Eng Int Syst (2020) 2: 89–98

Engineering
Intelligent Systems

Cluster Head Selection Algorithm and Simulation Based on Energy Change in Internet of Things Wireless Sensor Networks

Jingwei Xie* **and Huiyi Li**

Hunan Mass Media College, 410100, China

Internet of Things is based on the Internet and connects objects and objects to achieve information sharing. Wireless sensor network (WSN) is multi-hop ad-hoc network systems composed of a large number of sensor nodes deployed in the monitoring area, which is an important technical of the underlying network of the Internet of Things. The traditional low energy adaptive clustering hierarchy (LEACH) protocol for WSN has the disadvantages of large cluster head overhead, uneven cluster size distribution, limited energy and real-time replenishment. Therefore, how to balance energy consumption and extend network life have become the key points and difficulties of designing wireless sensor and network routing protocols. In this paper, based on summarizing and analyzing the traditional LEACH protocol and the improvement of this algorithm, a cluster head selection algorithm based on energy change is proposed. The algorithm calculates the ideal cluster radius by considering the trust level of the node and the residual energy level when electing the cluster head. This ensures that the number of cluster heads in each round is within the expected range; then the distance weight, neighbor the number of nodes and the size of the energy consumption are optimized to complete the competitive selection of the cluster head. The simulation results show that the proposed algorithm can further reduce energy consumption, extend the life time of the network.

Keywords: Cluster head selection algorithm; Wireless sensor networks; Internet of things; Algorithm simulation

1. INTRODUCTION

In recent years, wireless sensor networks (WSN) has been widely used in many Internet of Things fields such as military, environmental observation and forecasting, medical care, smart home, and building state detection[1]. WSN is a distributed sensor network based on wireless sensors; it is characterized by small size, low power consumption, and high integration. It combines wireless sensing technology, embedded technology and distributed processing technology to monitor objects and collect object data and information in real time. WSN needs to sense, collect and transmit object data, and their nodes have the dual characteristics of perception and reception. The WSN realizes the collection and transmission of network information through sensor nodes. The Internet of Things can realize the information transmission between objects. The applications of WSN ensure the security of information transmission. Therefore, effective measures are needed to improve the security of WSN.

At present, several WSN routing algorithms have been proposed [2–8]: LEACH algorithm is a traditional low-energy adaptive clustering routing algorithm. The cluster head node and the base station adopt a single-hop routing communication protocol.The energy consumption of the cluster head far away from the base station is too large and prematurely dies, affecting the life cycle of the entire network, and the single-hop transmission is not easy to expand the WSN; The basic idea of the protocol is that the choice of the cluster head is mainly based on the primary and secondary parameters. The residual energy of the node is taken as the condition, and the

*e-mail:43911919@qq.com

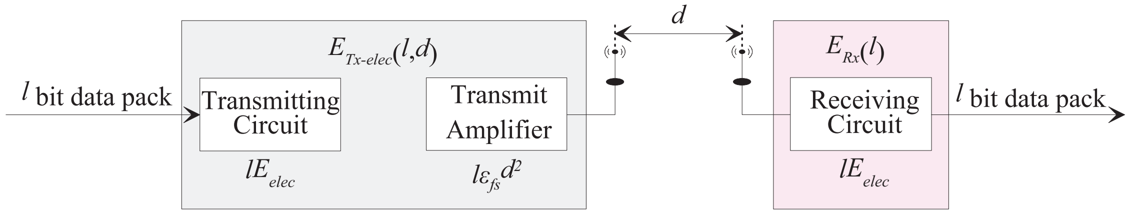

Figure 1 First-order radio model

secondary parameter is the main parameter of the cluster head. The secondary parameter depends on the communication cost within the cluster. For the LEACH algorithm, the cluster head far away from the base station is prematurely dying due to excessive energy consumption of the transmitted data. The predecessor adopted a multi-hop communication method, but the cluster head closer to the base station bears more data forwarding. The task consumes too much energy and is easy to form a hot zone. The cluster head selects a robust member node among the member nodes in the cluster. When the cluster head energy is insufficient or invalid, the robust node becomes a new cluster head to ensure the normal operation of the network. However, this robust node is just a spare node and does not really take on the work of the cluster head. The clustering algorithm generates a cluster head node and a member node in the cluster through a certain mechanism. The cluster head node coordinates and manages the work of the member nodes in the cluster, and is responsible for the collection, fusion and forwarding of information within the cluster. Many protocols have problems that are not further optimized after the number of cluster heads and cluster heads are selected.

In this paper, based on the classical LEACH protocol and the improvement of this algorithm, a cluster head selection algorithm based on energy change is proposed. In the research of WSN, energy efficiency has always been a hot issue. In order to extend the network life cycle, related software and hardware have been extensively studied. The clustering algorithm can effectively extend the network life cycle. In the clustering algorithm, the choice of cluster head has a great impact on the network energy consumption. The algorithm proposed in this paper considers the trust level of the node and the residual energy to calculate the ideal cluster radius when electing the cluster head. This ensures that the number of cluster heads in each round is within the expected range; Moreover, the number of neighbor nodes and the size of the energy consumption are optimized to complete the competitive selection of the cluster head. The simulation results show that the algorithm can further reduce energy consumption, and extend the life time of the network.

2. TRADITIONAL LEACH ROUTING PRINCIPLE

The LEACH protocol is a traditional WSN clustering protocol that consists of two main steps: an initialization phase and a steady state phase, where the initialization phase includes a cluster head selection portion and a clustered portion. Once the selected cluster head node receives the join message, a time division multiple access (TDMA) schedule is established to perform collision-free information interaction with the non-cluster head node. In LEACH, the cluster head node is selected based on the probability value of the sensor node, so the number of cluster head nodes and their distribution cannot be guaranteed. Therefore, based on the consideration of optimizing energy consumption, The cluster head nodes are not selected randomly.

When the cluster head is elected, each node will generate a random number in the interval [0, 1], and set a threshold $T(n)$, and if the number is less than $T(n)$, it will issue itself as a cluster head announcement [2].

$$T(n) = \begin{cases} \frac{p}{1-p\left(r \cdot mod\left(\frac{1}{p}\right)\right)}, n \in G \\ 0, n \notin G \end{cases} \tag{1}$$

Where p is the percentage of the cluster head node in all nodes, r is the round of the current cluster head election, and G is a set of nodes, including all of the most recent $1/p$ cluster head elections that have not been elected as the node of the cluster head.

The LEACH protocol uses a first-order radio model, as shown in Figure 1 [9]. In this model, the energy consumed by the transmitting circuit and the receiving circuit for wireless communication is E_{elec}, the free space fading channel mode amplification index is E_{fs}, the multipath fading channel mode amplification index is E_mp.

The energy consumed by the sensor node in the wireless transmission process is composed of two parts: the transmission circuit loss and the power amplification loss, that is, the energy consumed by transmitting the 1-bit data over the distance d is [10]:

$$\begin{aligned} E_{T\mathrm{X}\text{-}elec}(l,d) &= E_{Tx-elec}(l) + E_{Tx-amp}(l,d) \\ &= \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, d < d_0 \\ lE_{elec} + l\varepsilon_{fs}d^4, d \geqslant d_0 \end{cases} \end{aligned} \tag{2}$$

The sensor node receives 1 bit of data and the energy consumed is:

$$E_{Rx}(l) = lE_{elec} \tag{3}$$

Where d_0 is the distance threshold; $\varepsilon_{fs}d^2$ is the power amplifier power consumption using the free-space channel model when the transmission distance is less than d_0; and $\varepsilon_{mp}d^4$ is the power of the multipath fading channel model when the distance is greater than or equal to d_0 Amplifier power consumption.

It can be seen from the above wireless communication model that when the transmission distance exceeds a certain value, the energy consumption suddenly increases. The large transmission distance introduces much invalid information, so most of the energy is consumed for invalid information. If the distance that the node sends data is limited to an appropriate range, to reduce the amount of invalid information, which can save a lot of energy consumption, and extend the life of the WSN.

The basic idea of the LEACH algorithm is to randomly select the cluster heads by strictly equal probability methods, so that the energy load of the entire network is evenly distributed to each sensor node, to achieve the purpose of reducing network energy consumption and extending the network lifetime. However, the LEACH algorithm has some shortcomings. First, when selecting the cluster head node, the current residual energy of the node is not considered. To hold the cluster head strictly equal probability, the probability of the cluster head node with large energy gap is exactly the same. Thus, there is residual energy at the node. Secondly, the LEACH algorithm uses a mechanism for randomly generating cluster heads, and does not limit the positional relationship between the generated clusters, thus possibly causing multiple cluster heads [11].

3. NETWORK SYSTEM MODEL

3.1 Network Model

It is assumed that Na sensor nodes are deployed in a square area, and each node obeys a Gaussian distribution around the center of the area; all nodes have the same structure, and each node has a unique identifier (ID); the node periodically send the collected data to the sink node. The sensing area of the sink node has a certain distance and is responsible for receiving data and processing; all nodes are stationary and their position can be known; the network has narrow bandwidth and flat Rayleigh fading channel and transmission reception completely synchronization [12].

Considering a multi-hop sensor network in which n nodes are randomly distributed in the sensing area of $A \times B$. At the beginning of the deployment, each sensor node has certain data processing and communication capabilities. Suppose S is the set of all sensor nodes in the network, as shown[13]:

$$\mathrm{S} = \{S_1, S_2, \cdots, S_i\} \tag{4}$$

Where: S_i represents the $i-$th sensor node.

At the same time, the entire network is divided into grids $a_c \times b_c$ that are uniquely overlapped. Suppose Q represents all grid sets in the network, as in equation (5):

$$Q = \{Q_i | j = 1, 2, \cdots, q\} \tag{5}$$

Where: q represents the number of grids in the network. And there are p_j nodes in each grid. And the number of sensor nodes in each grid can be expressed as $\{S_{i|}i = 1, 2, \ldots, p_j\}$. Therefore, the number of nodes in the network can be expressed as:

$$n = q \sum_{j-1}^{q} p_j \tag{6}$$

In addition, assuming that the sensor node set in the j-th grid is Q_j, its definition formula is:

$$Q_i = \left\{ S_i^j | i = 1, 2, \cdots, p_j \right\} \tag{7}$$

Where S_i^j represents the i-th sensor node in the j-th grid.

Each node is selected as a cluster head in each grid, and other nodes in the grid are called member nodes, and are coordinated by the cluster head and the member nodes. The cluster head first receives the data from the member nodes, then performs data fusion and forwards them to the data center. The algorithm effectively selects the cluster head from the energy point of view, and thus can effectively adapt to the dynamic changes of the network topology.

The nodes of the WSN are assumed as follows [14]:

1) The nodes are randomly placed in a square area with a side length L, and the position information is unknown. Node energy is limited and can sense its own residual energy.
2) The node can send and receive data, and the strength of the transmitted signal is variable. The node can sense the strength of the received signal and the distance to the transmitting node can be estimated by comparing the strength of the transmitted signal.
3) Nodes work independently of each other, and each node has a unique ID, which can be a cluster head. The cluster head sends information directly to the sink node.
4) The location of the aggregation node is fixed and has strong processing capability. The energy of a sink node can be considered infinite relative to a general node.

3.2 Energy Consumption Model

The sensor generally performs two main processes of receiving and forwarding data. The energy consumption model uses the same model. Assuming that the sensor energy consumption is caused by receiving and forwarding data, $E_{tx}(d, b)$ and $E_{rx}(b)$ respectively represent the energy consumed to receive a b-bit information transmitted from a distance d meters and the energy consumed to forward b-bit data. Both can be expressed as follows [4]:

$$E_{tx}(d, b) = (\varepsilon d^a + E_{elec})b \tag{8}$$

$$E_{rx}(b) = E_{elec} b \tag{9}$$

It has the same initial energy difference as all node configurations in the LEACH routing protocol. The node energy is configured to be second-level heterogeneous, that is, only the Advanced Node and the Normal Node are included. Assume the total number of nodes in the network is N, the configuration ratio of the advanced node is m, and the energy of the advanced node is a times higher than that of the ordinary node. If the initial energy of the ordinary node is E_o, the total energy of the node of the entire heterogeneous wireless network is [15]:

$$E_{total}(b) = n(1-m)E_o + nm(1+a)E_o = nE_o(1+am) \tag{10}$$

All nodes of LEACH are configured with the same initial energy E_o, so in the second-level heterogeneous network, the total energy of the network is increased by $(1 + am)$ times. According to the LEACH protocol, the node generates a random number between 0–1. The probability threshold determines whether to become a cluster head; if the number is less than the threshold $T(n)$, the node broadcasts a message and it is a cluster head to the surrounding nodes. The LEACH protocol is based on a homogeneous network and is extended to a secondary heterogeneous network. The $T(n)$ of the advanced node and the ordinary node are respectively calculated as follows (2) and (3), where r is the number of rounds performed by the current network; G is the set of nodes in the last $1/(p(r/mod(1)/p)))$ round that have not been elected as the cluster head.

Let the node transform from the busy state (signal emission state) to the idle state and set the transition time from the idle state to the busy state as $\left[T_m^-, T_m^+\right]$, then [16]:

$$r_m^{on} = \left[r_1^- - r_1^+, r_2^- - r_2^+, \cdots, r_n^- - r_n^+\right] \tag{11}$$

$$r_m^{off} = \left[r_1^+ - r_1^-, r_2^+ - r_2^-, \cdots, r_n^+ - r_n^-\right] \tag{12}$$

Calculate the node energy consumption cost function: $r_m = \left[r_m^{on}, r_m^{off}\right]$, $P(r_m)$ is the probability density function of rm. The amount of entropy is:

$$H(r_m) = -\int P(r_m)\log_2\left[P(r_m)\right]dr \tag{13}$$

Equation (13) is the total energy entropy of the node energy consumption, including the entropy of the non-transmitted signal period and the transmitted signal period.

Define the node energy consumption cost formula:

$$E_m = C_m^{entrop} H_m + C_m^{occupy} O_m \tag{14}$$

Where C_m^{entrop} and C_m^{occupy} is the weight, which can be determined online or offline according to real-time conditions. The larger the value of the cost formula indicates that the future node is in a busy state, and the smaller the cost formula value, the more likely the point is to be idle.

4. CLUSTER HEAD SELECTION ALGORITHM

4.1 Node Trust Calculation

In the WSN environment, trust comes primarily from a comprehensive assessment of all or part of the communication, data, and energy. Here, considering the trust value of the node from the communication trust, data trust and energy trust of the node, the communication trust value C_i, the data trust value D_i and the energy trust value E_i of the node i are respectively calculated, and then the comprehensive trust value T_i of the node is calculated.

In WSN, control commands and data transmission are the main behaviors that can be observed by nodes. Malicious nodes will appear to discard, tamper with data packets, etc. Self-serving nodes may drop packets due to energy saving, so the communication behavior of nodes is needed to observe. Identifying a malicious node or a selfish node is a common mechanism for conducting trust evaluation. Using the simplified β trust model for communication trust calculation, the probability density function of the $\beta(a, b)$ distribution can be used to establish the reputation model of the node. The mathematical form is as follows:

$$f(x|a, b) = \frac{1}{B(a, b)} x^{a-1}(1-x)^{b-1} \tag{15}$$

Where: $0 \leqslant x \leqslant 1$, $a>0$, $b>0$.

Nodes N_i and N_j are neighbor nodes, and N_i's direct trust TD_{ij} to N_j is measured by N_j's packet forwarding rate for N_i (d_r is the number of packets received and forwarded, and d_t is the number of packets actually forwarded) [1]:

$$TD_{ij} = \frac{d_t}{d_r} \tag{16}$$

The indirect trust of the node N_i to N_j is obtained by monitoring the neighbor node N_m of N_j. Define P as the set of other recommended nodes, P={N_m| neighbor nodes of N_i, $T_{ij} \geqslant$ trust recommendation value Th}, all nodes in P weighted average of the trust value of N_j, get the indirect trust of N_i to N_j

$$M_{ij} = \frac{\sum_{m \in P} T_{im} T_{mj}}{\sum_{m \in P} T_{im}} \tag{17}$$

The trust value of node N_i to N_j

$$T_{ij} = \alpha TD_{ij} + (1-\alpha) M_{ij} \tag{18}$$

Where α denotes the weight of TD_{ij} in the calculation of the trust value, and the ratio of TD_{ij} and M_{ij} can be changed by changing the size of α.

The comprehensive trust value T_j of the node N_j is calculated by calculating the mean value of the trust value of all neighbor nodes of N_j:

$$T_j = \frac{\sum_{m \in P} T_{ij}}{k} \tag{19}$$

Where k is the total number of neighbor nodes of node N_j.

4.2 Cluster Head Selection

The production of non-uniformly distributed cluster heads is generated by changing the p-values in different regions of the WSN. In this regard, the LEACH algorithm is used to generate the cluster heads, and then the non-uniformly distributed sub-cluster heads are generated according to the distance between the nodes and the base stations. For the communication method, multiple hops are used between the clusters, and a single hop is used in the cluster. The nodes in the cluster are responsible for collecting data, and the primary cluster is responsible for fusion and forwarding, and forwards it to the secondary cluster head.

In order to successfully receive packet data, the received signal power at the receiver must be higher than the minimum threshold power $p_r(th)$. Therefore, the transmitted signal power at the transmitter must exceed the threshold $p_t(th)$ is [17]:

$$p_t(th) = \frac{p_r(th)16\pi^2 d^\mu Ls}{g_r g_t \varphi^2} \tag{20}$$

The total energy consumed by the transmitter is

$$\begin{aligned} E_{trans} &= \left[\frac{p_r(th)16\pi^2 R^\mu Ls}{g_r g_t \varphi^2 \upsilon_R} + \frac{p_r(th)16\pi^2 Ls}{g_r g_t \varphi^2}\right] N_p \\ &= \left(E_\mu R^\mu + E_e\right) N_p \end{aligned} \tag{21}$$

Where υ_R represents the rate at which the network node sends or receives data. R represents the node sensing radius, and N_p represents the number of packets. E_e is the amount of energy consumed in the transmitter circuit, and E_u is the amount of energy consumed by the power amplifier circuit in the transmitter.

The smaller the communication range from node $_i$ to cluster head *CHi*, the smaller the energy consumption between the cluster head and the node. Similarly, if the communication range from node i to base station *BS* is smaller, the energy consumption of data transmission is smaller. According to the free propagation energy model, the integrated distance weight can be expressed as [18]:

$$D\left(v_i\right) = 1 - \frac{d(i, CH_i)^2}{\sum\limits_{i=1}^{N(m,r-1)} d(i, CH_i)^2} - \frac{d(i, BS)^2}{\sum\limits_{i=1}^{N(m,r-1)} d(i, BS)^2} \tag{22}$$

Combined with the formula (22), the integrated distance weight is added to make the cluster head as close as possible to the base station, reducing the energy consumed by data transmission. The probability that node i becomes a cluster head is calculated by equation (23).

$$P_{i-ch} = \alpha F\left(E_{ir}\right) + \beta D\left(v_i\right) \tag{23}$$

Let α and β be the probability ratios of the residual energy weight of the adjustment node and the comprehensive distance weight in the cluster head competition, and $\alpha + \beta = 1$.

When the network is just established, there is no history of traffic statistics between the nodes, and its neighbor nodes have no trust value information about this node. Similarly, this node does not have trust value information about its neighbor nodes. Its neighbor node first assigns its value T_m to its trust value, and starts monitoring this node to count its traffic data. Initially, since there is no monitoring information of the node N_j by other nodes, the indirect trust value $R_{ij} = 0$, T_j=TD_{ij} during this time. If the node N_j is not a malicious node, the trust value of the other neighbor nodes to the node N_j can quickly reach a higher trust value as the node N_j forwarding rate increases. Some nodes may refuse to forward due to accidental interruption of the link, etc., so that some of the neighbor nodes may reduce the trust value of them quickly. If no measures are taken, they may be considered as untreatable by some neighboring nodes. So for such a node, a mechanism is needed to correct the trust value of the node. The equation above can effectively prevent this because some neighbor nodes have reduced their trust value, and other nodes maintain a high trust value state [19].

Figure 2 shows the flow chart of cluster head selection algorithm. The node selects the cluster head node with strong broadcast signal strength to join, then the cluster head node uses the TDMA policy to allocate the channel, and the node selects different according to the allocated channel. After the common nodes in the cluster collect information and transmit them to the cluster head, the cluster head performs data fusion, and then passes to the base station node. In the process of receiving the member data packet requesting to join the node and the member node waiting for the allocation mechanism, the cluster head node and the common nodes have the moment to open the receiver, and then the data transfer takes a while to enter a new round of clustering process.

4.3 Cluster Head Optimization

The node energy consumption in the network is mainly caused by the cluster head node in communication with the base station. Therefore, considering whether the node can serve as the cluster head node, not only the energy of the node itself but also the node head node is considered. Thus considering the energy loss caused by communication with the base station, the remaining energy of the computing node should be the current energy of the node minus the energy consumption caused by the communication between the node and the base station, and the distance from the node to the base station is combined to obtain the residual energy factor $E_p(i)$ of the node calculation [8]:

$$E_p(i) = \frac{E_{cur}\left(i\right) - LE_{cost}(i)}{l(i)} \tag{24}$$

Where $E_{cur}(i)$ represents the current energy of the node; L represents the data of the L bit; $l(i)$ represents the distance between the node and the base station; $E_{cost}(i)$ represents the energy consumption of the transmission of the 1-bit data node, where E_{cost}(i) is as follows [5].

$$E_{cost}(i) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2 + lE_{elec} \\ lE_{elec} + l\varepsilon_{fs}d^2 + lE_{elec} \end{cases} \tag{25}$$

Where E_{elec} is the circuit loss of transmission or reception; d_0 is the distance threshold; $\varepsilon_{fs}d^2$ is the power amplifier power consumption using the free-space channel model when the transmission distance is less than d_0; $\varepsilon_{mp}d^4$ is the distance greater than or equal to d_0, using more power amplifier power consumption of the path fading channel model.

When the node receives the signal of the pre-cluster header, it changes the pre-cluster header flag to the cluster head. If the node has not received the information of the pre-cluster header, but received the cluster head information of a neighbor's node, and apply to join the cluster. If the cluster head receives this signal, check the number of the cluster, and send a confirmation message. If no cluster head information is received, the node is set as the cluster head and the information is sent out. The sensor node sends a signal to the Sink node at regular intervals. The Sink node stores the size of the node and signal received by the signal in the table, and then passes the table to the node whose identifier is the cluster head.

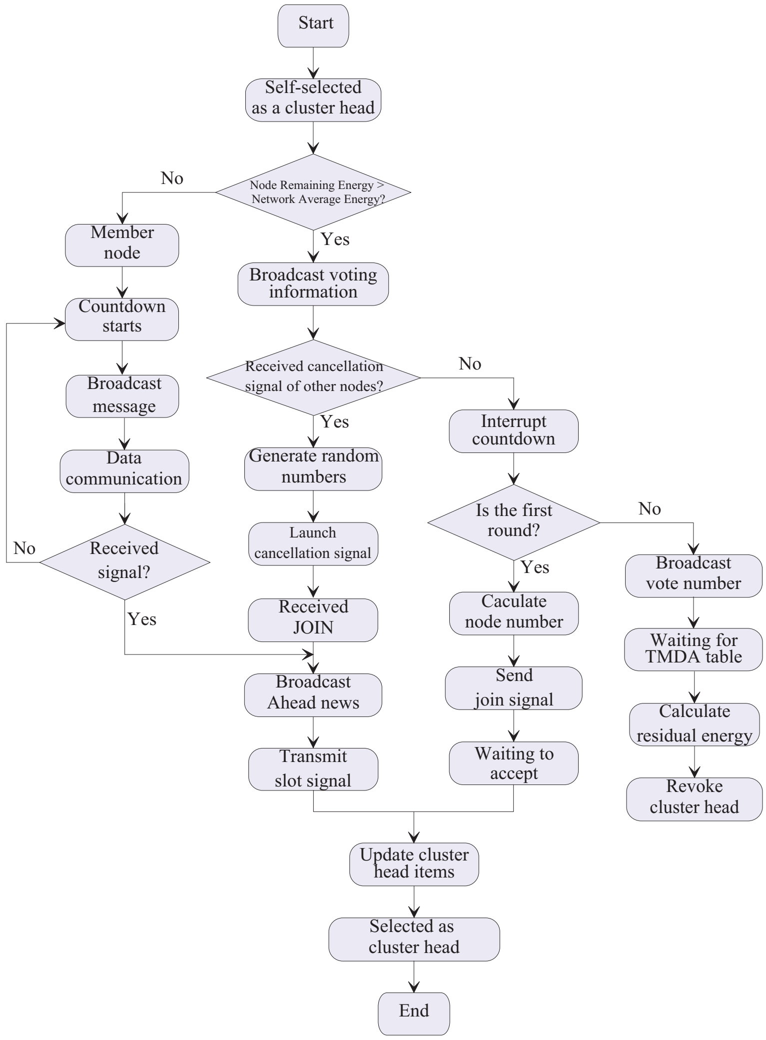

Figure 2 Cluster head selection algorithm flow

The cluster head node receives the table. The operators can find out whether they can directly pass the information to the Sink node. If they cannot directly pass the information to the cluster head, check whether there are nodes of the cluster member in the table. If there is the largest selection signal, set it as the gateway node. If there is no member of the cluster, then an adjacent cluster head node that can be connected is selected as the gateway node of the cluster [7].

5. ALGORITHM SIMULATION AND ANALYSIS

5.1 Simulation Parameters

Matlab simulation tool is used to simulate the LEACH algorithm, LEACH-F algorithm, HEED algorithm and the data analysis-based algorithm proposed in this paper. The simulation parameters are shown in Table 1. In the simulation scenario, the number of network nodes is 800, the network environment area is 200 m×200 m, the initial energy of the node is 4 J, the coordinates of the aggregation node are (300, 500), and the power threshold is $p_r(th)$=4×10^{-2} W, the number of network running rounds is 2000 rounds, and each packet length is 2.5×105 bps. The size of the data transmitted by the node is 400 Byte. The sink node is far from the sensing area and has a fixed position. For each network scale, the experimental result of each performance parameter is the average of 10 test results.

5.2 Simulation Result Analysis

The DAIC protocol divides the whole network into several layers. The simulated annealing algorithm is used to select cluster heads and form corresponding clusters in each layer. However, since the cluster head to the aggregation node still adopts the single-hop transmission mode, the hot zone problem cannot be fundamentally solved. The LEACH protocol proposes a cluster head rotation mechanism, which better solves the energy efficiency problem. However, the choice of the cluster head of this protocol depends entirely on the random number formed by the nodes in the cluster formation period, and the remaining parameters (such as residual energy and distribution position).

One of the biggest limitations is limited energy, in order to maximize the lifetime of the entire network, a good logical topology can effectively extend the lifetime of the entire

Table 1 Simulation parameter.

Parameter	Value	Parameter	Value
Area length/m	200	E_{elec}/(nJ·bit^{-1})	50
Area width/m	200	E_{amp}/(pJ·bit^{-1}·m^{-4})	0.0015
Node number	800	E_{fs}/(pJ·bit^{-1}·m^{-2})	15
Initial energy of node/J	2	Node initial trust value	1
Cluster head percentage%	5	Maximum round number/r	2000
Packet size/bit	400	Time spent in each round/s	40s
Transmission distance threshold/m	90	Base station location	(300, 500)
Energy averaging factor	1	Broadcast packet length/bit	100
Node induction radius/m	5	Expected cluster head number	50

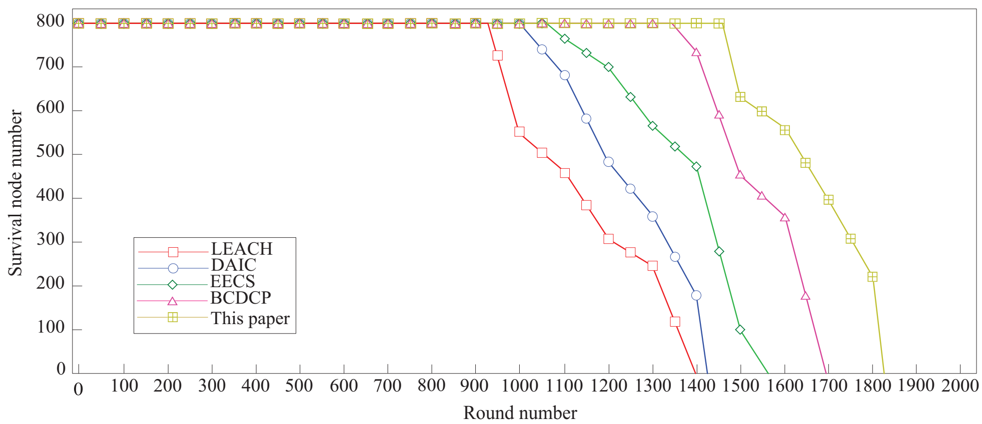

Figure 3 Comparison of survival nodes with different algorithms

network. WSN are characterized by small size, low power consumption, and high integration. There are also limitations such as limited power supply capacity, limited computing and storage capabilities, and limited communication capabilities. The deployment environment of sensor nodes is complex, and these restrictions make it vulnerable to attacks, and there are numerous security risks.

The choice of cluster heads in the protocol does not consider the initial energy and residual energy of the node, and may cause some nodes with small initial energy or nodes with small residual energy to be exhausted earlier. The EECS algorithm uses distributed working mode, and nodes only need local information to complete clustering and clustering. Full consideration is given to the residual energy of the node and the load balance between the cluster heads. Not considered, resulting in the following problems: The generation of the cluster head does not consider the geographical distribution of the nodes, and cannot guarantee the uniform distribution of the cluster head nodes, so that the size of the clusters cannot be guaranteed.

The BCDCP protocol evenly distributes energy consumption, and the minimum spanning tree is used between the sensor nodes to establish a multi-hop route between the base station and the cluster head. The comparison of survival nodes of LEACH, DAIC, EECS, BCDCP, and cluster head selection algorithm based on data analysis in this paper is shown in Figure 3.

The distribution of dependent node, main cluster head, secondary cluster head and ordinary node is shown in Figure 4. The cluster head node broadcasts the cluster head message in the wireless channel, and the non-cluster head node selects the cluster in which the cluster head with the strongest signal is located. The node transmits data to the cluster head through one-hop communication, and the cluster head also transmits the aggregated data to the sink node through one-hop communication. The protocol uses a random rotation to elect cluster heads, which avoids some nodes consuming excessive energy because they act as cluster heads multiple times. Since the base station is generally located far away from the sensing area, the communication between the node and the base station consumes a large amount of energy, and data aggregation can effectively reduce the communication between the node and the base station, thereby saving the energy of the entire network [20].

Figure 5 show the comparison of node survival rate with different algorithms. After the cluster head node is selected, the entire network is notified by broadcast, and other nodes in the network determine the dependent cluster according to the signal strength of the received information, and notify the corresponding cluster head node to complete the cluster establishment. Each round of execution can be divided into two phases, the cluster initialization phase and the stable data communication phase. The initialization phase of the cluster is mainly to elect the cluster head node, and other nodes decide which cluster to join. Immediately after the cluster is established, it enters a stable data communication phase, and the data communication phase lasts longer than the cluster initialization phase [21].

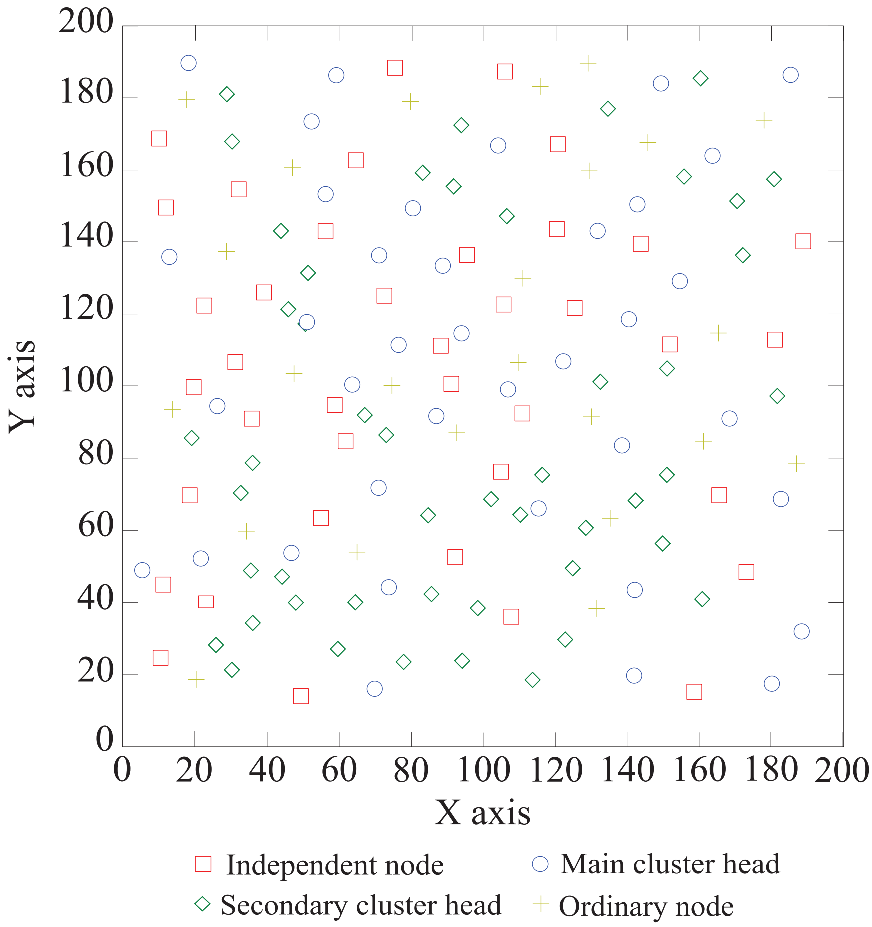

Figure 4 Distribution of different types of nodes with the cluster head selection algorithm in this paper

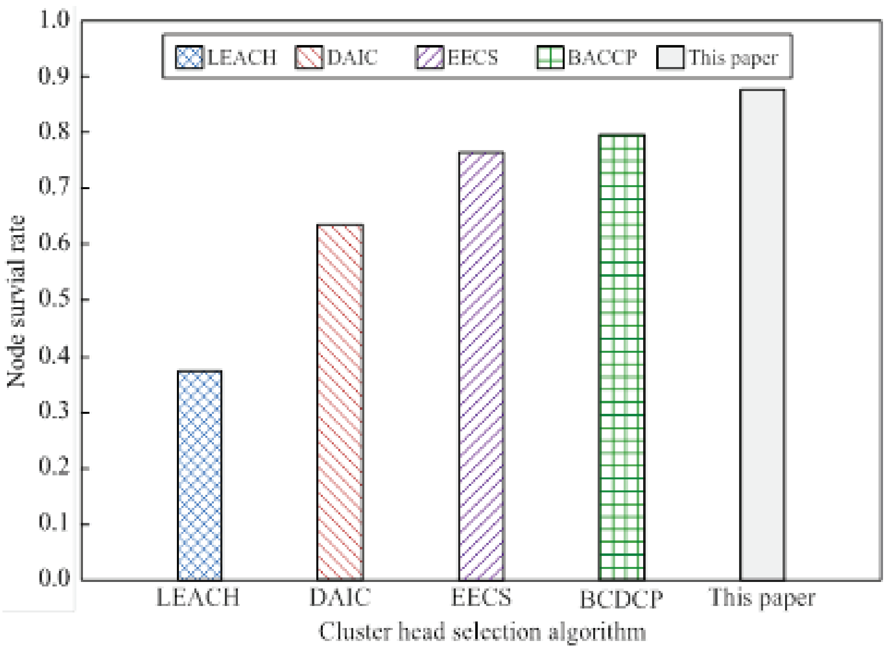

Figure 5 Comparison of node survival rate with cluster head selection different algorithms

When the weights are the same, the node with the smallest ID is selected as the main cluster head. Main cluster head node broadcasts the message including the main cluster head node ID and cluster head. If the node received more than one message, this node is a candidate gateway node, which sends a candidate gateway message to the cluster head, and repeats the above process until all nodes join a certain cluster [22].

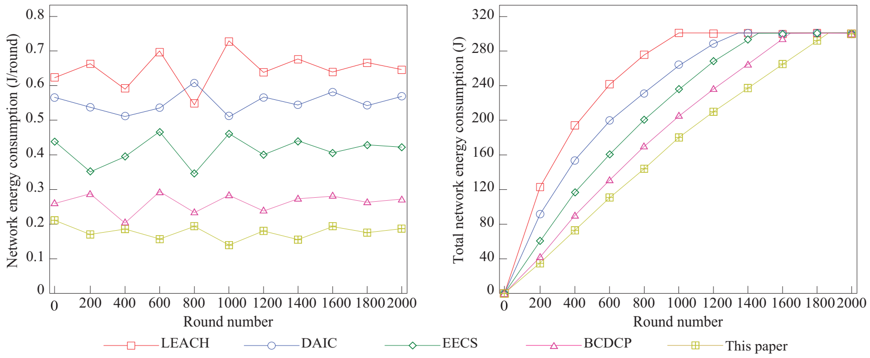

Figure 6 Comparison of network energy consumption of different cluster head selection different algorithms

The cluster head node uses TDMA to allocate a time slice to each node in the cluster to transmit data to it. In the LEACH protocol, the cluster head node has a large load and fast energy consumption compared with the ordinary member node. In order to balance the energy consumption of each node in the network and avoid premature death of the cluster head node, the TDMA is used to periodically elect the cluster head by round.

The non-cluster sensor node selects the nearest cluster head node through the signal strength of the broadcast information sent by the cluster head node, and sends a join request message to register the information. The predecessors also adopted a dual cluster head scheme. After the base station selects the cluster head from the sensor nodes, the cluster head node broadcasts a message, which includes the ID of the cluster head sensor node. Thereafter, the cluster head node broadcasts the schedule to its member nodes, and the member nodes begin transmitting their data to the cluster head node according to the schedule, and the data is transmitted to the base station through the cluster head node.

In the neighbor discovery phase, the node actively broadcasts the node ID and remaining energy. After receiving the information, the neighbor node updates its neighbor list, which records the neighbor ID, remaining energy, cluster head, and connectivity. After a broadcast, all nodes obtain all neighbor node ID, calculate connectivity and broadcast connectivity information to neighbor nodes. After secondary broadcast, all nodes obtain neighbor node connectivity information, and calculate the weight of each neighbor node. The node with the largest weight is elected as the main cluster head.

6. CONCLUSIONS

In this paper, based on the classical LEACH protocol and the improvement of its algorithm, a cluster head selection algorithm based on energy change is proposed. In the research of WSN, energy efficiency has always been a hot issue. In order to extend the network life cycle, related software and hardware have been extensively studied. The clustering algorithm can effectively extend the network life cycle. There is also a non-uniform distribution of cluster heads generated by the LEACH algorithm. Although the LEACH algorithm has certain problems, it is still cited as a classical algorithm. For example, the HEED algorithm is improved on the LEACH algorithm in which the selection mechanism of the sub-cluster head adopts the method of adaptive value function. The three key factors of the node meat energy, the distance of the node from the cluster head and the distance between the node and the base station are comprehensively considered in the fitness function. In this way, the secondary cluster head node can be a good task for data reception and forwarding.

In the clustering algorithm, the choice of cluster head has a great impact on the network energy consumption. The algorithm proposed in this paper considers the trust level of the node and the residual energy to calculate the ideal cluster radius when electing the cluster head. In the data transmission, multi-hop communication is adopted, and simple correlation multi-path routing is adopted, and a main path is established based on the minimum communication cost. When the range of its energy change exceeds the value, it broadcasts an energy message to the neighbor cluster head, indicating that it is no longer a next hop route, thus avoiding multiple cluster heads selecting the same relay cluster. This ensures that the number of cluster heads in each round is within the expected range; the number of neighbour nodes and the size of the energy consumption are optimized to complete the competitive selection of the cluster head. The simulation results show that the algorithm can further reduce energy consumption, increase data transmission and prolong network lifetime.

REFERENCES

1. Wang J. (2016). Research on the Internet of Things for wireless sensor networks. Telecom Power Technology, 33(6): 152–154.
2. Zhang D.X., Qin Z. (2015). Cluster head selection algorithm in wireless sensor network considering multi-factors. Journal of Chengdu University of Information Technology, 30(2): 160–165.
3. Men S.Z., Sun S.Y., Xu B.G. (2014). Wireless sensor networks non-uniform clustering and cluster Heads routing algorithm based on PSO. Chinese Journal of Sensors and Actuators, 27(9): 1281–1286.
4. Dang X.C., Hao Z.J. (2011). Multi-cluster header clustering algorithm in wireless sensor networks based on energy matrix. Computer Engineering and Applications, 47(19):71–74.
5. Xu X.L., Qiu J.N. (2009). Energy efficient cluster-head selection for heterogeneous wireless sensor networks. Chinese Journal of Sensors and Actuators, 22(3): 395–400.

6. Wang J., Shao Y.B., Long H., Du Q.Z. (2014). WSN cluster head selection algorithm based on energy and distance weighted. Transducer and Microsystem Technology, 33(5): 132–134.
7. Zhang Y., Song R.Y., Chen Z., Hu Q.L., Yue W.J. (2011). Research on topology control algorithm of mobile sensor networks based on cluster head selection. Chinese Journal of Sensors and Actuators, 24(11): 1602–1606.
8. Xu J.Z., Yao L.J., Yuan G.M. (2012). An improved cluster head election algorithm based on LEACH protocol. Computer Applications and Software, 29(11): 262–268.
9. Luo Q., Wang X.M. (2015). A wireless sensor networks cluster head selection algorithm based on fuzzy logic. Microelectronics and Computer, 32(9): 164–168.
10. Liu B., Pei B.N., ChenW.W., Yang Z.F. (2011). Improved clustering routing algorithm based on energy control of cluster-heads. Computer Engineering and Applications, 47(12): 69–71.
11. Zou C.W., Wang X.L. (2010). Clustering algorithm of WSN based on energy and velocity of consuming energy. Computer Technology and Development, 20(11): 14–17.
12. Zhao Z.C., Deng S.G., Jiang D.J. (2011). Algorithm in choice of head-node on WSN based on Gaussian distribution. Computer Engineering and Applications, 47(16): 83–85.
13. Liang P., He W. (2017). Grid dynamic energy threshold-based on cluster header algorithm in wireless sensor network. Chinese Journal of Sensors and Actuators, 30(10): 1583–1588.
14. Jia Y.L., Zhang C.Y., Liang K.W. (2018). Multiple competitions cluster head selection algorithm for WSN. Transducer and Microsystem Technologies, 37(2): 140–146.
15. . Ding F., Zhou Y.M. (2012). A energy-balanced cluster head selection protocol for heterogeneous wireless sensor networks. Journal of Shaoguan University (Natural Science), 33(10): 36–39.
16. Yang M., Xu R.C., Jiang T. (2011). A cluster head selection mechanism based on historical information. Communications Technology, 44(11): 97–99.
17. Liu W.J., Bai Y.Y. (2017). Energy level clustering and cluster head selecting algorithm for cellular clustering networks. Computer Engineering and Design, 38(7): 1759–1763.
18. Lai C., Jiang W.X. (2015). Energy-balanced routing algorithm based on cluster heads expectation for wireless sensor networks. Journal of Chinese Computer Systems, 36(12): 2685–2689.
19. Zhou L.P., Moydin K. (2012). Cluster head selection algorithm in WSN considering security and remaining energy. Computer Engineering and Applications, 48(3): 88–90.
20. Ni W.Y, Liu Q.W., Liu M.D. (2015). Wireless sensor network based on energy and distance. Journal of East China University of Science and Technology (Natural Science Edition), 41(1): 84–88.
21. Liu Y., Xu Z., Yang R. (2014). Research on Internet of things –based routing protocol for wireless sensor networks. Study on Optical Communications, 181(1): 67–69.
22. Xu D.D., Zhang Y. (2008). A two-cluster-heads clustering algorithm based on MAXD. Chinese Journal of Sensors and Actuators, 21(11): 1909–1912.