ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

Improved LEACH algorithm based on Cuckoo search mechanism

Huichen Liu[1,a], Zhijun Tang[1,2,b,*], Yingjun Deng[1,c]

[1]*School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan, China*
[2]*Hunan Software Vocational and Technical University, Xiangtan, China*
[a]*1450639063@qq.com*, [b]*zjtang@hnust.edu.cn*, [c]*329041699@qq.com*
[*]*Corresponding author*

Abstract: *The traditional Low-Energy Adaptive Clustering Hierarchy (LEACH) has the problem of selecting cluster heads by random numbers, not considering the residual energy of nodes in the cluster head election which can easily lead to the premature death of the low-energy nodes, as well as point-to- point communication over long distances which can lead to data distortion caused by insufficient energy supply. Aiming at the above problems, a cluster routing algorithm for Wireless Sensor Networks (WSN) based on a cuckoo search mechanism (CS-LEACH) is proposed. The proposed algorithm determines the optimal number of cluster heads for the network by optimizing the energy consumption of the entire network by derivation and sets the cluster head selection threshold using an equalization mechanism. Secondly, the algorithm shares the network energy load by setting first and second sub-cluster heads. Finally, in the data transmission phase, the algorithm makes use of the good global search capability of the cuckoo search mechanism to reasonably plan routes for the cluster heads and selects the optimal relay nodes to optimize the transmission routes. Simulation results show that compared with the classical LEACH algorithms, the proposed algorithm improves the network utilization.*

Keywords: *Cluster routing algorithm, Wireless sensor networks, Cuckoo search mechanism, Energy consumption equalization, Network utilization*

1. Introduction

A Wireless Sensor Network (WSN) consisting of all the sensor nodes can monitor a specified area and the data collected by the base station can be made available to the user for further use [1]. Currently, wireless sensor networks (WSNs) have been widely used in various fields, including military, agriculture, energy, healthcare, ocean monitoring, and environmental monitoring. In the 5G era, people require higher coverage of the network, and it is especially necessary to deploy the network in geographically complex and remote places, but due to the impossibility of human manipulation in this type of area, the nodes in the sensor network can only rely on a single battery to achieve the energy supply, and the system stops working once all the nodes' limited energy is exhausted. Therefore, energy consumption is a major constraint for wireless sensor networks.

The main goal of routing protocols is to reduce the energy consumption of nodes and prolong their survival time. Researchers have developed many routing protocols to address these constraints and have proposed several optimization techniques to define the optimal path between the transmitter and receiver nodes. Cluster routing protocols have the advantages of high energy utilization and simple data fusion operations, which have become the focus of prolonging the survival cycle of wireless sensor networks nowadays[2]. Heinzelman first proposed the Low Energy Adaptive Clustering Hierarchy [3], which effectively reduces the rate of energy consumption, prolongs the network lifecycle, and improve data transmission efficiency. It groups the sensor nodes. Each cluster consists of common nodes and is controlled by a cluster head (CH) which is selected by the cluster members to collect their data and send it to the base station (BS) which in turn reduces the energy consumption of the wireless sensor network. Literature [4] proposes an improved algorithm based on energy balanced and efficient LEACH protocol, which fully considers the residual energy factor of the nodes in the election of the cluster head using threshold formulae, so that the cluster head is in the case of sufficient energy to improve the survival cycle of the network, however, the algorithm does not consider the number of clusters, which is prone to uneven distribution of clusters. Literature [5] proposes a reliable topology control algorithm for cluster head into chain based on LEACH, which is a chain-based protocol in which each node communicates only with its immediate neighbors and takes turns transmitting to the base station in PEGASIS. This method improves the robustness of the network but increases the delay and the total energy consumption of the network is too fast.

For better energy optimization, many studies have combined the advantages of multiple optimization techniques using hybrid algorithms in order to address the complex and dynamic nature of these networks. Literature [6] has dealt with nodes with low dynamic trust in a single way without considering multiple factors such as whether they are malicious or abnormal nodes. As well as the reselection of cluster heads after the network has been running for a period of time, there may be a large difference in the gradient of dynamic trust between localities. Literature [7] combines LEACH and compressed sensing theory to improve the LEACH algorithm based on the spatiotemporal correlation of nodes after clustering the data is adopted from the cluster head using the compressed sensing theory and is transmitted to the distant aggregation node through a multi-hop routing method, which accepts the data from the cluster head and adopts a reconstruction algorithm to accurately recover the original signal. However, this algorithm sacrifices the complexity and ignores the cluster head routing planning during data transmission and no further planning is done on the selection of relay nodes. Based on this, this paper proposes a WSN cluster routing algorithm based on a cuckoo search mechanism to solve the problems of network survival period and low utilization of traditional clustering routing algorithms.

2. System model

2.1. Network model

The 100 WSN nodes are randomly distributed in a 100m×100m area to form the monitoring area of the WSN, the structure of which is shown in Figure 1. To facilitate analysis, in this paper, the network model is set as follows: nodes are randomly distributed in a square area with a unique base station; The location of all nodes (including the base station) remains unchanged after deployment; The initial energy of each sensor node is the same and the residual energy of the node at any given moment is known; the base station energy is not restricted; The monitoring area is free of obstructions. After determining the cluster head, ordinary member nodes will make cluster selections based on the strength of the received signals, which naturally leads to the difference in the size of each cluster eventually formed. If the number of member nodes in a cluster is too large, the cluster head will face the challenge of dramatically increasing energy consumption when managing, which in turn affects the energy utilization during subsequent data transmission, resulting in unnecessary energy wastage. To solve this problem, Chapter 3 introduces the method of primary and secondary cluster head collaboration mechanism, where a secondary cluster head is elected to equalize the energy consumption of the cluster head. However, this method is only effective in the stable period of the wireless sensor network, when the network enters into the turbulent period when the main cluster head or vice-cluster head appears to die, the cluster becomes ineffective, and can only continue to operate when a new round arrives.

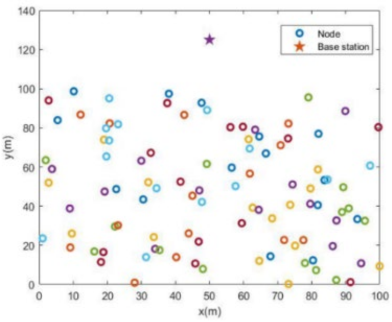

Figure 1: Random distribution of 100 nodes in the WSN

In this paper, we propose a WSN cluster routing algorithm based on the cuckoo search mechanism (CS-LEACH), whose network topology. In the CS-LEACH protocol, a dual sub-cluster head scheme is introduced, i.e., the method of "one main cluster head + two sub-cluster heads".

2.2. Energy consumption model

This paper adopts the energy consumption model of sensor nodes in the energy consumption model of wireless communication, which mainly includes the energy consumption generated by the node sending data and the energy consumption generated by the node receiving data. Setting the distance between the transmitting node and the receiving node is the length of the transmitted data, and is the energy consumption required to send and receive a unit bit of data packet. The energy consumption formula when sending a bit of data packet is shown in Equation (1):

$$E_{Tx}(k,d)=\begin{cases} lE_{elec}+l\varepsilon_{fs}d^2 & d<d_0 \\ lE_{elec}+l\varepsilon_{amp}d^4 & d>d_0 \end{cases} \tag{1}$$

The energy consumption equation for a node receiving a packet of 1 bit is shown in Equation (2):

$$E_{Rx}(l)=lE_{elec} \tag{2}$$

Where ε_{fs} and ε_{amp} are the communication energy parameters; ε_{fs} is the energy consumption required by the node to transmit in the free-space model, and ε_{amp} is the energy consumption required by the node to transmit in the multi-channel fading model; d_0 is the threshold for the transition between the free-space model and the multipath fading model. The transmission model is divided into free-space transmission model and multi-channel fading model in WSN depending on the distance between the aggregation center (base station) and the cluster head. When the communication distance of nodes transmitting data is less than d_0, the free space communication model is used, and when the transmission communication distance is greater than d_0, the multipath fading channel model is used. The threshold value d_0 is calculated using Equation (3):

$$d_0=\sqrt{\varepsilon_{fs}/\varepsilon_{amp}} \tag{3}$$

3. CS-LEACH Routing Algorithms

3.1. Optimal number of cluster heads

Let the number of cluster heads be k. Using the energy consumption model, the optimal number of cluster heads K_{opt} can be derived. Let nodes be randomly and uniformly distributed in a region size of M ×M. If there are k clusters, then each cluster contains N/k nodes on average, including 1 cluster head and $N/k-1$ ordinary nodes. The energy consumption of the cluster head consists of three main components: the energy consumed by the cluster head to broadcast the clustering message, to receive the normal node incoming information, to fuse the information and to send the processed packet to the base station. Therefore, the energy consumption of the cluster head when processing 1 frame of information is shown in Equation (4):

$$E_{CH}=lE_{elec}\left(N/k-1\right)+lE_{DA}\cdot N/k+lE_{elec}+l\varepsilon_{amp}d_{toBS}^4 \tag{4}$$

Where l is the length of the transmitted data in bits; d_{toBS} is the distance between the cluster head and the base station. The main task of the member nodes within a cluster is to transmit the information collected and sensed to the corresponding CH. In this paper, the member nodes are set closer to the cluster head, which means that the transmission distance within the cluster is small, so the free-space energy consumption model is used within the cluster. Therefore, the energy consumption of the member nodes is shown in Equation (5):

$$E_{non-CH}=lE_{elec}+l\varepsilon_{fs}d_{toCH}^2 \tag{5}$$

Where d_{toCH} is the distance between the ordinary node and the cluster head. The area of each cluster is approximately M^2/k . The general perceptual area shape is randomly distributed with a distribution density of $\rho(x,y)$. Here so that the cluster head is located at the center of the cluster, the following equation is shown in Equation (6):

$$E\left(d_{toCH}^2\right)=\iint\left(x^2+y^2\right)\rho\left(x,y\right)dxdy=\iint r^2\rho\left(r,\theta\right)drd\theta \tag{6}$$

Assume that this area radius $R=(M-\sqrt{\pi k})$, and x is constant for y and z, so Equation (7) can be

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

simplified as:

$$E\left(d_{toCH}^{2}\right)=\rho\int_{\theta=0}^{2\pi}\int_{r=0}^{M/\sqrt{\pi k}}r^{3}drd\theta=\frac{\rho}{2\pi}\frac{M^{4}}{k^{2}} \tag{7}$$

If the cluster node density is constant, then $\rho=\left(\frac{1}{M^2}/k\right)$ and $E(d_{toCH}^2)=\frac{1}{2\pi}\frac{M^2}{k}$, so Equation (8) can be simplified as:

$$E_{non-CH}=lE_{elec}+l\varepsilon_{fs}d_{toCH}^{2} \tag{8}$$

The energy consumption of each cluster is shown in Equation (9):

$$E_{cluster}=E_{CH}+\left(\frac{N}{k}-1\right)E_{non-CH}\approx E_{CH}+\frac{N}{k}E_{non-CH} \tag{9}$$

Therefore, the total energy consumption is shown in Equation (10):

$$E_{total}=kE_{cluster}=l\left[E_{elec}N+E_{DA}+k\varepsilon_{amp}d_{toBS}^{4}+E_{elec}N+\varepsilon_{fs}\frac{1}{2\pi}\frac{M^{2}}{k}N\right] \tag{10}$$

Finally, by using the derivative of k for E_{total}, the optimal number of WSN cluster heads is calculated as shown in Equation (11):

$$K_{opt}=\sqrt{\frac{N}{2\pi}}\times\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{amp}}}\times\frac{M}{d_{toBS}^{2}} \tag{11}$$

Where N is the total number of nodes; ε_{amp} is the multipath fading model coefficient; ε_{fs} is the free space model coefficient; M is the side length of the arrangement area; d_{toBS} is the distance between the cluster head and the base station.

3.2. Improvement of the threshold formula

Cluster establishment is the first key stage in LEACH protocol, in the traditional LEACH protocol, the cluster head selection adopts the random election mechanism, which leads to some sensor nodes may be frequently selected as cluster heads even if their remaining energy is limited or their location is remote. These nodes may consume energy rapidly in a short period of time, thus failing prematurely and causing serious "hotspot" problems. In this scheme, we first improve the threshold formula by introducing two key factors, namely, the residual energy of nodes and the distance between nodes and base stations, and construct a new threshold formula, which is shown in Equation (12) and aims to ensure that the selection of cluster heads is more reasonable and efficient, so as to optimize the performance of the whole network.

$$T(n)=\begin{cases}\dfrac{p}{1-p\left[r\bmod(1/p)\right]}*\dfrac{E_{res}\cdot d_{BStoO}}{E_{avg}\cdot d_{mtoBS}} & n\in G\\ 0 & n\notin G\end{cases} \tag{12}$$

Where p is the percentage of nodes that are cluster heads; G is the set of nodes that have not been elected as cluster heads in the last *1/p* rounds; r is the current number of rounds; E_0 is the initial energy of the system node; E_{res} is the remaining energy of the nodes i. d_{mtoBS} is the distance between the node and the base station; d_{BStoO} is the distance of the base station from the centre of the network. Theoretically, the node's own energy reserve is directly proportional to the working hours it can maintain, when the distance between the base station and the center of the network is kept constant, the shorter the distance between the node and the base station, the closer it converges to the aggregation node, the less energy it needs to consume in the data transmission process. Therefore, the probability of a node becoming a cluster head will be significantly higher if the node satisfies the threshold of generating a random number less than the output of the new threshold formula, a node that has sufficient energy and is located close to the base station. The newly designed threshold formula shows more reasonable characteristics in the cluster head election process, effectively breaks the original equal probability model of cluster head selection, and makes the selection of cluster head more in line with the actual needs and performance optimization objectives of the network.

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

3.3. Election of the dual deputy cluster heads

The first deputy cluster head is elected by introducing a deputy cluster head evaluation function by considering four factors: residual energy of the cluster members, distance of the cluster member node from the corresponding cluster head, distance between the cluster member node and the base station, and the average distance of the cluster member node to all the nodes within its communication radius. Assuming that e_1, e_2, e_3, e_4 are all equilibrium regulators belonging to [0, 1] and $e_1+e_2+e_3+e_4 = 1$, the subcluster head evaluation function set r is as shown in Equation (13):

$$Q_{SCH} = e_1 \frac{E_0 - E_{res}}{E_0} + e_2 \frac{d_{toCH}}{m} + e_3 \frac{d_{avg}}{d_{\max}} + e_4 \frac{d_{mtoBS}}{d_{BStoO}} \tag{13}$$

Where E_0 is the initial energy of the system node i; E_{res} is the remaining energy of the nodes i ; d_{toCH} is the distance between the common node and the cluster head; m is the number of nodes in the cluster (excluding the cluster head); d_{avg} is the average distance from the node to all nodes in its communication radius; d_{max} is the maximum distance from the node to all nodes in its communication radius; d_{mtoBS} is the distance between the node and the base station; d_{toBS} is the distance from the base station to the center of the network. From Equation (13), at a certain initial energy of a node, the larger the residual energy of the node, the smaller the factor e_1; At a certain number of nodes in the cluster, the smaller the distance between the node and the cluster head, the smaller the factor e_2; Within the communication radius of a node, when the maximum distance from the node to all its neighboring nodes remains constant, if the average distance from the node to these neighboring nodes is smaller, the smaller the value of factor e_3; When the distance from the base station to the center of the network remains constant, the shorter the distance between the node and the base station, the smaller the value of factor e_4. The above four conditions together act on the probability of a node being the first subcluster (SC) head, within each cluster, each member node corresponds to a SC head evaluation value, and the node with the smallest SC head evaluation value is finally elected as the SC head of the cluster, and when the first SC head is elected in the cluster, its role is to equalize the cluster head energy consumption, and the main and SC heads have the same management capability. The cluster is partitioned into two sub-clusters, which are managed by the main cluster head and the first sub-cluster head, respectively, and are responsible for collecting information from the sub-cluster member nodes and performing data fusion in their respective sub-clusters. The sub-base cluster where the first sub-cluster head is located aggregates the data and sends the information to the main cluster head, which forwards the information from the two sub-clusters to the base station. Only when one of the main cluster heads and the first sub-cluster head dies, the second sub-cluster head will act as the cluster head or the first sub-cluster head to ensure normal communication between the two sub-clusters, until then the first sub-cluster head will only act as a normal intra-cluster member node. At the beginning of a new round, the new cluster will be rebuilt in the setup phase, and the second sub-cluster head will act as the cluster head only when the new cluster head or sub-cluster head dies to ensure the reliable transmission of the data collected by the cluster nodes. The second vice cluster head election is calculated as shown in Equation (14):

$$\lambda = E_{\text{res}} / d_{toCH} \tag{14}$$

3.4. Data Transmission Phase

In the data transmission phase, after the cluster head is determined, the cluster head communicates through a single hop and allocates TDMA time slots to the cluster members. In order to reduce the energy consumption of the nodes, the member nodes transmit information like the cluster head during their respective time slots and remain dormant at other times. However, in the face of long-distance communication, even though the dual sub-cluster head mechanism has been set up to alleviate the pressure, there may still be insufficient energy supply due to the long point-to-point distance between the cluster head and the base station, resulting in data distortion. In order to solve this problem and achieve load balancing, an inter-cluster multi-hop communication strategy is adopted, which combines the powerful optimization capability of the cuckoo search mechanism to elect relay routing nodes within the monitoring area to plan the routes between cluster heads and aggregation nodes, where the relay nodes do not have data fusion capability. After the cluster division is completed, the position of the current node is established as the initial exploration point for the cuckoo search. In each subsequent iteration, the cuckoo explores new nest locations and lays eggs using the Levi flight mechanism. At the end of each flight, a new location will appear and be compared with the optimal individual location

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

recorded in previous iterations. If the new location exhibits superior performance, it is replaced with the historical individual optimal location; otherwise, the historical optimal location is kept unchanged and the individual optimal location continues to be updated through an iterative process. However, it is worth noting that when a nest owner discovers an alien bird egg, the nest loses its eligibility as an optimal location. Ultimately, the global optimal position is filtered from all the individual optimal positions, which serves as the basis for the final decision. This meticulous process ensures the efficiency and accuracy of the cuckoo search algorithm in routing optimization. In order to simulate and simplify the cuckoo search mechanism, three ideal states are assumed for the cuckoo search algorithm:

(1) Each cuckoo produces only one egg at a time and randomly selects nests for incubation, stipulating that a nest can only hold one egg.

(2) The selected nests are continuously updated to ensure that the highest quality nests are available for the next generation.

(3) The probability that the host bird discovers that the egg is an exotic bird is Pa (Pa is between 0 and 1), and once an exotic egg is discovered, the host bird will either discard the egg or abandon the nest in search of a new nest site.

Based on the three ideal states described, the cuckoo bird searches for the next nest site, which is calculated as shown in Equation (15):

$$X_k^{t+1} = X_k^t + \alpha \otimes Levy(\beta) \tag{15}$$

Where X_k^t denotes the *k*th solution in the *t*th iteration; the step factor α is a random number between 0 and 1, which is used to control the random search range; β=1.5; $\otimes$ denotes the point-to-point multiplication; $\alpha \otimes Levy(\beta)$ is the Lévy flight step; the step factor α and the Lévy random number are specified as shown in Equation (16) to Equation (19).

$$\alpha = \alpha_0 \cdot \left(X_k^t - X_{best}\right),\quad \alpha_0 = 0.001 \cdot r_{\max} \cdot e^{\frac{-r}{r_{\max}}},\quad Levy(\beta) \sim \frac{\phi \cdot \mu}{|v|^{1/\beta}} \tag{16-18}$$

$$\phi = \left(\Gamma(1+\beta) \cdot \sin\left(\pi \cdot \frac{\beta}{2}\right) / \left(\Gamma\left(\frac{1+\beta}{2}\right) \cdot \beta \cdot 2^{\frac{\beta-1}{2}}\right)\right)^{1/\beta} \tag{19}$$

Where X_{best} denotes the current optimal solution; $\mu \sim N(0,1)$; $v \sim N(0,1)$; r_{max} denote the current iteration and the maximum iteration, respectively; the introduction of α_0 is to improve the performance of the algorithm's local search. After the nest location is updated, a random number *rand* is generated and its size is compared with Pa. If $rand > Pa$, the host bird discovers with the probability that the nest is an alien egg, eliminates the egg, and randomly updates the nest location once in the neighborhood in the manner of Equation (20), otherwise the location remains unchanged.

$$X_k^{t+1} = X_k^t + v \cdot \left(X_j^t - X_h^t\right) \tag{20}$$

Where X_j^t and X_h^t denote the t-th two random solutions. The steps of cluster head routing planning based on the CS algorithm are as follows:

Step 1: Randomly initialize the bird's nest location and set other and other parameters.

Step 2: Update the position and state of each bird's nest using the Levy flight pattern.

Step 3: Compare the size of random number $rand$ with Pa. If $rand > Pa$, update the bird's nest position using the formula and vice versa unchanged.

Step 4: When the maximum number of iterations is satisfied, output the best position of the bird's nest, which is also the position of the relay node of the best inter-cluster multi-hop route. The process ends, if not satisfied repeat Step2-Step4.

The cluster head of the cluster first performs the data fusion process on the accepted data and later sends the processed information to the relay node. If the relay node is the cluster head itself, the data is sent directly to the aggregation node, otherwise, the data will be sent by the cluster head to the relay node. When each cluster head has found the relay node, the inter-cluster multi-hop forwarding route is established.

3.5. Verification Functions

The cluster head, when acting as a relay node, takes on the task of accepting data from the relay node of the fused previous hop, resulting in increased load. In order to better balance the load, this paper takes advantage of the good global search capability of the CS algorithm to improve the algorithm. In this paper, the cuckoo search mechanism is used to search all the nodes in the monitoring area to elect the optimal relay node by taking advantage of its strong optimization capability and superior search path and reducing the probability of the cluster head node as a relay node. Combined with the intra-cluster data transmission method, becoming a relay node has the following characteristics:

(1) Relay nodes are moderately distant from the cluster head and base station to ensure that less energy is consumed to transmit data along the path.

(2) The residual energy of the relay node ensures stability of data transmission.

(3) The election of relay nodes should not be restricted to cluster heads to reduce the burden on cluster heads.

(4) Relay nodes should reduce the possibility of being overexploited.

According to the above feature requirements, in order to verify the reasonableness as well as the importance of planning cluster head routes using the cuckoo search mechanism, a criterion is needed to comment on the good or bad, and the objective function F is set by considering the distance between cluster head and relay node, the distance between relay node and base station, as well as the residual energy of relay node. In multi-hop transmission mode, the energy consumption of routing mainly contains: (1) Energy consumed by the cluster head for transmitting data of length l bit to the next hop E_{ch}; (2) Energy consumed by relay node j to perform data fusion, including the energy consumed by accepting data of l bit sent from the node of the previous hop and the energy consumed by sending data of l bit to the relay node of the next hop $E_{relaynode}$. The formulae are given in Equation (21) to Equation (23).

$$E_{ch} = lE_{elec} + l\varepsilon_{fs}d_{chtoj}^2 \tag{21}$$

$$E_{relaynode} = 2lE_{elec} + lE_{DA} + l\varepsilon_{fs}d_{next}^2 \tag{22}$$

$$E_{total} = E_{ch} + E_{relaynode} \tag{23}$$

Where: d_{chtoj} is the distance from the cluster head to the next hop relay node j, d_{next} is the distance from the relay node to the next hop, and ε_{fs} is the free space model coefficient; E_{total} is the total energy consumption of the route from the cluster head to the base station. From the above equation, it can be seen that the smaller the distance from the cluster head to the next hop relay node as well as the distance from relay node j to the next hop, the smaller the total energy consumption of the cluster head to base station routing, in addition, the relay node has sufficient residual energy, which is also the key to the data transmission phase. In summary, this paper constructs the objective function F as follows:

$$F = \gamma_1(E_0 - E_{relay\_res}) \ / (E_0) + \gamma_2 \cdot E_{total} / E_{total\_\max} \tag{24}$$

Where E_0 is the initial energy of the node, E_{relay_res} is the residual energy of the relay node; E_{total_max} is the maximum value of the total energy consumed in traversing all the bird's nests inside the cluster head to the base station route; γ_1 and γ_2 are weighting factors of size between 0 and 1 and $\gamma_1 + \gamma_1 = 1$. From the equation, it can be seen that: when the initial energy of the node is certain, the larger the residual energy of the relay node, the smaller the distance from the cluster head to the next hop relay node as well as the distance from the relay node to the next hop, the smaller the energy overhead of routing and the smaller the value of the objective function, the more reasonable the route planning is elected as a relay node and combined with the cuckoo search mechanism.

3.6. Algorithm Flow

The specific steps of the WSN cluster routing algorithm based on the cuckoo search mechanism can be described as follows:

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

(1) Initialization phase: initialize network parameters and deploy sensor nodes.

(2) Cluster formation phase: the optimal number of cluster heads K_{opt} is calculated based on each known parameter. if the random number between 0 and 1 generated by a node in the WSN is less than the threshold value produced by the new threshold formula, and it is elected as a cluster head within the first 1/p rounds, the node becomes a cluster head, and conversely, it becomes a regular member node waiting to join the cluster. After the cluster head is selected, the cluster head node broadcasts its cluster message to the whole network, and the ordinary nodes receive the message, select the cluster head according to the strength of the received signal, and send the "application for joining the cluster message", and when all the ordinary nodes join the cluster, the cluster is established.

(3) Dual vice-cluster head election phase: The first vice-cluster head node is elected by setting the vice-cluster head evaluation function, and the second vice-cluster head is elected by setting the value, so as to achieve balanced energy consumption of cluster heads. When the first vice cluster head is elected in the cluster, its main function is to share the energy load for a single cluster head, and the main cluster head and the first vice cluster head have the same function. The cluster is partitioned into two sub-clusters, which are managed by the main cluster head and the first sub-cluster head, which is responsible for collecting information from the member nodes of the sub-clusters in their respective sub-clusters and performing data fusion. The sub-base cluster where the first sub-cluster head is located aggregates the data and sends the information to the main cluster head, which forwards the information from the two sub-clusters to the base station. When one of the main cluster heads and the first sub-cluster head dies, the second sub-cluster head will act as the cluster head or the first sub-cluster head to ensure the normal communication of the two sub-clusters, and before that, the second sub-cluster head only acts as an ordinary intra-cluster member node.

(4) Data stable transmission phase: member nodes collect external target data and transmit them directly to the cluster head or the first sub-cluster head of their cluster. The cluster head first implements fusion processing on the accepted data. Subsequently, the routes from the cluster head to the aggregation nodes are optimized by the cuckoo search method. Whenever a round of inter-cluster routing relay node selection is completed, the node's position is adjusted according to the predefined position update rule. If the performance of the node under the new position is better than the optimal value in the previous round, the position update is adopted; otherwise, the original position is kept unchanged. Similarly, if the current global optimal position shows superiority over the historical global optimal position, the position update is also performed; otherwise, the original position is maintained.

(5) If the preset maximum number of iterations is not reached, steps (2) to (4) are repeated. For every round of the iterative cycle, the cluster head node needs to be re-elected to maintain the balance of energy consumption in the network.

4. Performance evaluation

4.1. Simulation parameters settings

To verify the effectiveness of the WSN Cluster Routing Algorithm Based on Cuckoo Search Mechanism (CS-LEACH), it is necessary to carry out simulation experiments to analyze it. For this purpose, *N* sensor nodes are randomly deployed in a 100m×100m area, where the aggregation nodes located in the area with coordinates (50,125), and the other simulation parameters are set as shown in Figure 2.

Parameter name and variable name	Parameter values
Total number of nodes N	100
Initial energy of nodes E_0	0.5J
Energy consumption per bit of data sent and received E_{elec}	5×10^{-8}J
Free space model parameters ε_{fs}	10pJ/bit/m^2
Multipath decay spatial model parameters ε_{amp}	0.0013pJ/bit/m^4
Energy consumption in data fusion E_{DA}	5×10^{-9}J
Data message length DM	4000bits
Control message length CM	32bits

Figure 2: Simulation parameters

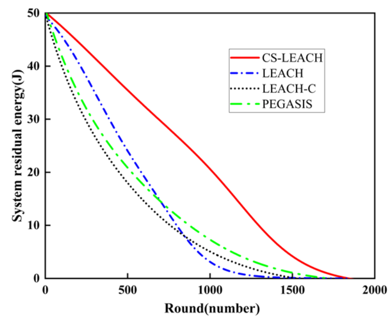

Figure 3: Trend of system residual energy

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

4.2. Analysis of simulation results

The trend of system residual energy with the number of cycle rounds for the four algorithms LEACH, LEACH-C, PEGASIS, and CS-LEACH is shown in Figure 3. From the figure, it can be seen that as the number of cycle rounds increases, the remaining energy in the network is less and the curve shows a decreasing trend. During the whole network activity period, the residual energy trend of the proposed scheme in this chapter tends to be obviously stable compared to the other three algorithms, which is due to the fact that the easy algorithm is more perfect in the consideration of the energy consumption point of view, and selecting the cluster head based on the residual energy and the distance and adopting the double sub-cluster head method can better shorten the transmission distance and reduce the burden of the cluster head. The survival period of the CS-LEACH algorithm is 1914 rounds, and the comparison of the survival period of the CS-LEACH algorithm is 1914 rounds, which is 55.6%, 28.5%, and 12.6% higher than that of LEACH, LEACH-C, and PEGASIS algorithms, respectively, and it can be seen that the algorithms proposed in this chapter can better save the system energy.

The number of surviving network nodes gives a more intuitive view of how long the network is working, and with limited energy, as the number of rounds increases, each node completes its own work task until the energy runs out and the node dies. The relationship between the number of surviving nodes and the number of rounds of the cycle for the four algorithms is shown in Figure 4. From the figure, it can be seen that in the algorithm proposed in this paper, the first node starts to die at 1150 rounds, which extends the number of rounds where the first node dies by almost double the number of rounds compared to the other three algorithms, especially comparing the LEACH algorithm. The effectiveness of the proposed algorithm can be verified due to the fact that the energy and transmission distance of the nodes are taken into account in the cluster head selection and dual sub-cluster heads are set up to share the energy load of the cluster heads, on the other hand, in the data transmission phase, the good global search capability of the cuckoo search mechanism is utilized to plan the routes for the cluster heads and to select the optimal relay nodes to optimize the transmission routes. The proposed optimized cluster routing algorithm based on the cuckoo search mechanism can effectively avoid the phenomenon of "energy hole" and improve the network lifetime to a certain extent.

Figure 5 depicts the fold plots of the number of rounds experienced by the four algorithms, LEACH, LEACH-C, PEGASIS, and CS-LEACH, for different node survival rates. From the figure, it can be seen that the slopes of the folds corresponding to the four algorithms are similar, but the folds corresponding to the CS-LEACH algorithm are more stable, with a small degree of slope variation. When the node survival rate is between 80% and 100%, the slope of the fold line corresponding to the CS-LEACH algorithm has a small degree of change, indicating that the algorithm is relatively more stable than the other three algorithms. It can be seen intuitively through the graphs that, under a certain node survival rate, the number of cycle rounds corresponding to the CS-LEACH algorithm is significantly higher than that of the other three algorithms, which verifies the superiority of CS-LEACH in improving the network utilization, and ensures the stability of the network operation to a certain extent.

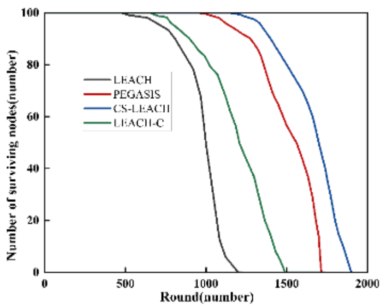

Figure 4: Node Number versus rounds

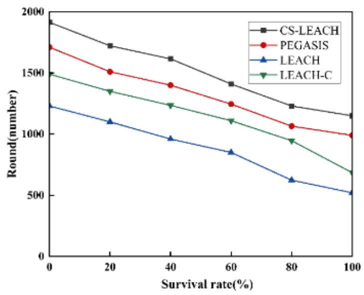

Figure 5: Comparison of node survival rate

5. Conclusion

A WSN cluster routing algorithm is proposed in this paper based on the cuckoo search mechanism, which introduces energy factor and distance factor to select cluster head in the threshold formula of cluster head selection to ensure the optimality of each cluster head selection. In the face of the cluster head due to too much forwarding data, the task is too heavy, considering the distance and energy

ISSN 2616-5775 Vol. 7, Issue 5: 55-64, DOI: 10.25236/AJCIS.2024.070507

factors, a double sub-cluster head shares the cluster head load and expands the network utilization. In the data transmission phase, point-to-point long-distance communication between cluster head nodes and aggregation nodes will have excessive energy consumption and sudden death of network nodes, resulting in shortening the network survival cycle. Multi-hop communication is proposed to plan routes for cluster heads using the good global search capability of the cuckoo search mechanism and select the optimal relay nodes to optimize the transmission routes. Simulation results show that the WSN cluster routing algorithm based on the cuckoo search mechanism can balance the network energy consumption and prolong the network survival cycle.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61875054).

References

[1] Zygowski C, Jaekel A. Optimal path planning strategies for monitoring coverage holes in Wireless Sensor Networks[J]. Ad Hoc Networks. 2020, 96: 101990.

[2] HuangT, Yi K., Cui G, Wang, Y. Hierarchical routing protocol based on non-uniform clustering for wireless sensor network[J]. Journal of Computer Applications. 2016, 36(1), 66.

[3] Heinzelman, WB. An Application-Specific Protocol Architecture for Wireless Microsensor Networks[C]. IEEE transactions on wireless communications google schola. 2002, 2.

[4] Huang LX, Wang H, Yuan LY, et al. Improved LEACH protocol algorithm for WSN based on energy balance and high efficiency[J]. Journal of Communications. 2017, 38(Z2): 164-169.

[5] Lindsey S, Raghavendra CS. PEGASIS: Power-efficient gathering in sensor information systems. In IEEE aerospace conference.2002, Vol. 3: 3."

[6] Zhang D, Wang L. WSN Dynamic Trust Double Cluster Head Selection Algorithm Based on K-Means[J]. Journal of Nanjing University of Posts and Telecommunications: Natural Science Edition. 2020, 40(2), 108-114.

[7] Xie X, Wang JN, Jiang N, et al. Improved Clustering Algorithm for Data Acquisition Based on Com- pressed Sensing in WSN[J]. Journal of East China Jiaotong University. 2018, 35(2): 113-119.