Research Article

Secure Clustering Strategy Based on Improved Particle Swarm Optimization Algorithm in Internet of Things

Zhanbiao Bao (ID)

Center of Education Technology, Henan University of Economics and Law, Zhengzhou, Henan 450046, China

Correspondence should be addressed to Zhanbiao Bao; bao@huel.edu.cn

This paper proposes a secure clustering strategy based on improved particle swarm optimization (PSO) in the environment of the Internet of Things (IoT). First, in the process of cluster head election, by considering the residual energy and load balance of nodes, a new fitness function is established to evaluate and select better candidate cluster head nodes. Second, the optimized adaptive learning factor is used to adjust the location update speed of candidate cluster head nodes, expand the local search, and accelerate the convergence speed of global search. Finally, in the stage of forwarding node election and data transmission, in order to reduce the energy consumption of forwarding nodes, each cluster head node elects a forwarding node among the ordinary nodes in its cluster, so that the elected forwarding nodes have the optimal energy and location relationship. Experiments show that the proposed method effectively prolongs the network lifetime compared with the comparison methods. The average node degree of the proposed method is less than 2.5.

1. Introduction

With the development of information society, the information industry is more and more vigorous, and the scope of information application is more and more extensive [1]. Especially in the process of the big data era, the acquisition of information has attracted more and more attention, and become one of the important hotspots of scientific research and value-added industry [2–4]. Effective information acquisition is the basis of information transmission, processing, mining, and application. Sensors connect the physical world and the digital world by capturing and revealing the physical phenomena of the world and transforming them [5]. Wireless sensor network (WSN) is a self-organized network composed of a large number of sensor nodes with information collection, processing, and transmission functions. Each sensor node can realize the data collection, processing, and transmission of the sensing area, and the whole network can cooperate to complete the information collection in the target area [6–8]. Through the deployment and use of the network, the ability of human society

to obtain information can be enhanced. It has changed the situation that people only rely on their own feelings and simple tools to perceive information since ancient times and greatly improved the accuracy and sensitivity of human access to data and information [9, 10].

WSN is often used to monitor events of interest in the surrounding environment. A typical WSN monitoring system consists of sensor nodes, base stations, and monitoring centers [11–13]. The remote user or monitoring center can send data acquisition, processing, and transmission commands to the sensor nodes in the network through the base station, and finally, receive the interested data through the base station [14–16].

In the environment of IoT, aiming at the problems of large node energy consumption and short network life cycle caused by the unbalanced location distribution of cluster head nodes screened by clustering routing protocol of wireless sensor networks and unreasonable data transmission path of forwarding nodes, a secure clustering strategy based on improved PSO algorithm in the IoT environment is proposed. The innovations of the proposed method lie in the following:

(1) In the process of cluster head election, considering the residual energy and load balance of nodes, a new fitness function is established to evaluate and select better candidate cluster head nodes, which effectively shortens the communication distance.

(2) The optimized adaptive learning factor is used to adjust the position update speed of candidate cluster head nodes, and the concave function nonlinear decreasing strategy is used to improve the inertia weight, which speeds up the convergence speed of global search.

(3) For each cluster head node, a forwarding node is selected from the ordinary nodes in its cluster, so that the selected forwarding node has the optimal energy and location relationship, and the node communication overhead is reduced.

The remaining chapters of this paper are arranged as follows. Section 2 introduces the related research in this field; Section 3 introduces the system model; Section 4 introduces the proposed clustering routing algorithm based on improved PSO algorithm; Section 5 is the experiment, and Section 6 summarizes this study.

2. Related Works

Compared with plane routing, cluster routing saves the energy consumption of the network to the greatest extent, facilitates data fusion, and is easier to expand [17–19]. As a simple and effective random search algorithm, intelligent algorithm has good application prospects in finding the optimal solution of function and neural network recognition [20]. In the late 19th century, James Kennedy of the United States proposed PSO for the first time. It is famous for its simplicity and effectiveness. Therefore, scholars at home and abroad choose PSO algorithm in cluster head selection. Reference [21] defined a new cost function to minimize the communication distance between cluster members and cluster heads and used the PSO to select cluster heads with higher energy to optimize network clustering. Reference [22] proposed a new clustering mode EECS for wireless sensor networks, which specified the size of clusters in the network. Clusters far from the base station have a larger radius, and clusters close to the base station have a smaller radius, so as to balance the energy consumption of cluster heads. Reference [23] proposed a new hierarchical clustering protocol PSO-HC, which balanced the network energy consumption by reducing the number of cluster heads and transmitted the information collected by cluster heads to the base station by multihop. Based on the clustering of the LEACH protocol, reference [24] used the ant colony algorithm to optimize the distance location between each node and the base station and obtained the best path. Through this path, the cluster head transmitted the information to the next cluster head in sequence until the information was transmitted to the base station.

Reference [25] proposed an improved PSO algorithm to optimize the running time of the network. The algorithm comprehensively considered the factors of energy and distance to realize the selection of target nodes and proposed the introduction of relay nodes to share the transmission

work of cluster head nodes. Reference [26] combined harmony search (HS) algorithm and PSO to select cluster head nodes. The algorithm not only had the efficient search efficiency of harmony search algorithm but also had the dynamic ability of the PSO algorithm, which could better improve the performance of the network. According to the uneven clustering algorithm, reference [27] used ant colony optimization algorithm to calculate the optimal distance between the node and the base station and transmitted the data of each cluster head node to the base station through the selected optimal path. The disadvantage of this algorithm is that the path from each cluster head to the base station needs to be calculated, so it will lead to excessive energy consumption. Reference [28] proposed a clustering strategy based on FBECS (fuzzy based enhanced cluster head selection). In the clustering stage, fuzzy logic algorithm was used to calculate the qualification index of each node according to the residual energy of the node and the distance between the node density and the base station. The node generated the cluster head by comparing the randomly generated random number with the qualification index. Reference [29] proposed a clustering method based on energy center search using PSO (EC-PSO), which used geometric method to select CHs and searched energy centers in CHs selection.

3. System Model

3.1. Network Model. In order to better optimize the algorithm, this model makes the following assumptions.

(1) The nodes are randomly deployed in the target area. After deployment, the nodes will no longer move, and each node has a unique ID.

(2) All nodes have the same initial energy, computing power, and communication ability and can automatically adjust the transmission power.

(3) Each node can communicate with the base station independently, and the location of the base station is known.

(4) The node can calculate its relative position according to the angle of arrival and signal strength.

3.2. Energy Model. The same energy model as LEACH is used for communication between nodes. According to the distance $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ between node i and node j, the channel model is selected. The energy consumed by a node to send k bit data to the node with a distance of d is

$$E_{TX}(k,d) \begin{cases} k\left(E_{\text{elec}} + \alpha_{fs}d^2\right), d < d_0, \\ \\ k\left(E_{\text{elec}} + \alpha_{mp}d^2\right), d > d_0, \end{cases}$$

$$d_0 = \sqrt{\frac{\alpha_{fs}}{\alpha_{mp}}}. \tag{1}$$

The energy consumed by the node to receive and fuse k bit data is

$$E_{RX}(k) = kE_{elec},$$
$$E_{DA}(k) = kE_{da}, \qquad (2)$$

where E_{elec} is the energy consumed to receive or send 1 bit data and E_{da} is the energy consumed by the node to fuse 1 bit data. The transmission model selected by the node for communication is determined by d_0. α_{fs} and α_{mp} are the energy required for power amplification of free space channel model and multipath fading model. When the distance between nodes $d < d_0$, choose the free space channel model; otherwise, choose the multipath fading model.

4. Clustering Routing Algorithm Based on Improved PSO Algorithm

4.1. Cluster Head Election

4.1.1. Cluster Head Initialization. Because the calculation of the PSO algorithm is relatively complex, in order to avoid increasing the energy consumption of nodes in election calculation, the clustering and routing calculation based on the PSO algorithm will be completed by the base station with unlimited energy. In the initialization process, all nodes send their residual energy, location, and number information to the base station, and the base station receives and saves the information of each node. After the base station completes the clustering calculation based on the PSO algorithm, the base station broadcasts the calculation results. Each surviving node obtains the specific location information of the elected node and routing node according to the received broadcast information. PSO algorithm selects particles randomly, which is easy to fall into local optimization, resulting in uneven distribution of nodes in the cluster. To solve this problem, the energy of nodes participating in cluster head node election is limited.

Suppose there is N surviving nodes in the WSN, the energy of node i is $E(i)$, and the base station calculates the average residual energy of all nodes in the WSN as $E_a = 1/N \sum_{i=1}^{N} E(i)$.

In order to ensure that the selected cluster head nodes have sufficient energy to process the data in the cluster, the base station forms a set EA of all nodes with energy greater than or equal to E_a, then randomly selects K nodes from the EA, and stores them in the candidate cluster head node set to form a particle. After initially determining a group of candidate cluster head nodes, other noncluster head nodes join the cluster head nodes closest to them, respectively, to complete the establishment of initial clustering. Then, continue to select K nodes in EA in a random way, conduct a total of M times of screening, and finally generate the M groups of initial cluster head node set; that is, M particles, and form the M groups of clusters.

4.1.2. Fitness Function. The cluster head node is responsible for most activities and needs to consume more energy, which may lead to premature failure due to excessive energy consumption and reduce the performance of the whole network. Therefore, when selecting the cluster head node, it should consider the following two aspects: select the cluster head node with large residual energy and balance the load of the cluster head node at the same time.

First, consider selecting the cluster head node set with large residual total energy. Suppose that K is the number of cluster head nodes selected by the current particle. CH represents the cluster head set and CH_k represents the cluster head node of k cluster. The ratio f_1 of the current residual energy of all nodes to the residual energy of the cluster head node set is

$$f_1 = \frac{\sum_{i=1}^{N} E(i)}{\sum_{i=1}^{N} E(CH_k)}, \qquad (3)$$

where $E(i)$ represents the residual energy of the sensor node i in the particle.

After the current particle selects the cluster head node, the set CS_k is used to represent the member node set of the cluster k. C_k is used to count the number of member nodes in the cluster k. If the node j joins the cluster k, it can be expressed as $j \in CS_k$, and the C_k plus 1.

After clustering the current particles, it should further consider the load of the current cluster. The load of the cluster mainly includes the load in the cluster and the load between clusters. For the load balancing in the cluster, it is considered from two aspects: the number of member nodes in the cluster and the distance from the member nodes in the cluster to the cluster head.

Let f_2 represent the average of the deviation between s and the maximum distance between each cluster head node and its nodes in the cluster.

$$f_2 = \sum_{k=1}^{K} \frac{1}{K} \left| \max \left(\sum_{j=1}^{N} d(j, CH_k) - s \right) \right|. \qquad (4)$$

Let f_3 represent the difference between the number of member nodes in each cluster and the average number of members in the cluster. The smaller the difference, the closer the number of members between clusters, and the more effective it is to balance the energy consumption between cluster heads, so as to prolong the life cycle of the network:

$$f_3 = \sum_{k=1}^{K} \frac{1}{K} \left| \frac{N}{C_k} - K \right|, \qquad (5)$$

where $d(x, y)$ refers to the distance between node x and node y. r is the radius of clusters when the whole network area can be covered. Ideally, the area of each cluster is evenly distributed in the network area, and there is no coverage area between each cluster. The area of each cluster should be $A/(pro \times N)$, where A is the area of the whole network area, pro is the preset proportion of cluster heads and $pro \times N$ is the number of cluster heads, so the ideal r value is $\sqrt{A/(pro \times N)}$.

f_4 represents the average distance between cluster heads. The smaller the average distance between cluster heads, the smaller the load between clusters, which can reduce the communication energy consumption of the network.

$$f_4 = \frac{2\left(\sum_{k=1}^{K-1} \sum_{j=k+1}^{K} d\left(CH_k, CH_j\right)\right)}{\left((K-1)/K\right)}. \tag{6}$$

In the fitness function, the node residual energy and load balance are comprehensively considered, and the multi-objective optimization problem is transformed into a single objective optimization problem by means of weighted sum:

$$F = \alpha f_1 + \beta f_2 + \chi f_3 + \delta f_4, \tag{7}$$

where α, β, χ, and δ are weighting factors.

4.1.3. Update Method of Speed and Position.

According to the initial fitness calculation and the initially generated local optimal location and global optimal location, start a round of iterative calculation. First, update the location of the candidate cluster head node set, and then calculate the fitness of the candidate cluster head node after location update. In order to complete the location update of the candidate cluster head node set and obtain the optimization results, let the velocity components of the candidate cluster head nodes in the x and y directions be v_x and v_y, respectively, and the position components be x_x and x_y, respectively.

The calculation of the two velocity components is initially generated randomly, but in each subsequent iteration, it is determined according to the change relationship between the velocity component of the previous round of the candidate cluster head node set and the local optimal position (p_{xi}, p_{yi}) and the global optimal position (p_{xg}, p_{yg}). The specific calculation is as follows:

$$\begin{cases} v_x(t) = w v_x(t-1) + c_1 l_1 \left(p_{xi}(t-1) - x_{xi}(t-1)\right), \\ + c_2 l_2 \left(p_{xg}(t-1) - x_{xi}(t-1)\right), \\ v_y(t) = w v_y(t-1) + c_1 l_1 \left(p_{yi}(t-1) - x_{yi}(t-1)\right), \\ + c_2 l_2 \left(p_{yg}(t-1) - x_{yi}(t-1)\right), \end{cases} \tag{8}$$

where w is the inertia weight, which indicates the influence of the speed of the previous $t-1$ iteration of the candidate cluster head node set on the speed of the candidate cluster head node set in this t iteration. c_1 is a cognitive learning factor and c_2 is a social learning factor, which, respectively, represent the acceleration weight of the candidate cluster head node set close to the local optimal position and the global optimal position. $l_1, l_2 \in (0, 1)$ are the random numbers.

Based on the two velocity components, the calculation method of the position components $x_{xi}(t)$ and $x_{yi}(t)$ of the candidate cluster head node in the direction x and y is as follows:

$$\begin{aligned} x_{xi}(t) &= x_{xi}(t-1) + v_{xi}(t), \\ x_{yi}(t) &= x_{yi}(t-1) + v_{yi}(t). \end{aligned} \tag{9}$$

4.1.4. Improved Inertia Weight.

Among the parameters that can be adjusted by the PSO algorithm, researchers often improve the important parameter of inertia weight. A larger inertia weight value is conducive to improve the global search ability of the PSO algorithm, and a smaller weight value will enhance the motion ability of particles.

Under the same convergence accuracy, the concave function decreasing strategy greatly improves the convergence speed of PSO compared with other improved strategies, and can more effectively control the balance between global search and local search.

The rule for improving inertia weight by using concave function nonlinear decreasing strategy is as follows:

$$w = (w_{\max} - w_{\min})\left[\left(\frac{C_{\text{Cur}}}{C_{\text{Loop}}}\right)2 + 2\frac{C_{\text{Cur}}}{C_{\text{Loop}}}\right] + w_{\max}, \tag{10}$$

where $w \in [0.5, 0.9]$ and C_{Cur} represent the number of rounds of the current iteration and C_{Loop} represents the total number of iterations.

4.1.5. Location Mapping Strategy.

After each round of iteration, the location of the candidate cluster head node set will be updated, and the updated node location may not find a matching surviving node in the WSN. At this time, location mapping is needed. The basic idea is to adopt the proximity principle to map the updated location to the location of the nearest surviving node. Take x_{xi} and x_{yi} as the updated node coordinates and CM_{nx} and CM_{ny} as the coordinates of the surviving node CM_n in the network. The location mapping process is as follows:

$$CM_n$$

$$= \left\{ \left(CM_{nx}, CM_{ny}\right) \middle| \min \sqrt{\left(CM_{nx} - x_{xi}\right)^2 + \left(CM_{ny} - x_{yi}\right)^2} \right\}. \tag{11}$$

The location mapping strategy solves the mismatch between the updated location and the actual surviving node location caused by the discrete distribution of network nodes. When the updated position coordinates of multiple nodes are the same, it is necessary to set a flag while updating the node. When updating and mapping the positions of other nodes, first check whether the flag has been identified as the cluster head node. If so, select the next closest node position in turn for mapping.

After completing the location mapping, take the candidate cluster head node set after the location update as the optimization result, and calculate the fitness value of each candidate cluster head node set. According to the calculation results, the local optimal location of each group of candidate cluster head node set and the global optimal location of M groups of cluster head node sets in this round are updated. If the iteration is not finished, continue to update and map the location of candidate cluster head nodes. Otherwise, the candidate cluster head node set of the global optimal position is calculated as the optimal cluster head node set, and the cluster head election is completed.

After completing the cluster head election, the base station further calculates the distance from the noncluster head node to each cluster head node and adds the noncluster

head node to the nearest cluster head node to complete clustering. Ordinary nodes only send data. After receiving the data from ordinary nodes, the cluster head node performs data fusion and then sends it to the corresponding forwarding node. The forwarding node needs to be further selected from ordinary nodes. It receives the data from the cluster head, selects the optimal path to the base station, and sends the data to the base station.

4.2. Election and Data Transmission of Forwarding Nodes

4.2.1. Election of Forwarding Nodes. After the cluster head node election is completed, the corresponding forwarding node will be selected for the cluster head node. If there are too few forwarding nodes in the network, multiple cluster head nodes send data to the base station through a small number of forwarding nodes, which will increase the energy consumption of forwarding nodes. To solve this problem, the existing WSN routing protocol adopts the way that one cluster head node corresponds to one forwarding node to increase the number of forwarding nodes, but it selects all nodes randomly without considering whether the residual energy and location of the selected nodes are balanced.

In the forwarding node election, the improved PSO algorithm of the above cluster head node election is adopted to elect a forwarding node for each cluster head node among the ordinary nodes in its cluster, so that the elected forwarding node has the optimal energy and location relationship, and avoid speeding up its energy consumption due

to too few forwarding nodes in WSN. In the calculation and evaluation of forwarding nodes, because forwarding nodes can only be elected in one cluster, the calculation methods of energy factor and location balance factor are different from those of cluster head node election. Similarly, N represents the number of surviving nodes in the WSN. After selecting K cluster head nodes, the WSN is divided into K clusters. During initialization, according to the cluster head node screening method, candidate forwarding nodes with high residual energy are screened out in each cluster to form a set of candidate forwarding nodes. Each candidate set contains K nodes. After initialization, the number of ordinary nodes is $N - 2K$. Let $E_{RN}^r(i)$ represent the residual energy of candidate forwarding node RN_i in round r, and $E_{CN}^r(i)$ represent the residual energy of ordinary node CN_j in round r. The calculation formula of energy factor fit_1 is

$$\text{fit}_1 = \frac{(N - 2K)\sum_{j=1}^{N-2K} E_{RN}^r(i)}{K\sum_{j=1}^{N-2K} E_{CN}^r(i)}. \tag{12}$$

Let $d(CN_k, CH_j)$ represent the distance between the ordinary node CN_k and the corresponding cluster head node CH_j, $d(RN_i, BS)$ represent the distance between the forwarding node RN_i and the base station BS, $d(RN_i, CH_j)$ represent the distance between the forwarding node RN_i and the corresponding cluster head node CH_j, and $d(RN_i, RN_m)$ represent the distance between the forwarding nodes RN_i and RN_m. The calculation method of the location equalization factor fit_2 of the forwarding node is as follows:

$$\text{fit}_2 = \frac{K^2 \sum_{j=1}^{N-2K} d(CN_k, CH_j)}{(N - 2K)\left(\sum_{i=1}^{K} d(RN_i, BS) + \sum_{i=1}^{K} d(RN_i, CH_j) + \sum_{i=1}^{K-1} \sum_{m=i+1}^{K} d(RN_i, RN_m)\right)}, \tag{13}$$

where the ordinary node CN_k is located in the cluster where the cluster head node CN_k is located and the candidate forwarding node RN_i corresponds to the cluster head node CH_j. The closer the candidate forwarding node set is to the base station, the smaller the distance between the cluster head node and the forwarding node in each cluster; the smaller the distance between the forwarding nodes in the candidate forwarding node set, the more balanced the location distribution of the candidate forwarding node set, and the greater the fit_2 value.

Based on the ability factor and location balance factor of candidate forwarding node election, the fitness function F2 for the evaluation of candidate forwarding node set constructed by weighting method is as follows:

$$F_2 = b\text{fit}_2 + (1 - b)\text{fit}_2, \tag{14}$$

where $b \in (0, 1]$ is the weight. The higher the residual energy of the candidate forwarding node set and the more balanced the location, the greater the fitness value of the candidate forwarding node set, indicating that the candidate forwarding node set is better. In the iterative election process of candidate forwarding node set, the

speed update and location mapping method similar to cluster head node election is used to select the optimal forwarding node set.

4.2.2. Data Transmission. WSN forward routing mainly selects relay nodes for cluster heads far away from the base station, and relay nodes reduce the load of cluster heads by forwarding data packets from cluster heads. Assuming that the data packet needs to be sent to the base station through n hop; that is, through $n - 1$ relay nodes, the optimal relay node should have the following characteristics:

(1) The distance between relay nodes should be as balanced as possible.

(2) The relay node should have high energy to meet the data transmission requirements.

(3) The total distance should be as close as possible to the direct communication distance between the cluster head and the base station.

(4) The selection of relay nodes should avoid selecting other cluster heads.

(5) The load of relay nodes should be as small as possible to avoid over development.

According to the above requirements, the corresponding fitness function is designed, which can be expressed as

$$F_3 = \varepsilon\left(\frac{E_{\min}}{E_{\text{init}}} + \frac{1}{\text{load}_{\max}}\right) + \phi\left(\frac{d_{\min}}{d_{\max}} + \frac{d_{\text{init}}}{d_{\text{total}}}\right), \qquad (15)$$

where $E_{\min}$ is the node with the smallest energy among the relay nodes; the load of the node is determined by the number of data packets received by the node. $\text{load}_{\max}$ is the node with the largest load in the relay nodes; $d_{\min}$ and $d_{\max}$ are the minimum and maximum distance between adjacent nodes in the forward routing path, respectively; d_{init} and d_{total} are the sum of the distance between the cluster head and the base station and the sum of distance of the whole forward routing path; $\varepsilon = 0.5 + 0.5E_{\text{loss}}/E_{\text{total}}$; $\phi = 1 - \varepsilon$. The total energy consumption of the nodes in the cluster is changed into the energy consumed by the cluster head that needs to plan the route, and the total energy in the cluster is changed into the initial energy of the cluster head. With the operation of the network, the parameter weight ratio is dynamically adjusted.

5. Experiments and Analysis

5.1. Experimental Setup. Simulating the generation of WSN in MATLAB and setting the base station outside the network, in this environment, the secure clustering strategy based on improved PSO algorithm is tested. The experimental computer is configured with 2.3 GHz, Intel Core I7, 8 GB RAM, and 64 bit Windows10. The experimental environment and parameter settings related to WSN initialization, cluster head node, forwarding node election, and data processing are shown in Table 1.

5.2. Performance Index Comparison of Topology Generated by Different Schemes. In order to prove the performance of the proposed method under the same experimental conditions, the proposed method is compared with the methods in references [28, 29], and the comparison results are shown in Figure 1. In order to compare the algorithms effectively without losing generality, it is assumed that n sensor nodes are randomly deployed in the target monitoring area of 200×200.

As can be seen from Figure 1(a), the average link length of the three methods continues to decrease with the growth of n. The proposed method can maintain the global connectivity of the network with a short average link length and has strong advantages in reducing the overall energy consumption of the network.

Figure 1(b) shows the comparison results of the average node degree of the three control schemes. It can be seen from the figure that the average node degree of the other two methods tends to rise with the increase of the network scale, while the topology obtained by the proposed method has a low average node degree, which is basically maintained below 2.5, and the average node degree is relatively stable

TABLE 1: Experimental parameters setting.

Parameter name	Value
Number of nodes	150
Base station location	(50,100)
Network scale	200 m × 200 m
E_{fs} (pJ.(bit·m^2)$^{-1}$)	10
E_{mp} (pJ.(bit·m^4)$^{-1}$)	0.002
E_{elec} (nJ.bit^{-1})	60
$d0$ (m)	70
Maximum number of rounds	2500
The number of iterations of the algorithm	200
Optimal number of clusters	10
Number of particle	20
Data packet length/bit	5000
Control packet length/bit	100
Node initial energy $E0$ (J)	0.7

and will not change greatly with the change of the network scale. It shows that this method has good stability and low interference.

As can be seen from Figure 1(c), when the network size is 70, the average number of joint points of the methods in references [28, 29] is 3.0 and 2.3, respectively, and the number of joint points of the proposed method is 0. The number of joint points of the methods in references [28, 29] decreases with the continuous growth of the network size, but the method in this paper has better stability and can ensure that there are basically no joint points under different network sizes. This is because the proposed method comprehensively considers the residual energy and load balance of nodes, establishes a new fitness function, and can select better candidate cluster head nodes. Therefore, compared with the other two methods, the proposed method can better ensure the connectivity of topology when clusters are connected.

5.3. Location Balance Test of Cluster Head Nodes. The location balance factor reflects the distribution balance of cluster head node set and forwarding node set in WSN. The more balanced the node distribution is, the smaller the total communication distance in WSN communication is, and the lower the node energy consumption will be. Based on the above experimental environment, the total distance from all ordinary nodes to cluster heads and the total distance from all cluster head nodes to base stations in multiple iterations of the methods proposed in this paper, references [28, 29] are tested, respectively. The experimental results are shown in Figures 2 and 3, respectively.

As can be seen from Figures 2 and 3, with the increase of the number of iterations, the total distance from all ordinary nodes to cluster heads and the total distance from all cluster heads to base stations in the three methods become smaller and smaller. In the method of reference [28], the energy and location of nodes are not considered in cluster head election. Initially, the two distances in Figures 2 and 3 are the largest, but the two distances drop rapidly after 2600 rounds. This is because their communication distance is large, energy consumption is large and unbalanced, and node death is fast.
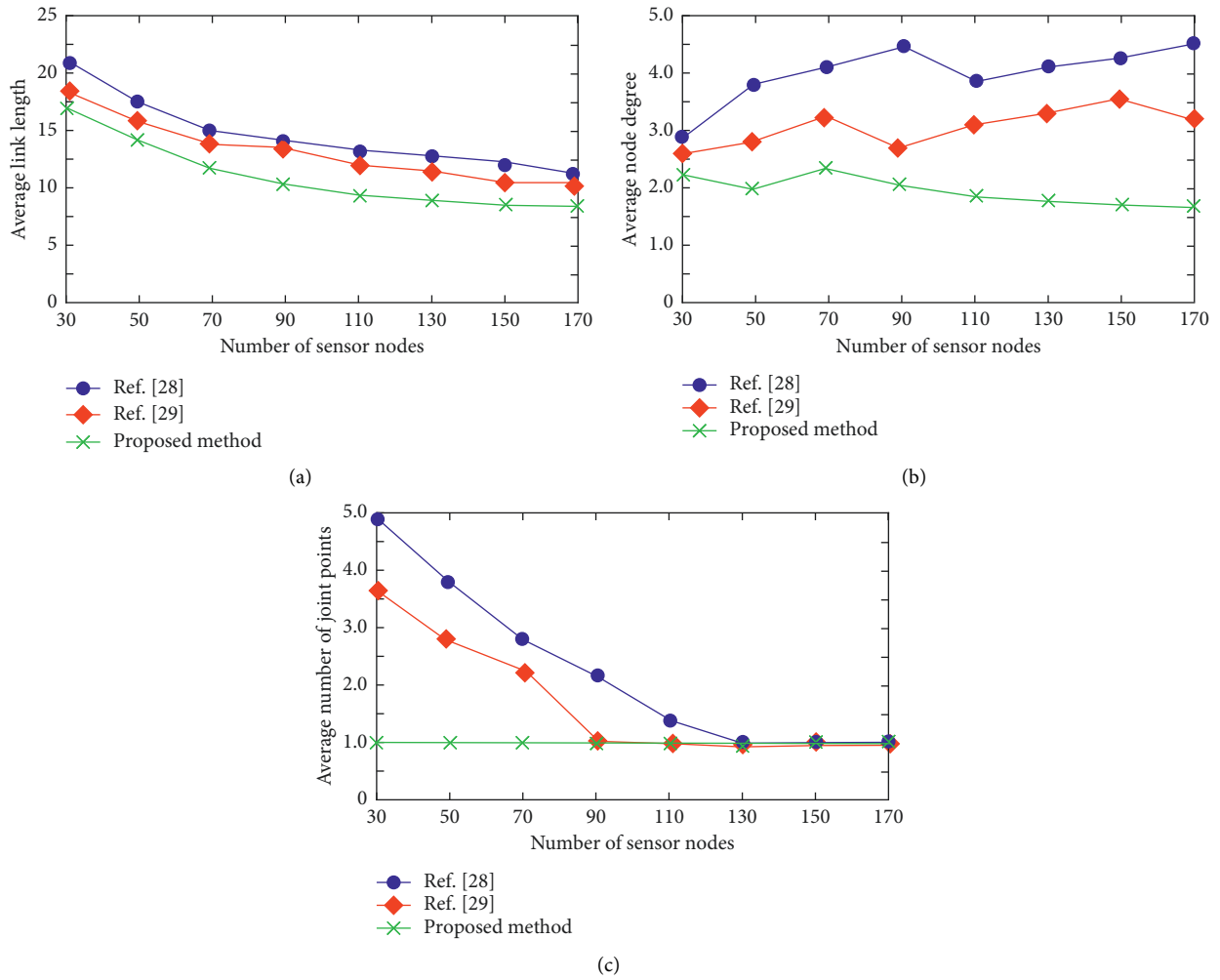
(a)

(b)

(c)

FIGURE 1: Comparison of performance indices of topologies generated by different schemes: (a) comparison of average link length; (b) comparison of average node degree; (c) comparison of average number of joints.
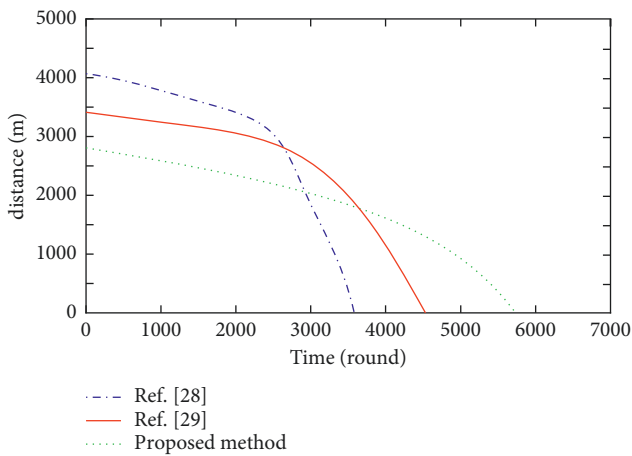

FIGURE 2: Sum of distances from ordinary nodes to cluster heads in different methods.
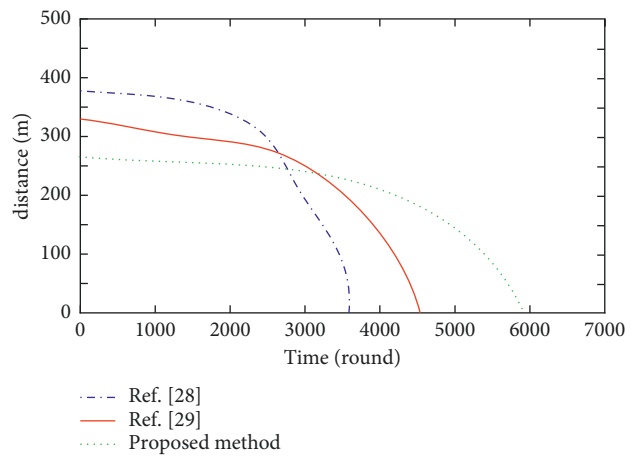

FIGURE 3: Sum of distances between cluster heads and base station in different methods.

With the large reduction of surviving nodes, their communication distance decreases, and the life cycle of WSN is greatly shortened. This is because, the proposed method elects a forwarding node for each cluster head node among the ordinary nodes in its cluster so that the selected forwarding node has the optimal energy and location relationship, which effectively reduces and balances the node energy consumption. The number of dead nodes decreases linearly and slowly, the change is relatively stable and prolongs the life cycle of WSN.

The method in reference [29] considers the location information. At the beginning, the two distances are relatively reduced, the energy consumption is reduced, and the life cycle of WSN is prolonged. The optimization of the inertia weight in the method of reference [29] makes the elected cluster head nodes close to the global optimization. The two distances measured at the beginning are lower than those in the method of reference [28], the node energy consumption is reduced, and the life cycle of WSN is longer than that in the method of reference [28]. The methods in references [28, 29] have fewer surviving nodes due to the proliferation of dead nodes in the later stage of iteration. The measured two distances are smaller than the proposed method, but the life cycle is smaller.

6. Conclusion

A secure clustering strategy based on improved PSO algorithm in the environment of IoT is proposed. Considering the residual energy and load balance of nodes, a new fitness function is established, the better candidate cluster head nodes are evaluated and selected. The location update speed of candidate cluster head nodes is adjusted through the optimized adaptive learning factor, so as to elect a forwarding node for each cluster head node among the ordinary nodes in its cluster for reducing the energy consumption of forwarding nodes. Experiments show that under the same conditions, the average node degree of the proposed method is less than 2.5, which is better than the comparison methods. Compared with the comparison methods, the proposed method can significantly improve the network life cycle of clustering method and optimize node energy consumption.

In WSN clustering routing protocol, the selected cluster head nodes usually accounts for only a small part of the total nodes, so it has sparse characteristics. In future research work, sparse theory can be introduced to solve the problem of cluster head node selection in WSN clustering routing. In addition, in practical applications, the location of sensor nodes is not fixed, so energy optimization can be studied for large-scale dynamic networks.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

References

[1] N. Li, J. Chen, and Y. Yuan, "A Wi-Fi indoor localization strategy using Particle Swarm Optimization based artificial neural networks[J]," *International Journal of Distributed Sensor Networks*, vol. 12, no. 3, pp. 458–467, 2016.

[2] C. Li and S. Yang, "A Clustering Particle Swarm Optimizer for Dynamic optimization[C]," in *Proceedings of the 2009 IEEE congress on Evolutionary Computation*, pp. 439–446, IEEE, Trondheim, Norway, 2009.

[3] X. Liang, W. Li, Y. Zhang, and M Zhou, "An adaptive particle swarm optimization method based on clustering," *Soft Computing*, vol. 19, no. 2, pp. 431–448, 2015.

[4] G. T. Pulido and C. A. Coello, "Using clustering techniques to improve the performance of a multi-objective particle swarm optimizer," in *Proceedings of the Genetic and Evolutionary Computation - GECCO 2004*, pp. 225–237, Springer, Berlin, Heidelberg, 2004.

[5] A. A. A. Esmin, R. A. Coelho, and S. Matwin, "A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data," *Artificial Intelligence Review*, vol. 44, no. 1, pp. 23–45, 2015.

[6] S. Yang and C. Li, "A clustering particle swarm optimizer for locating and tracking multiple optima in dynamic environments," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 6, pp. 959–974, 2010.

[7] C. L. Huang, W. C. Huang, H. Y. Chang, Y. C. Yeh, and C. Y Tsai, "Hybridization strategies for continuous ant colony optimization and particle swarm optimization applied to data clustering," *Applied Soft Computing*, vol. 13, no. 9, pp. 3864–3872, 2013.

[8] B. Jarboui, M. Cheikh, P. Siarry, and A Rebai, "Combinatorial particle swarm optimization (CPSO) for partitional clustering problem," *Applied Mathematics and Computation*, vol. 192, no. 2, pp. 337–345, 2007.

[9] S. Paterlini and T. Krink, "Differential evolution and particle swarm optimisation in partitional clustering," *Computational Statistics & Data Analysis*, vol. 50, no. 5, pp. 1220–1247, 2006.

[10] M. Omran, A. P. Engelbrecht, and A. Salman, "Particle swarm optimization method for image clustering," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 03, pp. 297–321, 2005.

[11] S. Rana, S. Jasola, and R. Kumar, "A review on particle swarm optimization algorithms and their applications to data clustering," *Artificial Intelligence Review*, vol. 35, no. 3, pp. 211–222, 2011.

[12] L. Y. Chuang, C. J. Hsiao, and C. H. Yang, "Chaotic particle swarm optimization for data clustering," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14555–14563, 2011.

[13] Q. Cai, M. Gong, L. Ma, S. Ruan, F. Yuan, and L Jiao, "Greedy discrete particle swarm optimization for large-scale social network clustering," *Information Sciences*, vol. 316, no. 8, pp. 503–516, 2015.

[14] X. Cui, T. E. Potok, and P. Palathingal, "Document Clustering Using Particle Swarm Optimization[C]," in *Proceedings of the 2005 IEEE Swarm Intelligence Symposium*, pp. 185–191, IEEE, Pasadena, CA, USA, 2005.

[15] M. G. H. Omran, A. Salman, and A. P. Engelbrecht, "Dynamic clustering using particle swarm optimization with application

in image segmentation," *Pattern Analysis & Applications*, vol. 8, no. 4, pp. 332–344, 2006.

[16] Y. Lu, S. Wang, S. Li, and C Zhou, "Particle swarm optimizer for variable weighting in clustering high-dimensional data," *Machine Learning*, vol. 82, no. 1, pp. 43–70, 2011.

[17] R. Kuo and L. M. Lin, "Application of a hybrid of genetic algorithm and particle swarm optimization algorithm for order clustering," *Decision Support Systems*, vol. 49, no. 4, pp. 451–462, 2010.

[18] G. Kannan and T. Sree Renga Raja, "Energy efficient distributed cluster head scheduling scheme for two tiered wireless sensor network," *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 167–174, 2015.

[19] J. Wang, Y. Cao, B. Li, Hj Kim, and S Lee, "Particle swarm optimization based clustering algorithm with mobile sink for WSNs," *Future Generation Computer Systems*, vol. 76, no. 6, pp. 452–457, 2017.

[20] R. Liu, Y. Chen, L. Jiao, and Y Li, "A particle swarm optimization based simultaneous learning framework for clustering and classification," *Pattern Recognition*, vol. 47, no. 6, pp. 2143–2152, 2014.

[21] P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: particle swarm optimization approach," *Engineering Applications of Artificial Intelligence*, vol. 33, no. 1, pp. 127–140, 2014.

[22] M. Ye, C. Li, and G. Chen, "EECS: an energy efficient clustering scheme in wireless sensor networks[C]," in *Proceedings of the PCCC 2005. 24th IEEE International Performance, Computing, and Communications Conference 2005*, pp. 535–540, IEEE, Phoenix, AZ, USA, 2005.

[23] R. S. Elhabyan and M. C. E Yagoub, "Particle swarm optimization protocol for hierarchical clustering in Wireless Sensor Networks[C]," in *Proceedings of the 10th IEEE international conference on collaborative computing: networking, applications and worksharing*, pp. 417–424, IEEE, Miami, FL, USA, 2014.

[24] A. Salehpour, B. Mirmobin, and A. Afzali-Kusha, "An Energy Efficient Routing Protocol for Cluster-Based Wireless Sensor Networks Using Ant colony optimization[C]," in *Proceedings of the 2008 International Conference on Innovations in Information Technology*, pp. 455–459, IEEE, Al Ain, United Arab Emirates, 2008.

[25] Y. Zhou, N. Wang, and W. Xiang, "Clustering hierarchy protocol in wireless sensor networks using an improved PSO algorithm," *IEEE Access*, vol. 5, no. 3, pp. 2241–2253, 2017.

[26] T. Shankar, S. Shanmugavel, and A. Rajesh, "Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks," *Swarm and Evolutionary Computation*, vol. 30, no. 4, pp. 1–10, 2016.

[27] J. Du and L. Wang, "Uneven Clustering Routing Algorithm for Wireless Sensor Networks Based on Ant colony optimization[C]," in *Proceedings of the 3rd IEEE International Conference on Computer Research and Development*, pp. 67–71, IEEE, Shanghai, China, 2011.

[28] P. S. Mehra, M. N. Doja, and B. Alam, "Fuzzy based enhanced cluster head selection (FBECS) for WSN," *Journal of King Saud University Science*, vol. 32, no. 1, pp. 390–401, 2020.

[29] J. Wang, Y. Gao, W. Liu, A. Sangaiah, and H. J Kim, "An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network," *Sensors*, vol. 19, no. 3, p. 671, 2019.