IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 3 February 2023, accepted 26 February 2023, date of publication 3 March 2023, date of current version 9 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3252027

RESEARCH ARTICLE

WSN Clustering Routing Algorithm Combining Sine Cosine Algorithm and Lévy Mutation

XIAOLING GUO, YONGFEI YE, LING LI, RENJIE WU, AND XINGHUA SUN
School of Information Science and Engineering, Hebei North University, Zhangjiakou 075000, China

Corresponding author: Yongfei Ye (583053006@qq.com)

This work was supported in part by the Medical Science Research Project of Hebei Province under Grant 20200488; in part by the Natural Science Foundation of Hebei Province under Grant H2022405021; in part by the Research Project of Basic Scientific Research Business of Provincial Colleges and Universities of Hebei North University under Grant JYT2022020; in part by the College Students' Innovation and Entrepreneurship Training Program under Grant 202210092006 and Grant S202210092031; in part by the Key Research Projects of Humanities and Social Sciences in Colleges and Universities in Hebei Province under Grant SD2022048; and in part by the Research and Practice Project of Municipal School Deeply Integrated Education and Teaching Reform, in 2022, under Grant SXRHJG202231.

ABSTRACT Directed against the disadvantages of relatively short life-cycle and unbalanced energy utilization among nodes in WSN, a clustering routing algorithm combining sine cosine algorithm and Lévy mutation is developed. During the cluster head election stage, the amount of cluster heads is dynamically calculated according to the surviving nodes for keeping it at a reasonable value; taking full account of the current energy of nodes, only nodes with high energy can be candidate cluster heads; the fitness function is constructed according to intra-cluster distance, so that the distribution structure within the cluster are relatively uniform; the Sine Cosine Algorithm with improved step size search factor is used for cluster head election, and Lévy mutation is introduced to realize the variation of population. The group of individuals with the lowest fitness function value is used as final election scheme for current round. In the data transmission phase, for the sake of avoiding long-distance transmission, the relay node is designed to forward data. The proposed algorithm effectively extends network life-cycle and well equalizes the load of network nodes.

INDEX TERMS Clustering algorithm, energy balance, life cycle, Lévy mutation, sine cosine algorithm, WSN.

I. INTRODUCTION

Wireless Sensor Network (WSN) is compounded from many micro intelligent sense nodes. The nodes can collect, handle and transfer data in monitoring area cooperatively, then relay them to base-station. In WSN, initial energy carried by every sensor node is generally limited. Therefore, the primary goal of WSN is to reduce sense node energy consumption [1], [2], [3]. Cluster division and data transmission in cluster routing algorithms are closely related to the energy consumption. Therefore, clustering routing algorithm has always been a significant research content of WSN. Clustering routing algorithm is generally divided into distributed routing algorithm [4], [5], [6], [15] and centralized clustering routing algorithm [7], [8], [9], represented by LEACH and LEACH-C respectively, which have different application scenarios. The distributed algorithm requires higher computing, storage and power capability of sensor nodes, and the centralized clustering routing algorithm can select the relatively optimal cluster head or path scheme in each round. On the division of clusters, some scholars first select cluster heads and then divide clusters [10], [11], some first divide clusters, and then select cluster heads [12], [13]. In order to save energy, someone also put forward the concepts of main cluster head, secondary cluster head or double cluster head [14]. During the dynamic selection of cluster heads, some factors such as clustering uniformity, base station (BS) location, residual energy, distance, cluster head load and cluster head selection times are generally considered [15], [16], [17]. In the data transmission stage, considering BS location, there are generally two types of transmission data: single-hop [18], [19] and multi-hop [20], [21]. During the process

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

of inter-cluster multi-hop transmission, constructing the least energy-consuming path from cluster heads to BS becomes the core. Inspired by PEGASIS protocol, other algorithms introduce the idea of chaining [22], [23] to construct chain structure within or among clusters to save energy consumed by nodes. Some experts consider some factors which affecting network energy consumption from the perspective of system heterogeneous [16], [24], [28] or nonuniform clustering [6], [14], [19], [20], [25]. In addition, in recent years, many experts have introduced various innovative algorithms in WSN and improved the algorithm prototype, such as swarm intelligence optimization algorithm [26], [27], [28], artificial intelligence algorithm [8], [29], genetic evolution algorithm [11], [30], [31] or combined with fuzzy logic algorithm [32], trying to find new ways and new ideas to solve the problem of WSN clustering routing algorithm.

The paper proposes a WSN clustering routing algorithm (SCA-Lévy Clustering Routing Algorithm) which combines sine and cosine and Lévy mutation. Sine and cosine algorithm with improved step-size search factor is added to select cluster heads. Lévy mutation is used to increase population diversity and improve global exploration ability. The fitness function is constructed on account of the sum of intra-cluster distance variance. The number of calculated surviving nodes dynamically determine the amount of cluster heads in this round. Nodes with high residual energy are more likely to become candidate cluster heads, saving node energy consumption as much as possible. Its main contributions are as follows:

1. The sine-cosine algorithm and Lévy mutation are fused and applied to the WSN clustering routing algorithm. We improve search step of SCA algorithm and increase the global search ability. The optimal cluster-head scheme in each round is selected to save energy. At the same time, Lévy variation was introduced to make the population diversity better.
2. Considering the location relationship between the BS and the monitoring area, an election scheme combining residual energy and intra-cluster distance is proposed, which makes the member nodes uniformly distributed around cluster heads and ensures the minimum inter-cluster communication cost.
3. During data transmission stage, the relay node rules are designed to avoid long-distance transmission. In the selection of relay nodes, the location of nodes, the remaining energy of the BS location are used to find the relatively optimal relay node to extend the network life cycle.

II. RELATED WORK

Using one or two improved new swarm intelligence optimization algorithms to solve the clustering routing problem of WSN shows some adaptability and effectiveness, but there is still room for improvement in saving energy consumption and extending life cycle. Zhao et al. [33] proposed that the good global search ability of beetle antennae search (BAS) algorithm can be used to improve the position update of the contraction and encirclement phase of the whale optimization algorithm (WOA). And a clustering selection function in view of energy and distance has been proposed for making clustering more reasonable. This paper effectively combines the advantages of BAS algorithm and WOA algorithm, and applies them to WSN clustering routing algorithm, which shows good performance in saving node energy. Zhu and Sun [12] proposed that the cluster area can be firstly divided. The cuckoo algorithm was used to select the initial cluster center. K-means method was taken to divide uniform clusters. During data transmission phase, routing for cluster heads was also designed to save energy. This paper considered more comprehensively, but from the simulation results, compared with the [33], this paper still has some room for improvement in extending the network life cycle. Yang and Lu [34] proposed that fuzzy neural network can be used to select cluster heads. And particle swarm optimization algorithm was introduced to improve routing. For studying the fitness function, the number of relay nodes, the relay load factor, the distance from the gateway to the sink node were comprehensively considered. This algorithm was more complex, and the proposed fuzzy neural network content needed further improvement. Sun et al. [35] proposed that, the location update formula of sparrow search algorithm can be improved by using adaptive t-distribution, and its effectiveness and feasibility were verified. Then, the improved algorithm was used to elect cluster heads in LEACH algorithm, which making LEACH cluster head election more optimized. However, it had a relatively long process from the first dead node to 80% of dead nodes, which indicating the energy consumption among nodes is not uniform. Sun et al. [36] proposed that, firefly algorithm can be used to cluster, and chaos theory was introduced to balance the distribution of clustering centers. Double cluster head mechanism was adopted. Main cluster head was in responsible for information reception and data fusion, and sub cluster head was responsible for data transmission. In process of data transmission, Bellman-Ford multi hop path algorithm was adopted. But the first dead node appears earlier in this algorithm. Chang et al. [37] proposed that, firstly, the improved artificial fish swarm algorithm can be used to iteratively output initial center of FCM; secondly, the optimal number of cluster heads were determined. Clustering could be reasonably completed. Thirdly, energy and distance are calculated for cluster head election. In this network, the death time of the first node and half of the nodes were used to measure the network life-cycle and effectiveness. From the simulation results, the algorithm can be further improved in balancing energy consumption. He et al. [38] proposed improved sine cosine algorithm (SCA) adapted to WSN. The inertia weight factor was introduced, and the fitness function was constructed by residual energy and distance among nodes. The improved sine cosine algorithm (CRISCA) was applied to WSN clustering routing algorithm, and compares it with the classic LEACH algorithm. However, the performance in terms of extending the network life cycle was not very ideal. Based on the accumulation and thinking of previous

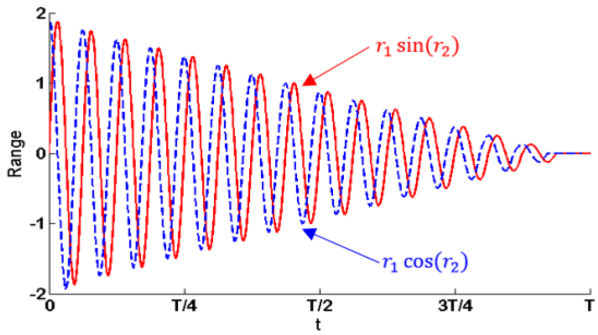


FIGURE 1. The fluctuation range of $r_1\sin(r_2)$ or $r_1\cos(r_2)$.

work, we propose SCA-Lévy clustering routing algorithm that combines sine and cosine algorithm and Lévy mutation.

III. SINE COSINE ALGORITHM

Sine Cosine Algorithm (SCA) [39] randomly generates several candidate solutions and uses the mathematical models of sine and cosine to update the position in the solution space. In it, X_i^t is the position in i-th dimension of t-th iteration, P_i^t is the current optimal individual in the i-th dimension, iterate calculation process is executed according to formula (1):

$$X_i^{t+1} = \begin{cases} X_i^t + r_1 \times \sin(r_2) \times \left| r_3 P_i^t - X_i^t \right|, r_4 < 0.5 \\ X_i^t + r_1 \times \cos(r_2) \times \left| r_3 P_i^t - X_i^t \right|, r_4 \geq 0.5 \end{cases} \quad (1)$$

In formula (1), r_1 is the search step size, which controls the search process of the algorithm. In general, $r_1 = a(1 - \frac{t}{T})$. In it, a is usually set as 2. t represents current iteration number. T is maximum iteration number. r_2, r_3 and r_4 are random parameters, $r_2 \in [0, 2\pi]$, $r_3 \in [0, 2]$, $r4 \in [0, 1]$.

In SCA algorithm, r_1 decreases from 2 to 0 with the increase of t. Moreover, r_2 is a random number between 0 and 2π. It is not difficult to see that when $0 \leq t \leq T/2$, r_1 decreases from 2 to 1, at this time, the sine function value $r_1\sin(r_2)$ or cosine function value $r_1\cos(r_2)$ is in $[1, 2]$ or $[-2, -1]$. According to formula (1), the location update of the solution fluctuates greatly, which can be explored in a large range of space to search for more possible solutions. When $T/2 \leq t \leq T$, r_1 decreases from 1 to 0, the sine function value $r_1\sin(r_2)$ or cosine function value $r_1\cos(r_2)$ must be in $[-1, 1]$. At this moment, the position update fluctuation of the solution is slight, and the possible solutions can be found in the local range. The fluctuation range of $r_1\sin(r_2)$ or $r_1\cos(r_2)$ was shown in the Figure 1.

In SCA, the search step size factor r_1 decreases linearly to 0 with the increases of iterations t, in the later period of iteration, the update change of the solution has been very light, and the solution could easily be local optimum. Inspired by the train of thought [39], the improvement for the search step factor r_1 is made, as shown in formula (2). Before and after improvement, the relationship between r_1 and t is shown in the Figure 2.

$$r_1 = a\sin(\frac{\pi}{2}(1 - \frac{t}{T})) + b \quad a = 2, b = 0.5 \quad (2)$$

After improvement, it is easy to see that the search step size factor r_1 is in $[0.5, 2.5]$ from Figure 2, and the sine

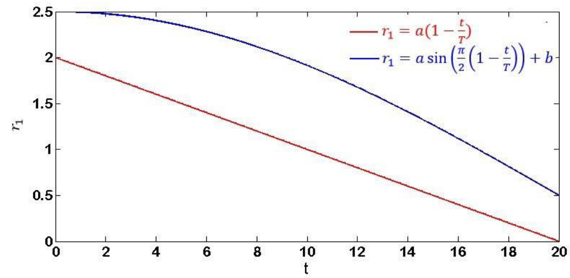


FIGURE 2. Comparison Chart of the search step factor r_1 Value.

function value $r_1\sin(r_2)$ or cosine function value $r_1\cos(r_2)$ is in $[-2.5, 2.5]$. The solution space is wider. In addition, When the improved search step size factor r_1 decreasing, it is slower than linear decreasing, which is also conducive to global development.

IV. LÉVY MUTATION

Lévy flight is a random walk. There is a relatively high probability of large steps in the process of random walk. It is very suitable for the application scenario of jumping out of local search and increasing global exploration. The random walk in Lévy flight is called Lévy distribution [40], and its model is shown in formula (3):

$$L(s) \sim |s|^{-1-\beta} \quad (3)$$

In the formula (3), s deputies the random walk step size, β is an exponential parameter that determines the distribution shape. Usually, Mantegna algorithm is used to generate random steps of Lévy flight, as shown in formula (4)-(6):

$$s = \frac{u}{|v|^{1/\beta}} \quad (4)$$

$$u \sim N(0, \sigma_u^2), v \sim N(0, \sigma_v^2) \quad (5)$$

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \sigma_v = 1 \quad (6)$$

u, v are the parameters subject to normal distribution, Γ is gamma function, $0 < \beta < 2$, generally β is 1.5.

By using Lévy variation for the population individuals, the diversity of the population may be expanded, the solution domain may be fully traversed, and the global search ability can be promoted. Therefore, the Lévy mutation operation is introduced. First, the average fitness function value is calculated. For those individuals whose fitness function value is less than average, Lévy mutation is performed. Lévy mutation based on Lévy flight is shown in formula (7):

$$X_{ij}(t+1) = X_{ij}(t) + L(s) * \left| X_j^*(t) - X_{ij}(t) \right| \quad (7)$$

In formula (7), $X_{ij}(t+1)$ is the location after Lévy mutation, $X_{ij}(t)$ is the current location, $X_j^*(t)$ is the current optimal position.

V. SCA- LÉVY CLUSTERING ROUTING ALGORITHM

In SCA-Lévy algorithm, the BS runs the SCA and Lévy mutation for election. The group of cluster heads with the

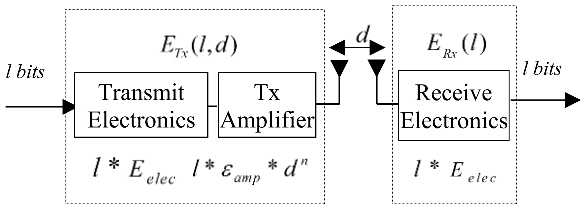


FIGURE 3. Energy consumption model.

lowest communication cost in each round is elected. Other nodes choose closest cluster head to become their member nodes. In data transmission phase, member nodes collect information. Cluster heads execute local fusion and forward to BS directly or by the relay node.

A. NETWORK MODEL

Suppose the WSN monitoring area is $M \times M$ in size, and BS is located above it. Sensor nodes are deployed for data sensing and collection. Energy carried by every node at the beginning is the same and will not be supplemented later. During network operation, the energy is constantly consumed. If it is below zero, the node is considered dead. To make the initial distribution of nodes more uniform, Tent chaotic map is used for deployment, as shown in formula (8) and (9):

$$x^{i+1} = \begin{cases} 2x^i, x^i \in [0, 0.5] \\ 2(1-x^i), x^i \in (0.5, 1] \end{cases} \tag{8}$$

$$z = x^{i+1}(ub - lb) + lb \tag{9}$$

Formula (8) is used to generate [0,1] chaotic sequence, and formula (9) is used to map the initial position of the node. lb, ub are the lower and upper bounds of node position respectively. Tent chaotic map can increase the space coverage and make the distribution of nodes relatively uniform.

B. ENERGY CONSUMPTION MODEL

During the data transmission, the energy consumption model [31] is shown in Figure 3.

The energy consumed for sending data is calculated according to formula (10):

$$E_{TX}(l, d) = \begin{cases} lE_{elec} + lE_{fs}d^2, d < d_0 \\ lE_{elec} + lE_{mp}d^4, d \geq d_0 \end{cases} \tag{10}$$

The energy consumed for receiving data is calculated according to formula (11):

$$E_{Rx}(l) = lE_{elec} \tag{11}$$

In formula (10) and (11), l represents data and d represents distance. E_{elec} is the energy consumption for per-unit data. ε_{fs} and ε_{mp} are known numbers, $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \approx 87$.

The energy consumed for data fusion is calculated according to formula (12):

$$E_{Fx}(n, l) = nlE_{DA} \tag{12}$$

In formula (12), n means total node numbers in every cluster. l represents data total transferred to cluster heads, $E_{DA} = 5nJ/bit/signal$.

C. DYNAMIC CLUSTER HEAD

The selected cluster heads should not be too many or too few and need to be kept reasonable. Too many or too few will increase energy consumption [41]. The number is closely related to surviving nodes. When dead nodes appear, the cluster heads should be reduced accordingly. Assume that the elected cluster heads are k_{opt}, calculated according to formula (13):

$$k_{opt} = round(N_{alive} * p) = round[(N - D_{dead}) * p] \tag{13}$$

In formula (13), N_{alive} represents surviving sensor nodes in current network, p represents cluster head election probability. N represents the total nodes. D_{dead} represents dead nodes.

In each round, the average residual energy is calculated by BS. Candidate cluster head set is composed of nodes with high residual energy. k_{opt} nodes randomly are selected from the candidate set. The initial population is formed by randomly selecting m times.

D. FITNESS FUNCTION

In order to minimize energy consumption and communication cost, fitness function is constructed with the communication distance within the cluster [42]. The population with the minimum fitness function value would develop into cluster head election scheme. When the fitness function value is smaller, the intra-cluster distance is smaller. The cluster heads are relatively evenly distributed around the monitoring area. Member nodes consume less energy.

$d_{toCH}(i)$ represents the distance from member nodes to its own cluster head. There are k_{opt} clusters in every round. The member node number in each cluster is uncertain, which may be n, m or q. D is the sum of the squares of the distances within the cluster of all clusters. The calculation is shown in formula (14). The fitness function is the sum of the absolute values of the difference between the sum of the squares of the distances within the cluster of each cluster and the average of the squares of the distances within the cluster of all clusters. And the calculation is shown in formula (15).

$$\begin{cases} D = \sum_{i=1}^{n} d_{toCH}^2(i) + \sum_{i=1}^{m} d_{toCH}^2(i) + \cdots + \sum_{i=1}^{q} d_{toCH}^2(i) \\ n + m + \cdots + q = N_{alive} - k_{opt} \end{cases} \tag{14}$$

$$f = \left| \sum_{i=1}^{n} d_{toCH}^2(i) - \frac{1}{k_{opt}} * D \right| + \left| \sum_{i=1}^{m} d_{toCH}^2(i) - \frac{1}{k_{opt}} * D \right| + \cdots + \left| \sum_{i=1}^{q} d_{toCH}^2(i) - \frac{1}{k_{opt}} * D \right| \tag{15}$$

Figure 4 shows the distribution of cluster heads when each algorithm runs to 400 rounds. It can be seen from the figure that the clustering is not uniform, some clusters are large, and some are small. Compared with Figures (8), it can be seen that because the fitness function of LEACH-SCA uses formula (15), the clustering is relatively uniform.

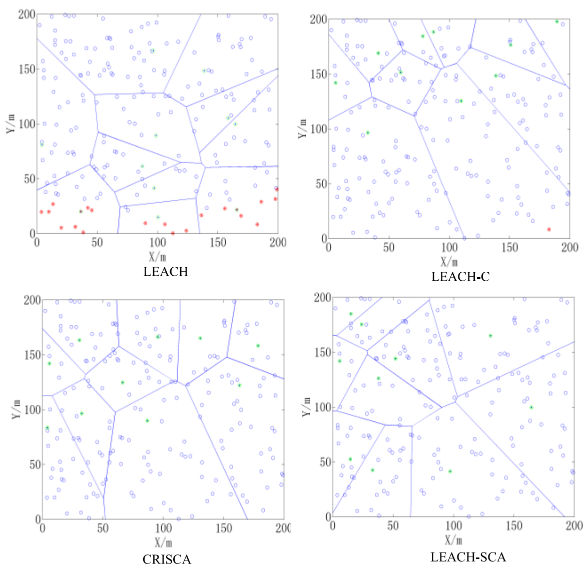


FIGURE 4. Distribution of cluster heads when each algorithm runs to 400th rounds.

E. CLUSTER HEAD ELECTION ALGORITHM

The process of cluster head election is shown in Figure 5, and the specific description is as follows:

Step 1: parameter initialization: population number m, iteration number T, etc.;

Step 2: The k_{opt} for the number of cluster heads elected are calculated by using formula (13) in this round;

Step 3: The candidate cluster head collection is built with those nodes owned high residual energy;

Step 4: In the candidate cluster head set, randomly select k_{opt} cluster heads for m elections, the initial population is constituted;

Step 5: Location is updated by formulas (1) and (2) to search solution;

Step 6: All individual fitness function values are calculated according to formula (15);

Step 7: The average fitness function value of all individuals is also calculated. Lévy mutation is performed according to formula (7) when individuals whose fitness function value smaller than average value. The method can increase population diversity and expand the global search capability;

Step 8: A group of individuals with the lowest fitness function value is recorded;

Step 9: Steps 5-8 are repeated until the maximum iteration number T, and the group of individuals with the smallest fitness function value is recorded.

F. RELAY NODE

During transmitting from cluster heads to BS, if the distance is greater than d_0, a relay node may be found to forward in the region; if the distance is less than d_0, directly sending data to the BS can be used.

The definition of Relay Node (RN):

Among the surviving nodes in the region, some nodes can act as relay nodes if they meet the following conditions: if

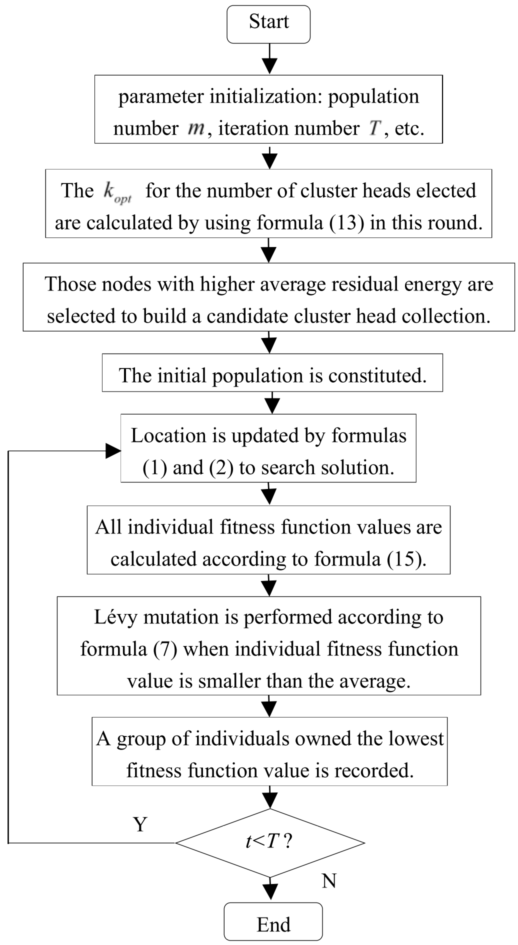


FIGURE 5. The flow chart of cluster head election.

the sum of the square of the distance from cluster head to relay node (d^2_{CH-RN}) and the square of the distance from relay node to base-station (d^2_{RN-BS}) is less than the square of the distance from the cluster head to the BS (d^2_{CH-BS}); and the distance from cluster head to relay node (d_{CH-RN}) and the distance from the relay node to the BS (d_{RN-BS}) are less than $1\big/\sqrt{2}$ times of d_{CH-BS}; at the same time, the value $\max(E_i\big/(d^2_{CH-RN} + d^2_{RN-BS}))$ is maximum when node energy is divided by the sum of square of the distance, as shown in formula (16):

$$RN = \begin{cases} E_{RN} > 0 \cap d^2_{CH-RN} + d^2_{RN-BS} < d^2_{CH-BS} \\ d_{CH-RN} < 1/\sqrt{2} * d_{CH-BS} \\ d_{RN-BS} < 1/\sqrt{2} * d_{CH-BS} \\ \max(E_i/(d^2_{CH-RN} + d^2_{RN-BS})) \end{cases} \tag{16}$$

According to the proven of multi hop forwarding in LEACH algorithm, the selected relay node RN should meet $d^2_{CH-RN} + d^2_{RN-BS} < d^2_{CH-BS}$. In Figure 6, the square $ABCD$ represents the monitoring area, H represents any cluster head, and F represents BS. Take HF as the diameter, and make a circle with I as the center. Any point G on the circle and its diameter form a right-angled triangle, and the right-angled triangle satisfies $FG^2 + HG^2 = FH^2$. According to the triangle rule

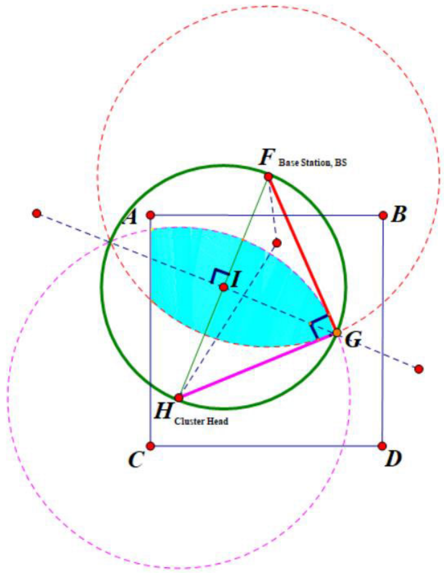


FIGURE 6. Range of relay node.

and Pythagorean theorem, for achieving $d^2_{CH-RN}+d^2_{RN-BS} < d^2_{CH-BS}$, the relay node should be found within the circle with I as the center. The triangle formed by any point outside the circle and diameter FH is $d^2_{CH-RN} + d^2_{RN-BS} > d^2_{CH-BS}$. In the triangle composed of cluster head node, relay node and BS, d_{CH-RN} and d_{RN-BS} must be limited for avoiding the long distance. Taking the right semicircle as an example, FH is the diameter. The only isosceles right triangle ΔFHG is made. $FG = HG = 1/\sqrt{2}FH$. The preferred location of the relay node is any point on the IG (except for point G). The closer the relay node is to point I, the shorter the distance d_{CH-RN} and d_{RN-BS}, and the better the location. However, in the actual node distribution, there may be no such point. In this paper, we will design two circles. One is with F as the center and FG as the radius, the other is with H as the center and HG as the radius. The intersection of two circles and square $ABCD$ is the optional range of relay nodes. Finally, a node is calculated as relay node if the value $E_i\big/(d^2_{CH-RN} + d^2_{RN-BS})$ is Maximum, when its energy is divided by the sum the square of the distance. The key pseudocodes of relay node algorithm are as follows:

Algorithm 1 Relay Node Algorithm

```
Input: Distance from cluster heads to BS (d_{CH-BS}) and
residual energy of cluster heads (E_i)
Output: Relay Node
Begin:
if d_{CH-BS} > d_0
for i=1: 1: NodeNumber
  if E_RN > min E & d^2_{CH-RN} + d^2_{RN-BS} < d^2_{CH-BS} & RNflag == 0
    if d_{CH-RN} < 1/√2 * d_{CH-BS} & d_{RN-BS} < 1/√2 * d_{CH-BS}
      if E_i/(d^2_{CH-RN} + d^2_{RN-BS}) > RNE
      RNE = E_i/(d^2_{CH-RN} + d^2_{RN-BS})
      end
    end
  end
end
end
Terminate
```

VI. SIMULATION AND ANALYSIS

MATLAB is used for programming simulation. The network life-cycle is specified as the appearing time of the first dead node. The operation stops when dead nodes reach 80%. Taking clustering and relay forwarding, energy consumption, network life cycle, throughput and applicability as performance evaluation indicators, the SCA-Lévy algorithm is compared with LEACH, LEACH-C, CRISCA [38] and LEACH-SCA algorithm (using the standard SCA algorithm to replace the simulated annealing algorithm).

A. SIMULATION PARAMETER SETTING

Table 1 shows the parameter and value used in the experimental simulation.

B. SCA- LÉVY CLUSTERING AND RELAY FORWARDING

In the SCA-Lévy algorithm, the BS is located above the region. The blue hollow circle represents common node, and the green meter circle represents selected cluster heads. The Voronoi Diagram is taken to draw clustering situation. The black line between nodes indicates that the cluster head has found a relay node.

Figure 7-9 are the schematic diagram of clustering and relay forwarding when the SCA-Lévy algorithm runs to 300, 400 and 500 rounds.

It can be seen that cluster heads are relatively evenly distributed, and member nodes in each cluster are also relatively evenly around the cluster heads. This is because when building the fitness function, the direct relationship between energy consumption and distance was fully considered. The fitness function was constructed by the distance within the cluster. Fitness function of each round is the smallest, that is, the square of the intra-cluster distance is the closest to the average value all clusters. Therefore, member nodes are relatively compact around cluster head. In addition, relay node found for cluster heads is also reasonable. The relay node is roughly located in the middle range from cluster head to BS. Two distances are both in an appropriate range, avoiding long-distance transmission.

C. ENERGY CONSUMPTION

Figure 10 is a comparison chart of cumulative energy consumption. Cumulative energy consumption line graph of SCA-Lévy algorithm is almost at the bottom, the least, and close to linear. When the initial energy of each algorithm is the same, it is not difficult to see that SCA-Lévy algorithm not only consumes the least energy per round, but also consumes almost uniform energy per round. In addition, SCA-Lévy algorithm runs more rounds than others, so its energy utilization rate is significantly higher than other algorithms.

Figure 11 is a comparison chart of the energy consumption of each round.

In Figure 11, the energy consumption of each round algorithm is the minimum in SCA-Lévy, which is about 0.105J

TABLE 1. Simulation parameter and value.

Parameter	Value
Region $M \times M$	200×200
Node Number N	200
Base station (BS)	100,250
Node initial energy	0.5 J
Packet size l	4000 bits
Cluster head selection probability p	0.05
E_{elec}	$E_{elec} = 50nJ/bit$
ε_{fs}	$\varepsilon_{fs} = 10pJ/bit/m^2$
ε_{mp}	$\varepsilon_{mp} = 0.0013pJ/bit/m^4$
Population m, Maximum rounds T	$m = 15, T = 20$
（SA, Simulated Annealing algorithm） used in LEACH-C	In SA algorithm, the length of Markov chain Len=15, the initial temperature=100, the control temperature * 0.5 for each drop of temperature, and exit condition Temperature<=0.0001

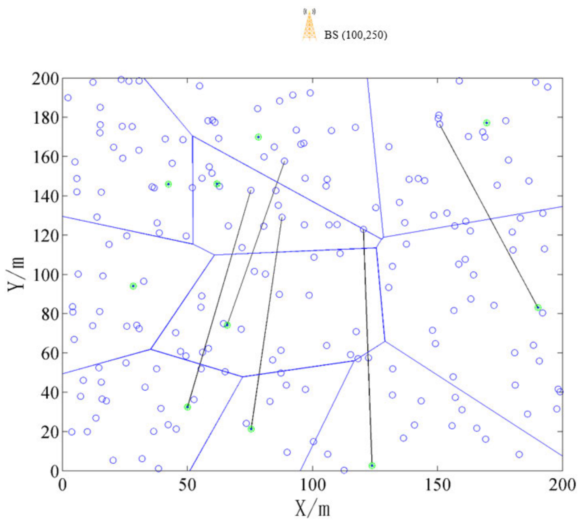


FIGURE 7. Schematic diagram of clustering and relay forwarding when SCA-Lévy algorithm runs to 300 rounds.

per round, lower than other algorithms in the comparison. In addition, the energy consumption of each round was very uniform and maintained at nearly a horizontal line before the first death node appearing, while the energy consumption of other algorithms fluctuated differently and was very uneven. The energy consumption of SCA-Lévy algorithm is balanced, which can ensure the minimum cost of each round of communication, greatly improve the energy utilization, put off the appearing time of the first dead node.

D. NETWORK LIFE CYCLE

Figure 12 is the comparison chart of cumulative dead nodes for. The first dead node of other algorithms appears relatively

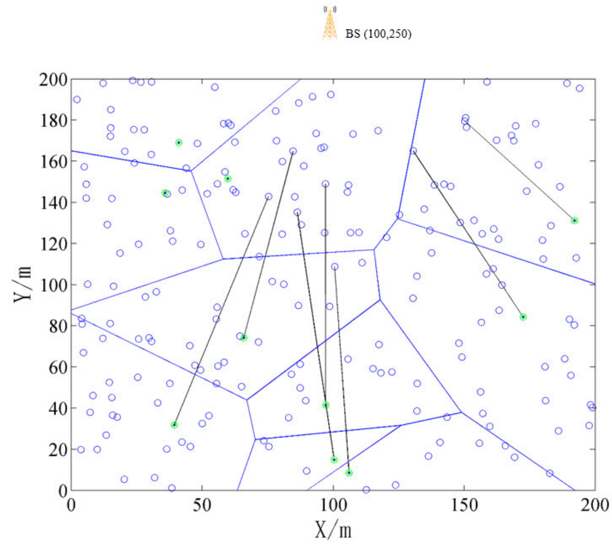


FIGURE 8. Schematic diagram of clustering and relay forwarding when SCA-Lévy algorithm runs to 400 rounds.

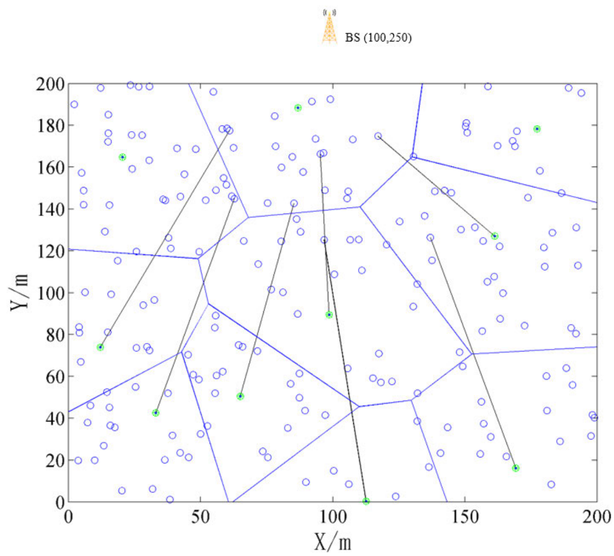


FIGURE 9. Schematic diagram of clustering and relay forwarding when SCA-Lévy algorithm runs to 500 rounds.

early, and the process from the first dead node to 80% node death lasts a long time. The first dead node of the SCA-Lévy algorithm appears the latest, and since the first dead node appears, other nodes also die quickly, and the broken line graph rises in a straight line, which indicating that the SCA-Lévy algorithm shows good performance in balancing node energy consumption and extending life-cycle.

Table 2 is a comparison table of the specific rounds of each algorithm when there are death nodes of different proportions. The specific rounds of the first death node of each algorithm can been seen clearly. LEACH is at 303 rounds, LEACH-C is at 398 rounds, CRISCA is at 487 rounds, LEACH-SCA is at 782 rounds, SCA-Lévy is at 925 rounds. SCA-Lévy has a good life cycle, and 20%, 50% and 80% of the dead nodes have more rounds than other

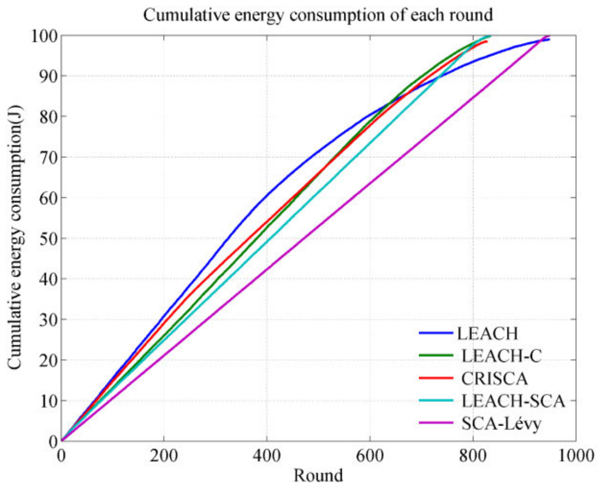


FIGURE 10. Comparison chart of cumulative energy consumption.

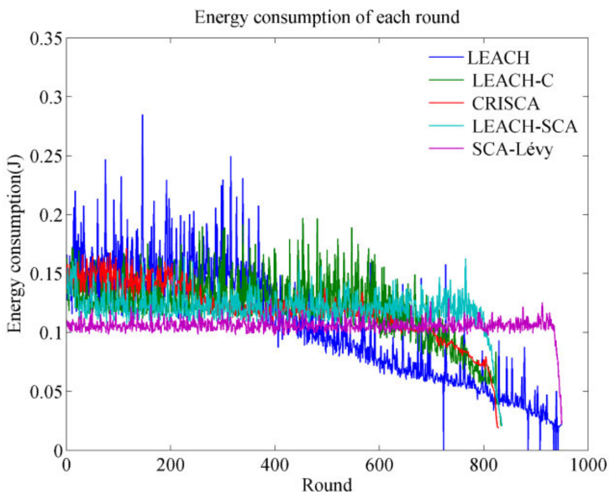


FIGURE 11. Comparison diagram of energy consumption of each round.

algorithms. It takes only 25 rounds from the first dead node to 80%, effectively avoiding network segmentation problem caused by premature death of individual nodes.

E. THROUGHPUT

Figure 13 and 14 is a statistical comparison chart. SCA-Lévy algorithm is relatively optimal. The largest number of packets are sent to BS. That's because SCA-Lévy algorithm uses a more reasonable clustering algorithm and sets relay nodes to forward data. The communication cost in each round of the network is the lowest. The energy consumption among nodes is relatively balanced. The longer the life-cycle, the more data it sends to BS. So SCA-Lévy algorithm has the strongest information collection capacity and throughput.

F. APPLICABILITY

We set different monitoring areas, total number of nodes and BS locations, and keep the initial carrying energy and average node density unchanged. Average node density refers to the

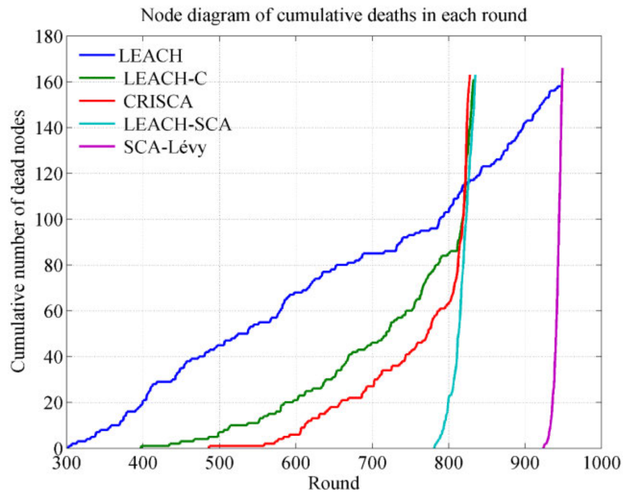


FIGURE 12. Comparison diagram of cumulative dead nodes.

TABLE 2. Comparison of appearing round of dead nodes.

	the round of dead nodes			
	1th	20%	50%	80%
LEACH	303	473	788	949
LEACH-C	398	669	819	833
CRISCA	487	744	819	828
LEACH-SCA	782	811	823	835
SCA- Lévy	925	940	945	949

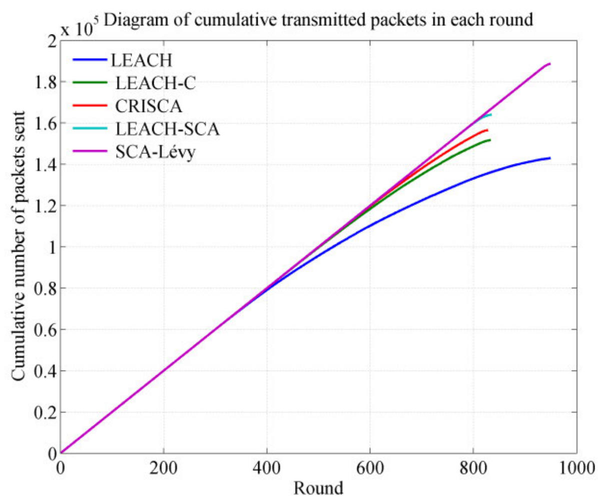


FIGURE 13. Diagram of cumulative transmitted packets in each round.

number of nodes per unit square area. Table 3 shows the changes of life cycle in the different algorithms. It can be seen that the SCA-Lévy algorithm shows applicability in different environments. When the area of the region is 100 × 100, the difference between the first dead node of each algorithm is not very obvious, LEACH is 882 rounds, and SCA-Lévy is 1025 rounds. When the area of the region becomes larger, the

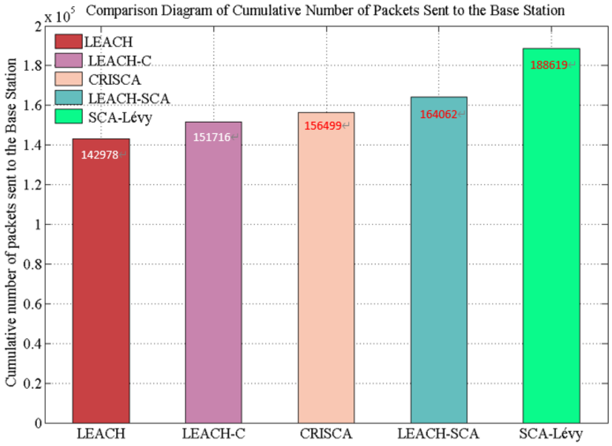


FIGURE 14. Statistical comparison of cumulative packets received by the BS.

TABLE 3. Comparison of life cycle in different scenarios.

Different Scenarios	100×100,50 (50,150)		200×200,200 (100,250)		300×300,450 (150,350)		400×400,800 (200,450)	
Algorithm/Round	1th	80%	1th	80%	1th	80%	1th	80%
LEACH	882	1175	303	949	81	806	23	690
LEACH-C	963	975	398	833	150	741	87	655
CRISCA	978	994	487	828	82	757	16	506
LEACH-SCA	999	1013	782	835	208	655	84	622
SCA-Lévy	1025	1036	925	949	669	710	459	531

advantages of SCA-Lévy algorithm become more obvious. In the area of 300 × 300, the first death node of LEACH appears in 81 rounds, and the death node appears prematurely, while the first death node of SCA-Lévy is in the 669 rounds. In other algorithms, the process from the first dead node to 80% of the dead nodes is relatively long, while in SCA-Lévy, 80% of the dead nodes are in 710 rounds, which only takes about 40 rounds. It is proved that its energy consumption is relatively uniform.

VII. CONCLUSION

For improving the energy utilization of WSN, a clustering routing algorithm combining sine cosine algorithm and Lévy mutation is raised. SCA algorithm with improved step size search factor is utilized to select cluster heads, and Lévy mutation is used for increasing global exploration ability and population diversity. The SCA-Lévy algorithm proposed is superior to other algorithms in terms of avoiding premature node death, improving network throughput, balancing energy consumption and prolonging network life cycle. However, the design and application of relay nodes have certain requirements for the area size, node number and BS location. When the area or node number continue to increase, network life cycle is limited. The next step will consider inter-cluster multi-hop and multi-hop of relay nodes in bigger network workspace.

REFERENCES

[1] S. Snowndeswari and E. Kavitha, "A comparative study on energy efficient clustering based on metaheuristic algorithms for WSN," *Int. J. Adv. Technol. Eng. Explor.*, vol. 9, no. 86, pp. 111–126, Jan. 2022.

[2] R. Sharma, V. Vashisht, and U. Singh, "Metaheuristics-based energy efficient clustering in WSNs: Challenges and research contributions," *IET Wireless Sensor Syst.*, vol. 10, no. 6, pp. 253–264, 2020.

[3] W. He, "Energy-saving algorithm and simulation of wireless sensor networks based on clustering routing protocol," *IEEE Access*, vol. 7, pp. 172505–172514, 2019.

[4] 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. 3005–3014.

[5] B. Wang and D. S. Fu, "Improved for LEACH routing protocol in wireless sensor network," *Instrum. Technique Sensor.*, no. 8, pp. 71–73, Aug. 2016.

[6] C.-F. Li, G.-H. Chen, M. Ye, and J. Wu, "An uneven cluster-based routing protocol for wireless sensor networks," *Chin. J. Comput.*, vol. 30, no. 1, pp. 27–36, Jan. 2007.

[7] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[8] L. L. Pang, J. Y. Xie, and Q. Q. Xu, "Neural network-based routing energy-saving algorithm for wireless sensor networks," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, Jul. 2022, doi: 10.1155/2022/3342031.

[9] C. H. Lin and F. Jiang, "Research of multidimensional optimization of LEACH protocol based on reducing network energy consumption," *J. Electr. Comput. Eng.*, vol. 2021, pp. 1–9, Feb. 2021, doi: 10.1155/2021/6658454.

[10] H. B. Li et al., "Clustering routing algorithm for wireless sensor network based on hierarchical neighboring nodes," *Comput. Eng.*, vol. 46, no. 6, pp. 187–195, Jun. 2020.

[11] J. Bhola, S. Soni, and G. K. Cheema, "Genetic algorithm based optimized leach protocol for energy efficient wireless sensor networks," *J. Ambient Intell. Hum. Comput.*, vol. 11, no. 3, pp. 1281–1288, Mar. 2020.

[12] K. L. Zhu and A. J. Sun, "WSN clustering routing algorithm based on Cuckoo search algorithm optimized K-means," *Chin. J. Internet Things.*, vol. 6, no. 1, pp. 73–81, Mar. 2022.

[13] A. J. Sun and K. L. Zhu, "WSN clustering routing algorithm based on improved S-FCM and PSO algorithm," *J. Xi'an Univ. Posts Telecommun.*, vol. 27, no. 2, pp. 7–16, Mar. 2022.

[14] H. Chen and Y. Gao, "Uneven clustering routing algorithm for WSNs based on double cluster heads," *Comput. Eng.*, vol. 48, no. 10, pp. 184–192, Oct. 2022.

[15] J. Pradeep and M. P. M. Kumar, "Distributed entropy energy-efficient clustering algorithm for heterogeneous wireless sensor network based chaotic firefly algorithm cluster head selection," *J. Crit. Rev.*, vol. 7, no. 8, pp. 1208–1215, 2020.

[16] W. Osamy, A. Salim, and A. M. Khedr, "An information entropy based-clustering algorithm for heterogeneous wireless sensor networks," *Wireless Netw.*, vol. 26, no. 3, pp. 1869–1886, Apr. 2020.

[17] N. Moussa and A. El Belrhiti El Alaoui, "An energy-efficient cluster-based routing protocol using unequal clustering and improved ACO techniques for WSNs," *Peer Peer Netw. Appl.*, vol. 14, no. 3, pp. 1334–1347, May 2021.

[18] C. Nakas, D. Kandris, and G. Visvardis, "Energy efficient routing in wireless sensor networks: A comprehensive survey," *Algorithms*, vol. 13, no. 3, pp. 1–65, Mar. 2020.

[19] J. X. Miao et al., "Non-uniform clustering routing algorithm for balanced energy consumption in WSN," *Comput. Eng. Des.*, vol. 43, no. 2, pp. 301–307, Feb. 2022.

[20] X. T. Liu, Z. P. Chen, and Y. R. Huang, "An uneven clustering routing algorithm based on energy consumption balance," *Microelectron. Comput.*, vol. 36, no. 2, pp. 36–40, Feb. 2019.

[21] A. Rezaeipanah, P. Amiri, H. Nazari, M. Mojarad, and H. Parvin, "An energy-aware hybrid approach for wireless sensor networks using re-clustering-based multi-hop routing," *Wireless Pers. Commun.*, vol. 120, no. 4, pp. 3293–3314, Oct. 2021.

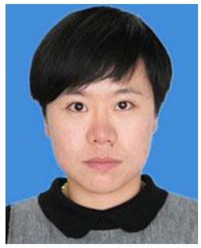
[22] A. M. Khedr, A. Aziz, and W. Osamy, "Successors of PEGASIS protocol: A comprehensive survey," *Comput. Sci. Rev.*, vol. 39, pp. 1–24, Feb. 2021.

[23] W. P. An and Y. F. Shao, "Wireless sensor network routing protocol based on improved gray wolf optimization algorithm," *Chin. J. Sens. Actuators.*, vol. 35, no. 5, pp. 676–682, May 2022.

[24] J. Wang, J. T. Liu, and F. Q. Zhang, "Heterogeneous clustering routing protocol based on energy and trust," *Comput. Eng. Des.*, vol. 42, no. 4, pp. 927–933, Apr. 2021.
[25] X. L. Tang, M. Zhang, P. P. Yu, W. Liu, N. Cao, and Y. F. Xu, "A nonuniform clustering routing algorithm based on an improved K-means algorithm," *Comput., Mater. Continua*, vol. 64, no. 3, pp. 1725–1739, Apr. 2020.
[26] J. Zhang and J. S. Wang, "Improved salp swarm algorithm based on Lévy flight and sine cosine operator," *IEEE Access*, vol. 8, pp. 99740–99771, 2020.
[27] H. J. Gao, Y. T. Li, P. Kabalyants, H. Xu, and R. Martinez-Bejar, "A novel hybrid PSO-K-means clustering algorithm using Gaussian estimation of distribution method and Lévy flight," *IEEE Access*, vol. 8, pp. 122848–122863, 2020.
[28] X. Q. Zhao et al., "Heterogeneous wireless sensor network routing protocol based on simulated annealing algorithm and modified grey wolf optimizer," *Chin. J. Internet Things.*, vol. 5, no. 2, pp. 97–106, Jun. 2021.
[29] J. Y. Dai, X. H. Deng, B. Wang, and H. H. Wang, "Clustering routing protocol for WSNs based on neural network optimization by improved firefly algorithm," *J. Beijing Univ. Posts Telecommun.*, vol. 43, no. 3, pp. 131–137, Jun. 2020.
[30] M. Kaedi, A. Bohlooli, and R. Pakrooh, "Simultaneous optimization of cluster head selection and inter-cluster routing in wireless sensor networks using a 2-level genetic algorithm," *Appl. Soft Comput.*, vol. 128, Oct. 2022, Art. no. 109444, doi: 10.1016/j.asoc.2022.109444.
[31] F. Z. Dong et al., "WSN clustering routing algorithm based on genetic algorithm and fuzzy C-means clustering," *J. Comput. Appl.*, vol. 39, no. 8, pp. 2359–2365, Aug. 2019.
[32] P. S. Mehra, M. N. Doja, and B. Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN," *J. King Saud Univ.-Sci.*, vol. 32, no. 1, pp. 390–401, 2020.
[33] F. Zhao, N. Gao, and K. Y. Zhang, "WSNs clustering routing protocol based on whale optimization algorithm and beetle antennae search," *Transducer Microsyst. Technol.*, vol. 41, no. 9, pp. 42–45, 2022.
[34] M. L. Yang and C. Lu, "Routing Optimization algorithm based on entropy weight adaptive clustering and improved PSO in WSN," *Comput. Appl. Softw.*, vol. 39, no. 7, pp. 109–116, Jul. 2022.
[35] Q. C. Sun, Z. L. Zhu, and F. Q. Zhang, "Optimizing WSN clustering protocol based on improved SSA," *Radio Commun. Technol.*, vol. 48, no. 1, pp. 132–139, Jun. 2022.
[36] A. J. Sun and S. P. Zheng, "WSN clustering algorithm based on chaos optimized firefly algorithm," *Chin. J. Sens. Actuators.*, vol. 34, no. 9, pp. 1224–1230, Sep. 2021.
[37] Y. F. Chang, B. J. Song, X. P. Chen, and Y.-X. Zhu, "WSN clustering algorithm based on improved artificial fish swarm algorithm and fuzzy c-means," *J. Inf. Eng. Univ.*, vol. 23, no. 1, pp. 18–23, Feb. 2022.
[38] X. He, Z. Z. Ning, and X. Y. Yang, "WSN clustering routing protocol based on improved sine cosine algorithm," *J. Xi'an Univ. Posts Telecommun.*, vol. 26, no. 2, pp. 15–20, Mar. 2021.
[39] L. Q. Yong, Y. H. Li, and W. Jia, "Literature survey on research and application of sine cosine algorithm," *Comput. Eng. Appl.*, vol. 56, no. 14, pp. 26–34, Jun. 2020.
[40] Y. T. Li, T. Han, H. Zhao, and X. F. Wang, "An improved sine cosine optimization algorithm with self-learning strategy and Levy flight," *J. Chongqing Univ.*, vol. 42, no. 9, pp. 56–66, Sep. 2019.
[41] C. A. Hu and J. Ye, "Clustering routing algorithm for wireless sensor networks with whale algorithm," *Comput. Eng. Des.*, vol. 40, no. 11, pp. 3067–3072, Nov. 2019.
[42] Z. S. Wang et al., "Clustering routing protocol based on artificial bee colony algorithm in WSN," *Microelectron. Comput.*, vol. 38, no. 4, pp. 74–80, Apr. 2021.

XIAOLING GUO received the master's degree in computer application from Northeastern University, in 2011. Since 2011, she has been with the College of Information Science and Engineering, Hebei North University. She is committed to studying computer networks, sensor networks, and the Internet of Things.

YONGFEI YE received the M.S. degree in computer software and theory from Yanshan University, in 2008. She is currently an Associate Professor with the School of Information Science and Engineering, Hebei North University, Zhangjiakou, China. Her research interests include information security, big data analysis, and precision agriculture.



LING LI received the M.S. degree in control science and engineering from Hebei University of Technology, in 2012. She is currently a Lecturer with Hebei North University, Zhangjiakou, China. Her current research interests include WSNs and wireless network security.

RENJIE WU received the master's degree in engineering from Beijing University of Technology, in 2009. He is currently an Associate Professor with the School of Information Science and Engineering, Hebei North University. He has been studying computer networks and related security issues. His research interests include computing networks, cyberspace security, and the Internet of Things application development.

XINGHUA SUN received the M.S. degree from the School of Mathematical Information, Shanghai Normal University, in 2007. He is currently a Professor with the School of Information Science and Engineering, Hebei North University, Zhangjiakou, China. His current research interests include mobile communication and lighting control networks.