

Dynamic Cluster Routing Algorithm for WSN Suppression of Energy Holes based on Diffusion Theory of Multi-component Liquid Mixtures

Jian-Po Li, Min He
School of Computer Science
Northeast Electric Power University, Jilin 132012, China
jianpoli@163.com, 2302089043@qq.com

Bin-Feng Jiang
Big Data Business School
Jilin Railway Technology College, Jilin 132299, China
53436707@qq.com

Ting-Wen Yu
Guangdong Provincial Key Laboratory of Digital Grid Technology
Digital Grid Research Institute China Southern Power Grid, Guangzhou 510000, China
ytw2966@163.com

En-Yuan Zhang
State Grid Gansu Electric Power Company
Longnan Power Supply Company, Longnan 742500, China
526144234@qq.com

Yue-Hua Yang*
School of Computer Science
Northeast Electric Power University, Jilin 132012, China
yyh0504@126.com

*Corresponding author: Yue-Hua Yang
Received May 13, 2024, revised October 7, 2024, accepted January 15, 2025.

ABSTRACT. *The formation and evolution mechanism of energy holes in wireless sensor networks are discussed in depth, and the characteristics of energy holes are analyzed in detail. On this basis, a series of definitions are proposed to describe energy holes and their related states, including fusion energy hole, energy hole, independent energy hole, hole edge domain, multi-energy hole distance, and death node discretization. By comprehensively considering the key parameters such as node density of surviving nodes in the edge region of the energy hole, residual energy, data transmission, and distance between multiple energy holes, the concept of energy hole diffusion coefficient is obtained, and accordingly, the angle and velocity of the energy hole in the direction of diffusion are calculated, to construct a dynamic evolution model of the energy hole. Based on this model, a dynamic clustering routing algorithm (DCRA) is designed, which selects cluster heads by comprehensively considering the energy hole diffusion coefficient, the node degree change rate, and the distance between the node and the edge of the hole and dynamically adjusts the cluster radius of the nodes in the region of the edge of the hole, to effectively minimize the impact of the hole. The algorithm also employs the transmission strategies of intra-cluster node transmission priority and inter-cluster forwarding cost, as well as cluster head reselection, alternate dormancy of low-energy nodes, and dormancy scheduling strategies for nodes within the edge zone of the hole. Simulation results verify the significant advantages of the algorithm in terms of network lifetime and energy equalization compared to algorithms such as LEACH, UCDS, and EHSRA.*

Keywords: Wireless sensor networks, Energy hole dynamic evolution model, Dynamic clustering routing, Diffusion of multi-component liquid mixtures

1. **Introduction.** Wireless Sensor Networks (WSN), as an important part of the field of the Internet of Things, have received extensive attention in scientific research and practical applications in recent years. With the rapid development of science and technology, the application scenarios of WSN are constantly broadening, from environmental monitoring to military reconnaissance, from intelligent transportation to health monitoring, and its role is becoming more and more prominent [1]. WSN are distributed networks consisting of a large number of low-cost, miniaturized sensor nodes, which are equipped with the ability of environment sensing, data acquisition, and wireless transmission. They build the network through self-organization to achieve monitoring of the target area and real-time data collection. The emergence of WSN has greatly improved our ability to obtain environmental information and provided strong support for the development of many fields [2].

However, WSN also face many technical challenges, among which the energy hole problem is particularly prominent. During multi-hop transmission, each node performs different functions. Some nodes not only need to process their data but also need to forward data generated by other nodes. Nodes close to the aggregation nodes undertake heavier forwarding tasks and hence consume more energy. As some of the nodes are depleted prematurely due to over-consumption of energy, it leads to the emergence of one or more regions where the energy is completely depleted and these regions are defined as energy holes [3]. Especially during data transmission, nodes close to the data aggregation point are responsible for forwarding a large amount of data, their energy consumption rate is much faster than other nodes, which in turn is more likely to generate energy holes. The existence of energy holes not only affects the data transmission efficiency and reduces the effective coverage area of the network, but also leads to partial or total loss of network functionality.

When energy holes appear, most of the surrounding surviving nodes have relatively high residual energy, how to fully utilize these nodes to delay the generation of energy

holes as much as possible and reduce their adverse effects deserves in-depth study. Since sensor nodes are usually deployed in unattended areas with limited energy supply, how to efficiently utilize energy and extend the network lifecycle becomes a key issue for research [4]. Therefore, when designing and managing wireless sensor networks, it is necessary to consider how to appropriately allocate nodes, optimize energy management strategies, and implement routing algorithms to minimize the impact of energy holes, extend the network life cycle, and improve performance. Although researchers have proposed how to effectively inhibit and control the formation and spread of energy holes, eliminating this problem is still a challenge due to the uncertainty and complexity of the network [5].

1.1. Realted work. In recent years, some progress has been made in the research on the energy hole problem of WSN, and some suppression methods and strategies have been proposed from the aspects of energy hole characteristics and evolution laws, energy-saving routing, transmission range adjustment, and node scheduling optimization.

At present, the characterization of energy holes is mainly based on the two dimensions of time and space, and qualitatively described in terms of its formation causes, definition, and characteristics in different periods. References [6] random movement of nodes and long time dormancy leads to frequent changes in the network topology and nodes need to consume a lot of energy and time to discover all their neighbors leading to the death of nodes and energy holes. In the GS-MAC protocol [7] it is found that nodes do not have a strict schedule to maintain data transmission as well as node addresses are too long leading to energy wastage. In some networks, repetitive route searches and routing loops also speed up node energy consumption and form energy holes when location and topology information is missing [8]. When a node in a network is the only node connecting two parts of the network, then when the node dies, a network partition or energy hole occurs [9]. CPPE algorithm [10] introduce chaotic mapping into the PPE algorithm, the chaotic map replaces the initialization population of the original PPE algorithm to enhance performance and convergence. TOA algorithm [11] is the newest metaheuristic optimization algorithm that mimics the growth and reproduction of tumbleweeds, allowing the algorithm to jump out of the local optimum, maintain population diversity, and improve global search ability.

For energy efficient routing, Cluster head selection is multifunctionally derived via the Adaptive Remora Optimization Algorithm [12], and the multifunction formulation takes into account known factors such as energy consumption, distance, throughput, packet delivery rate, and path loss. Butterfly Optimization Algorithm [13] is used to select the optimum number of cluster heads from the dense nodes. The parameters to be considered for selecting the cluster heads are the residual power of the node, distance from other nodes in the network, distance from the base station, node centrality, and modality. Particle Swarm Optimization is used for the formation of the cluster heads by selecting certain parameters. In the Dynamic Cluster Head Optimal Path for Energy Hole Suppression Algorithm [14], an algorithm based on a minimum lifetime of cluster head with traffic threshold is proposed using different cluster radius and dynamic cluster head movement design, which aims to equalize the cluster head load and suppresses energy hole attacks. The RBM protocol [15] divides the network into two regions: region 1 near the BS communicates directly, while Region 2 away from the BS has routing nodes to communicate with the BS, the routing nodes are not involved in the sensing function but will only move in region 2 to collect data and forward it to the BS.

By modeling the working state of networked sensors [16], energy parameters are added as influencing factors to the actual prediction calculations, and the clustering mechanism and particle swarm optimization algorithm are used to improve the path selection and

increase the fault-tolerance of wireless network transmission in the face of large amounts of dynamic data. References [17] derived the optimal transmission distance between any two nodes through a mathematical analytical model with the design criterion of selecting the farthest forwarding node within the communication range and the direction to the homestay, thus forming a multichain structure by precise distance, direction, and multiple paths. Combinatorial data prediction models [18] are used to construct prior data to control delays and predict future data, thus reducing unnecessary data transfers. Bidirectional spacing Figure [19] for connectivity maintains strong connectivity while keeping the minimum number of links required and ensures that each node is assigned an appropriate power spacing to maintain network communication.

GS-MAC protocol [7] to minimize power consumption, nodes sleep periodically, but GS-MAC avoids periodic node synchronization, nodes use UTC to maintain a strict schedule to avoid energy wastage and maintain a constant low-duty cycle even if the node density is increasing, and a competitive approach is used on reserved time slots for scheduling between communication rounds to maintain scalability. Reference [20] divides intermediate nodes into two types: those responsible only for collecting data and those responsible only for sending data, so that redundant nodes can enter sleep mode directly after completion. The CSSP protocol [21] uses the residual energy of the nodes, distance to neighbors, and coverage area parameters to select the working and sleeping nodes in the network to minimize energy consumption.

1.2. Motivation and contribution. In the research related to WSN routing algorithms, the analysis of energy hole characteristics is often one-sided and lacks an in-depth study of the overall evolution process of the network. Aiming at the lack of sufficient analysis of the evolutionary trend of energy holes in current wireless sensor networks, we discuss and propose a dynamic cluster routing algorithm for suppressing energy holes in wireless sensor networks.

The main innovations and contributions of this work include:

(1) Characterization of energy holes: to make up for the inadequacy of the existing methods in comprehensively analyzing energy holes, a series of basic definitions are proposed to characterize energy holes in the diffusion process, and the influence mechanisms of different parameters on the diffusion of energy hole is revealed. The angles and velocities of energy holes in the diffusion direction are calculated, and an energy hole dynamic evolution model based on diffusion of multi-component liquid mixtures.

(2) WSN clustering method based on energy hole diffusion coefficient: the cluster head is selected by comprehensively considering the energy hole diffusion coefficient, the node degree change rate, the distance of nodes from the edge of the hole, and the cluster radius of nodes in the edge region of the hole is dynamically adjusted to effectively reduce the impact of the hole.

(3) Optimization method of WSN data transmission for suppressing energy hole: intra-cluster node transmission ordering strategy based on the rate of change of node degree and inter-cluster transmission strategy based on relay node forwarding cost are used.

(4) Node dormancy scheduling strategy to suppress energy holes: when a node is successfully elected as a cluster head, the node will be prohibited from being elected as a cluster head again in the next three rounds of cluster formation; the node's residual energy drops below a preset threshold and adopts alternate dormancy; dormancy scheduling strategy for nodes within the edge zone of the hole.

2. Modeling of the dynamic evolution of energy holes.

2.1. Characterization of energy holes. Through the careful study of the increase and decrease in the number of changes in the position of death nodes, the generation of energy holes, and the diffusion mechanism, to understand their inherent laws, a series of basic definitions describing the states related to energy holes are proposed, and the death node discretization is introduced to measure the degree of discretization of the death nodes. The specific definitions are as follows:

Definition 1: A region in the network is a quasi-energy hole when there is only one dead node in the region and no other dead node within two hops of its neighborhood. A quasi-energy hole indicates a higher probability of energy hole formation within its neighborhood.

Figure 1(a) shows the distribution of nodes in a certain area of the network, where a potential energy hole area, namely a quasi-energy hole, is defined within the one-hop range of node D_1 .

Definition 2: If two or more dead nodes are present in a region and the minimum number of hops between these dead nodes is less than or equal to 2, and also the distance between them is less than or equal to the communication radius of the node, then these dead nodes, along with all the surviving nodes on the transmission paths between them, together constitute an energy hole.

Figure 1(b) shows the distribution of dead nodes in a region of the network. The number of hops between dead nodes D_1 and D_2 is 2, then D_1 and D_2 and the surviving node N_1 on the transmission path between them form an energy hole $H_1 = \{D_2, N_1, D_1\}$. dead nodes D_3 and D_4 are adjacent to each other and the number of hops between the nodes is 1, forming an energy hole $H_2 = \{D_3, D_4\}$.

Definition 3: An independent energy hole is an energy hole that does not have any intersection between that hole and other energy holes in the network.

As shown in Figure 1(b), both energy holes $H_1 = \{D_2, N_1, D_1\}$ and $H_2 = \{D_3, D_4\}$ are independent energy holes.

Definition 4: When two or more energy holes intersect, the original independent energy holes fuse together to form a fusion energy hole.

As shown in Figure 1(c), based on Figure 1(b), when a death node occurs between the original energy holes $H_1 = \{D_2, N_1, D_1\}$ and $H_2 = \{D_3, D_4\}$, a new fused energy hole $H_3 = \{D_2, N_1, D_1, D_5, D_6, N_6, D_3, D_4\}$ is formed.

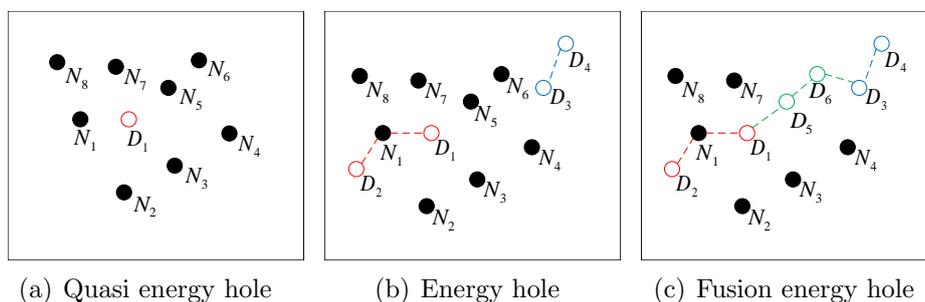


FIGURE 1. Diagram of energy hole changes

Definition 5: Within an energy hole of the network, if there are surviving nodes within one hop of the dead nodes, these dead nodes are called void edge dead nodes and these surviving nodes are called void edge surviving nodes. So nodes within one hop of all hole edge dead nodes in the same energy hole form the void edge dead domain and so nodes within one hop of all hole edge surviving nodes form the hole edge surviving domain.

As shown in Figure 2, in energy hole $H_1 = \{D_2, N_1, D_1\}$, there are surviving nodes $N_1, N_3, N_5,$ and N_7 within one hop of D_1 , and there is a surviving node N_1 within one hop of D_2 , then the dead nodes at the edge of the hole are D_2 and D_1 , and the surviving nodes at the edge of the hole are N_1, N_3, N_5 and N_7 . In energy hole $H_2 = \{D_3, D_4\}$, there are surviving nodes N_4 and N_6 within one hop of D_3 and no surviving nodes within one hop of D_4 , the dead nodes at the edge of the hole are D_3 and the surviving nodes at the edge of the hole are N_4 and N_6 .

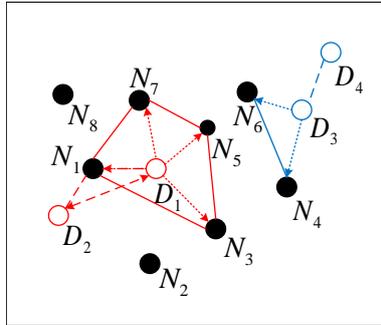


FIGURE 2. Diagram of the energy hole edge domain

Definition 6: Multiple energy hole distance is the shortest distance, denoted by d_H , between the edge dead nodes located in two energy holes, and the maximum distance between two energy holes, denoted by d_{H-max} .

As shown in Figure 3, d_H is the distance between the hole edge death node D_1 of energy hole H_1 to the hole edge death node D_3 of energy hole H_2 .

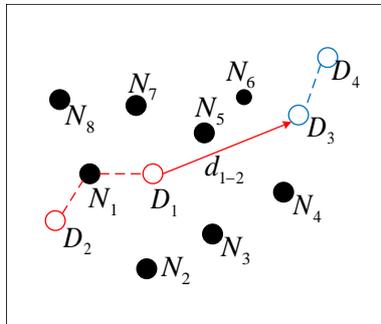


FIGURE 3. Diagram of the distance between multi energy hole

Definition 7: Death node discretization is used to measure the degree of death node discretization, defined by the minimum distance between death nodes, the average distance between death nodes, and the number of death points, denoted by f_{Dnd} , as shown in Equation (1):

$$f_{Dnd} = 0.2\sqrt{d_{death-min}} + 0.1 \log_5 d_{death-avg} + \frac{0.7}{N_{death}} \tag{1}$$

where, $d_{death-min}$ is the minimum distance between dead nodes; $d_{death-avg}$ is the average distance between dead nodes; and N_{death} is the number of dead nodes.

The number of dead nodes is the core element that determines the degree of its distribution, while the size of the minimum and average distance between dead nodes affects the degree of its distribution to different degrees. To effectively analyze the distribution characteristics of the dead nodes, a differentiated weight allocation strategy is adopted, whereby the number of dead nodes is given the maximum weight of 0.7 to highlight its

criticality, while the other two factors are given a weight of 0.1 and 0.2 respectively to reflect their relatively minor influence. To better illustrate this concept, it is useful to give a few specific examples.

If the minimum and average distances between death nodes are large and the number of death nodes is small, it indicates a large death node discretization. As shown in Figure 4(a), $d_{death-min-1} = d_{11}$, in Figure 4(b) $d_{death-min-2} = d_{21}$, Since, $d_{death-min-1} < d_{death-min-2}$, $d_{death-avg-1} < d_{death-avg-2}$, and $N_{death-1} > N_{death-2}$. Therefore, the death node discretization of Figure 4(b) is larger than that of Figure 4(a).

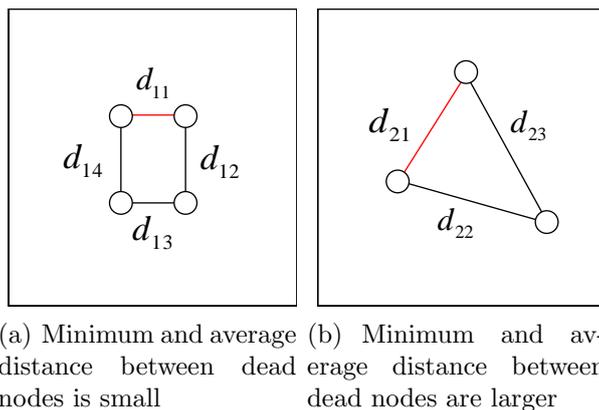


FIGURE 4. Plot of the effect of different values of minimum and average distances on death node discretization if the number of dead nodes is small

If the minimum distance between the dead nodes is small, the average distance is large, and the number of dead nodes is large, it indicates that the death node discretization is small. As shown in Figure 5(a), $d_{death-min-3} = d_{32}$, in Figure 5(b) $d_{death-min-4} = d_{41} = d_{42}$, Since, $N_{death-3} < N_{death-4}$, $d_{death-min-3} > d_{death-min-4}$ and $d_{death-avg-3} < d_{death-avg-4}$. Therefore, the death node discretization of Figure 5(b) is less than that of Figure 5(a).

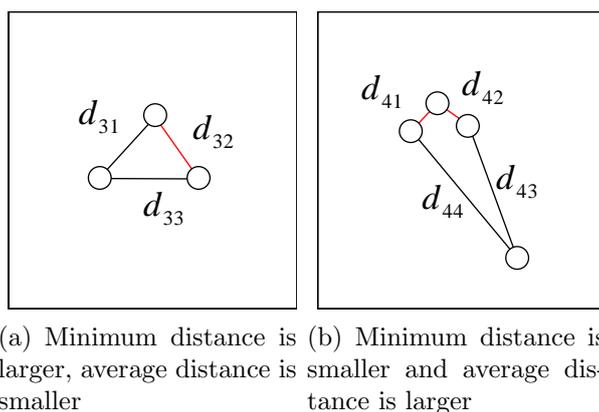


FIGURE 5. Plot of the effect of different values of minimum distance and average distance on death node discretization if the number of dead nodes is high

The minimum distance between dead nodes and the number of dead nodes did not differ significantly, and if the average distance between dead nodes is larger, it indicates that these nodes are more discrete. In Figure 6(a), $d_{death-min-5} = d_{52}$ and in Figure

6(b), $d_{death-min-6} = d_{61} = d_{62}$. Since, The minimum distance between dead nodes and the number of dead nodes did not differ significantly, and $d_{death-avg-5} < d_{death-avg-6}$, therefore, the death node discretization of Figure 6(b) is greater than that of Figure 6(a).

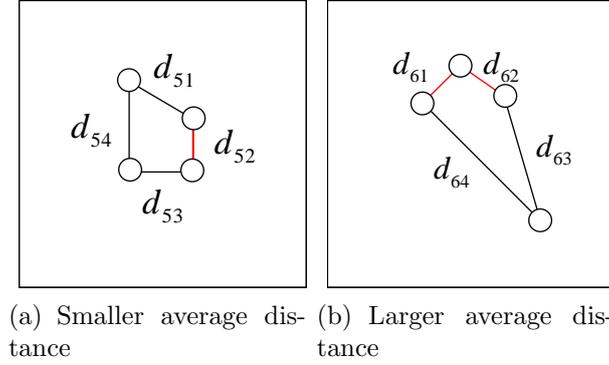


FIGURE 6. Plot of the effect of average distance value on death node discretization when there is no significant difference between the minimum distance between death nodes and the number of death nodes

2.2. Diffusion coefficient of energy holes. To gain a deeper understanding of the dynamic evolution of the energy hole, it is possible to analyze it by analogy with the diffusion phenomenon of multicomponent liquid mixtures [22]. In the construction of the simulation environment, the following assumptions are set:

(1) Once the first dead node appears in the network, this can be analogized as a drop of a white diffusing substance into the liquid.

(2) Although the death of a node in the network is a discrete event and diffusion in the liquid is a continuous process, a random uniform distribution of multiple nodes in the wireless sensor network space can be viewed as a multi-component liquid mixture model. In this model, the nodes in the network maintain their unique properties and functions.

(3) The dead nodes in the network are regarded as diffusing substances in the multi-component liquid mixture, while the surviving nodes in the network are equivalent to the diffusing medium, and together they constitute the dynamic evolution of the whole network system.

The dynamic evolution process of energy hole fusion is analyzed analogously by dropping some diffusion medium into different concentrations of diffusing substances. The focus will be on analyzing the effects of node density, residual energy, energy consumption in the previous round, amount of transmitted data, and distance between multiple energy holes on energy hole diffusion of surviving nodes within the survival zone at the edge of the hole. The energy hole diffusion coefficient is calculated based on the above key parameters as shown in Equation (2) to quantitatively assess the diffusion trend of energy holes in the network.

$$K = \ln \left(\frac{N_{alive}}{N_{S_H}} \right) \cdot \left(\frac{E_{max}}{E_{S_H-res}} \cdot \frac{E_{S_H-con}}{E_{con-avg}} \right) \cdot \left| \ln \frac{R_{datas-avg}}{R_{S_H-datas}} \right| \cdot \frac{d_{H-max}}{d_H} \quad (2)$$

where, N_{alive} is the number of nodes surviving; N_{S_H} is the number of neighboring nodes in the hole edge surviving zone; E_{max} is the initial energy of the nodes; E_{S_H-res} is the remaining energy of the nodes in the hole edge surviving zone; E_{S_H-con} is the energy consumption of the hole edge surviving zone in the previous round; $E_{con-avg}$ is the average energy consumption of all the nodes in the previous round; $R_{datas-avg}$ is the average

amount of data transmitted by the network; $R_{S_H-datas}$ is the transmitted data in the hole edge surviving zone in the previous round volume.

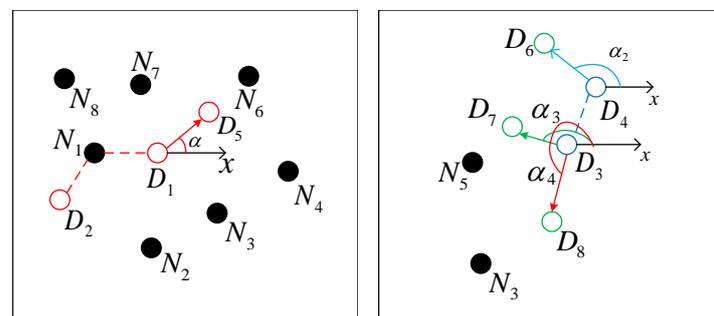
In the energy hole diffusion coefficient calculation formula, when the number of surviving nodes at the edge of the hole is more, a smaller cluster radius is used to form the cluster, at this time, the number of nodes in the cluster is less, the communication distance is shorter, the energy hole needs to cross more nodes to expand to other regions, and the diffusion rate of the energy hole is slowed down. The diffusion coefficient of a single energy hole can be adaptively adjusted according to the energy situation of the surviving nodes at the edge of the hole when the remaining energy in the surviving nodes at the edge of the hole is low and the rate of energy consumption is high, they will lose their function earlier due to the rapid depletion of the energy, and the diffusion speed of the energy hole is accelerated. Void edge surviving nodes with a high amount of transmitted data will likely fail earlier and the spread of energy holes is accelerated. Null hole edge survivor nodes are closer to multiple failure regions and these edge survivor nodes are responsible for collecting regional data as well as maintaining communication connectivity across these energy hole regions, resulting in faster diffusion as edge nodes consume more energy in maintaining communication.

2.3. Direction and velocity of energy hole diffusion. As shown in Figure 7(a), after completing one round of network data transmission, the dead node D_5 is generated in the next hop range of the dead node, the coordinate positions of the two nodes are known, and the direction of diffusion of the energy hole is shown in Equation (3):

$$\theta_1 = \alpha_1 \tag{3}$$

As shown in Figure 7(b), the last round of dead nodes D_3 and D_4 generate new dead nodes D_6 , D_7 , and D_8 in the new round. α_2 is the angle of direction from D_4 to the dead node D_6 in its one-hop range, and α_3 and α_4 are the angle of direction from D_3 to the dead nodes D_7 and D_8 in one of its ranges, and the direction of diffusion of the energy hole is calculated as shown in Equation (4):

$$\theta_2 = \frac{1}{2} \cdot \left(\frac{2}{3} \cdot \frac{\alpha_3 + \alpha_4}{2} + \frac{1}{3} \cdot \alpha_2 \right) \tag{4}$$



(a) Energy hole generates a death node (b) Energy hole generate multiple death nodes

FIGURE 7. Diagram of hole diffusion

From this derivation, it is obtained that for the occurrence of multiple dead nodes in the dead zone at the edge of the hole, the angle θ_H of the diffusion direction of the energy hole is calculated according to Equation (5).

$$\theta_H = \frac{1}{g \cdot v} \left(\sum_{i=1}^{u_1} \alpha_i + \sum_{j=1}^{u_2} \alpha_j + \dots + \sum_{z=1}^{u_g} \alpha_z \right) \quad (5)$$

where, g is the number of dead nodes at the edge of the previous round, v is the number of new dead nodes in the new round. u_1 is the number of new dead nodes within one hop of the first dead node of the previous round in the new round; α_i is the angle of the direction of the first dead node of the previous round to the i -th new dead node in its range of one hop in the new round; u_2 is the number of new dead nodes within one hop of the second dead node of the previous round in the new round; α_j is the angle from the second dead node of the previous round to the j th new dead node direction in its range of one hop in the new round; u_g is the number of new dead nodes within one hop of the second dead node of the previous round in the new round; α_z is the angle from the second dead node of the previous round to the j th new dead node direction in its range of one hop in the new round;

According to the liquid diffusion theory, the diffusion velocity of an energy hole is shown in Equation (6).

$$v_H = \frac{d_{\Delta t}}{\Delta t} \cdot K \quad (6)$$

where, $d_{\Delta t}$ is the farthest distance from the location of the center of the energy hole in the previous round to the newly generated death node. Δt is the network round data transmission time.

By constructing the dynamic evolution model of energy holes, it can comprehensively present the real-time state of nodes in the whole network, the distribution of energy consumption, the trend of hole changes, and other key information, to more accurately formulate energy hole suppression strategies.

3. WSN clustering method based on energy hole diffusion coefficient.

3.1. Cluster head selection method. By comparing the change of neighboring nodes of the node to be selected as cluster head nodes in two consecutive rounds, we calculate the rate of change of the node degree as a way to determine whether new dead nodes have been added around it. This process is carried out based on Equation (7) and aims to determine whether the node is located in the edge region of the cavity. If it is indeed located at the edge of the cavity, the probability of the node becoming a cluster head is reduced accordingly to optimize the network layout and performance.

$$rate = \frac{N_R}{N_{R-1}} \quad (7)$$

where, N_R is the number of neighbor nodes at round R ; N_{R-1} is the number of neighbor nodes at round $R - 1$.

Secondly, the energy hole diffusion coefficient of the nodes located within the survival zone at the edge of the hole is also taken into account, including node density, residual energy, energy consumption in the previous round, amount of data transmitted, and distance between multiple energy holes. If the diffusion coefficient of the node to be selected as cluster head is large, the probability of selection as cluster head is reduced. The cluster head selection probability is denoted as:

$$P_H = \max \left(\frac{\eta_1 \cdot rate}{K} + (1 - \eta_1) \cdot \frac{d_{toBS-MAX} - d_{toBS}}{d_{toBS-MAX} - d_{toBS-MIN}}, p_{min} \right) \quad (8)$$

where, η_1 is the weight; $d_{toBS-MAX}$ is the maximum value of the distance from all nodes to the base station; $d_{toBS-MIN}$ is the minimum value of the distance from all nodes to the base station, d_{toBS} is the node-to-base station distance; and p_{min} is the minimum election convergence probability.

3.2. Cluster radius calculation method. The rate of change of node degree is first utilized as an important indicator to determine whether a new dead node has been added around the cluster head node and whether the node is located in the edge domain of the cavity. By analyzing the changes of neighbor nodes in the two rounds before and after the node, the dynamic changes of the network topology are predicted and responded to. If the cluster head node is located at the edge of the cavity, it will reduce its cluster radius accordingly, secondly, the effect of the distance from the node to the base station on the cluster radius is also considered, and when the node is closer to the base station, its cluster radius will be reduced accordingly. The cluster radius is denoted as:

$$R_H = \left(1 + \frac{d_{CHtoBS} - d_{CHtoBS-avg}}{\eta_2 \cdot (d_{CHtoBS-MAX} - d_{CHtoBS-MIN})} \right) \cdot \left(1 - \frac{1}{\eta_3 \cdot rate} \right) \cdot R_L \quad (9)$$

where, d_{CHtoBS} is the cluster head to base station distance, $d_{CHtoBS-avg}$ is the average of all cluster head to base station distances; $d_{CHtoBS-MAX}$ is the maximum of all cluster head to base station distances; $d_{CHtoBS-MIN}$ is the minimum of all cluster head to base station distances; η_2 and η_3 are the weight; $N_{CH-neighbor}$ is the number of neighboring nodes of the cluster head; and R_L is the radius of the cluster when uniformly clustered.

4. Optimization method for WSN data transmission with energy hole suppression.

4.1. Transmission ordering strategy for intra-cluster nodes based on the rate of change of node degree. After the network is clustered, it will enter a stable data transmission phase, when multiple nodes may simultaneously request to receive data from the same cluster head node. However, at the same moment, a cluster head node can only receive data from one node, which may lead to an inter-cluster node access conflict problem. While assigning transmission order to intra-cluster nodes by the cluster head node, the possible selection of nodes with less residual energy or more neighboring nodes and whether the node is located in the edge domain of the hole or not have been taken into account in the past.

To solve these problems, the concept of intra-cluster node transmission priority is proposed in the intra-cluster data transmission phase, where the cluster head node calculates the data transmission priority of the intra-cluster nodes based on the residual energy of the nodes in the cluster as well as the rate of change of the node degree, and when the residual energy of the node is smaller, its transmission priority is higher. In addition, the rate of change of node degree is introduced to determine whether the node is located at the edge of the cavity or not, and the smaller the rate value is, the higher the transmission priority is.

$$P_{ij} = \eta_4 \cdot \left(1 - \frac{E_{j-res}}{E_{max}} \right) + (1 - \eta_4) \cdot \frac{1}{rate} \quad (10)$$

where, η_4 is the weight and E_{j-res} is the residual energy of the nodes in the cluster.

4.2. Inter-cluster transmission policy based on relay node forwarding cost. In the inter-cluster data transmission phase, the cluster head node selects the next hop as a relay node from its neighboring cluster head nodes. To optimize the transmission path and increase the transmission rate, an inter-cluster transmission scheme is proposed from the cluster head node m to its neighboring cluster head node n . If the distance from node m to the base station is larger than the communication radius of node m , a relay node with less forwarding cost needs to be selected for data transmission. The relay node should have high energy and the shortest effective path, the formula for the relay node forwarding cost f_{mn} is shown in Equation (11).

$$f_{mn} = \begin{cases} \left(1 - \frac{E_{n-res}}{E_{max}}\right) \cdot \left(\frac{d_{ntoBS} - d_{toBS-MIN}}{d_{toBS-MAX} - d_{toBS-MIN}}\right) \cdot \left(\frac{1}{rate}\right), & d_{mtoBS} \geq d_{ntoBS}, \\ 2E_{max}, & \text{else.} \end{cases} \quad (11)$$

where, E_{n-res} is the remaining energy of the cluster head node to be selected as the next hop; d_{mtoBS} is the distance from the cluster head node to be searched for the next hop to the base station, and d_{ntoBS} is the distance from the cluster head node to be selected as the next hop node to the base station.

5. Node Dormant Scheduling Strategy for Suppressing Energy Hole. To characterize the energy hole phenomenon, several node dormancy strategies are proposed to be combined with routing protocols to solve the energy hole problem:

5.1. cluster head reselection strategy. In the process of cluster structure construction for wireless sensor networks, to maintain the stability of the network and avoid frequent cluster head changes, a specific cluster head re-election strategy is proposed, whereby when a node is successfully elected as a cluster head, the node will be prohibited from being elected as a cluster head again in the next three rounds of the cluster formation process. The implementation of this strategy relies on the list of neighboring nodes maintained by each node. During the cluster head election process, an elected cluster head node broadcasts a message informing its surrounding nodes that it has been elected as a cluster head along with its node ID. Neighboring nodes that receive this message update their neighbor node list to record the node as a cluster head in the most recent round. In the following rounds of cluster head election, each node refers to the record in its neighbor node list. Any node ID marked in the list as having served as a cluster head in the last three rounds will be excluded as a candidate cluster head, thus preventing these nodes from serving as cluster heads continuously for a short period.

5.2. Low energy node alternate dormancy strategy. In wireless sensor networks, each node continuously monitors its residual energy status. When the residual energy of a node drops below a preset threshold value of 20%, the node switches between its working and dormant states according to an alternating dormancy strategy to ensure that it can participate in network communication and data transmission normally when it has enough energy.

5.3. Node dormancy scheduling strategy within the edge survival zone of an energy hole. In networks, coverage usually refers to the communication range of a node, the range within which a node can communicate directly with other nodes. The coverage overlap ratio is defined as the ratio of the coverage overlap area between the communication range of a node and the communication range of its neighboring nodes to the communication area of that node, and the formula for the coverage overlap ratio C_R is expressed as:

$$C_R = \frac{A_{overlap}}{A_{cover}} \quad (12)$$

where, $A_{overlap}$ is the area of coverage overlap between nodes and neighboring nodes within the edge survival zone of the energy hole and A_{cover} is the area of coverage of nodes within the edge survival zone of the energy hole.

The node dormancy time T_{S_H} within the survival zone of the energy hole edge is shown in Equation (13).

$$T_{S_H} = \frac{K}{rate} \cdot \frac{C_{R-1}}{C_R} \quad (13)$$

where, C_R is the R -th round coverage overlap rate; and C_{R-1} is the $R-1$ th round coverage overlap rate.

6. Simulation experiment and analysis.

6.1. Experimental simulation environment. The NS2 platform is chosen to be used in Linux environment for simulation and validation. The simulation is set up with the presence of 100 sensor nodes in the network, randomly distributed in a square area of 100×100 , and the initial energy of each node in the network E . To validate the effectiveness of the proposed DCRA algorithm, the algorithm is compared with the LEACH [23], UCDS [24], and EHSRA [25] algorithms.

6.2. Simulation results analysis. Figure 8 shows the comparison of DCRA with LEACH, UCDS, and EHSRA algorithms in terms of number of surviving nodes. The horizontal axis represents the working time of the network and the vertical axis represents the number of surviving nodes. During the cluster head election, the network intensifies the energy consumption because of determining whether the nodes in the network are located in the null edge domain or not at the initial stage of the network as well as calculating the energy hole diffusion coefficient. And the first dead node in the network is generated in the 390th round, compared to the LEACH algorithm, the appearance of the first dead node in DCRA algorithm is delayed by about 95%, and DCRA algorithm is advanced by about 9.3% and 7.14% compared to UCDS and EHSRA algorithms, respectively. The DCRA algorithm successfully extends the life cycle of the network by determining whether a node is located in the edge region of the hole, avoiding premature death of nodes located in the edge region of the hole to equalize the energy consumption of the network, and slowing down the spreading of the hole by the energy diffusion coefficient, which extends the life cycle of the network by 137.96%, 46.02%, and 21.23% compared to the LEACH, UCDS, and EHSRA algorithms, respectively.

Figures 9 and 10 show the comparison of the total energy consumption percentage and the average energy consumption percentage per round for the four algorithms, respectively. In the cluster head selection process, the presence of dead nodes in the network is considered, and the single energy hole diffusion influencing factor is added for cluster head selection and cluster radius calculation, which ensures the optimal cluster head node selection and prevents the cluster head nodes from dying prematurely. A node dormant scheduling strategy is also proposed to slow down the death of surviving nodes to extend the network lifetime. The percentage of total network energy consumption of the DCRA algorithm is significantly lower than the total network energy consumption of LEACH, UCDS, and EHSRA algorithms. The average percentage energy consumption of the DCRA algorithm fluctuates steadily in the range of 1%-2% per round as compared to the LEACH, UCDS, and EHSRA algorithm networks.

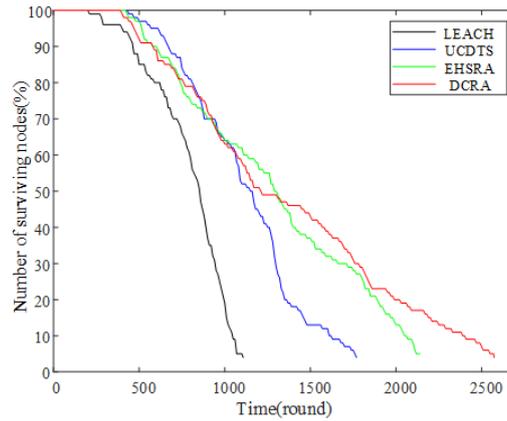


FIGURE 8. Comparison of the number of surviving nodes in algorithms

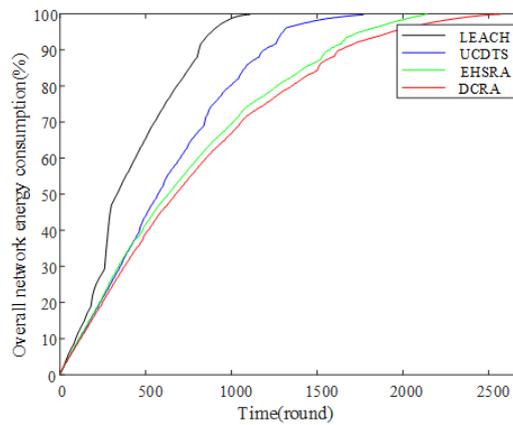


FIGURE 9. Comparison of overall energy consumption of algorithmic networks

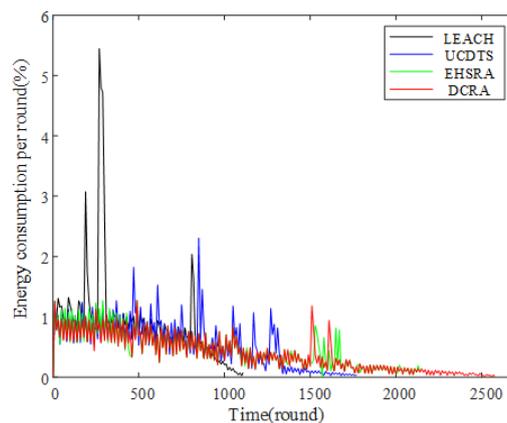


FIGURE 10. Comparison of energy consumption per round of the algorithm network

The properties of energy holes are deeply analyzed through the statistical concepts of energy coefficient of variation, energy quartile spacing, and energy skewness. These metrics provide a more comprehensive and nuanced view compared to the simple average

residual energy and energy standard deviation. Using these performance metrics, the node residual energy values at different numbers of dead nodes are explored in detail.

The energy coefficient of variation is the ratio of the standard deviation of the node residual energy to the average residual energy, which is used to characterize the degree of dispersion of the data distribution and is calculated as shown in Equation (14):

$$CV_{E_{res}} = \frac{\sqrt{\frac{1}{n} \cdot \sum_{k=1}^n (E_{k-res} - \frac{1}{n} \sum_{k=1}^n E_{k-res})^2}}{E_{avgres}} \quad (14)$$

where E_{k-res} is the node residual energy.

TABLE 1. Energy coefficient of variation of the algorithm in the presence of different number of dead nodes

	LEACH	UCDTS	EHSRA	DCRA
10%	0.5225	0.5263	0.4976	0.3811
20%	0.6101	0.6080	0.5474	0.3834
30%	0.5710	0.6555	0.5489	0.3152
40%	0.6073	0.6877	0.6090	0.3252
50%	0.6801	0.7262	0.5934	0.1062

Table 1 shows the energy coefficient of variation of the four algorithms in the presence of different numbers of dead nodes calculated according to Equation (14), the DCRA algorithm has the smallest energy coefficient of variation in the presence of 10% to 50% dead nodes. The energy coefficient of variation of the DCRA algorithm is also reduced by at least 23.41% than the other three algorithms which are the most effective, and it can be reduced up to about 82.10%. The lower energy coefficient of variation indicates that the relative fluctuation of the residual energy of the nodes in the network is small, which allows effective comparison even between networks of different sizes or with different average energies.

Energy quartile spacing is a statistic that describes the stability and degree of dispersion of the data distribution, especially when analyzing the node residual energy values. A set of node residual energy data is first arranged from smallest to largest and then divided into equal quartiles. Where Q1 is the first quartile, the energy value is located at 25% position and Q3 is the third quartile, the energy value located at 75% position. The energy quartile spacing can effectively represent the degree of variability of the data after removing the extreme values, reflecting the middle 50% range of the node's residual energy distribution, the distance from the first quartile to the third quartile. The energy quartile spacing is calculated as shown in Equation (15):

$$IQR_{E_{res}} = Q_{E_{res-3}} - Q_{E_{res-1}} \quad (15)$$

where, $Q_{E_{res-3}}$ denotes the third quartile of the residual energy data of a set of nodes; $Q_{E_{res-1}}$ denotes the first quartile of the residual energy data of a set of nodes.

Table 2 shows the energy quartile spacing of the four algorithms in the presence of different numbers of dead nodes calculated according to Equation (15), the DCRA algorithm also improves the energy quartile spacing by at least 39.16% over the other three algorithms that are the most effective in the presence of 10% to 50% dead nodes, and the maximum can be about 98.88%. Smaller energy quartile spacing implies that most of the nodes in the network have very close residual energy, i.e., the distribution of residual energy in the middle 50% of the nodes is more concentrated. This indicates that the energy

consumption of the network is relatively balanced and does not fluctuate significantly. Energy quartile spacing can help to identify anomalous nodes in the network, i.e., those nodes whose residual energy is much lower or much higher than that of the majority of the nodes.

TABLE 2. Energy quartile spacing of the algorithm in the presence of different number of dead nodes

	LEACH	UCDTS	EHSRA	DCRA
10%	0.8633	0.9526	1.0186	0.6136
20%	0.7174	0.7234	1.3627	0.5837
30%	0.6092	0.7215	1.3966	0.0152
40%	0.4463	0.5341	1.182	0.0143
50%	0.2299	0.3847	0.8675	0.0097

The energy skewness is used to help understand how the energy distribution is skewed concerning the mean value and is calculated as shown in Equation (16):

$$b_{E_{res}} = \frac{\frac{1}{n} \sum_{k=1}^n (E_{k-res} - \frac{1}{n} \sum_{k=1}^n E_{k-res})^2}{\left[\frac{1}{n-1} \sum_{k=1}^n (E_{k-res} - \frac{1}{n} \sum_{k=1}^n E_{k-res})^2 \right]^{\frac{3}{2}}} \quad (16)$$

Table 3 demonstrates the energy skewness of the four algorithms when a different number of dead nodes occur as calculated according to Equation (16), the DCRA algorithm has energy skewness less than zero during the occurrence of 10% to 50% dead nodes, indicating that there are more nodes in the network with residual energy above the average i.e., the tail of the energy distribution extends towards the lower values and there are a few nodes in the network with fast energy depletion. Whereas, other algorithms appear to have energy skewness greater than zero, indicating that there are more nodes in the network with residual energy below the average, i.e., the tail of the energy distribution extends to higher values and some nodes consume energy faster due to overloading or poor location.

TABLE 3. Energy skewness of the algorithm in the presence of different number of dead nodes

	LEACH	UCDTS	EHSRA	DCRA
10%	-0.51	-0.28	-0.78	-1.32
20%	0.01	0.19	-0.5	-1.4
30%	0.08	0.59	-0.23	-1.65
40%	0.53	1.05	0.39	-1.85
50%	1.37	1.52	0.6	-5.12

Figure 11 shows the contour plot of the DCRA algorithm against LEACH, UCDTS, and EHSRA algorithms for 50% of dead nodes. The contour plot is a series of curves formed by connecting nodes having the same energy level, each curve represents the same node energy level, the value of the color bar indicates the percentage of remaining energy of the node, the larger the value and the darker the color indicates the higher the remaining energy of the node, the nodes within the white area are the nodes which are dead or about to be dead.

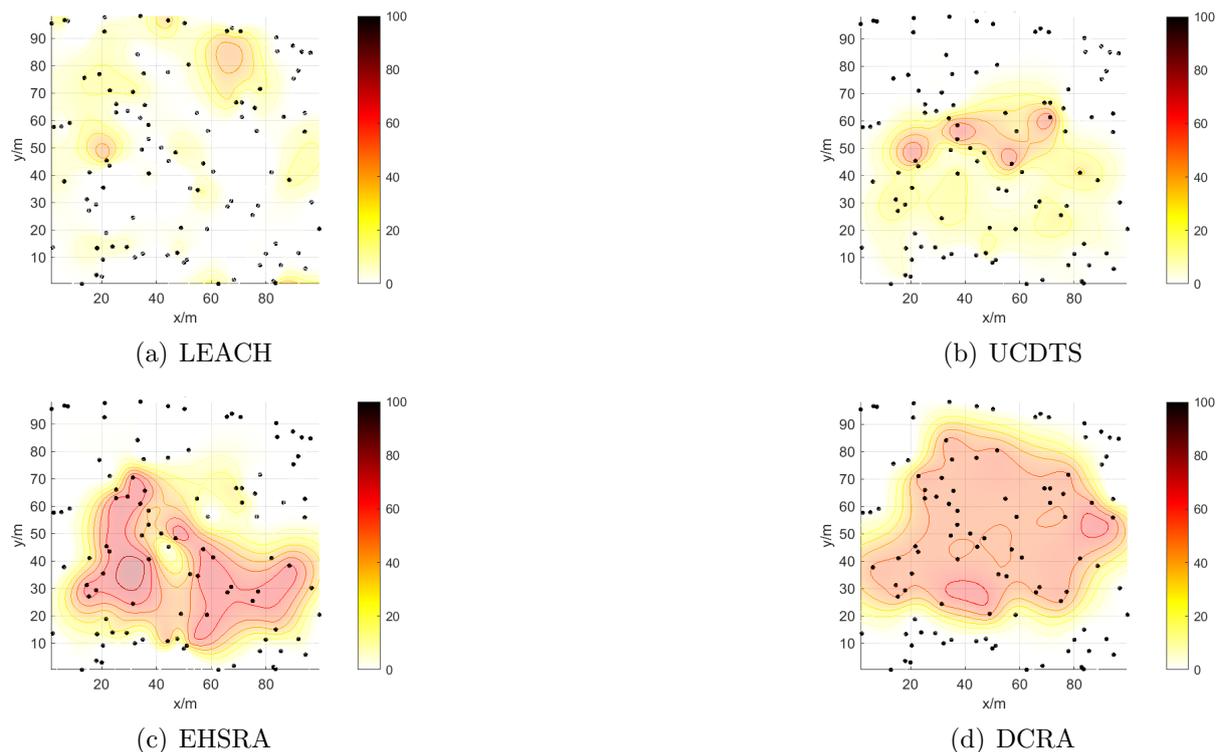


FIGURE 11. Comparison of contour plots of different algorithms at 50% of dead nodes

For in-depth analysis of the network performance, the contour plots when the percentage of dead nodes reaches 50% are specially selected for comparison to compare the performance of the DCRA algorithm with LEACH, UCPTS, and EHSRA algorithms in terms of energy balancing. From the contour plots, it can be seen that the surviving nodes of LEACH and UCPTS algorithms are scattered in the network, showing obvious energy inequality, and even some areas are close to “dead”. Although the EHSRA algorithm has more energy reserves for the surviving nodes, its communication area only accounts for 27.5% of the network area, on the contrary, the DCRA algorithm covers 52.5% of the network area, and the DCRA algorithm shows an obvious marginalization trend in the distribution of the dead nodes in the early stage, i.e., the dead nodes are concentrated in the peripheral area of the network. In the core area of the network, on the other hand, the gradient change of node energy is controlled within 10%, indicating that the energy level changes relatively gently in the network and there is no sharp drop in energy, suggesting that the DCRA algorithm, by optimizing the communication and collaboration strategies among nodes, successfully assigns high-energy-consuming tasks to the edge nodes of the network, while retaining the nodes in the core area as a back-up force. This not only reduces the pressure on the network center and extends the overall lifetime of the network, but also effectively inhibits the expansion of the energy hole and improves the stability and reliability of the network.

7. Conclusion. In this paper, a dynamic evolution model of energy holes for the diffusion mechanism of multi-component liquid mixtures is constructed, which accurately captures the characteristics of energy holes at different stages of development and provides an effective tool for real-time monitoring of the trend of node energy consumption in wireless sensor networks. Based on this model, firstly, by accurately analyzing the rate of change of

the diffusion coefficient and node degree of the energy hole, the structure of the cluster can be more finely tuned to achieve more efficient data transmission and energy utilization. Second, the problem of priority setting for intra-cluster transmission is successfully solved, and the calculation method of inter-cluster node forwarding cost is optimized, which effectively improves the network performance and node energy efficiency. Finally, a cluster head reselection strategy, an alternate dormancy mechanism for low-energy nodes, and a dormancy scheduling scheme for nodes within the survival zone at the edge of the energy hole are proposed, which significantly extends the network lifetime and realizes a balanced distribution of energy consumption. The simulation experiments fully verify the excellent performance of the DCRA algorithm in enhancing the network lifetime and balanced energy consumption. However, the dynamic evolution model of the energy hole still needs to be strengthened in terms of comprehensiveness and depth, and therefore, the subsequent studies will address further improvement and in-depth exploration of the model.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (No. 61501106). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

- [1] J. Li, G.-C. Li, S.-C. Chu, M. Gao, and J.-S. Pan, "Modified parallel tunicate swarm algorithm and application in 3d wsns coverage optimization," *Journal of Internet Technology*, vol. 23, no. 2, pp. 227–244, 2022.
- [2] J. Li, J. Lv, P. Zhao, Y. Sun, H. Yuan, and H. Xu, "Research and application of energy-efficient management approach for wireless sensor networks," *Sensors*, vol. 23, no. 3, p. 1567, 2023.
- [3] T. Zhang, "An intelligent routing algorithm for energy prediction of 6g-powered wireless sensor networks," *Alexandria Engineering Journal*, vol. 76, pp. 35–49, 2023.
- [4] A. N. Pathak and A. R. Yadav, "Scheduling based on residual energy of sensors to extend the lifetime of network in wireless sensor network," *Journal of Engineering and Applied Science*, vol. 71, no. 1, p. 100, 2024.
- [5] L. Han, X. Chen, M. Leng, Y. Deng, T. Hao, and H. Yang, "Research on energy consumption and identifying critical nodes of weighted scale-free topology in wireless sensor network," *IEEE Access*, 2024.
- [6] H. Sun, Z. Meng, D. Wang, and H. Li, "Fedab: A low-latency energy-efficient proactive neighbor discovery protocol in mlde-wsn," *IEEE Access*, vol. 11, pp. 22 843–22 854, 2023.
- [7] M. Majham, M. P. Mwaimu, T. Kivevele, and R. S. Sinde, "Gs-mac: A scalable and energy efficient mac protocol for greenhouse monitoring and control using wireless sensor networks," *IEEE Access*, 2023.
- [8] D. Xiao, M. Zhao, N. Jia, T.-R. Peng, Y. Chen, and L. Ma, "A novel routing control technique for the energy hole in the underwater acoustic distributed network," *Complexity*, vol. 2021, pp. 1–10, 2021.
- [9] N. Sharmin, A. Karmaker, W. L. Lambert, M. S. Alam, and M. S. A. Shawkat, "Minimizing the energy hole problem in wireless sensor networks: A wedge merging approach," *Sensors*, vol. 20, no. 1, p. 277, 2020.
- [10] T.-Y. Wu, A. Shao, and J.-S. Pan, "Ctoa: toward a chaotic-based tumbleweed optimization algorithm," *Mathematics*, vol. 11, no. 10, p. 2339, 2023.
- [11] T.-Y. Wu, H. Li, and S.-C. Chu, "Cppe: An improved phasmatodea population evolution algorithm with chaotic maps," *Mathematics*, vol. 11, no. 9, p. 1977, 2023.
- [12] S. Kaviarasan and R. Srinivasan, "Developing a novel energy efficient routing protocol in wsn using adaptive remora optimization algorithm," *Expert Systems with Applications*, vol. 244, p. 122873, 2024.
- [13] R. Mishra and R. K. Yadav, "Energy efficient cluster-based routing protocol for wsn using nature inspired algorithm," *Wireless Personal Communications*, vol. 130, no. 4, pp. 2407–2440, 2023.

- [14] J. Li, X. Hou, Y. Wang, and J. Huo, "Quantum genetic uneven clustering algorithm for wireless sensor networks." *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 3, pp. 596–605, 2018.
- [15] M. F. Mukhtar, M. Shiraz, Q. Shaheen, K. Ahsan, R. Akhtar, and W. Changda, "Rbm: Region-based mobile routing protocol for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–11, 2021.
- [16] T. S. Panag and J. Dhillon, "Predator–prey optimization based clustering algorithm for wireless sensor networks," *Neural Computing and Applications*, vol. 33, no. 17, pp. 11 415–11 435, 2021.
- [17] K. Wang, C.-M. Yu, and L.-C. Wang, "Dora: A destination-oriented routing algorithm for energy-balanced wireless sensor networks," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 2080–2081, 2020.
- [18] K. Jain, A. Agarwal, and A. Abraham, "A combinational data prediction model for data transmission reduction in wireless sensor networks," *IEEE Access*, vol. 10, pp. 53 468–53 480, 2022.
- [19] R. Kavra, A. Gupta, and S. Kansal, "Interval graph based energy efficient routing scheme for a connected topology in wireless sensor networks," *Wireless Networks*, vol. 27, pp. 5085–5104, 2021.
- [20] H. Zhang, "Wireless sensor network node sleep scheduling algorithm," in *Journal of Physics: Conference Series*, vol. 1976, no. 1. IOP Publishing, 2021, p. 012038.
- [21] P. Rawat and S. Chauhan, "Particle swarm optimization based sleep scheduling and clustering protocol in wireless sensor network," *Peer-to-Peer Networking and Applications*, vol. 15, no. 3, pp. 1417–1436, 2022.
- [22] V. Kossov and H. Altenbach, "Diffusion mechanisms of convective instability in liquid and gas mixtures," *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 103, no. 11, p. e202300801, 2023.
- [23] S. B. Khan, A. Kumar, A. Mashat, D. Pruthviraja, M. K. Imam Rahmani, and J. Mathew, "Artificial intelligence in next-generation networking: Energy efficiency optimization in iot networks using hybrid leach protocol," *SN Computer Science*, vol. 5, no. 5, p. 546, 2024.
- [24] J. Li and Z. Dong, "Uneven clustering and data transmission strategy for energy hole problem in wireless sensor networks." *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 2, pp. 500–509, 2017.
- [25] J. Li, Q. Han, and W. Wang, "Characteristics analysis and suppression strategy of energy hole in wireless sensor networks," *Ad Hoc Networks*, vol. 135, p. 102938, 2022.