



# WSN routing algorithm based on node classification and multi-layer transport

Xiuwu Yu<sup>1,2,3</sup> · Yinhao Liu<sup>1</sup> · Yong Liu<sup>1,2</sup>

Published online: 5 October 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

In recent years, wireless sensor networks (WSNs), as the underlying basic interface of the Internet of Things, have developed rapidly, and the performance requirements of WSNs are constantly improved in various application environments. However, in the practical application, the traditional wireless sensor network has the problem of uneven and limited energy consumption. To solve this problem, a non-uniform clustering algorithm based on node classification and multi-level data transmission (NCMLT) is proposed in this paper. The algorithm divides the nodes into S1 and S2 levels according to the distance from the base station, and transmits data in chain and cluster structures respectively, effectively combining the advantages of the two structures. In the stage of cluster head election, the threshold formula of low energy adaptive clustering hierarchy (LEACH) is improved, and the remaining energy of nodes and other factors are added into the candidate cluster head election function. When clustering, considering the rationality of the competitive radius, the distance and energy factors are adjusted to form non-uniform clustering, which reduces the energy cost of the network. Based on the greedy algorithm, the chain structure is constructed according to the defined chain head election formula, and the optimal chain node is selected by calculating the relay value of the cluster chain as the relay node for data transmission between the cluster head that is closer to the base station and the base station, and the data transmission distance between the nodes is optimized by means of hierarchical multi-hop. Through simulation experiments, compared with the energy-efficient uneven clustering algorithm (EEUC) and the uneven clustering routing algorithm based on ant colony optimization (URACO), the network lifetime of NCMLT algorithm is increased by 47.43% and 5.02%, respectively, which proves that NCMLT algorithm can effectively balance node energy consumption and extend network lifetime.

**Keywords** Wireless sensor network · Non-uniform clustering · Balanced energy consumption · Classification of nodes

## 1 Introduction

Wireless sensor network is a kind of distributed sensor network, low cost, self-organization ability, easy to deploy. As an intelligent system, it has the functions of data collection, fusion and transmission, and can be monitored in real time.

Because it can adapt to complex and even harsh environments, it has been used in various industries [1, 2]. While the sensor nodes are usually powered by the battery, the energy is limited, and it is not easy to supplement. How to balance and reduce energy consumption, efficiently use network energy, and extend the life of the network has become the key point of research [3, 4].

Most of the energy consumed by wireless sensor networks is used to transmit data packets [5]. Compared with plane routing, the hierarchical routing protocol can effectively reduce the amount of data transmission between nodes, reduce energy consumption more significantly, and reduce the impact of dynamic changes in topology on network performance through cluster management, which has better research value [6]. The LEACH algorithm [7, 8] is considered to be the earliest hierarchical routing

✉ Yinhao Liu  
2200919415@qq.com

<sup>1</sup> University of South China, Hengyang, China

<sup>2</sup> Hunan Engineering Research Center for Uranium Tailings Decommission and Treatment, Hengyang, China

<sup>3</sup> Hunan Province Engineering Research Center of Radioactive Control Technology in Uranium Mining and Metallurgy, Hengyang, China

algorithm based on uniform clustering, in which each node randomly serves as the cluster head with the same probability. But the number of the cluster head is unstable, the cluster is unbalanced, and the cluster head transmits data to the sink node in a single-hop manner. The energy consumption is too large. Power efficient gathering in sensor information systems (PEGASIS) [9] uses greedy algorithms to connect all nodes into a chain structure according to the distance from the base station, and randomly select the chain head to communicate with the base station, which effectively saves energy. But because the nodes transmit data one by one, the network delay is too long and the real-time performance is poor.

Based on the above analysis, a non-uniform clustering routing algorithm for wireless sensor networks based on node hierarchical and multi-level data transmission is proposed in this paper. The main contributions of this paper are as follows:

- (1) According to the distance between the node and the base station, the node is divided into S1 and S2 levels. The S1 level nodes transmit data in a chain structure, and the S2 level nodes transmit data in clusters.
- (2) In the S1 nodes, the optimal chain node is selected as the chain head to directly communicate with the base station. In S2 nodes, a variety of parameters are combined into an adaptation function and inserted into the cluster head threshold formula of LEACH algorithm, and the weight of each parameter is adjusted according to the network operation, so as to increase the probability of high-quality nodes becoming cluster heads.
- (3) In the data transmission phase, multi-level multi-hop routes are constructed based on relay nodes. In the S2 nodes, each cluster head selects the next hop according to the cost function. At the critical point of S1 nodes and S2 nodes, the chain node with the best relay value is selected according to the relay function as the relay node for the communication between the end cluster head and the base station.

In this algorithm, the distance of data transmission is optimized by node classification. Considering the different characteristics and tasks of nodes in different locations, different structures are used to construct routes. Improve the traditional chain and cluster routing, improve their role in energy balance. Through the experimental demonstration, the best parameter selection is obtained. The purpose of the research is to improve the service time of the network. In the simulation part, the feasibility of the algorithm is verified.

## 2 Related works

In view of the problem that the LEACH protocol has unreasonable cluster head election, the LEACH-improve protocol proposed by Huang Lixiao et al. [10] integrates the residual energy factor, spacing factor and density factor into the traditional LEACH threshold calculation formula, so that the cluster head of the election is more reasonable. But there will still be extreme scale clusters. Li Chengfa et al. [11] introduced the concept of competitive radius (EEUC). The sensor network was clustered unevenly, and the cluster size closer to the base station was smaller, which reduced the energy consumption of the cluster head near the base station, thereby balancing the network energy consumption. However, it does not get the optimal parameter value. Liu Wei et al. [12] used the temporary cluster head competition radius combined with multi-hop routing between clusters to optimize the energy consumption of network nodes. It does not consider the actual situation of inter-cluster transmission path.

Aiming at the problem of route selection, Huang Ying et al. [13] proposed a wireless sensor network routing optimization with energy and path constraints, constructed a high-speed transmission path, and ensured the reasonable distribution of energy according to energy factors. The path construction mode is based on a single node, which is easy to be affected by extreme nodes, resulting in path failure and poor energy balance effect of the whole network. Bi Xiaojun et al. [14] used the gravitational search algorithm to plan the communication links of the network cluster head, so as to reduce the load energy consumption of the communication between the cluster head nodes. This method is based on the design of high-energy nodes, and there will be energy imbalance in the network region without high-energy nodes. Liu Sanyang et al. [15] proposed a routing algorithm based on energy consumption area perception, introduced the concept of “pre-sensing area”, reduced the energy consumption caused by “energy backhaul”, and solved the problem of invalid propagation of information by excluding dead nodes.

Some scholars divide the network area according to some indicators. The fuzzy c-means algorithm [16] divides the network into M clusters, minimizing the distance between nodes and the center of the cluster. Hu Yuan et al. [17] divided the monitoring area into rings, so that the cluster radius within the same ring was equal. The closer to the base station, the smaller the cluster radius, but the inner span of the outer ring cluster was too large, resulting in excessive energy consumption of the outer link points. Liu Zhuang et al. [18] solved the problem of energy imbalance in boundary region based on adjustable grid, but did not give a clear election mechanism in the cluster head

election. Jiang Bing et al. [19] limited the communication between adjacent clusters under the same sector, which reduced the cluster head reconstruction, data transmission energy consumption and communication distance. Li Hongbin et al. [20] classified cluster heads according to distance to avoid selecting cluster heads too densely, and classified adjacent nodes according to energy, so as to determine relay nodes, but the effect was not great.

In recent years, researchers have combined various algorithms or models to apply WSNs routing protocol. Literature [21] and Literature [22] respectively studied WSNs clustered routing based on game theory model and fuzzy logic algorithm, which effectively improve the network life cycle, but the effect is not satisfactory. Wang et al. [23] used a cluster routing protocol based on compressed sensing to determine the communication radius of the cluster to extend the network life. Fu et al. [24] improved the energy utilization, reliability, and real-time performance of the network through the node hibernation mechanism. Liu Hong et al. [25] introduced ant colony algorithm to determine routing (URACO), selected routing nodes through pheromone concentration to optimize path quality, and introduced the concept of superiority in the process of pheromone update to achieve path quality optimization and effectively improve network energy utilization. However, the goal of this algorithm is to satisfy the shortest path, and it fails to comprehensively consider other energy measures, such as average energy, etc. In the planned optimal path, some nodes with low energy will run out of energy prematurely, resulting in insufficient choice of path quality. Meanwhile, in the routing protocol optimized by ant colony algorithm, neighbor nodes need to exchange information such as residual energy, pheromones and heuristic factors frequently, so the update speed is slow and the network energy cost is increased sharply.

### 3 Network and energy models

#### 3.1 Network model

Suppose WSNs have the following characteristics:

- (1) Nodes are randomly deployed in a fixed monitoring area, and each node has its unique ID;
- (2) All nodes are isomorphic and fixed, the positioning algorithm has been completed, the location coordinates are known, and the error influence is small;
- (3) All nodes have the same initial energy, the energy cannot be replenished, and the current residual energy can be obtained;

- (4) The base station (Sink node) is located outside the monitoring area, the position is fixed, and the energy is infinite;
- (5) All nodes can adjust their own transmission power according to the information transmission distance;
- (6) Wireless communication energy consumption accounts for a large proportion, ignoring the rest of the energy consumption.

#### 3.2 Energy dissipation model

The wireless communication energy consumption model consists of two parts: a transmitting circuit and an amplification circuit, and the energy consumed by the node to transmit  $k$  bit data to the node at  $d$  distance can be shown in Eq. (1):

$$E_{tx}(k, d) = \begin{cases} kE_{elec} + k\varepsilon_{fs}d^2, & d < d_{init} \\ kE_{elec} + k\varepsilon_{amp}d^4, & d \geq d_{init} \end{cases} \quad (1)$$

Among them:  $\varepsilon_{fs}$  is the power amplification factor under the free space model, and the  $\varepsilon_{amp}$  is the power amplification factor under the multipath attenuation model;  $d_{init} = \sqrt{\varepsilon_{fs}/\varepsilon_{amp}}$ , which is the critical value. If  $d < d_{init}$ , the free space model is used, and if  $d \geq d_{init}$ , the multipath attenuation model is used;  $E_{elec}$  is the energy consumed by the node to send or receive 1 bit of data.

The energy consumed by the node to receive  $k$  bit data is shown in Eq. (2):

$$E_{rx}(k) = kE_{elec} \quad (2)$$

### 4 NCMLT routing algorithm

#### 4.1 Node classification

The monitoring area of the wireless sensor network is usually large, the number of sensor nodes is large, widely distributed, and the nodes far from the base station will bear too much energy burden if they directly transmit data to the base station. This paper according to the distance from the sensor nodes to the base station, the nodes are divided into S1, S2 two levels. The data transmission method and the next hop node are selected according to the node level. The transmission path is reasonably optimized. The classification formula is shown in Eq. (3).

$$S_l = \begin{cases} S_1, & d_l \leq d_T \\ S_2, & d_l > d_T \end{cases} \quad (3)$$

Formula:  $d_l$  is the actual distance from the node to the base station;  $d_T$  is the hierarchical distance threshold. In order to ensure the number of nodes at all levels, make the

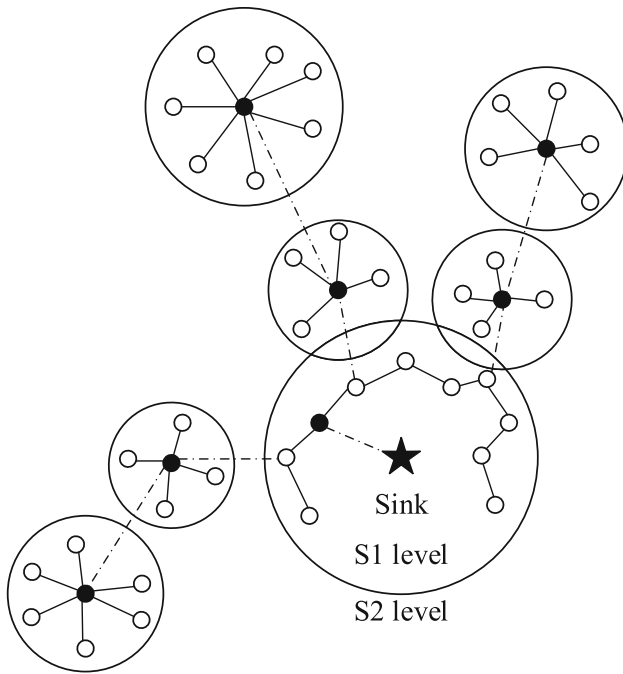


Fig. 1 Node hierarchy

nodes transmit data in a free space model, so that  $d_T = R_s$ .  $R_s$  is the node's perceived distance. The specific grading chart is shown in Fig. 1.

## 4.2 Build the chain structure and chain head

In S1-level nodes, the greedy algorithm is used to build a chain structure, which is carried out on a round-by-round basis. The node farthest from the base station is used as the starting node of the chain construction, and the node closest to it is selected in each round to join the chain as the next hop, and the node is used as the first node of the next round until all nodes join the chain to complete the construction of the chain, so as to ensure the shortest data communication distance between S1-level nodes.

In the classic chain algorithm, PEGASIS algorithm, the first node of the chain communicating with the base station adopts the method of taking turns, and the node with less residual energy or far away from the base station is prone to die faster, resulting in the premature end of the entire network life cycle. In order to solve this problem, the NCMLT algorithm prioritizes the node that is closer to the base station and has more residual energy as the chain leader, so that the chain head node can always bear the energy consumption of the entire chain structure to forward data, and the specific election formula is shown in Eq. (4).

$$E_c = E_r/d_l \quad (4)$$

Formula:  $E_r$  is the remaining energy of the current node.

## 4.3 Cluster election and clustering

S2 nodes also take “round” as the collection period to build a cluster structure. Each round is divided into cluster-head election, non-uniform clustering, route construction and data transmission stages. Since the cluster-head node is responsible for data fusion and forwarding, and consumes much more energy than other nodes in the cluster, factors such as its remaining energy and distance from the Sink node need to be taken into account during the cluster-head election. Before the election, the Sink node collects the location and energy information of all nodes. Based on the threshold formula of the LEACH protocol, the sink node introduces the heart rate, the remaining energy of the node, and the distance between the node and the base station as factors to build an adaptation function into the threshold formula, so as to improve the rationality of the elected cluster head.

The closer the cluster head is to the neighbor node, the less energy is consumed in data transmission. The degree of proximity between the cluster head and the neighbor node is defined as the centripetal heart rate, expressed as:

$$H(i) = \left[ \left( x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 + \left( y_i - \frac{1}{n} \sum_{j=1}^n y_j \right)^2 \right]^{-\frac{1}{2}} \quad (5)$$

Due to the different distribution locations of nodes, the energy consumed by each node to receive and send data is also different, so the residual energy factor is expressed as:

$$E(i) = \frac{E_r(i)}{E_a(i)} \quad (6)$$

The distance factor is expressed as:

$$d_i = \frac{d_{\max} - d_l}{d_{\max} - d_{\min}} \quad (7)$$

The adaptation function is expressed as:

$$F(i) = \tau H(i) + \mu E(i) + \varphi d(i) \quad (8)$$

The improved threshold formula is shown in Eq. (9):

$$T(n) = \begin{cases} \frac{P}{1 - P[r \bmod (1/P)]} F(i), & n \in G \\ 0, & n \notin G \end{cases} \quad (9)$$

Formula:  $P$  is the ratio of the number of cluster heads to the total number of nodes;  $r$  is the current number of campaign rounds;  $(x_i, y_i)$  is the position of the head of a cluster, and  $(x_j, y_j)$  is the position of its neighbor node;  $n$  indicates a node at the head of the cluster that was not

elected in round  $r$ ;  $E_a(i)$  is the average energy of nodes;  $d_{max}$  indicates the maximum distance from the S2 nodes to the base station;  $d_{min}$  represents the minimum distance from the S2 nodes to the base station;  $\tau$ ,  $\mu$ , and  $\varphi$  are parameter factors that control the weight of their corresponding influencing factors. The value are 0–1, and the value meet the requirement that  $\tau + \mu + \varphi = 1$ .  $G$  is the set of nodes that have not been elected as the head of the cluster in the previous  $1/P$  round.

After the network deployment is completed, the nodes calculates the probability of becoming the candidate cluster leader according to Eq. (9). The nodes randomly generates a number in the range of 0–1, which is compared with the calculated probability  $T(n)$ . When the number is less than  $T(n)$ , this node becomes the candidate cluster leader, otherwise this node goes to sleep and is awakened in the cluster formation stage.

Since nodes are immovable, there is not much difference in node energy at the initial stage of network operation, but energy is the most important factor of the network, so parameter factors  $\mu$  are set to 0.4,  $\tau$  and  $\varphi$  to 0.3. With the running of the network, the energy of the whole network decreases continuously, and the energy is more critical for cluster head selection, and the parameter factor  $\mu$  should increase with the increase of the running time. Let:

$$P(r) = \frac{\sum_{i=1}^n E_o - \sum_{i=1}^n E_r}{\sum_{i=1}^n E_o} \tag{10}$$

where,  $\sum_{i=1}^n E_o$  represents the initial energy of the entire network, and  $\sum_{i=1}^n E_r$  represents the remaining energy of the round  $r$  of the entire network. When  $P(r) = 0.2$ , that is, the energy consumption of the network is 20%, the parameter factor  $\mu$  increases to 0.5,  $\tau$  and  $\varphi$  are 0.25; When  $P(r) = 0.4$ , that is, the energy consumption of the network is 40%, the parameter factor  $\mu$  increases to 0.6,  $\tau$  and  $\varphi$  are 0.2. When  $P(r) \geq 0.6$ , that is, when the network energy consumption is more than 60%, the parameter factor  $\mu$  increases to 0.8, and  $\tau$  and  $\varphi$  are 0.1.

The nodes close to the sink node should not only be responsible for collecting and transmitting the data of the cluster, but also forwarding the data sent by other clusters. In order to avoid premature death of the nodes, the area closer to the base station should generate more clusters and make the cluster have a smaller scale. Considering the current remaining energy of the candidate cluster head, the competition radius calculation formula for the cluster head is shown in Eq. (11).

$$R_c = \left[ 1 - c \frac{d_{max} - d_l}{d_{max} - d_{min}} - \frac{(1 - c)}{E(i)} \right] R_0 \tag{11}$$

Formula:  $R_0$  is a pre-defined maximum competitive radius that controls the competitive radius of the candidate cluster head within a reasonable range. Due to the operation of the network, the residual energy is reduced.  $E(i)$  is used as a regulator to reduce the cluster radius, which can reduce the energy consumed by each node.  $c$  is the control parameter.

Candidate cluster heads check if there are other candidate cluster heads within its race radius. If not, then itself directly becomes the final cluster head; On the contrary, compare the adaptation function values obtained from formula (8) and wait for the candidate cluster head with large adaptation value to make a decision on whether to become the final cluster head. If the candidate cluster head with large adaptation value becomes the final cluster head, other candidate cluster heads within the competition radius will withdraw from the race.

After the final cluster head is generated, each elected cluster head broadcasts the recruitment message within its competition radius, and other nodes are awakened and select the cluster head with the strongest signal strength to join.

### 4.4 Data transfer phase

The algorithm adopts the mode of single hop within the cluster and multiple hop between clusters for data transmission. The nodes in the cluster directly send data to the cluster head of the cluster to which it belongs. The cluster head receives the data of all member nodes, and transmits the data packet to the next hop node after integration. In terms of inter-cluster transmission, the energy consumption of wireless data transmission is proportional to the distance, the multipath attenuation model is  $d^4$ , and the free space model is  $d^2$ . Therefore, the nodes consume much less data to transmit data under the free space model. In order to ensure that the cluster head transmits data in the free space model, the cluster head does not directly transmit data with the base station, but selects the neighbor cluster head as the next hop according to the cost function. If the cluster head is close to the base station and there is no suitable next-hop cluster head, the optimal S1-level node is selected as the relay node according to the cluster chain relay function, and the data is transmitted to the base station through the chain structure.

The cost function is shown in Eq. (12).

$$cost(i, j) = \alpha \frac{E_{neighbor}(s_i)}{E_{current}(s_j)} + \beta \frac{N_{non-CH}(s_j)}{N_{non-CH}(s_i)} + \gamma \frac{d_{si-sj}^2 + d_{sj-BS}^2}{d_{sj-BS}^2} \tag{12}$$



Formula:  $\alpha, \beta$  and  $\gamma$  are weighted coefficients and satisfy  $\alpha + \beta + \gamma = 1$ ;  $\overline{E_{neighbor}(s_i)}$  represents the mean remaining energy of the neighbor cluster head of the cluster head  $s_i$ ;  $E_{current}(s_j)$  represents the remaining energy of the cluster head  $s_j$ ;  $N_{non-CH}(s_j)$  represents the number of member nodes at cluster head  $s_j$ ;  $\overline{N_{non-CH}(s_i)}$  represents the mean number of neighbor cluster leader member nodes of cluster head  $s_i$ ;  $d_{si-sj}$  represents the distance from cluster head  $s_i$  to cluster head  $s_j$ ;  $d_{sj-BS}$  indicates the distance from the cluster head  $s_j$  to the base station.

The cluster chain relay function is shown in Eq. (13).

$$G = \lambda E(i) + \theta \frac{d_a}{d_i} + \frac{\eta}{L(i)} \tag{13}$$

Formula:  $\lambda, \theta$  and  $\eta$  are parameter factors, and  $\lambda + \theta + \eta = 1$ ;  $d_a$  is the average distance from all chain nodes to the base station;  $L(i)$  is the distance from the cluster head to the relay chain node.

When each cluster locates the next hop node, the cluster chain multi-hop route is established.

After receiving the data sent by the cluster leader, the S1-level nodes send the data packets to the chain head node in turn along the chain structure. The chain head consolidates the data and sends it to the base station. At this point, the data transfer for the current round is completed. The algorithm flowchart is shown in Fig. 2.

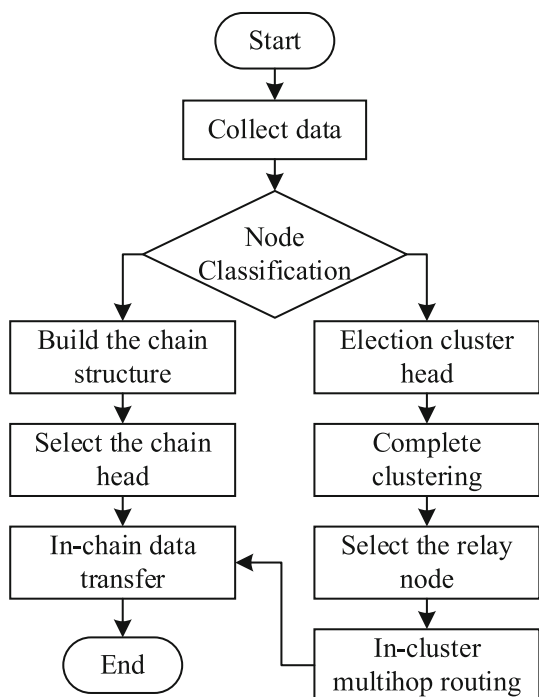


Fig. 2 Flow chart

## 5 Simulation experiment and result analysis

### 5.1 Different from the comparison algorithm

The research of EEUC algorithm focuses on the construction of optimization cluster, but it does not explain the selection of parameters and data transmission path in detail. In EEUC algorithm, the next hop node set is defined as:

$$S_i R_{CH} = \left\{ S_j \mid \begin{matrix} d(S_j, DS) < d(S_i, DS) \\ d(S_i, S_j) \leq kR_c \end{matrix} \right\} \tag{14}$$

That is, the cluster-head  $S_i$  limits the selection range of the next hop candidate node to a region closer to the Sink node than itself, and then selects the node with more remaining energy in this set as the relay node. In this method, distance and energy are progressive. In the NCMLT algorithm, distance and energy are level, and the next hop node selected by the cost function has a stronger comprehensive ability.

URACO algorithm selects cluster heads based on each index and its weight, as shown in Eq. 15, where the index does not include the distance between the node and the base station.

$$C(m) = \sum_{n=1}^3 \omega_n F_{mn} \tag{15}$$

where  $\omega_n$  is the corresponding weight of each indicator, and  $F_{mn}$  is the function of each indicator.

This method always makes the best node in the network become the cluster head, because the distance factor is not considered, if these nodes are distributed together, the cluster will be out of balance. On the contrary, the index selected by NCMLT algorithm is more comprehensive, and the election of cluster heads has a certain randomness. URACO improves the ant colony algorithm and applies it to path optimization, but the resulting path is greatly affected by node energy, as shown in formula 16.

$$P_s = \delta_1 \frac{E_{con}}{E_{min}} + \delta_2 \frac{P(S)}{P(M)} \tag{16}$$

where  $E_{min}$  is the minimum value of node energy in the ergodic path;  $E_{con}$  is the sum of energy consumption of path nodes.  $\delta_1$  and  $\delta_2$  are the influence coefficients.  $P(S)$  and  $P(M)$  are the energy standard deviation and mean of  $x$  nodes on the path, respectively.

The information interaction in the ant colony algorithm also increases the energy consumption of data transmission between nodes, affecting the network update speed.

The algorithm in this paper draws on the advantages of previous researches, considers clustering and routing comprehensively. The most important thing is to make

nodes with different degrees of distance from the base station play different roles through the classification of nodes. The combination of cluster and chain optimizes the network structure.

### 5.2 Introduction to simulation

Verifies its performance by comparing with EEUC algorithm and URACO algorithm through simulation experiments. In this paper, the NCMLT algorithm is simulated by using MATLABR2018b software. In order to ensure comparability, the basic simulation parameters are selected from EEUC algorithm, and 400 nodes are arranged in an area of 200 m × 200 m, and the base station is located in the regional center, the specific parameters are shown in Table 1. Node distribution is shown in Fig. 3.

This paper verifies the rationality and effectiveness of NCMLT algorithm by studying the optimal value of parameters and the influence of network scale on stability, and comparing it with other algorithms in terms of life cycle, energy consumption and data transmission performance.

### 5.3 Parameter influence

In the clustering of S2 nodes, the number and scale of the generated clusters are determined by the initial competition radius  $R_0$ , the value of  $R_0$  is related to the maximum scale of a single cluster. In order to explore the specific impact of the value of  $R_0$  on the algorithm,  $R_0$  is set to start from 20 and increase in unit of 10 to 80, and the death time of the first node in the network is observed, as shown in Fig. 4.

As can be seen from Fig. 4, when  $R_0$  is 50, the first dead node appears at the latest time. This means that the initial radius set to 50 m is the most appropriate in the NCMLT algorithm. When the initial radius is less than 50 m, the

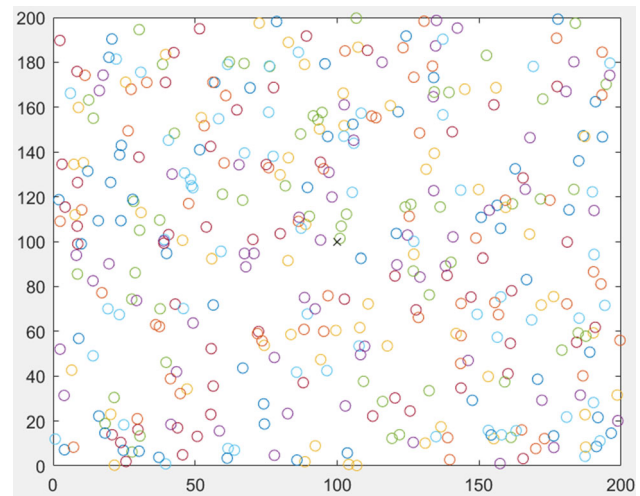


Fig. 3 Nodes distribution

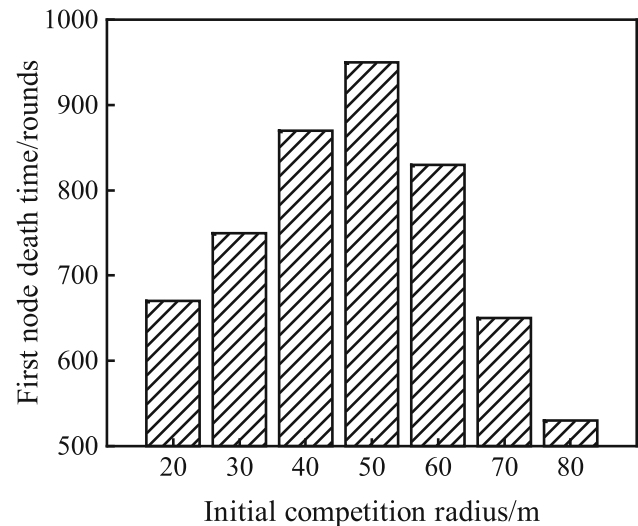


Fig. 4 Value of  $R_0$

cluster size will be small. In order to achieve better data transmission, more clusters need to be generated, that is, more cluster heads are required, which increases the burden on some nodes. When the initial radius is greater than 50 m, the cluster size will be larger, and the requirements on the cluster head will be higher, which is not conducive to energy balance.

Figure 5 shows the influence of the control parameter  $c$  of the competition radius on the time of first node death in the network. Taking  $c$  from 0 to 1 means changing the influence weights of the two factors, distance and energy. When  $c$  increases, the proportion of distance factor in the competition radius increases, and correspondingly, the proportion of energy factor decreases. This changes the size of clusters in the network, and thus changes the life cycle of the network. As can be seen from Fig. 5, when  $c$  increases from 0 to 0.4, the number of death rounds of the

Table 1 Simulation parameters

Parameters	Value
Region size/m <sup>2</sup>	200 × 200
Number of nodes	400
Sink location	(100, 100)
Node initial energy/J	0.5
$R_s$ /m	50
Packet size/bit	4000
$E_{elec}/(nJ*\text{bit}^{-1})$	50
$\epsilon_{fs}/(pJ*\text{bit}^{-1}*m^{-2})$	10
$\epsilon_{amp}/(pJ*\text{bit}^{-1}*m^{-4})$	0.0013
$R_0$ /m	50
$\tau, \mu, \varphi$	0.3, 0.4, 0.3

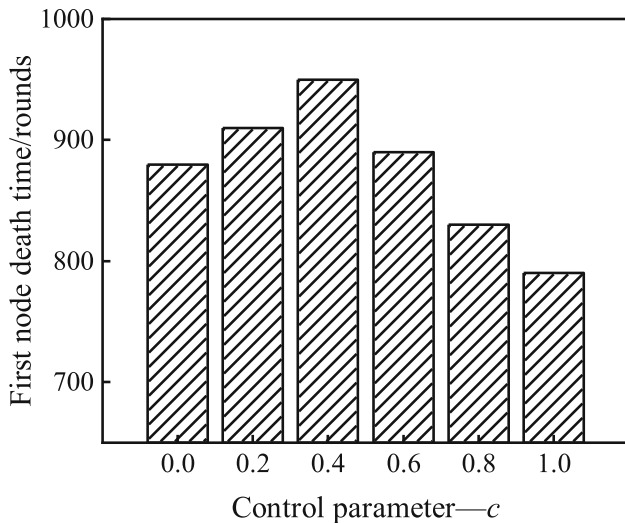


Fig. 5 Value of  $c$

first node gradually delays, indicating that the distance factor has a balanced effect on node energy consumption. When  $c$  exceeds 0.4, the number of death rounds of the first node gradually advances, indicating that the degree of energy balance decreases. This is mainly because the nodes pay too much attention to the distance factor, which leads to the neglect of the energy factor when planning the scale of the cluster, so that the node load is too high. Therefore, in this algorithm, the optimal value of  $c$  should be 0.4.

### 5.4 Effect of area size

In the simulation experiment, the number of cycles experienced by the sensor network from the start of operation to the death of the first node is called the stability period of

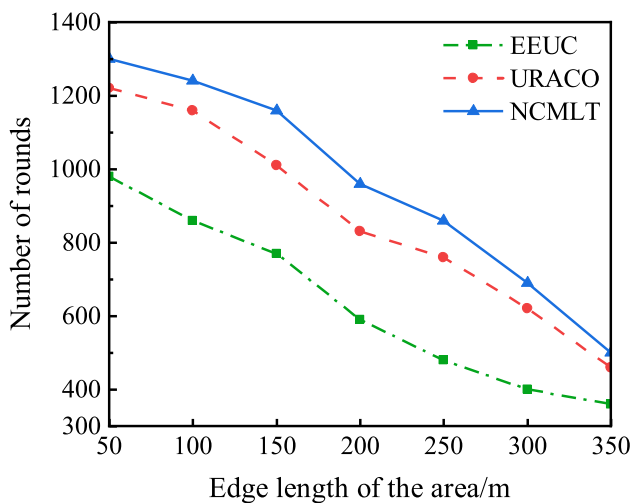


Fig. 6 Network area size

the network. Figure 6 shows the steady-state periodic change curves of the three routing algorithms under different network area sizes. The experimental results show that the stability period of each algorithm decreases with the increase of the region area. Under the same region size, NCMLT algorithm has the longest period and EEUC algorithm has the shortest period, so NCMLT algorithm has better stability.

### 5.5 Network lifetime

The number of surviving nodes in a network is typically used in WSN to evaluate the lifetime of the network. In the simulation, a node is considered dead when it consumes 99% of the initial energy. The number of cycles in which 90% of nodes die is usually defined as the network lifetime. In order to facilitate the statistics of experimental data, this paper defines the number of rounds in which all nodes die as the network lifetime. The comparison chart of the network survival time of EEUC algorithm, URACO algorithm and NCMLT algorithm is shown in Fig. 7. In terms of the first node death time and all node death time, NCMLT algorithm has obvious delay when compared with EEUC algorithm and URACO algorithm. In the NCMLT algorithm, all nodes die in about 1150 rounds, while in the EEUC algorithm, all nodes die in 780 rounds, and in the URACO algorithm, all nodes die in 1095 rounds. The performance of NCMLT algorithm is improved by 47.43% compared with the EEUC algorithm and 5.02% compared with the URACO algorithm, indicating that the NCMLT algorithm effectively balances the network energy consumption and prolongs the network survival cycle by transmitting data in the form of cluster chain combination.

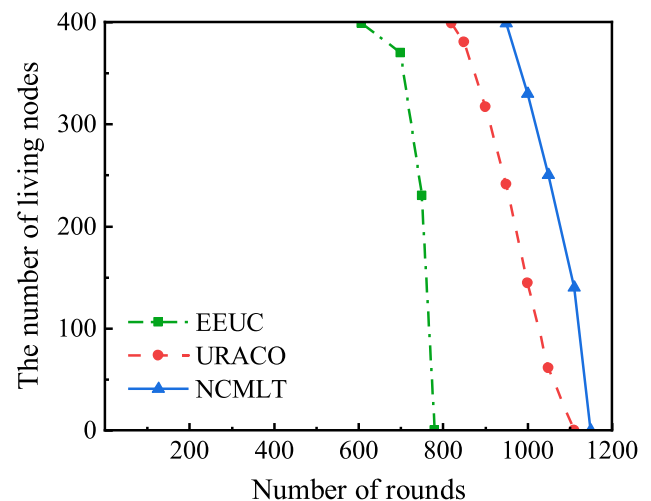


Fig. 7 Comparison of network time-to-live



### 5.6 Cluster heads consume energy

In order to prove that the NCMLT algorithm reduces the energy consumption of the cluster head, during the running of the simulation experiment, 10 rounds of the three algorithms without a death node were randomly selected to calculate the energy consumption variance of the cluster head of the three algorithms, and the energy consumption equilibrium of the cluster head of the three algorithms was compared. It can be seen from Fig. 8 that the variance values of EEUC algorithm, URACO algorithm and NCMLT algorithm are small, indicating that the three algorithms use their own methods to balance the energy consumption of the cluster head. But the variance curve of the NCMLT algorithm is significantly lower than that of the EEUC algorithm and the URACO algorithm, indicating that the non-uniform cluster structure designed by the NCMLT algorithm better balances the energy consumption of the cluster head and plays an optimization effect.

### 5.7 Network energy consumption

The network energy consumption of EEUC algorithm, URACO algorithm and NCMLT algorithm is shown in Fig. 9. With the increase of the number of rounds, the network energy consumption of the three algorithms continues to increase, but the rising trend of NCMLT algorithm is slower than that of EEUC algorithm and URACO algorithm. After the network runs for a period of time, some nodes die, but the network can still be maintained by using other nodes to rebuild routes. At this time, due to the reduction of the number of nodes, the growth rate of energy consumption slows down. When most of the nodes are dead and the data cannot be transmitted smoothly, most of

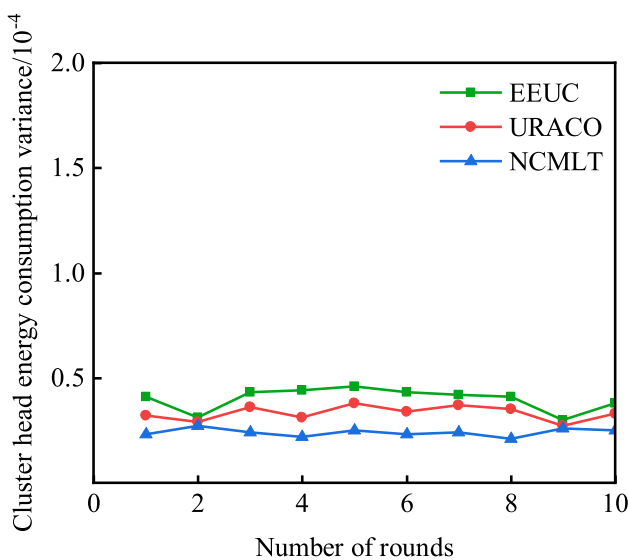


Fig. 8 The variance of the energy consumption of the cluster head

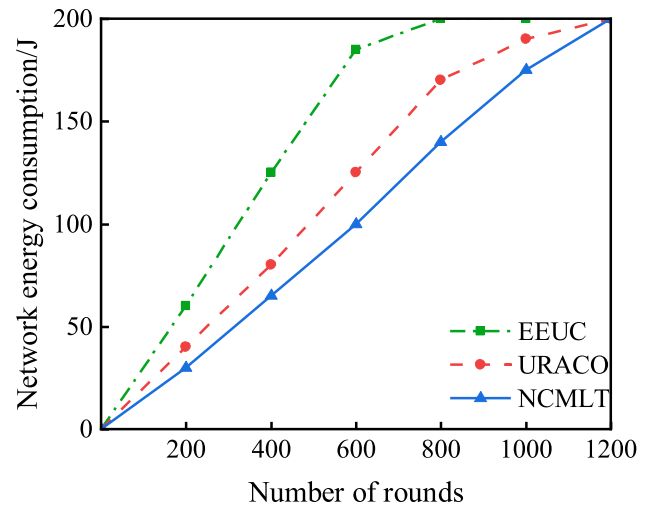


Fig. 9 Network energy consumption

the remaining energy of the non-dead nodes is used to send out signals to find other nodes that can communicate until the energy is exhausted. In the whole process, under the same number of rounds, the network energy consumption of the NCMLT algorithm is the least, which indicates that the NCMLT algorithm effectively allocates the node data transmission distance through the node classification, reduces the energy consumption and prolongs the network life.

### 5.8 Average latency and throughput rate

Data transmission delay time is one of the important indexes to measure the performance of routing algorithms. The smaller the delay time, the better the stability of data transmission. The relationship between the average transmission delay time of the three algorithms and the number of nodes is shown in Fig. 10. As can be seen from the

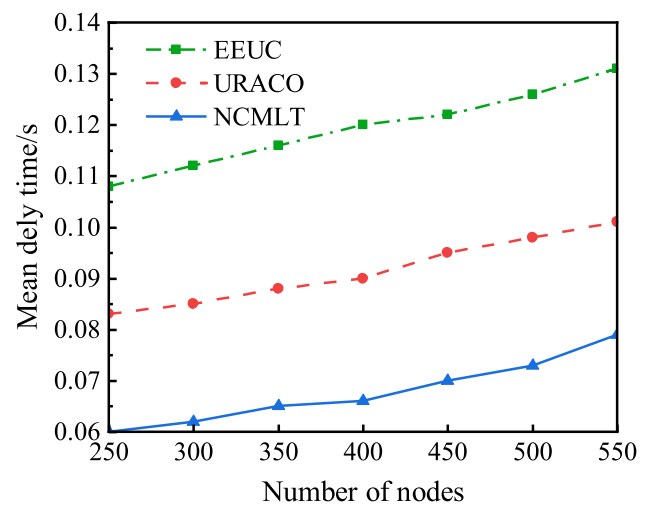


Fig. 10 Mean delay time

figure, the average delay time of EEUC algorithm is the largest, followed by URACO algorithm and NCMLT algorithm. This is because the initial information flooding of the EEUC algorithm increases the transmission time. URACO algorithm and NCMLT algorithm establish a clear optimized transmission path and reduce time-consuming information flooding, but there are many data factors in URACO algorithm that need to be forwarded between nodes, which also causes some delays in data transmission. The NCMLT algorithm transmits less data on the optimal path and reduces latency. Through simulation data analysis, the average data delay of NCMLT algorithm is reduced by 45% compared with EEUC algorithm and 26.67% compared with URACO algorithm.

The main purpose of deploying WSNs nodes is to collect various data in the monitoring area, so the success of data receiving is an important basis for evaluating the effectiveness of the algorithm. The throughput rate is defined as the percentage of data successfully received by the receiving node. The throughput rates of the three algorithms are shown in Fig. 11. It can be seen from the figure that the network throughput of the three algorithms increases with the increase of the number of network nodes. This is mainly because as node density increases, more sensor nodes participate in forwarding data, which further reduces the probability of data retransmission. Compared with EEUC algorithm and URACO algorithm, NCMLT algorithm optimizes the selection of relay nodes, and the construction of cluster chain hybrid makes the network more balanced, so that the Sink node can receive data more smoothly.

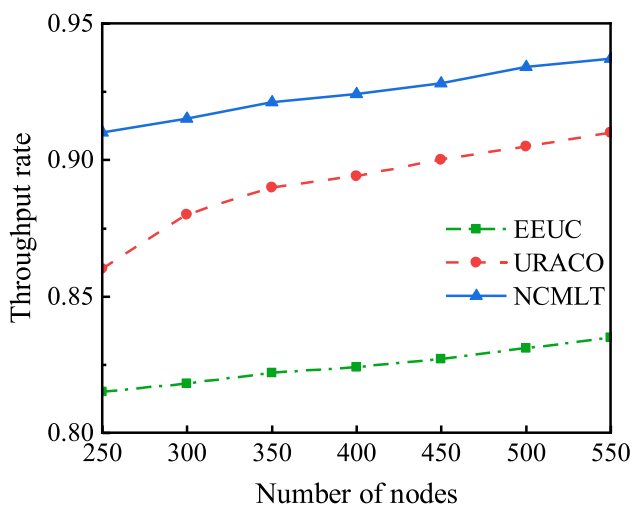


Fig. 11 Throughput rate

## 6 Conclusion

Aiming at the problem of uneven energy consumption of nodes in wireless sensor network, this paper proposes a non-uniform clustering routing algorithm based on node classification and multi-level data transmission, which classifies nodes and transmits data in different structures, and reasonably optimizes the distance of data transmission by each node. The threshold formula for selecting cluster heads and the calculation of the competition radius refer to the regulators of energy and distance, which effectively balances the cluster head load. By calculating the remaining energy and the distance between the two hops, the cost function and the cluster chain relay function are established, and the multi-hop and cluster chain combination between clusters are realized. Simulation results show that NCMLT can effectively balance the energy consumption of wireless sensor network nodes and prolong the network life.

**Acknowledgements** This work was in part supported by Hunan Provincial Natural Science Foundation of China (2021JJ50093); the National Natural Science Foundation of China (No. 11875164).

## References

- Zheng, M., Chen, L., Liang, W., et al. (2017). Energy-efficiency maximization for cooperative spectrum sensing in cognitive sensor networks. *IEEE Transactions on Green Communications and Networking*, 1(1), 29–39.
- Zhao, X. Q., Cui, Y. P., Guo, Z., et al. (2022). Energy-efficient clustering routing protocol for wireless sensor networks based on virtual force. *Journal of Software*, 33(2), 622–640.
- Wang, J. H., & Shi, W. X. (2018). Survey on cluster-based routing protocols for cognitive radio sensor networks. *Journal on Communications*, 39(11), 156–169.
- Yu, X. W., Li, P., Liu, Y., et al. (2021). A non-uniform clustering routing protocol based on dynamic competitive radius. *Chinese Journal of Sensors and Actuators*, 34(3), 400–406.
- Yan, J., Zhou, M., & Ding, Z. (2016). Recent advances in energy-efficient routing protocols for wireless sensor networks: A review. *IEEE Access*, 4, 5673–5686.
- Mehrabi, A., & Kim, K. (2016). Maximizing data collection throughput on a path in energy harvesting sensor networks using a mobile sink. *IEEE Transactions on Mobile Computing*, 15(3), 690–704.
- Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4), 660–670.
- Wang, B., & Fu, D. S. (2016). Improvement of LEACH routing protocol in wireless sensor networks. *Instrument Technique and Sensor*, 126, 71–74.
- Lindsey, S. (2002). PEGASIS: Power efficient gathering in sensor information systems. In *Proceedings of IEEE aerospace and electronic systems society*, pp. 1125–1130. IEEE.
- Huang, L. X., Wang, H., Yuan, L. Y., et al. (2017). Improved LEACH protocol algorithm for WSN based on energy balance

- and high efficiency. *Journal on Communications*, 38(S2), 164–169.
11. Li, C. F., Chen, G. H., Ye, M., et al. (2007). An uneven cluster-based routing protocol for wireless sensor networks. *Chinese Journal of Computers*, 30(1), 27–36.
  12. Liu, W., Du, J. H., Jia, S. L., et al. (2019). Energy efficient clustering routing protocol for wireless sensor networks. *Journal of Beijing University of Aeronautics and Astronautics*, 45(1), 50–56.
  13. Huang, Y., & Hua, Y. Q. (2020). Routing optimization of the wireless sensor networks with energy and path constraints. *Journal of Xidian University*, 47(3), 113–120.
  14. Bi, X. J., & Diao, P. F. (2016). Routing and clustering algorithm heterogeneous wireless sensor networks based on gravitational search algorithm. *Control and Decision*, 32(3), 563–569.
  15. Liu, S. Y., Zheng, Y. L., & Bai, Y. G. (2019). Energy balance routing algorithm based on forward-aware for wireless sensor networks. *Control and Decision*, 34(7), 1425–1432.
  16. Bhatti, D. M. S., Saeed, N., & Nam, H. (2016). Fuzzy c-means clustering and energy efficient cluster head selection for cooperative sensor network. *Sensors*, 16(9), 1–17.
  17. Hu, Y., Niu, Y. G., & Zou, Y. Y. (2017). A zone-based unequal multi-hop clustering algorithm in WSNs. *Control and Decision*, 32(9), 1695–1700.
  18. Liu, Z., Feng, X., Zhang, J. F., et al. (2016). Improved cross region GPSR routing algorithm based on adjustable grid. *Journal of Jilin University (Science Edition)*, 54(4), 852–856.
  19. Jiang, B., Mao, T., Tang, D., Wu, Z., et al. (2017). Clustering routing algorithm based on farmland wireless sensor network. *Transactions of the Chinese Society of Agricultural Engineering*, 33(16), 182–187.
  20. Li, H. B., Liu, Z. L., Chen, Q., et al. (2020). Cluster routing algorithm for wireless sensor networks based on hierarchical neighboring nodes. *Computer Engineering*, 46(6), 187–195.
  21. Dwivedi, B., Patro, B., Srivastava, V., et al. (2021). LBR-GWO: Layered based routing approach using grey wolf optimization algorithm in wireless sensor networks. *Concurrency and Computation: Practice and Experience*, 12(11), 12–21.
  22. Ahmed, S., Gupta, S., Suri, A., et al. (2021). Adaptive energy efficient fuzzy: An adaptive and energy efficient fuzzy clustering algorithm for wireless sensor network-based landslide detection system. *IET Networks*, 10(1), 1–12.
  23. Wang, Q., Lin, D. Y., Yang, P. F., et al. (2019). An energy-efficient compressive sensing-based clustering routing protocol for WSNs. *IEEE Sensors Journal*, 19(10), 3950–3960.
  24. Fu, J. S., Liu, Y., Chao, H. C., et al. (2016). Green alarm systems driven by emergencies in industrial wireless sensor networks. *IEEE Communications Magazine*, 54(10), 16–21.
  25. Liu, H., & Li, H. W. (2018). Uneven clustering routing algorithm based on ant colony optimization. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 46(8), 50–54.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Xiuwu Yu** Xiuwu Yu received his Ph.D. degree on wireless monitoring in Safety science and Engineering from Wuhan University of Technology, Wuhan China in 2013. Currently he is working as a Professor in School of Resource & Environment and Safety Engineering at University of South China, Hengyang China. His research interests include localization, routing, data fusion and coverage control in wireless sensor network.



**Yin hao Liu** is now working toward his Master Degree in the School of Resource & Environment and Safety Engineering, University of South China (USC), Hengyang, China. His research interest is wireless sensor network.



**Yong Liu** received his Ph.D. degree in Safety science and Engineering from University of South China, Hengyang China in 2015. In 2016, he held a post as the director of Hunan engineering research center for uranium tailings decommission and treatment, Hengyang, China. He was the head or member of more than 10 foundation projects. His current research interests include the safety comprehensive evaluation and uranium tailing monitoring.