

Research Article

Hybrid PSO and Evolutionary Game Theory Protocol for Clustering and Routing in Wireless Sensor Network

Yu Song ^{1,2}, Zhigui Liu ¹ and Xiaoli He³

¹School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

²Department of Network Information Management Center, Sichuan University of Science and Engineering, Zigong 643000, China

³School of Computer Science, Sichuan University of Science and Engineering, Zigong 643000, China

Correspondence should be addressed to Yu Song; songyu@suse.edu.cn and Zhigui Liu; zhigui_liu_swust@hotmail.com

Received 1 July 2020; Revised 18 September 2020; Accepted 9 October 2020; Published 31 October 2020

Academic Editor: Lei Zhang

Copyright © 2020 Yu Song et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Compared with traditional networks, WSNs have more limited resources such as energy, communication, computing, and storage. The problem of how to achieve energy saving, extend network life cycle, and improve network performance under these limited resources has always been an issue of great interest in WSN research. However, existing protocols do not consider that sensor nodes within the BS threshold may not be clustered. These nodes can directly transmit data to the BS. This simplifies the cluster routing process of the entire WSN and saves more energy. This paper introduces an efficient, and energy-efficient, clustering and equalization routing protocol called the PSOLB-EGT protocol. This protocol introduces a new approach by combining improved particle swarm optimization (PSO) and evolutionary game theory (EGT) algorithms to address the problem of maximizing the network lifetime. The operation of the wireless sensor network is divided into an initialization phase and a data transmission phase. In the initialization phase of the wireless sensor network, the improved PSO algorithm is used to establish clusters and select CHs in areas other than the BS threshold. Entering the data transmission phase, we analyze this problem from the perspective of game theory. We use improved noncooperative evolutionary game theory to build models to solve the problem of the energy waste caused by routing congestion. The proposed PSOLB-EGT protocol is intensively experimented with a number of topologies in various network scenarios, and the results are compared with the well-known cluster-based routing protocols that include the swarm intelligence-based protocols. The obtained results prove that the proposed protocol has increased 9%, 8%, and 5% compared with the ABC-SD protocol in terms of network life, network coverage, and amount of data transmitted, respectively.

1. Introduction

A wireless sensor network (WSN) consists of a large number of microscale, low-power-consumption, and energy-constrained sensor nodes with information sensing, data processing, and wireless communication functions. With the continuous development of network technology and wireless communication technology, WSNs have been widely used in many fields such as the military, environmental monitoring, medical care, and industry.

WSNs are often deployed in hostile environments and need to continuously sense and transmit data unattended. Because a WSN has characteristics such as a large number of nodes, a wide geographical distribution, and a complex

working environment, it is often not realistic to replace the battery to supplement energy after the completion of the layout. Therefore, one of the main challenges in WSNs is the energy consumption problem [1]. Usually, there are two ways to reduce network power consumption: One is to design the hardware equipment of the sensor network to have lower energy consumption, such as a low-power CPU or transmitter. The other is to use more reasonable energy-saving protocols such as a low-energy adaptive clustering hierarchy (LEACH) and the threshold-sensitive energy efficient sensor network protocol (TEEN) [2].

The main task of our study is to optimize the selection and formation of clusters at the initial stage by using an improved protocol. In the data transmission stage, we analyze this

problem from the perspective of game theory and use the improved noncooperative evolutionary game theory (EGT) to solve the routing problem of WSNs [3]. Our proposed method is based on a more reasonable network topology framework with a two-layer cluster structure. Before each round of communication, the improved particle swarm optimization (PSO) algorithm is adopted to reselect cluster-head nodes instead of randomly selecting cluster head (CH) patterns. The PSO algorithm is optimized to avoid local optimization and accelerate global convergence. Compared with previous protocols, this clustering protocol has the following advantages:

- (i) It enables data aggregation at the CH to discard redundant and uncorrelated data. Thereby, it saves energy in the sensor nodes
- (ii) Routing can be more easily managed because only the CH needs to maintain the local route setup of other CHs and thus requires little routing information. This in turn greatly increases the scalability of the network
- (iii) It also saves communication bandwidth because the sensor nodes only communicate with their CH, thus avoiding redundant message exchanges between them

The purpose of the routing algorithm is to select the optimal path to reduce communication delays and energy consumption. Once the optimal set of CHs is elected in the clustering phase, the next step is to find the optimal routing tree from the CHs to the BS while minimizing the total cost. Calculating a desirable route is a challenging problem. Assuming that one path is better than any other path, this path may be used more frequently, which may cause nodes on this path to run out of energy faster [4].

1.1. Related Works. Previous studies have shown that the use of cluster-based layered protocols has broad application prospects for improving the energy efficiency of sensor nodes. In hierarchical routing, when sensor nodes perform multihop communication and data aggregation or fusion in the cluster area, the energy consumption of the sensor nodes is greatly reduced, thereby reducing the amount of information sent to the cluster [5].

LEACH was the first hierarchical routing protocol proposed for WSNs. Many subsequent hierarchical routing protocols are based on LEACH. The LEACH core algorithm consists of three steps [6]. First, CH nodes are randomly selected so that the energy load of the whole network is evenly distributed to each sensor node. Then, data fusion technology is used to reduce the amount of data sent. Finally, the goal of reducing network energy consumption and improving the overall network lifetime is achieved.

$$T(n) = \begin{cases} \frac{p}{1 - p \times [r \bmod (1/p)]}, & \text{if } n \in G, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where p is the probability that the current node becomes the cluster head, r is the current number of rounds, n represents a node, and G is the set of CH nodes that have not been elected in the past $1/p$ rounds. When $r = 0$, each node has the same probability p as the CH. Once a node is elected as the CH, it will not be elected as the CH node in the next $1/p$ round. After $1/p$ rounds, all nodes will be elected as the CH.

Although LEACH can effectively reduce the energy consumption of the entire network, it also has many defects. The most obvious shortcomings are as follows: (i) Each CH communicates with the BS in a single-hop mode. A CH node that is far away from the BS consumes a large amount of energy. (ii) The residual energy of the current node is not considered when selecting the CH. If the energy of the randomly selected CH is too low, it may accelerate node death and reduce network lifetime. (iii) The number of CH nodes in LEACH is usually 5% of the number of nodes. When the distribution density of the sensors varies, a fixed number of CHs cannot optimize the network overhead [7].

It is precisely because of the obvious defects of the LEACH protocol that many scholars are trying to optimize it. Some experts have achieved good results using fuzzy logic methods [8].

WSN routing protocols can be classified into active and reactive types according to different application modes. Active sensor networks continuously monitor surrounding physical phenomena and send monitoring data at a constant rate, while responsive sensor networks only transmit data when observed variables change.

TEEN, which is also based on the LEACH protocol, is the first hierarchical WSN routing protocol for responsive networks. TEEN works in the same way as LEACH, except that it only sends data after the sensor node detects relevant data [9].

After TEEN recluses each cluster area, the CH node needs to broadcast the following three parameters to the members of the cluster: (i) Feature value: this is the physical parameter of the data that the user cares about. (ii) Hard threshold (HT): this is the absolute threshold of the monitored data eigenvalue. When the characteristic value monitored by the node exceeds this threshold, the transmitter is started and reports this value to the CH node. (iii) Soft threshold (ST): this monitors the small range change threshold of the feature value to trigger the node to start the transmitter to report data to the CH.

By setting hard and soft thresholds, TEEN effectively reduces the amount of data sent and is more energy efficient than LEACH. It is suitable for environments that require real-time monitoring of changes. Monitors can also balance the accuracy and requirements of the system of monitoring data by setting different soft thresholds [10].

Although the TEEN protocol is more energy efficient than the LEACH protocol, TEEN still has the following disadvantages: (i) When the threshold has not been reached, the user will not be able to obtain information. (ii) The receiver of the CH node is always active. To receive data from member nodes at any time, the receiver increases the burden on the CH nodes to some extent.

Swarm intelligence optimization originates from a specific phenomenon of group movement in nature. By studying

the group behavior of these real creatures, human beings seek the rules governing them through simulation and imitation, and we build artificial intelligence models to solve complex problems that cannot be solved by conventional methods in daily life and new problems that have not yet been solved. In recent years, some intelligent algorithms, such as ACO, PSO, and GA, have been studied to solve NP-hard optimization problems in WSNs. In 1995, Kennedy and Eberhart proposed and designed a general method for solving practical optimization problems, namely, the particle swarm optimization (PSO) algorithm [11]. By sharing and conveying information about the location information of a food source among groups, the speed of finding food can be accelerated, and the final goal can be achieved through joint effort. The PSO algorithm is easy to realize with high accuracy and is faster and more efficient than other algorithms [11, 12].

People often refer to two important kinds of information in their decision-making process. The first is their own experience, and the second is the experience of others in the group. Similarly, in the process of foraging in birds, each bird's initial state is in a random position, and the direction of flight is also random. However, over time, these initially random birds spontaneously organize into a colony by learning from each other, sharing information, and accumulating foraging experiences. Each bird remembers the best location it finds, which is called a local optimum. In addition, it can also remember that the optimal location searched so far in the flock is the global extreme value, which is called the global optimum. The foraging centers of the whole flock move toward the global optimum.

In the PSO model, each individual can be regarded as a particle, and the flock of birds as a swarm of particles. In a D -dimensional target space, for a group of M particles, the position of the i -th particle is expressed as x_{id} . In other words, the position of each particle is a potential solution. By substituting x_{id} into the objective function, its adaptive value can be calculated and evaluated according to the size of the adaptive value. The best position experienced by individual particles is denoted as $p_{best}(t)$, and the best position experienced by all particles in the whole population is denoted as $g_{best}(t)$. The velocity of the i -th particle is expressed as v_{id} . Standard PSO (SPSO) can be expressed with

$$\begin{aligned} v_{id}(t+1) &= w \times v_{id}(t) + c_1 \times r_1 \times (p_{best}(t) - x_{id}(t)) \\ &\quad + c_2 \times r_2 \times (g_{best}(t) - x_{id}(t)), x_{id}(t+1) \\ &= x_{id}(t) + v_{id}(t+1), \end{aligned} \quad (2)$$

where w is the inertia weight factor and parameters c_1 and c_2 are collectively referred to as learning factors and are, respectively, referred to as cognitive parameters and social parameters. r_1 and r_2 are random numbers in the range $[0, 1]$.

We usually use a linearly decreasing inertia weighting factor w to improve performance. At each update iteration, the value of w decreases linearly from approximately 0.9 to 0.4. Choosing the appropriate inertia weight w can provide a balance between global and local search, and it can result in finding a sufficiently optimal solution with fewer average iterations. Its value is set as follows:

$$w_c = w_{max} - (w_{max} - w_{min}) \left(\frac{t}{T_{max}} \right), \quad (3)$$

where T_{max} is the total number of iterations, t is the current iteration, $w_{max} = 0.9$, and $w_{min} = 0.4$. w_c is a linearly decreasing value. It decreases with the increase in the number of search rounds, which indicates that the effect of the inertial velocity of particles decreases gradually.

Traditional PSO algorithms do not have genetic and crossover operations, relying on the $p_{best}(t)$ and $g_{best}(t)$ of the particles to complete the search. It has the advantages of fewer parameters, simple structure, fast search speed, and easy convergence. However, due to the lack of dynamic adjustment of parameters, it can easily fall into a local optimum, resulting in low convergence precision.

Wireless sensor network routing is a challenging area of research. In general, when we try to optimize routing problems, there are many metrics that should be considered, for example, the distance between nodes, data delay, residual energy of nodes, transmission rate of each link, and distributed characteristics of wireless sensor networks.

In the past, game theory focused on the field of economics, and it was used to study the decision-making process of economic activities so that people could optimize outcomes for their economic interests. Since the 1980s, game theory has been improved and applied more widely. In addition to its applications in economics, it is also widely used in biology, computer science, public policy, and other disciplines and has had an important impact.

Biologists Maynard Smith and Price in 1973 introduced classical game theory into biological evolution analysis and put forward the basic equilibrium concept of EGT, the evolutionary stable strategies (ESSs). In 1978, Taylor and Jonker discovered the relationship between evolutionary stable strategies and replication dynamics, marking the birth of EGT.

In classical game theory, it is assumed that the players in the game are completely rational and the decisions they make are optimal. These players are known to be rational and fully aware of the game. However, in the actual maximum lifetime problem of WSNs, not all nodes participating in the game are completely rational, so the assumptions of classical game theory are not applicable to this problem.

Compared with classical game theory, EGT does not require participants to be completely rational and have sufficient information but only requires participants with limited rationality to learn from each other step by step to make the whole group composed of participants reach equilibrium. It studies the entire group of participants [12]. Therefore, EGT is selected in this paper to study the routing problem of improved protocol in WSNs.

There are two basic concepts in evolutionary game theory, i.e., ESSs and replication dynamic models. The ESS is defined as follows:

If there is a real number α greater than 0, for all the policy sets $m(m \neq n)$ such that

$$U(n, \beta m + (1 - \beta)n) > U(m, \beta m + (1 - \beta)n), \quad (4)$$

for any $\beta \in (0, \alpha)$, then n is called an ESS factor and α is the invasion bound, which is a constant associated with policy m . $\beta m + (1 - \beta)n$ is determined from groups that choose the ESS and groups that adopt a mutation strategy, and it belongs to a mixed group [13].

Replication dynamic theory analyzes the behavior of the whole group based on the principle of “survival of the fittest” in the theory of evolution and changes in the group behavior in the evolutionary game with this method. The expression that determines group behavior is as follows:

$$X(t) = \frac{dx}{dt} = X_i(t)[(U_i(t) - \bar{U}(t))], \quad (5)$$

where $X_i(t)$ represents the proportion of the number of individuals in the group choosing strategy i at time t in the whole group. $U_i(t)$ represents the benefit for individuals in the group who select strategy i at time t . $\bar{U}(t)$ is the average benefit received by each individual in the group at time t [14].

In this section, we review the game theory used to enhance energy conservation and extend the network lifetime. If the individuals who select strategy i in the group obtain more benefits than the average utility of each individual in the group, then the proportion of individuals who select pure strategy i will increase. Conversely, if the utility obtained by the individuals in the group choosing pure strategy i is smaller than the average utility obtained by each individual in the group, the individuals will change their choice to other strategies accordingly [15, 16].

For a better solution, there must be a balance between the WSN energy consumption and the stability of the system. This motivated us to combine PSO and EGT into a new improved protocol. The reasons for the superior performance of the proposed hybrid PSO-EGT protocol over existing protocols are given in what follows.

1.2. Contributions. This paper proposes a CH selection algorithm for WSNs based on improved PSO search and finds the best data transfer path between CHs and the BS through EGT. The main contributions of this paper include the following:

- (i) In clustering, the node within the distance threshold D_0 directly transmits data to the BS, which greatly reduces the energy consumption of the clustering and routing of these nodes
- (ii) A model for calculating the optimal CH number is proposed. Instead of determining the number of CHs as the percentage of the total number of nodes in the past, the model determines the optimal cluster head number according to the actual parameters in the WSN working environment
- (iii) In terms of routing, the working area of the WSN is sorted and planned into different concentric circles according to the threshold value D_0 , and the routing is layered

- (iv) When selecting the transmission path, the shortest path is changed from the previous simple choice to the path jointly determined by the shortest path, the residual power of the relay node, and congestion penalty factors

1.3. Organization. The rest of the paper is organized as follows: In Section 2, the energy and system models are introduced. Section 3 provides two algorithms (i.e., hybrid PSOLB and EGT) for the analysis and design of clustering and routing. The simulation results are presented in Section 4. Finally, Section 5 concludes this paper.

2. Energy and System Model

2.1. Energy Model. Generally, a wireless sensor network node is composed of four modules, i.e., an information sensing module, an information processing module, an information communication module, and an energy supply module. The information sensing module is responsible for collecting and transforming information about the perceived object. The information processing module is responsible for controlling the operation of the entire node and storing and processing the data it collects itself and the data sent by other nodes. The information communication module is responsible for communicating with other nodes, communicating via interactive control messages, and receiving and transmitting service data. The energy supply module is responsible for providing the sensor nodes with the energy required for operation, typically via a large capacity microbattery. It can be seen that each module needs to consume energy. Therefore, it is necessary to establish an energy model, also called an energy consumption model. In contrast to previous studies, we consider the energy consumption of each functional module of the wireless sensor node. The energy consumption model is shown in Figure 1. It can be seen that the energy consumption of the wireless sensor node is mainly determined by three modules. In this section, we will build an information sensing energy consumption model, an information processing energy consumption model, and an information communication energy consumption model. It is necessary to analyze the energy consumption of each module to establish the total energy consumption model of the node [17].

The energy consumption of the information sensing module is related to the power, sensing time, and sensing data quantity of the sensor. Generally, the sensing energy consumption E_s for a node can be expressed as follows:

$$E_s = I_s + V_s + L_s + T_s, \quad (6)$$

where I_s and V_s represent the current and voltage when sensing, respectively; L_s denotes the number of bits in the sensing data; and T_s represents the sensing time.

The information processing energy consumption is mainly determined by the energy consumption per unit byte of information processing and the total amount of data to be processed. The calculation is shown in

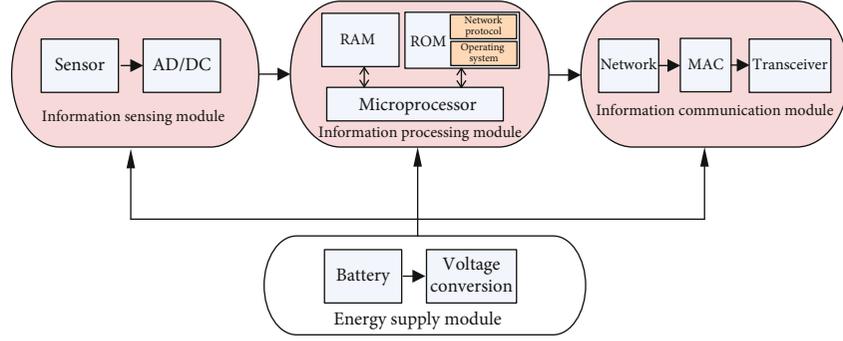


FIGURE 1: Composition of wireless sensor node energy consumption.

$$E_p = l \times E_{da}, \quad (7)$$

$$E_{da} = I_{read} \times V_{read} \times T_{read} + I_{write} \times V_{write} \times T_{write},$$

where l is the total amount of data to be processed and E_{da} represents the energy required for processing one bit of data information. I_{read} , V_{read} , and T_{read} , respectively, represent the current, voltage, and time for reading data when the sensor processes data. In the same way I_{write} , V_{write} , and T_{write} , respectively, represent the current, voltage, and time for writing data when the sensor processes data.

The energy consumption of sensor nodes during the process of data communication is mainly composed of two parts, i.e., the energy consumption of receiving information and the energy consumption of sending information sending [18].

Energy consumed by data transmission is

$$E_{tx}(l, d) = E_{tx-elec}(l) + E_{tx-amp}(l, d) = \begin{cases} l \times E_{elec} + l \times \epsilon_{fs} \times d^2, & d < d_{th}, \\ l \times E_{elec} + l \times \epsilon_{mp} \times d^4, & d \geq d_{th}, \end{cases} \quad (8)$$

and the energy consumed by data reception is

$$E_{rx}(l) = E_{rx-elec}(l) = l \times E_{elec}, \quad (9)$$

$$d_{th} = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}},$$

where l is the total amount of data to be transmitted or received. E_{elec} is the energy consumed to send one bit of data. The amplifier power consumptions ϵ_{fs} and ϵ_{mp} are determined by the transmission distance and the received bit error rate. d is the distance between two sensor nodes. The signal energy consumption model is divided into two categories according to distance (i.e., a free space model and a multipath attenuation model). When the transmission distance is less than the distance threshold d_{th} , the free space energy consumption model is adopted in the communication mode; otherwise, the multipath attenuation model is adopted. d_{th} is a constant, and the value depends on the network environment.

Thus, we can determine the total energy consumed by each sensor [19].

$$E_{total} = E_s + E_p + E_{tx} + E_{rx}. \quad (10)$$

2.2. System Model. In this section, we assume a WSN consists of N nodes (H CH nodes and $N-H$ non-CH nodes) and one BS. There is a set of sensors that are randomly distributed in a designated area. Once the deployment is complete, all sensor nodes become static nodes. In the model we propose, each sensor node can only be assigned to one cluster, and each CH node acts as the CH of exactly one cluster. Each node has local information, including its own unique ID, the CH node ID, the information collection and transmission round, the residual energy level, and the distance to its neighbors, etc. As discussed above, the entire operation of the WSN is divided into T rounds by time. In each round, a normal sensor node sends the monitored data to its cluster head. After receiving the data, the cluster head fuses the data to discard the redundant data and either sends the fused data directly to the BS or forwards it to the BS in multiple hops through other nodes (CH or member nodes) [20].

As shown in Figure 2, according to the general working principle of the WSN, this paper analyzes the initialization stage and data transmission stage. The focus is on optimizing the election of CHs and data transfer [21].

In this paper, the proposed PSO-EGT protocol is based on the classical LEACH protocol. The protocol uses a round as a unit; each round consists of an initialization phase and a data transmission phase for the purpose of reducing unnecessary energy consumption. The specific process of the two phases of each round is shown in Figure 2. The initialization phase is for calculating the optimal number of clusters, electing the CHs, and forming clusters. In the data transmission phase, the WSN performs data aggregation and data transmission and reception [22].

The proposed scheme, called the hybrid PSO-EGT protocol, is based on the common characteristics of both the PSO and EGT algorithms. In this paper, a centralized two-tier PSOLB-EGM protocol is proposed to solve the problem of clustering and routing in WSNs [23]. To implement the hybrid PSO-EGT protocol, the flows to be followed are shown in Figure 3.

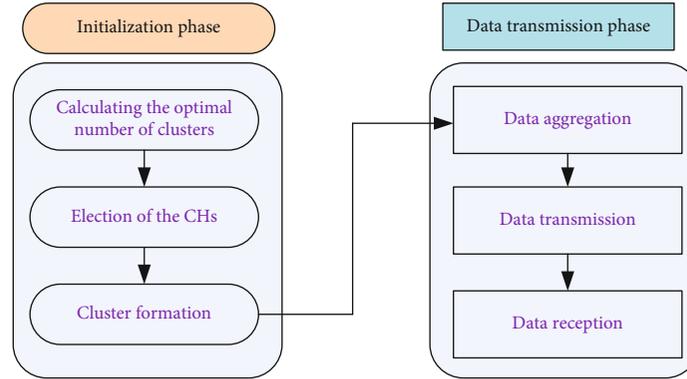


FIGURE 2: Structure of the WSN general working principle.

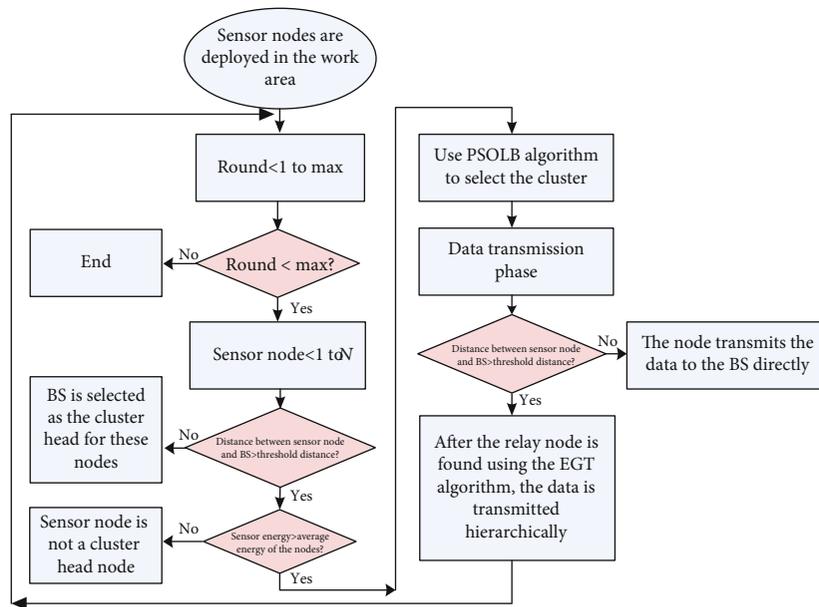


FIGURE 3: Flows of the PSOLB-EGT protocol.

The main task of the protocol in the initialization phase is to select the cluster head.

First, a node must satisfy two conditions at the same time to become a candidate cluster head: the distance from the node to the BS is greater than the threshold. The residual energy of the node is greater than the average residual energy of all nodes.

The cluster head is selected using our improved PSOLB algorithm for candidate cluster head.

In the data transmission phase, nodes within the BS threshold directly transmit data to the BS. Nodes outside the BS threshold use the improved EGT algorithm to transmit data to the BS.

3. Algorithm Design and Implementation

3.1. The PSOLB Algorithm and Parameter Setting. According to the energy consumption model of WSN, the energy consumption of nodes is affected by sensor data volume, sensor

time, node distance, and other factors. Among these influencing factors, the first to be considered is the distance between sensors after clustering, that is, the compactness of clusters [24].

In many existing clustering algorithms, the number of CHs is usually fixed. The algorithm proposed in this paper takes into account that in an actual working environment, sensor node power, cluster number, and other factors are usually related to the size of the monitoring area, BS location, and number of sensor nodes. If there are too many CHs in the network, redundant CH nodes will consume more energy in the network. If the number of CH nodes in the network is too small, remote nodes will not be able to find suitable CHs to join, resulting in the loss of monitoring data and excessive network delay [25].

It is assumed that N sensor nodes are randomly arranged in a monitoring area of size $M \times M$. Since the work of the CH node is mainly divided into receiving information, processing information, and sending information, the energy consumption

of a single CH node is the sum of the energy consumptions of the above three tasks. In the initial condition, we assume that the CH node is far from the BS. According to equations (8) and (12), we can determine the total energy consumption of the i -th CH node as follows:

$$E_{\text{CH}_i} = E_{tx} + E_{rx} + E_{dx} = (a \times l \times E_{\text{elec}}) + (l \times E_{\text{elec}} + l \times \varepsilon_{mp} \times d_{\text{toBS}}^4) + (a \times l \times E_{da}), \quad (11)$$

where $a = N/H$ represents the total number of nodes in the cluster (including one CH node and $a - 1$ non-CH nodes), l represents the total amount of transmitted data, E_{da} is the energy consumption of data aggregation, and d_{toBS} is the distance from the CH node to the BS [26].

Then, according to the Euclidean theorem, the distance between a CH and the BS can be expressed as follows:

$$d_{\text{toBS}} = \sqrt{x_1^2 + y_1^2}. \quad (12)$$

The function of a non-CH node is to send perceived information. We assume that the distance between all non-CH nodes and CH nodes is within the threshold d_{th} . The total energy consumption of the j -th non-CH node in the i -th cluster can be calculated by

$$E_{\text{non-CH}_j} = l \times E_{\text{elec}} + l \times \varepsilon_{fs} \times d_{\text{toCH}_i}^2. \quad (13)$$

Similarly, the distance between the j -th non-CH node and the i -th CH can be expressed as follows:

$$d_{\text{toCH}_i} = \sqrt{x_2^2 + y_2^2}. \quad (14)$$

The expected distance from the CH node to the sink node is

$$\begin{aligned} E(d_{\text{toBS}}^4) &= \int_0^M \int_0^M \left(\sqrt{x_1^2 + y_1^2} \right)^4 \rho(x_1, y_1) dx_1 dy_1 \\ &= \int_0^M \int_0^M (x_1^2 + y_1^2)^2 \rho(x_1, y_1) dx_1 dy_1, \end{aligned} \quad (15)$$

where $\rho(x_1, y_1)$ is the network node density function. Similarly, we can determine the mathematical expectation of $d_{\text{toCH}_i}^2$.

$$\begin{aligned} E(d_{\text{toCH}_i}^2) &= \int_0^M \int_0^M \left(\sqrt{x_2^2 + y_2^2} \right)^2 \rho(x_2, y_2) dx_2 dy_2 \\ &= \int_0^M \int_0^M (x_2^2 + y_2^2) \rho(x_2, y_2) dx_2 dy_2. \end{aligned} \quad (16)$$

The total energy consumption of the entire WSN can be obtained from equation (16).

$$E_{\text{total}} = \sum_{i=1}^H \left(E_{\text{CH}_i} + \sum_{j=1}^{a-1} E_{\text{non-CH}_j} \right). \quad (17)$$

According to equations (15) and (16), the expected value of the total energy consumption is

$$\begin{aligned} E(E_{\text{total}}) &= H \times E(E_{\text{CH}_i}) + (N - H) \times E(E_{\text{non-CH}_j}) \\ &= H \times ((a \times l \times E_{\text{elec}}) + (l \times E_{\text{elec}} + l \times \varepsilon_{mp} \times d_{\text{toBS}}^4) + (a \times l \times E_{da})) \\ &\quad + (N - H) \times (l \times E_{\text{elec}} + l \times \varepsilon_{fs} \times d_{\text{toCH}_i}^2) \\ &= H \times ((a \times l \times E_{\text{elec}}) + (l \times E_{\text{elec}} + l \times \varepsilon_{mp} \times E(d_{\text{toBS}}^4)) + (a \times l \times E_{da})) \\ &\quad + (N - H) \times (l \times E_{\text{elec}} + l \times \varepsilon_{fs} \times E(d_{\text{toCH}_i}^2)). \end{aligned} \quad (18)$$

We take the derivative of equation (18) with respect to H and set its derivative equal to zero, so that we can determine the optimal number of clusters containing the maximum lifetime of all nodes.

$$\frac{dE(E_{\text{total}})}{dH} = 0. \quad (19)$$

In contrast to the traditional PSO algorithm, we first propose that the best result of the last round particle search should be learned; this result is denoted as l_{best} , and we assign the learning weight c_3 to it [27].

$$\begin{aligned} v_{id}(t+1) &= w \times v_{id}(t) + c_1 \times r_1 \times (p_{\text{best}}(t) - x_{id}(t)) \\ &\quad + c_2 \times r_2 \times (g_{\text{best}}(t) - x_{id}(t)) \\ &\quad + c_3 \times r_3 \times \Omega \times (l_{\text{best}}(t-1) - x_{id}(t)), \end{aligned}$$

$$\Omega = \begin{cases} 1, & \text{otherwise,} \\ 0, & l_{\text{best}}(t-1) = p_{\text{best}}(t) \text{ or } l_{\text{best}}(t-1) = g_{\text{best}}(t), \end{cases}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1), \quad (20)$$

where parameters c_1 and c_2 are collectively referred to as learning factors and are, respectively, referred to as cognitive parameters and social parameters. c_3 is the learning factor of l_{best} . r_1 and r_2 are the intelligent learning factor in this paper. Similarly, r_3 is the intelligent learning factor of l_{best} . Ω is a control parameter. When l_{best} coincides with p_{best} or g_{best} , $\Omega = 0$, and $c_3 \times r_3 \times \Omega \times (l_{\text{best}}(t-1) - x_{id}(t))$ fails to prevent repeated learning of particles.

The PSOLB algorithm allows particles to learn from the best particles in the previous round of searching to avoid falling into a local optimization and to find the global optimization faster. Therefore, PSOLB can meet the requirements of the exploration and development of high-dimensional problems.

In each round, the particle updates its position through three extreme values (i.e., p_{best} , the individual historical optimal solution; g_{best} , the global historical optimal solution; and l_{best} , the global optimal solution in the last round). The

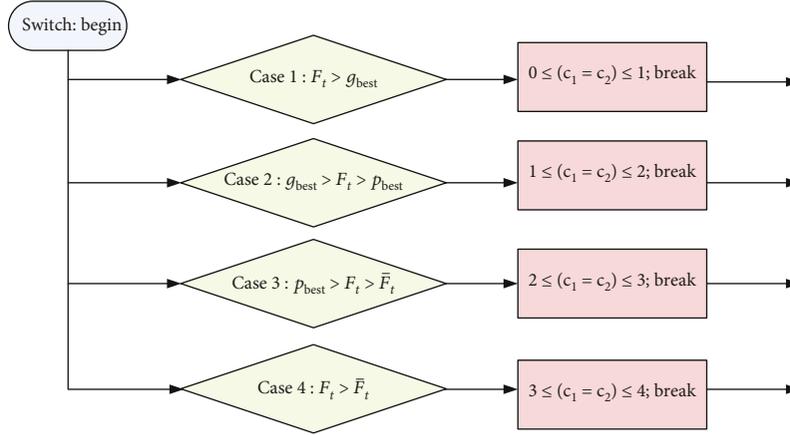


FIGURE 4: Specific improvement methods.

position of the particle is the optimal solution of the distance and is the only criterion for evaluating the particle.

Another improvement to the standard PSO algorithm is to change the learning factors c_1, c_2 from fixed values to dynamic adaptive learning factors. In the basic PSO algorithm, c_1, c_2 have the ability of self-learning and learning from excellent individuals, so that these particles are close to the best points in the group or field. c_1 and c_2 regulate the maximum stride length of particles flying in the optimal direction of individuals or groups, respectively. When the learning factor is small, a particle may wander in a region far away from the target region. When the learning factor is large, the particle can move rapidly to the target region or even beyond the target region. Therefore, the values of c_1 and c_2 will affect the performance of the PSO algorithm [28]. The specific improvement methods are as given in Figure 4.

From Figure 4, we can see that in the proposed method, the learning factors c_1 and c_2 are no longer fixed values of at 2 but are dynamically estimated by the fitness function of the particles. We divide them into four levels. When the particle fitness is greater than g_{best} , this indicates that the particle is close to the optimal solution. At this point, we reduce the value of the particle learning factor $[0, 1]$ and improve the local search ability in order to search for the optimal solution faster. When the fitness of a particle is low, the distance between the particle and the optimal solution is very far. We force particles to take steps out of the local area (c_1 and c_2 take larger values) to search for the best solution faster.

To overcome the shortcomings of the typical linear decrease strategy, the inertia weighting factor in this paper adopts the linear differential decrement method, and its expression is as follows:

$$\begin{aligned} \frac{dw}{dt} &= -\frac{2(w_{\max} - w_{\min})}{T_{\max}} \times t, \\ w_c &= w_{\max} - \frac{(w_{\max} - w_{\min})}{T_{\max}^2} \times t^2, \end{aligned} \quad (21)$$

where t is the current number of search rounds, T_{\max} is the maximum number of search rounds, w_{\max} is the initial

inertia weight value, and w_{\min} is the inertia weight value at the maximum number of search rounds. In this paper, w_{\max} and w_{\min} take values of 0.9 and 0.4, respectively. In the previous linear decrement strategy, the rate of decline of w_c is linear. That is, the presearch and postsearch decline speeds of w_c are the same. When w_c adopts the differential decrement method, the trend of decreasing this value in the PSO algorithm in the early search is slower, and the global search ability is stronger. The downward trend in the search period accelerates, and the local search ability is strengthened. To some extent, the differential decrement method overcomes the limitations of the typical linear decrement strategy and accelerates the convergence of the PSO algorithm. The specific optimization results will be analyzed with the simulation experiment results.

After the optimal number of CHs is determined, the optimization protocol will take the CHs as the core and optimize the minimum energy consumption of the overall network and the maximum life cycle of the nodes.

To achieve the goal of saving energy, the following three values are minimized: the residual energy of non-CH nodes, the distance between non-CH nodes and CH nodes, and the number of hops from a CH to the BS in multihop mode.

The fitness function of the CH node can be expressed as

$$\begin{aligned} F_t &= \mu_1 \times E_{\text{residual-non-CH}} + \mu_2 \times D_{\text{to CH}_i} + \mu_3 \times H_{\text{to BS}}, \\ \mu_1 + \mu_2 + \mu_3 &= 1, \end{aligned} \quad (22)$$

where μ_1, μ_2 , and μ_3 are the control parameters of the three parts of the fitness function. In this protocol, the equivalent values of the three control parameters represent the same effect of the three influencing factors. $E_{\text{residual-non-CH}}$ is the sum residual energy of the non-CH nodes, and its expression is

$$E_{\text{residual-non-CH}}(t) = E_{\text{residual-non-CH}}(t-1) - \sum_{i=1}^H \sum_{j=1}^{a-1} E_{\text{non-CH}_{ji}}(t), \quad (23)$$

where $E_{\text{non-CH}_j}$ is the energy consumption of the j -th non-CH node in the i -th cluster at time t .

$$D_{\text{to CH}_i} = \sum_{j=1}^{a-1} d_{\text{to CH}_j}, \quad (24)$$

where $d_{\text{to CH}_j}$ is the distance from the j -th non-CH node to the i -th CH node.

$$H_{\text{to BS}} = \sum_{i=1}^H H_{\text{to BS}_i}, \quad (25)$$

where $H_{\text{to BS}_i}$ represents the number of hops from the i -th CH node to the BS in multihop mode. If the distance from the CH node to the BS is beyond the threshold H_{th} , the information should be transmitted in multihop mode. Multihop mode is more energy consuming than the single-hop mode, so the number of such CH nodes should be minimized.

Because the CH node needs to undertake more work, the node with more residual energy has a greater chance of being selected as the CH node [29].

Moreover, in order to better complete the following tasks, the optimization protocol stipulates that only nodes with residual energy larger than the average residual energy have the opportunity to be selected as CH nodes. Before selecting CHs, we put the nodes with a residual energy greater than the average residual energy into a set $\Phi = \{E_{\text{residual}} \geq \bar{E}_{\text{residual}}\}$.

Then, the clustering results of each round are calculated by using the proposed PSOLB algorithm [30].

The BS will select the best CH node. Before each round, the WSN will complete the following steps:

- (i) First, neighbor nodes are discovered: each node in the network broadcasts its own ID, residual energy, distance from the neighbor node, distance from the BS, and other information to neighbor nodes
- (ii) Second, neighbor node information is updated according to received data packets
- (iii) Third, the cluster configuration is broadcast: after the BS completes the network configuration, the BS uses flood broadcast again to transmit the configuration to all nodes. It broadcasts the packet containing the configuration. Each node that receives the packet changes its state to the CH node

3.2. The EGT Algorithm. In a WSN, the choice of ideal routing is very challenging. If one path is better than the others, this may cause unbalanced competition between the paths. A well-behaved path may be used more frequently than other paths, resulting in a more crowded path and faster power consumption. Because of the limited energy resources, each node saves energy for its own benefit. Unbalanced competition between paths can lead to path congestion, higher latency, additional packet collisions, and shorter network lifetimes [31].

In this section, we analyze the unbalanced competition problem for paths from the perspective of EGT, and we model the routing in a WSN as an evolutionary anticonoordination routing game. Compared with classical game theory, EGT pays more attention to dynamic strategic changes. The decision-making process can be seen as a strategic evolution over time. The EGT algorithm can be divided into three steps for routing. The specific steps are as follows.

As shown in Figure 5, we simulate the layout of a WSN as a concentric circle. The innermost radius of the concentric circle is R , where $R = d_{th} = \sqrt{\epsilon_{fs}/\epsilon_{mp}}$. Therefore, all nodes within this circular range can directly transmit data to the BS in the single-hop mode. No clustering is carried out in this region. Thus, the overall energy consumption of the WSN is saved. In the concentric circle with a radius of $2R$, the nodes transmit data to the mounting head. The CH transmits data to the innermost concentric circle node through multiple hops and then transmits the data to the BS. Similarly, nodes in the outer concentric circle area first transfer data to the CH in the region, and then the CH in turn transfers the data to nodes in the inner concentric circle, up to the BS [32].

The method of choosing a reasonable route in multihop mode is very important. First, the node establishes a list of all the neighbor nodes, which are arranged from near to far according to the distance from the node. Each node stores five pieces of information: the node ID, the distance from the node to the neighbor node, the distance from the neighbor node to the BS, the residual energy, and the bonus value.

In general, the most efficient way for a CH to transmit data in multihop mode is to find the shortest relay node. To ensure that the global shortest path can be found every time, we use the A^* algorithm to select the relay node. The A^* (A -star) algorithm is the most effective direct search method to solve the shortest path in a static road network, and it is also a common heuristic algorithm for many other problems [33].

$$f(n) = g(n) + h(n), \quad (26)$$

where $f(n)$ is the cost estimation from the initial state to the target state via the state n , $g(n)$ is the actual cost from the initial state to the state n in the state space, and $h(n)$ is the cost estimation of the optimal path from the state n to the target state.

Now that we know the distance from the node to the neighbor and the distance from the neighbor to the BS, we can calculate the shortest path from the current node to the BS.

As shown in Figure 6, in many cases, we will face the problem of routing, that is, two CH nodes will transmit data through the same relay node. The relay node will transfer a large amount of data. When a node is a relay node of multiple CH nodes at the same time, it may easily die because of excessive energy consumption, which is not conducive to a good lifetime for the entire WSN.

EGT is a powerful mathematical tool that models strategic interaction and analysis of competition, conflict, and cooperation with multiple entities.

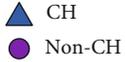
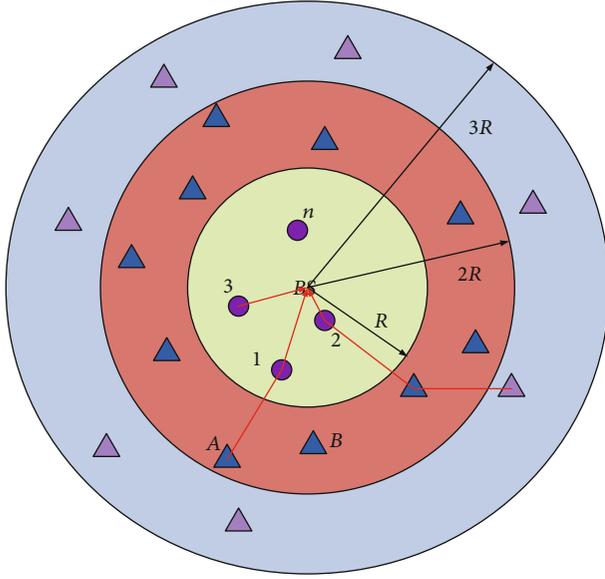


FIGURE 5: Combination of single hop and multihop.

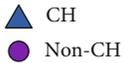
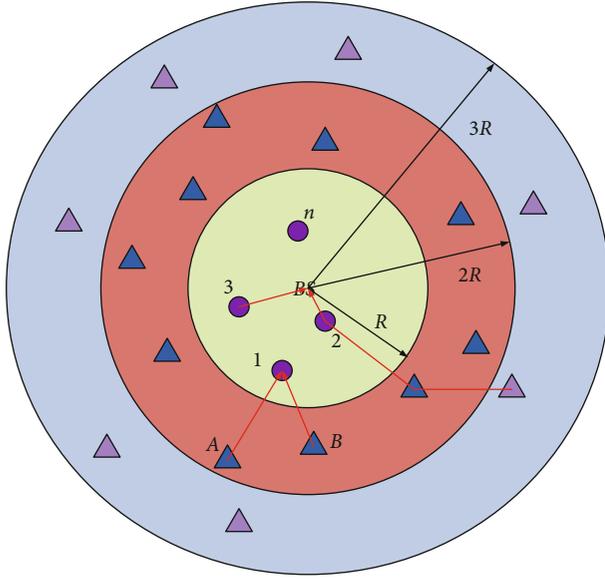


FIGURE 6: Two nodes find the same relay node.

In this section, EGT is selected to study routing congestion in the WSN. An evolutionary game does not require game participants to have complete rationality, but participants with limited rationality can make the whole group they compose reach equilibrium step by step through learning from each other [34].

It can be seen from the previous derivation that the difference form of the EGT replication dynamic process can be

expressed by equation (5). In this paper, $X_1(t)$ represents the proportion of CH nodes in the second concentric circle region that have node 1 selected as the relay node at the same time [35].

$$X_1^A(t) = \frac{N_{\text{node 1 as relay}}}{N_{A \text{ located}}}, \quad (27)$$

where $N_{\text{node 1 as relay}}$ represents the number of nodes that select node 1 as the relay node. N is the total number of nodes in the concentric circle area where node A is located.

$$X_i^N(t) = \frac{N_{\text{node } N \text{ as relay}}}{N_{A \text{ located}}}, \quad (28)$$

where $N_{\text{node } i \text{ as relay}}$ represents the number of nodes that selects node i as their relay node.

The utility function of node i at time t being selected as a relay node is expressed as

$$U_i^N(t) = q_1 \times \frac{1}{E_{tx_i}} + q_2 \times \frac{E_{\text{residual-non-CH}_i}}{E_{\text{initial}}} - q_3 \times \frac{N_{\text{node } i \text{ as relay}} - 1}{N_{\text{node } i \text{ as relay}}}, \quad (29)$$

where E_{tx_i} represents the energy of data transmission through node i : the shorter the transmission distance, the higher the benefit. $E_{\text{residual-non-CH}_i}/E_{\text{initial}}$ is the percentage of the residual energy of relay node i . E_{initial} is the initial energy of the node. The higher the percentage of residual energy, the higher the benefit. We encourage CH nodes to select nodes with high residual energy as the relay nodes. $(N_{\text{node } i \text{ as relay}} - 1)/N_{\text{node } i \text{ as relay}}$ is the penalty parameter. It can be seen from equation (27) that when more nodes select the same relay node at the same time, the penalty parameter increases exponentially. q_1 , q_2 , and q_3 are three weight values that can be adjusted according to actual conditions. In this section, we set the three weight values as follows:

$$q_1 = q_2 = q_3 \approx 0.33, \quad (30)$$

$$X_i^N = \frac{dx}{dt} = \frac{N_{\text{node 1 as relay}}}{N_{i \text{ located}}} \times (U_i^N(t) - \bar{U}(t)) = \frac{N_{\text{node 1 as relay}}}{N_{i \text{ located}}} \times \left(\left(q_1 \times \frac{1}{E_{tx_i}} + q_2 \times \frac{E_{\text{residual-non-CH}_i}}{E_{\text{initial}}} - q_3 \times \frac{N_{\text{node } i \text{ as relay}} - 1}{N_{\text{node } i \text{ as relay}}} \right) - \bar{U}(t) \right), \quad (31)$$

where $\bar{U}(t)$ is the average utility function, and it is expressed by

$$\bar{U}(t) = \sum_{N_{\text{relay}}=1}^{N_{\text{relay}}=i} X_i^N(t) \times U_i^N(t). \quad (32)$$

As seen from Figure 6, when the CH node A selects node 1 as the relay node, the revenue of the CH node is relatively large. When the utility value is greater than the average utility, the CH node A will not change its policy but will continue to

select node 1 as the relay node. When the CH node B also selects node 1 as the relay node, the utility values of CHs A and B are reduced. Therefore, CH A or B will change its policy and select another node as the relay node.

After several evolutions, the game will eventually converge to the evolutionary stable strategy ESS determined by the following equation.

$$X_i^N = 0. \quad (33)$$

3.3. Algorithm Convergence Analysis. After all CHs select a node as the relay node to transmit data, each CH evaluates its own utility according to its own utility function and then compares the obtained utility with the average utility of the group: if its own utility is smaller than the average utility of the group, it randomly chooses to access other relay nodes according to a certain probability; if not, it will keep the previous access network unchanged. See the appendix for a detailed description.

3.4. Analysis of Algorithm Time Complexity. Two important indicators to measure the performance of the optimization algorithm are the accuracy and the speed of the algorithm in solving the problem. The accuracy analysis of the PSOLB-EG protocol has been given in the previous sections. The speed of the algorithm is affected by the time complexity of the algorithm. Therefore, we will focus on the time complexity of each algorithm and give the overall time complexity of each algorithm.

Standard PSO algorithms include processes such as particle swarm initialization, search updates, and particle swarm updates. The computational complexity of the algorithm is $O(N \times I \times D)$, where N is the number of particles, I is the maximum update algebra, and D is the target dimension. However, the protocol we propose is to cluster at nodes other than the BS threshold. The fixed learning factor value is changed to the adaptive value, which increases the comparison time $O(N_{\overline{BS}} \times I \times D + N_{\overline{BS}} \times Ct)$ of the fitness function and the target value; $N_{\overline{BS}}$ represents nodes other than the BS threshold and Ct represents the comparison time. In the same way, all the particles in the process of each round of search have learned the best results of the last round, so the total PSOLB time complexity is $O(N_{\overline{BS}} \times I \times D + N_{\overline{BS}} \times (Ct + Lt))$, where Lt is the learning time of each particle. Because $N_{\overline{BS}}$ is less than N , the proposed protocol is less complex than the standard PSO algorithm.

In terms of data transmission routing, the time complexity of the general exhaustion algorithm is $O(N_{CH} \times N_{CH-neighbor} \times N_{multihop})$, where N_{CH} represents the number of cluster heads, $N_{CH-neighbor}$ represents the number of neighbors of the cluster head, and $N_{multihop}$ represents the number of hops from the cluster head to the BS. The proposed EGT algorithm establishes a cluster head route, and its time complexity becomes $O(N_{\overline{BSCH}} \times N_{upper} \times N_{multihop})$, where $N_{\overline{BSCH}}$ represents the number of cluster heads outside the BS threshold, and N_{upper} represents the number of nodes in the upper layer. Because $N_{\overline{BSCH}} < N_{CH}$ and $N_{upper} < N_{CH-neighbor}$, the time complexity

of the EGT algorithm is less than that of the general exhaustive algorithm.

Through the above analysis, the PSOLB algorithm takes slightly more time than the standard PSO algorithm, but it is better in terms of search accuracy, so the PSOLB algorithm is more energy efficient than the standard PSO algorithm. Regarding the routing algorithm, the EGT algorithm is superior to other search methods in solving accuracy and time. The overall PSOLB-EGT time complexity is superior to that of the standard PSO algorithm as well as the greedy routing algorithm and some other cluster routing algorithms. The above conclusions will be specifically analyzed in the simulation experiment.

3.5. Analysis of Algorithm Space Complexity. Space complexity is mainly used to measure the storage space occupied by the algorithm. We analyze the space complexity of the hybrid algorithm from two aspects: PSOLB algorithm and EGT algorithm. We assume that each population has N particles, the search space is D -dimensional, and there are a total of M populations. For the standard particle swarm algorithm, the storage space required to store the particle position and velocity is $O(2NDM)$, and the space occupied by the P_{best} of each particle is $O(NDM)$, so the space complexity of the standard particle swarm algorithm is $O(3NDM)$. For the PSOLB algorithm, in addition to the particle position, velocity, and storage space required, each particle needs to store the optimal position of the previous round, so the position space of $O(NDM)$ needs to be increased. In addition, consider that the space for storing the dynamic learning step length is $O(ND)$. Therefore, the total space complexity of the LB-PSO algorithm is $O(4NDM + ND)$.

In the WSN data transmission stage, the space complexity of the EGT algorithm is as follows: assuming that the number of nodes within the BS threshold is V , this part of the nodes directly transmits data to the BS, and its space complexity does not need to be calculated compared to the original algorithm. The current node uses the EGT algorithm to calculate the storage space of the fitness function value of each neighbor node as A , the storage space for the average fitness of the remaining nodes is B , and the storage space for the temporary variables of the improved EGT algorithm is C . Then, the worst space complexity of the improved EGT algorithm is $O((N - V) \times (A + B) + C)$. This has lower space complexity than the standard EGT algorithm.

Thus, it is the LB-PSO algorithm space rather than the standard PSO complexity that increases, but the EGT-improved algorithm reduces the complexity of the algorithm rather than the original space. Therefore, the space complexity of the hybrid algorithm is basically equivalent to the original standard algorithm.

4. System Simulation Analysis

In this section, we use numerical simulation to evaluate the performance of our algorithm PSOLB-EGT. Assume that the network coverage is $200 \text{ m} \times 200 \text{ m}$, where the BS coordinates are located at $(100, 100)$, and 100 to 400 sensor nodes are randomly distributed. The simulation parameters used

in this paper, which accord with those in references [14, 16, 22], are shown in Table 1, and all simulation models and algorithms are coded in MATLAB 2015b.

To evaluate the performance of the algorithm, we compare the clustering and routing of the PSOLB-EGT protocol with those of PSO-C, ABC-SD, TPSO-CR, and JCR. TPSO-CR is a two-tier PSO algorithm which is proposed to solve the problem of clustering and routing in WSN [36]. The ABC-SD algorithm exploits the biologically inspired fast and efficient searching features of the Artificial Bee Colony metaheuristic to build low-power clusters [37]. JCR in terms of energy consumption is the clustered routing protocol proposed in 2016 [30, 31]. Since PSO-C does not include a multihop routing protocol, it adopts a greedy routing algorithm. In the greedy routing algorithm, every node is assumed to know its distance to neighbor nodes. CH selects its nearest neighbor node as a relay node [38].

In addition, we use energy consumption, lifetime, residual energy (number of surviving nodes), packet loss rate, and node capacity (throughput) as indicators in simulating and comparing the four protocols. The throughput capacity of the WSN refers to the total data rate (or the total number of data) sent on the network, that is, the data rate (or the number of data) sent by the CHs to the BS and the data rate (or the number of data) sent from the nodes to the CHs. The size of the throughput is often used in network protocols to indicate the performance of the network. Therefore, throughput is often used as one of the evaluation criteria for routing protocols [39].

In the scenario where the network model is uniformly distributed for N sensor nodes, Michael and Martin in [40] have deduced the upper and lower bounds of the capacity in the application case. When the number of nodes N in the network tends to infinity, the upper and lower boundaries coincide and the conclusion is that the capacity is $\Theta(\log(N))$. The throughput capacity of a single node is $\Theta(\log(N)/N)$. For wireless sensor networks with limited node energy, throughput capacity and available throughput for a single node are important performance metrics [34, 35]. We will prove its importance in later simulations.

On the other hand, the node capacity of a WSN also refers to the ability of the node to forward the amount of data. An important cause of node congestion in WSNs is that the load is greater than the maximum capacity of the node. Usually the WSN node load refers to the amount of information that a node needs to forward at a certain moment. Since the WSN node is limited by hardware resources, the node has a fixed capacity, and if the load exceeds its capacity at a certain time, congestion will occur [41].

$$\omega = \begin{cases} 1, & \text{if } Co > Li = \sum_1^k lki, \\ 0, & \text{otherwise,} \end{cases} \quad (34)$$

$$\eta = \frac{L_i}{C_o}, \quad (35)$$

where Co represents the capacity of node i , L_i represents the load of node i , k represents the number of all neighbor nodes

TABLE 1: Relay node status list.

| Node | Relay node list | η | Relay node status |
|------|-----------------|--------|-------------------|
| A | 1 | 65% | Live |
| A | 2 | 70% | Live |
| A | 3 | 50% | Live |

TABLE 2: Relay node status list.

| Node | Relay node list | η | Relay node status |
|------|-----------------|--------|-------------------|
| A | 1 | 0 | Dead |
| A | 2 | 70% | Live |
| A | 3 | 50% | Live |

TABLE 3: Simulation input parameters.

| Parameters | Value |
|---------------------------------------|--------------------|
| Sensor field region (m ²) | 200 * 200 |
| BS location | 100, 100 |
| Number of sensors | 100, 200, 300, 400 |
| Optimal number of clusters | 5, 10, 15, 20 |
| Initial energy of the node (J) | 200 |
| Data packet length (bits) | 4096 |
| Energy data aggregation (nJ) | 5 |
| Number of rounds | 2000 |
| Swarm size | 25 |
| Particle position | [0, 200] |
| Node capacity (kbits) | 0.1 |
| E_{elec} (nJ/bit) | 50 |
| E_{amp} (pJ/bit) | 0.0012 |
| E_{fs} (pJ/bit) | 10 |
| d_{th} (m) | 30 |

of the node, l_{ki} indicates the amount of data transmitted by the neighbor node k to node i , and η indicates the node capacity occupancy. In the subsequent simulation experiments, we will analyze node capacity in detail [42].

In the PSOLB-EGT protocol, congestion is solved by the following steps: establishing a list of next-hop spare relay nodes, detecting node congestion, establishing an alternate data transmission path, and restoring the original transmission path after congestion is released.

In the PSOLB-EGT protocol, congestion is solved by the following steps: establishing a list of next-hop spare relay nodes, node congestion detection, establishing an alternate data transmission path, and restoring the original transmission path after congestion is released.

As shown in Figure 6, we use equation (29) to calculate the utility function values for each possible relay node of the node and sort them accordingly. For example, the utility function values of the next-hop relay nodes of node A are arranged in the order of node 1 > node 2 > node

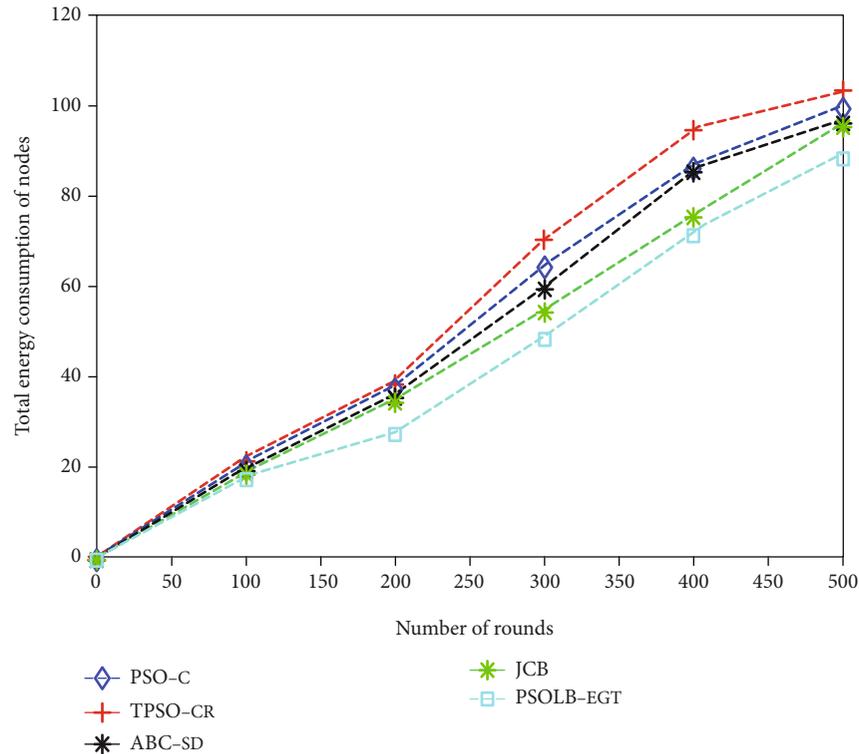


FIGURE 7: Comparison of node energy consumption after running the same cycle for each protocol.

We establish a list of the next-hop relay node statuses for node A. As shown in Table 1, the utility function values of each possible successor node in the next hop of node A are sequentially sorted, and information about their node capacity occupancy, survival, etc. is displayed

According to equations (33) and (35), it is judged whether the relay node in the list is congested. For example, node 1 is not congested and continues to be used as a relay node

If node 1 is congested, node A will automatically elect node 2 as the relay node [43]

In the next round of data transmission, if node 1 congestion is removed, node 1 continues to be used as the relay node. If a node consumes too much energy and dies, the node state changes from live to dead. As shown in Table 2, when relay node 1 dies, node A selects node 2 as the relay node [44]

To test our algorithm, we considered an initial population of 50 particles and let them evolve 2000 times. The parameters required for the experiment are shown in Table 3.

The comparison of the energy consumption of various protocols under the same running cycle is shown in Figure 7. Clearly, the PSOLB-EGT protocol shows the best energy optimization results for the WSN.

We changed the position of the BS in the sensor layout area, as shown in Figure 8. When the BS was in different locations, the total energy consumption of the WSN was compared after running various protocols at the same time for 100 rounds. It can be observed that the PSOLB-EGT protocol is superior to existing protocols in all network scenarios. This is because the PSOLB-EGT protocol uses a more suitable path to transfer aggregate data from the CH to the BS.

We ran five protocols and found that the PSOLB-EGT protocol is superior to the other four protocols when comparing the change in the number of alive nodes in each protocol from 100 to 0. As shown in Figure 9, the PSOLB-EGT protocol is later than other protocols in terms of the first and last node death times. This shows that the protocol can effectively extend the life cycle of WSN.

For the same layout area, with an increase in the number of sensors, each protocol runs for 500 rounds, and we observe that the total residual energy also increases with the increase in the number of sensors. The comparison of residual energy of various protocols is shown in Figure 10.

Through the above simulation experiment, we can observe the following:

- (1) As shown in Figure 7, all other conditions being equal, the PSOLB-EGT protocol consumes less energy than the other four protocols when running the same rounds. It can be inferred that when WSN first started 100 rounds, the energy consumption gap of several protocols was small. As time goes by, the PSOLB-EGT protocol consumes less energy than other protocols
- (2) The results in Figure 8 show that, by changing the location of the BS, the PSOLB-EGT protocol has better scalability than the other protocols
- (3) The PSOLB-EGT protocol has greater advantages than the other four protocols in terms of the first death of nodes. Figure 9 shows that after WSN runs the PSOLB-EGT protocol, its first dead node appears

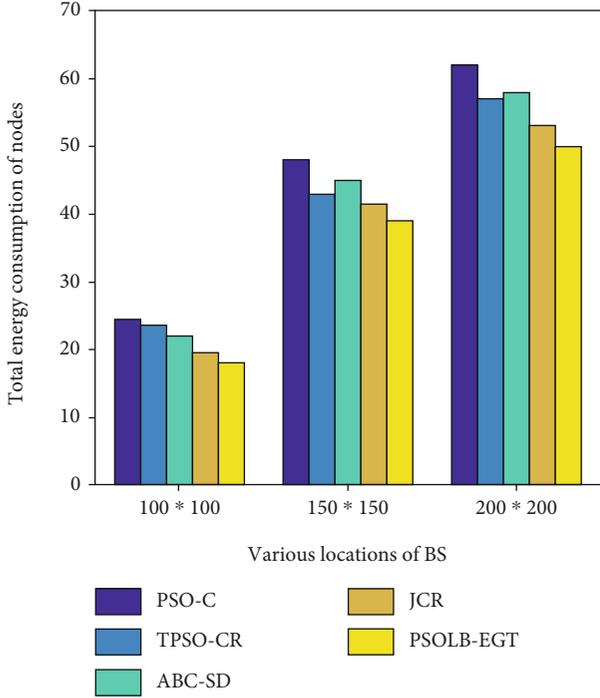


FIGURE 8: Node energy consumption of various protocols when the BS is in different locations.

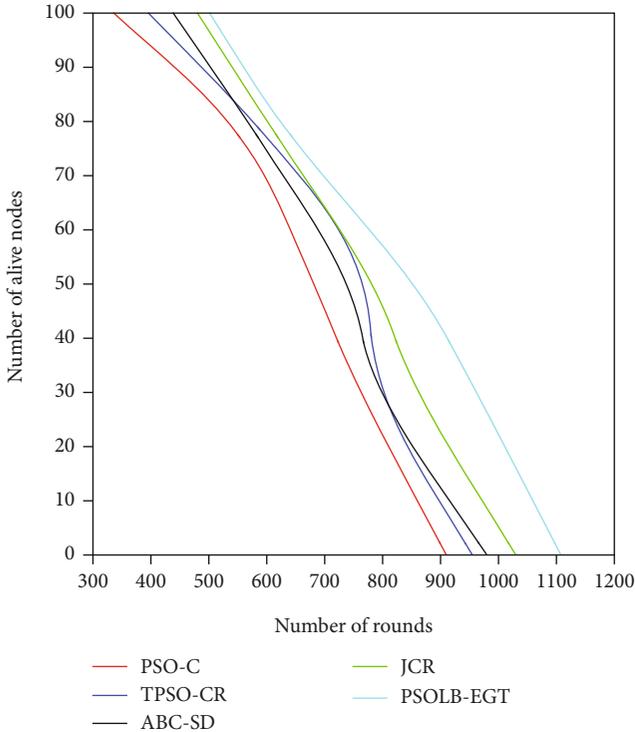


FIGURE 9: Live nodes in the network for various algorithms.

later than other protocols. As the running time of WSN goes by, the node mortality and total running time of the entire WSN are better than those of other protocols

- (4) As the runtime increases, the data transmitted by the WSN continues to increase. Under the same conditions, the WSN uses the traditional algorithm and the improved PSOLB-EGT protocol to perform the node capacity experiment. From the perspective of the average occupancy rate of node capacity, according to the analysis of the experimental results in Figure 11, the other four algorithms increase the average occupancy rate of node capacity as the amount of communication data per unit time increases. The average occupancy rate of the node capacity of the PSOLB-EGT protocol changes slightly
- (5) Figure 12 shows the network throughput comparison between PSOLB-EGT and the other protocols. Throughput is defined here as the number of data packets successfully received at the BS during simulation time. The results represent an average of three different runs for each network size, and it is observed that PSOLB-EGT is superior to the other protocols in network throughput

We implement all protocols to evaluate the rate of packet loss. It is clear from Figure 13 that in all the network scenarios considered, the packet loss rate of the PSOLB-EGT protocol is much lower than that of other protocols. These protocols generate high rates of data loss. This is due to the absence of hybrid-hop communication.

Simulation results in different environments show that the PSOLB-EGT protocol is superior to the existing network protocols in terms of data transmission energy.

We evaluated the number of unclustered sensors per round in each network protocol. The chart in Figure 14 shows the average number of unclustered sensors per round.

As shown in Figure 15, we analyze the node capacity congestion rates of the five protocols for increasing runtimes. In this paper, the total number of nodes successfully transmitting data to the aggregation point is used to describe the node capacity congestion rate of the network as the number of rounds changes. It only makes sense for the WSN to successfully transmit the data generated by the node to the aggregation point. The amount of data successfully delivered reflects the smoothness of the network to some extent. Therefore, the goal pursued by WSN routing is to consume as little energy as possible and transmit as much data as possible to the aggregation point. As seen from Figure 15, the algorithm we proposed has better node throughput capacity.

To verify the energy-saving effectiveness of the PSOLB clustering algorithm separately, we unified all five clustering algorithms (including PSOLB) with the greedy routing algorithm. The operating conditions of the five clustering algorithms are exactly the same. It can be seen from Figure 16 that after the same number of rounds of PSOLB, the remaining energy of the entire WSN is better than it is with any of the other four clustering algorithms. The main

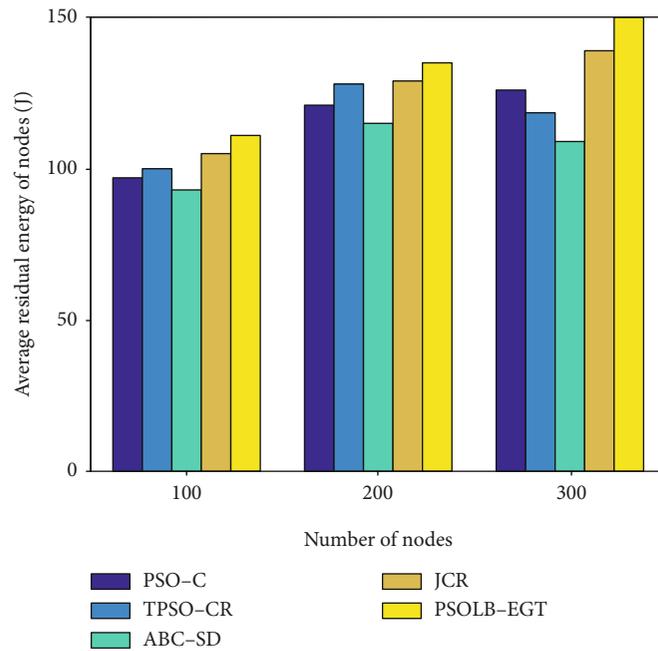


FIGURE 10: Comparison of changes in energy consumption when the number of nodes increases.

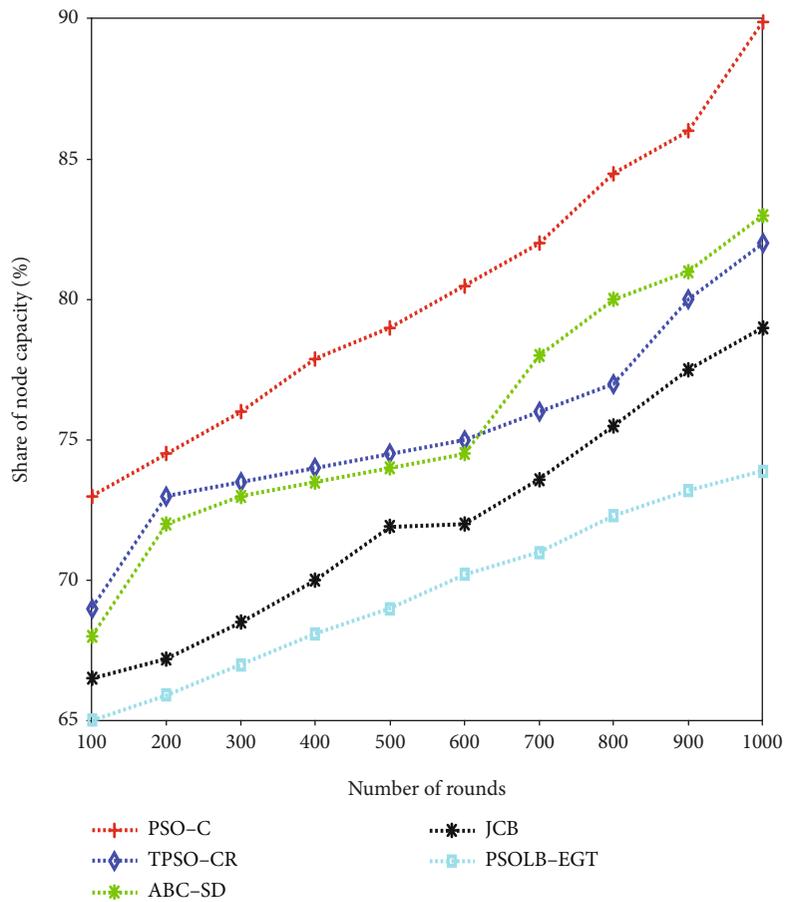


FIGURE 11: Comparison of average occupancy rates of different algorithm nodes.

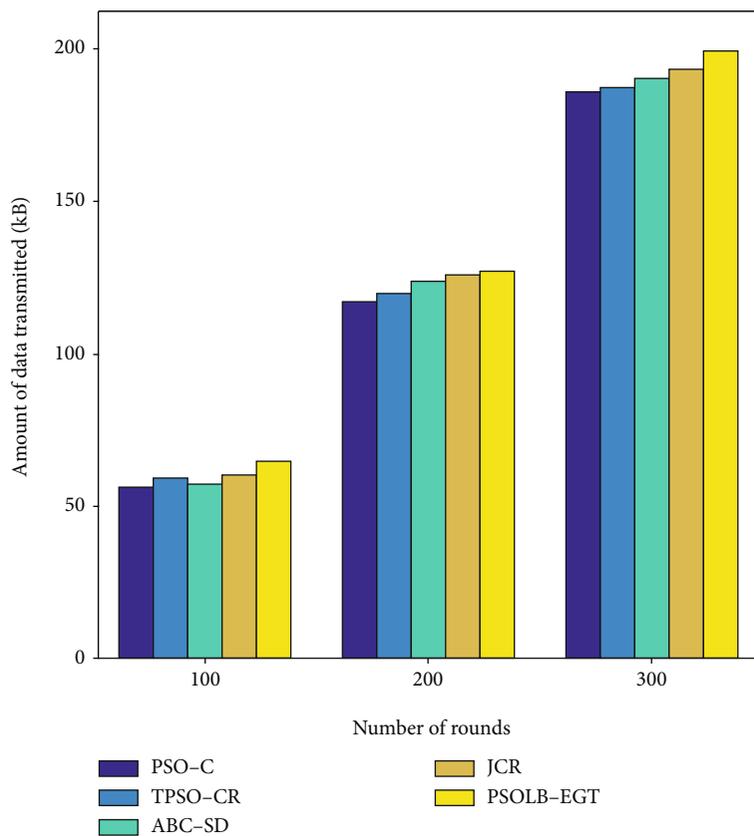


FIGURE 12: Comparison of data transmission amount under the same cycle condition.

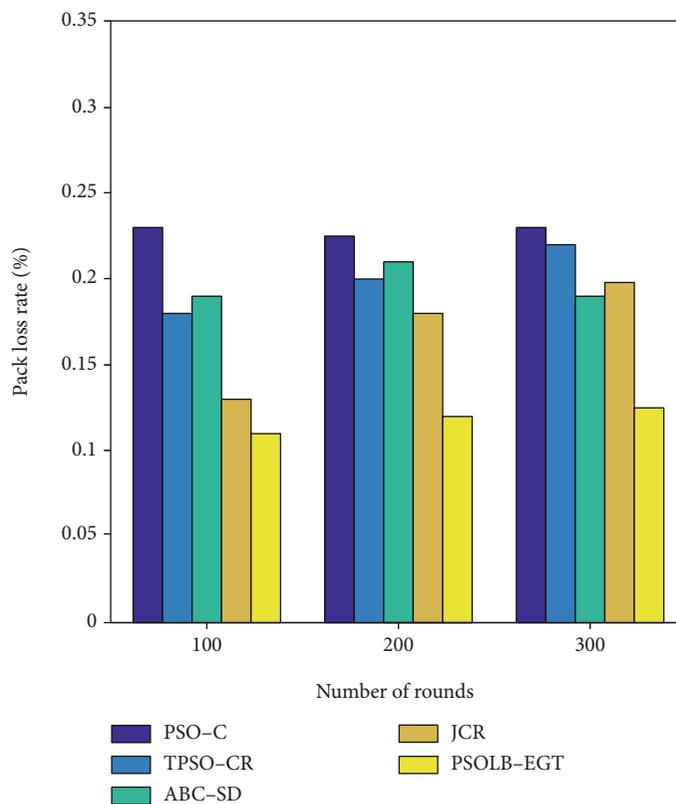


FIGURE 13: Pack loss rate.

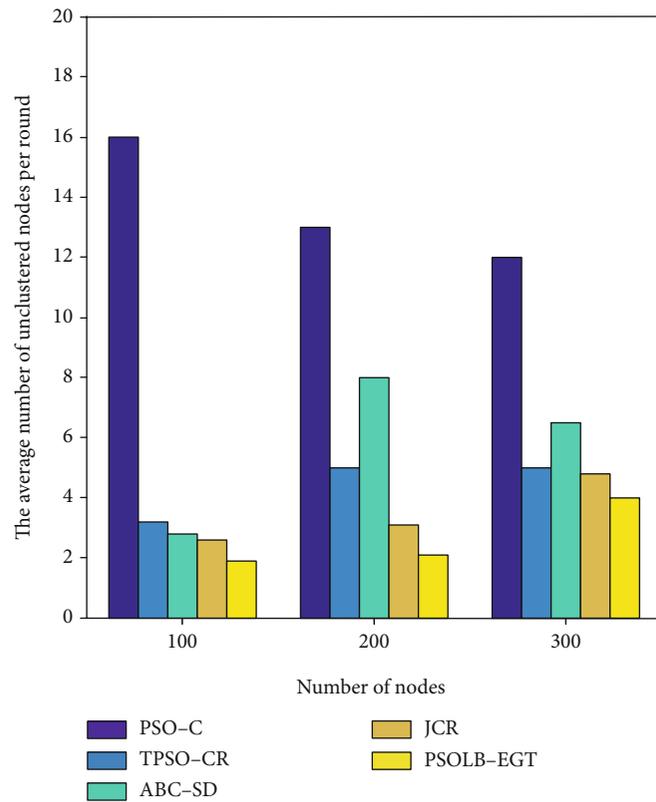


FIGURE 14: Amount of unclustered sensors.

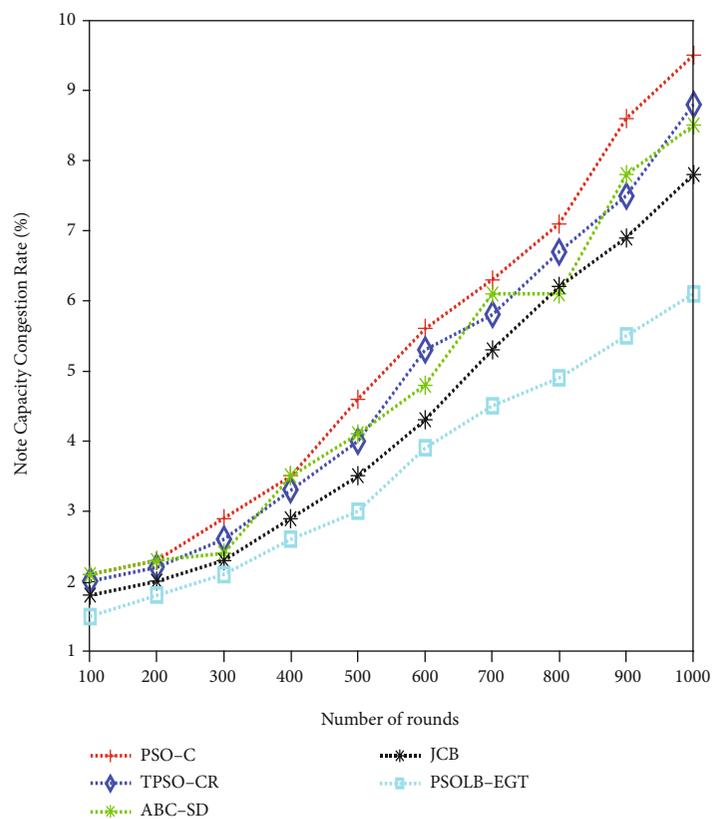


FIGURE 15: Node capacity congestion rate.

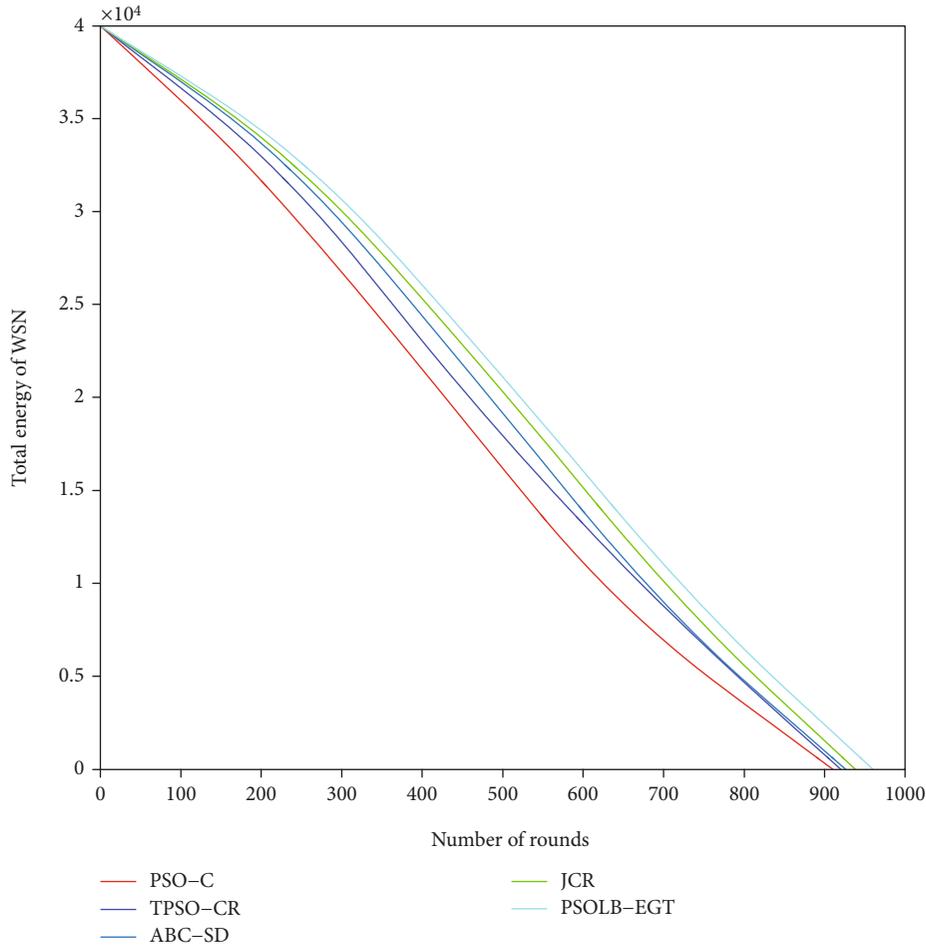


FIGURE 16: Comparison of total WSN energy consumption after using five clustering algorithms plus greedy routing.

effect is that we use adaptive learning factors and optimization measures for the last round of optimal results. However, it can be seen that when the clustering algorithms of JCR and PSOLB use the greedy routing algorithm, the energy consumption of the entire WSN is accelerated.

5. Conclusion

In this paper, we use the improved PSO algorithm and EGT, respectively, to solve two well-known optimization problems of WSN, namely, selection of CHs and routing between the CHs and the receiver. Then, we propose a clustering and routing protocol called PSOLB-EGT. The protocol incorporates an improved CH selection algorithm based on PSO search, which has a better fitness function. Next, an improved routing algorithm based on EGT is proposed to transmit aggregate data from CHs to the receiver for large-scale WSNs. This algorithm uses a novel routing function. The simulation results show that the proposed protocol is superior to existing protocols in the network life cycle, network coverage, and packet transmission capacity [45].

The existing research still has many shortcomings to be further addressed. The prospects for follow-up work include the following:

The routing protocols and positioning techniques studied in this paper are limited to two-dimensional plane space, but three-dimensional space is more in line with the actual application environments. It is important to study the routing and positioning technology of 3D wireless sensors. Therefore, it is necessary to extend the scope of research from two-dimensional space to three-dimensional space and perform further research on clustering routing technology for 3D WSNs.

In addition, a node needs to obtain its own global geographic information when clustering. Therefore, the proposed protocol requires the node to be equipped with a positioning device such as GPS, which increases the hardware requirements of the node to some extent.

The clustering routing algorithm proposed in this paper only assumes that the nodes in the cluster can synchronously send and receive data after receiving the cluster head broadcast and sending data packets to the CH. The corresponding synchronization mechanism is not designed. Therefore, the problem of how to design a reliable and practical mechanism that can synchronize communication between nodes and reduce the communication overhead is a subject that needs to be further studied in sensor network applications.

In the future, we will further optimize the PSOLB-EGT protocol according to the experimental results to improve

the energy saving effect of the protocol. We will study the energy saving optimization of data fusion, another important part of WSNs.

Appendix

Proof of Convergence

When the dynamic process of node routing replication continues, there will be a certain moment when the net utility of each cluster in the same group is the same, and the net utility at this moment is the average net utility of each individual in the group. That is, individual players no longer adjust their strategies to improve their net utility. To calculate the information of each participant and obtain the average utility of the group, it is necessary to set up a central control entity to maintain the information of the participants in the system and inform the participants of the average utility of their group at that time.

According to formulas (29) and (30), $X_i^N = 0$, that is, the equilibrium point of the evolutionary game.

$$\left[q_1 \frac{1}{E_{(Si)}} + q_2 \frac{E_{\text{residual-non-CH}_i}}{E_{\text{initial}}} - q_3 (n_r - 1)^{n_r - 1} \sigma \right] = \left(\sum_1^{n_r} \frac{n_{\text{to node } n_r} \times \delta_{n_r}}{N_{\text{all}}} \right). \quad (\text{A.1})$$

According to this equation (A.1), at this time, the fitness value of a single node is the average fitness value of all nodes in the WSN. That is, a single node no longer adjusts its own strategy to improve the fitness value, that is, the equilibrium point of the evolutionary game.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the Zigong City Key Science and Technology Program (No. 2019YYJC16), by the Artificial Intelligence Key Laboratory of Sichuan Province (Nos. 2017RYY05 and 2018RYJ03), and by the Horizontal Project (Nos. HX2017134 and HX2018264).

References

- [1] S. Yadav and R. S. Yadav, "A review on energy efficient protocols in wireless sensor networks," *Wireless Networks*, vol. 22, no. 1, pp. 335–350, 2016.
- [2] H. Jawad, R. Nordin, S. Gharghan, A. Jawad, and M. Ismail, "Energy-efficient wireless sensor networks for precision agriculture: a review," *Sensors*, vol. 17, no. 8, p. 1781, 2017.
- [3] W. Rehan, S. Fischer, and M. Rehan, "A critical review of surveys emphasizing on routing in wireless sensor networks-an anatomization under general survey design framework," *Sensors*, vol. 17, no. 8, p. 1713, 2017.
- [4] F. Yaghoubi, A. Abbasfar, and B. Maham, "Energy-efficient RSSI-based localization for wireless sensor networks," *IEEE Communications Letters*, vol. 17, no. 6, pp. 973–976, 2014.
- [5] H. Nabizadeh and M. Abbaspour, "IFRP: an intrusion/fault tolerant routing protocol for increasing resiliency and reliability in wireless sensor networks," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 14, no. 1, pp. 52–69, 2013.
- [6] Asaduzzaman and H. Y. Kong, "Energy efficient cooperative LEACH protocol for wireless sensor networks," *Journal of Communications and Networks*, vol. 12, no. 4, pp. 358–365, 2010.
- [7] F. Gao, H. Wen, L. Zhao, and Y. Chen, "Design and optimization of a cross-layer routing protocol for multi-hop wireless sensor networks," in *International Conference on Sensor Network Security Technology & Privacy Communication System IEEE*, Nangang, China, May 2013.
- [8] D. Gao, H. Lin, and X. Liu, "Routing protocol for k-anycast communication in rechargeable wireless sensor networks," *Computer Standards & Interfaces*, vol. 43, pp. 12–20, 2016.
- [9] S. Lee, Y. Noh, and K. Kim, "Key schemes for security enhanced TEEN routing protocol in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 9, no. 6, Article ID 391986, 2013.
- [10] Y. Ge, S. Wang, and J. Ma, "Optimization on TEEN routing protocol in cognitive wireless sensor network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, 27 pages, 2018.
- [11] Y. Zhou, N. Wang, and W. Xiang, "Clustering hierarchy protocol in wireless sensor networks using an improved PSO algorithm," *IEEE Access*, vol. 5, pp. 2241–2253, 2017.
- [12] D. R. Prasad, P. V. Naganjaneyulu, and K. S. Prasad, "Energy efficient clustering in multi-hop wireless sensor networks using differential evolutionary MOPSO," *Brazilian Archives of Biology & Technology*, vol. 59, no. spe2, pp. 1–16, 2016.
- [13] D. Lin, Q. Wang, D. Lin, and Y. Deng, "An energy-efficient clustering routing protocol based on evolutionary game theory in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 11, no. 11, 2015.
- [14] M. S. Kordafshari, A. Movaghar, and M. R. Meybodi, "A joint duty cycle scheduling and energy aware routing approach based on evolutionary game for wireless sensor networks," *Iranian Journal of Fuzzy Systems*, vol. 14, no. 2, pp. 23–44, 2017.
- [15] X. Liu, "A survey on clustering routing protocols in wireless sensor networks," *Sensors*, vol. 12, no. 8, pp. 11113–11153, 2012.
- [16] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados, "Energy-efficient routing protocols in wireless sensor networks: a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 551–591, 2013.
- [17] Y. Jin, L. Wang, Y. Kim, and X. Yang, "EEMC: an energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks," *Computer Networks the International Journal of Computer & Telecommunications Networking*, vol. 52, no. 3, pp. 542–562, 2008.
- [18] Y. Jin, L. Wang, Y. Kim, and X.-Z. Yang, "Energy efficient non-uniform clustering division scheme in wireless sensor networks," *Wireless Personal Communications*, vol. 45, no. 1, pp. 31–43, 2008.

- [19] K. Munusamy, "Distributed energy efficient clustering algorithm for wireless sensor networks," *Journal of Microelectronics, Electronic Components and Materials*, vol. 45, no. 3, pp. 180–187, 2015.
- [20] M. Nesa Sudha, M. L. Valarmathi, and A. S. Babu, "Energy efficient data transmission in automatic irrigation system using wireless sensor networks," *Computers & Electronics in Agriculture*, vol. 78, no. 2, pp. 215–221, 2011.
- [21] J. Hao, G. Duan, B. Zhang, and C. Li, "An energy-efficient on-demand multicast routing protocol for wireless ad hoc and sensor networks," in *GLOBECOM 2013 - 2013 IEEE Global Communications Conference*, pp. 4650–4655, Atlanta, GA, USA, December 2013.
- [22] X. L. He, X. W. Fu, and Y. S. Yang, "Energy-efficient trajectory planning algorithm based on multi-objective PSO for the mobile sink in wireless sensor networks," *IEEE Access*, vol. 7, pp. 176204–176217, 2019.
- [23] N. Gharaei, K. A. Bakar, S. Z. M. Hashim, and A. H. Pourasl, "Energy-efficient intra-cluster routing algorithm to enhance the coverage time of wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4501–4508, 2019.
- [24] F. Fanian and M. Kuchaki Rafsanjani, "Cluster-based routing protocols in wireless sensor networks: a survey based on methodology," *Journal of Network and Computer Applications*, vol. 142, pp. 111–142, 2019.
- [25] S. Chowdhury and C. Giri, "Energy and network balanced distributed clustering in wireless sensor network," *Wireless Personal Communications*, vol. 105, no. 3, pp. 1083–1109, 2019.
- [26] H. Fariborzi and M. Moghavvemi, "EAMTR: energy aware multi-tree routing for wireless sensor networks," *IET Communications*, vol. 3, no. 5, pp. 733–739, 2009.
- [27] J. Hong, J. Kook, S. Lee, D. Kwon, and S. Yi, "T-LEACH: the method of threshold-based cluster head replacement for wireless sensor networks," *Information Systems Frontiers*, vol. 11, no. 5, pp. 513–521, 2009.
- [28] S. Suganthi and S. P. Rajagopalan, "Multi-swarm particle swarm optimization for energy-effective clustering in wireless sensor networks," *Wireless Personal Communications*, vol. 94, no. 4, pp. 2487–2497, 2017.
- [29] J. Albath, M. Thakur, and S. Madria, "Energy constraint clustering algorithms for wireless sensor networks," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2512–2525, 2013.
- [30] Y. X. Peng, Y. H. Li, L. Shu, and W. B. Wang, "An energy-efficient clustered distributed coding for large-scale wireless sensor networks," *Journal of Supercomputing*, vol. 66, no. 2, pp. 649–669, 2013.
- [31] Q. W. Liu and M. D. Liu, "Energy-efficient clustering algorithm based on game theory for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 13, no. 11, 2017.
- [32] A. M. Shamsan Saleh, B. Mohd Ali, M. F. A. Rasid, and A. Ismail, "A self-optimizing scheme for energy balanced routing in wireless sensor networks using SensorAnt," *Sensors*, vol. 12, no. 8, pp. 11307–11333, 2012.
- [33] I. C. Paschalidis and R. M. Wu, "Robust maximum lifetime routing and energy allocation in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 8, no. 8, Article ID 523787, 2012.
- [34] A. A. M. Rahat, R. M. Everson, and J. E. Fieldsend, "Evolutionary multi-path routing for network lifetime and robustness in wireless sensor networks," *Ad Hoc Networks*, vol. 52, pp. 130–145, 2016.
- [35] K. Saleem, A. Derhab, M. A. Orgun et al., "Cost-effective encryption-based autonomous routing protocol for efficient and secure wireless sensor networks," *Sensors*, vol. 16, no. 4, p. 460, 2016.
- [36] R. S. Y. Elhabyan and M. C. E. Yagoub, "Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network," *Journal of Network and Computer Applications*, vol. 52, pp. 116–128, 2015.
- [37] A. A. A. Ari, B. O. Yenke, N. Labraoui, I. Damakoa, and A. Gueroui, "A power efficient cluster-based routing algorithm for wireless sensor networks: honeybees swarm intelligence based approach," *Journal of Network and Computer Applications*, vol. 69, pp. 77–97, 2016.
- [38] S. Tyagi and N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 623–645, 2013.
- [39] M. Sabet and H. Naji, "An energy efficient multi-level route-aware clustering algorithm for wireless sensor networks: a self-organized approach," *Computers & Electrical Engineering*, vol. 56, pp. 399–417, 2016.
- [40] M. Tekkalmaz and I. Korpeoglu, "Distributed power-source-aware routing in wireless sensor networks," *Wireless Network*, vol. 22, no. 4, pp. 1381–1399, 2016.
- [41] D. K. Sharma, S. K. Dhurandher, D. Agarwal, and K. Arora, "kROP: k-means clustering based routing protocol for opportunistic networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 4, pp. 1289–1306, 2019.
- [42] K. Haseeb, N. Islam, A. Almogren, and I. U. Din, "Intrusion prevention framework for secure routing in WSN-based mobile Internet of Things," *IEEE Access*, vol. 7, pp. 185496–185505, 2019.
- [43] K. Haseeb, N. Islam, T. Saba, A. Rehman, and Z. Mehmood, "LSDAR: a light-weight structure based data aggregation routing protocol with secure internet of things integrated next-generation sensor networks," *Sustainable Cities and Society*, vol. 54, article 101995, 2019.
- [44] K. A. Darabkh, M. Z. El-Yabroudi, and A. H. El-Mousa, "BPA-CRP: a balanced power-aware clustering and routing protocol for wireless sensor networks," *Ad Hoc Networks*, vol. 82, pp. 155–171, 2019.
- [45] A. Wang, W. B. Heinzelman, A. Sinha, and A. P. Chandrakasan, "Energy-scalable protocols for battery-operated micro-sensor networks," *Journal of Vlsi Signal Processing Systems for Signal Image & Video Technology*, vol. 29, no. 3, pp. 223–237, 2001.