula, "A survey on unequal clustering protocols in wireless sensor networks," *Journal of King Saud University - Computer and Information Sciences*, vol. 31, no. 3, pp. 304–317, 2019.
- [16] S. Soro and W. B. Heinzelman, "Prolonging the lifetime of wireless sensor networks via unequal clustering," in *19th IEEE International Parallel and Distributed Processing Symposium*, p. 8, Denver, CO, USA, 2005.
- [17] M. Ye, C. F. Li, G. H. Chen, and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," in *IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, pp. 597–604, Washington, DC, USA, 2005.
- [18] X. Wang, Y. Peng, and L. Huang, "An improved unequal cluster-based routing protocol for energy efficient wireless sensor networks," in *2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 165–169, Changsha, China, 2019.
- [19] C. J. Jiang, W. R. Shi, X. L. Tang, P. Wang, and M. Xiang, "Energy-balanced unequal clustering routing protocol for wireless sensor networks," *Journal of Software*, vol. 23, no. 5, pp. 1222–1232, 2012.
- [20] Y. Shui, *Research on Non-uniform Clustering Routing Protocol of Mine Wireless Sensor Network*, China University of Mining and Technology, 2020, <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD202101&filename=1020902481.nh&uniplatform=NZKPT&v=DDujhyWFD-mfqexAtpk0mXOBdYsy5t6IzGxBc2TN2ZQyYGuu97JdfD-oA7p5lInM>.
- [21] L. I. Xinchun and G. A. O. Baisheng, "WSN routing algorithm based on optimal number of clusters and improved gravitational search," *Journal of Computer Applications*, vol. 37, no. 12, pp. 3374–3380, 2017.
- [22] E. Zaraket, N. M. Murad, S. S. Yazdani, L. Rajaoarisoa, and B. Ravelo, "An overview on low energy wake-up radio technology: active and passive circuits associated with MAC and routing protocols," *Journal of Network and Computer Applications*, vol. 190, 2021.
- [23] M. Abdurrohman, Y. Supriadi, and F. Z. Fahmi, "A modified E-LEACH routing protocol for improving the lifetime of a wireless sensor network," *Journal of Information Processing Systems*, vol. 16, no. 4, pp. 845–858, 2020.
- [24] S. Mirjalili, "SCA: a sine cosine algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.
- [25] L. Aibin and L. Wenyi, "Reliability evaluation of distribution network with distributed generation based on Latin hypercube sequential sampling," in *2020 3rd International Conference on Electron Device and Mechanical Engineering (ICEDME)*, pp. 97–99, Suzhou, China, 2020.
- [26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, WA, Australia, 1995.
- [27] H. Yu, Y. Gao, and J. Wang, "A multi-objective particle swarm optimization algorithm based on competition mechanism and Gaussian variation," *Complexity*, vol. 2020, Article ID 5980504, 23 pages, 2020.
- [28] J. Zhao, H. Zhu, Y. Hu et al., "A gravitational search algorithm based on Levy flight," *Journal of Physics: Conference Series*, vol. 1865, no. 4, 2021.
- [29] S. Arora and S. Singh, "Butterfly optimization algorithm: a novel approach for global optimization," *Soft Computing*, vol. 23, no. 3, pp. 715–734, 2019.
- [30] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [31] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020.
- [32] J. S. Liu, M. M. Yuan, and F. Zuo, "Global search-oriented adaptive leader salp swarm algorithm," *Control and Decision*, vol. 36, no. 9, pp. 2152–2160, 2021.
- [33] L. Banteng, Z. Ying, and C. Yourong, "Research on the routing algorithm optimizing lifetime of wireless ad hoc networks," *Chinese Journal of Sensors and Actuators*, vol. 30, no. 3, pp. 463–466, 2017.