Computer Networks 167 (2020) 106994

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

# A clustering routing algorithm based on wolf pack algorithm for heterogeneous wireless sensor networks

Y.U. Xiu-wu[a,b], Y.U. Hao[a,*], Liu Yong[a,b], Xiao Ren-rong[a]

[a] University of South China, 28 Changsheng West Road, Hengyang, 421001, Hunan Province, China
[b] Hunan Engineering Research Center for Uranium Tailings Decommission and Treatment, 28 Changsheng West Road, Hengyang, 421001, Hunan Province, China

ARTICLE INFO

*Article history:*
Received 6 July 2019
Revised 23 October 2019
Accepted 12 November 2019
Available online 13 November 2019

*Keywords:*
Heterogeneous sensor networks
Wolf pack algorithm
Routing algorithm
Dynamic clustering

ABSTRACT

In order to maximize the network performance of heterogeneous sensor networks and effectively control the network cost, a clustering routing algorithm based on wolf pack algorithm (CLWPA) for heterogeneous wireless sensor networks is proposed. Firstly, the optimal deployment of heterogeneous nodes is transformed into a mixed integer programming problem. The approximate optimal solution of the problem is obtained by through the wolf pack algorithm (WPA) which improved by logistic function and levy flight, then a heterogeneous network routing algorithm based on the improved wolf pack algorithm (LWPA) is proposed. Secondly, in order to solve the problem of fixed path in LWPA routing algorithm, the concept of edge degree is introduced to improve DEEC algorithm. The improved DEEC algorithm (IDEEC) is used to dynamically cluster common nodes in heterogeneous networks, and the data transmission mode is carried out after the clustering mode set. Finally, through simulation analysis, compared with other three heterogeneous network routing algorithms, CLWPA algorithm effectively prolongs the network's stable period and lifetime, and the energy consumption is more balanced.



## 1. Introduction

Wireless Sensor Network (WSN) is a vital link in the Internet of Things project to undertake intelligent sensing. It monitors designated areas in real time through sensor nodes and sends the acquired data to sink nodes in a wireless multi-hop manner. It has a very wide application prospect in the fields of national defense security, environmental monitoring, disaster early warning and so on [1,2]. Among them, routing algorithm, as one of the core technologies of WSN, is mainly used to find the optimal path from sensor node to sink node and accurately forward monitoring data along the specified path, so as to reduce network energy consumption and prolong the lifetime. Due to the limitation of wireless sensor network resources, it is quite different from the traditional wired network, and many traditional wired routing technologies are not suitable for it. Therefore, it has higher scientific research value to study reasonable and effective WSN routing algorithm according to its own characteristics [3,4].

The design goals of routing algorithms in wireless sensor networks mainly include finding energy-efficient data transmission paths, maximizing the lifetime of the network, improving the robustness and reliability of routing, supporting data fusion and data forwarding, etc. In addition, according to different monitoring environments and application conditions, some other requirements are needed, such as network security and self-adaptability [5,6,7]. WSN can be divided into four different heterogeneous types according to the difference of nodes' computing capability, sensing capability, communication capability and energy factors: computing energy heterogeneous type, node energy heterogeneous type, link heterogeneous type and network protocol heterogeneous type [8].

Energy heterogeneous sensor networks with different initial energies have more direct effect and practical significance in reducing network energy consumption and improving network lifetime. Document [9] proves that using the small-world characteristics of sensor networks, deploying heterogeneous nodes that can communicate directly with sink nodes in node-intensive areas and forming super links will enable heterogeneous sensor networks to have smaller average path length and higher clustering coefficient at the same time, thus changing their communication distance by adjusting the transmission power of heterogeneous nodes, improving the energy utilization rate of nodes, reducing the network communication overhead, making the network energy consumption more balanced and prolonging the network life cycle.

* Corresponding author.
*E-mail address:* 616176246@qq.com (Y.U. Hao).

https://doi.org/10.1016/j.comnet.2019.106994
1389-1286/

In actual design, a heterogeneous sensor network can be formed by setting up power supply for appropriate sensor nodes, thus prolonging the network lifetime of WSN. This kind of design has more advantages in areas with harsh environment and is also the overall trend of wireless sensor network design. Even in homogeneous sensor networks, there are characteristics of sensor networks with heterogeneous energy. For example, when a sensor network works for a period of time, new sensor nodes will be added in order to prolong the service time of the network. At this time, the new node has more energy than the old node. However, the cost of heterogeneous nodes is relatively expensive. Considering the cost of sensor networks, heterogeneous nodes cannot be deployed in WSN without restriction [10,11]. Therefore, how to optimize the deployment of heterogeneous nodes to achieve optimal network performance of sensor networks is a problem that must be solved.

As a novel swarm intelligence optimization algorithm, wolf pack algorithm (WPA) has been successfully used to optimize many linear programming models and large-scale nonlinear programming models. WPA has also achieved good results in solving many different types of optimization problems [12]. However, the WPA still has some shortcomings, such as: in the process of solving optimization problems, the WPA is easy to fall into local optimization, resulting in premature algorithm; There are many parameters in the WPA, and the algorithm is sensitive to the setting of some parameters, so the setting of parameters will affect the performance of the algorithm [13].

Based on the above analysis, the main contributions of this paper are as follows:

(1) The global search capability of the WPA is used to find the optimal solution of the optimal deployment model of heterogeneous nodes, and use logistic function and Levy flight to improve the running step size and besieging step size of the WPA respectively to enhance the search capability of the algorithm and prevent the algorithm from entering local optimization.
(2) Dynamic clustering is carried out for sensor networks with heterogeneous nodes deployed. Edge degree is used to improve the threshold function of cluster head election in DEEC algorithm, so as to avoid a large number of cluster heads being distributed on the edge of monitoring area, thus consuming more energy for data transmission and accelerating the death of cluster heads.
(3) Propose a clustering routing algorithm based on WPA for heterogeneous wireless sensor networks (CLWPA). Simulation results show that CLWPA makes the energy consumption of common nodes in the network more balanced and effectively improves the life cycle and stability cycle of sensor networks.

The structure of this paper is as follows: In the Section 2, we introduce the literature review related to clustering heterogeneous network routing algorithms. We introduce the network model and energy consumption model of the system in the Section 3. In the Section 4, we introduce the optimal deployment of heterogeneous nodes, the WPA and the improvement of the WPA, and use the improved WPA to obtain the optimal solution of the optimal deployment problem. A clustering routing algorithm for heterogeneous sensor networks based on improved WPA is proposed in the Section 5. In the Section 6, we analyzed the complexity of the data and simulated the impact of the algorithm on the number of fixed nodes, specific network energy consumption and network lifetime. Finally, we made a summary in Section 7.

## 2. Related work

In order to reduce network energy consumption and improve network life cycle, many effective routing algorithms have been proposed [14-16]. Among them, clustering routing algorithm has become the focus of WSN routing algorithm research due to its good scalability and excellent network performance [17,18]. Typical clustering routing algorithms include LEACH [19], TEEN [20], SEP [21], DEEC [22], etc. SEP and DEEC are the basis of many heterogeneous WSN clustering routing algorithms, but SEP algorithm is only suitable for secondary heterogeneous WSN, and the cluster head uses a single hop method to send data to the base station, so the algorithm has great limitations. DEEC is commonly used in multi-level heterogeneous networks. The algorithm uses the ratio of average network energy and node residual energy to control the probability of nodes becoming cluster heads. Therefore, nodes with more initial energy and residual energy are more likely to become cluster heads.

Literature [23] has made a very detailed study on the performance comparison between heterogeneous sensor networks and homogeneous sensor networks. The research results show that sensor networks with appropriate addition of heterogeneous nodes have greatly improved performance and life cycle. Therefore, this paper will not compare the performance of heterogeneous networks and homogeneous networks, but mainly compare the performance with the following several routing algorithms of heterogeneous networks.

The EDDEEC algorithm in document [24] improves the cluster head election of the basic DEEC algorithm and takes into account the location distribution of nodes, making it easier for nodes close to the base station to become cluster heads, but without protecting nodes with low initial energy, it is easy to cause the nodes near the base station to accelerate death due to repeated becoming cluster heads. The K-means clustering routing algorithm (KCA) proposed in document [25] randomly selects $K$ node positions as initial clustering centers in the network, then calculates the best clustering center through the K-means algorithm and deploys heterogeneous nodes as cluster heads at the best clustering center. The network is clustered only once, and the network structure remains unchanged after clustering. K-means algorithm only calculates a local approximate solution and is greatly influenced by the initial value. The optimization degree of the solution obtained by improper selection of the initial value is not high. Document [26] proposes a heterogeneous network routing algorithm based on mixed integer programming (HRMIP). The average path length of nodes is shorter, and the overall energy consumption of the network is correspondingly reduced. However, after the deployment of heterogeneous nodes, the data transmission path will remain fixed and will not change, so that nodes with higher energy consumption will consume more energy, resulting in the rapid death of nodes with higher energy consumption and a short network stability period.

Although WPA is a new type of swarm intelligence algorithm proposed only in recent years, since it was proposed, many experts and scholars have conducted in-depth optimization research on it, and it has been widely used in practical production and life problems. For example: Three-Dimensional Underwater Path Planning [27]; Traveling salesman problem [28]; UAV route planning [29]; Multidimensional knapsack problem [30], etc. Compared with other swarm intelligence algorithms of the same type, such as Particle Swarm Optimization, Genetic Algorithm, Ant Colony Optimization, etc. the WPA has better optimization performance through optimization tests and comparative experiments on the same function [31].

## 3. System model

### 3.1. Network model

The sensor network with heterogeneous energy needs long-term operation and the node position is relatively fixed. In order not to lose generality, assuming that the monitoring area is a two-dimensional square area with an area range of $R \times R$. A certain number of sensor nodes are unevenly and randomly distributed in the monitoring area, and the nodes periodically transmit monitoring information to sink nodes. The heterogeneous sensor network has the following properties:

First: All nodes are stationary or move only slightly. sink nodes and heterogeneous nodes can be deployed anywhere; There is only one sink node in the monitoring area;
Second: The number of common nodes is $N$, Node location coordinates are known as $(x_i,y_i)$, and $1 \le i \le N$, $0 \le x_i \le R$-1, $0 \le y_i \le R$-1;
Third: Heterogeneous nodes directly communicate with sink nodes through a super link without bandwidth limitation after receiving monitoring data of surrounding common nodes;
Fourth: All nodes have certain data fusion capability and unique ID.

### 3.2. Energy consumption model

In this paper, the energy consumption model of radio frequency free space communication is adopted. See Eq. (1) to (4) for the energy consumption model [8,9]:

$$E_{Tx}(k,d) = \begin{cases} kE_{elec} + k\varepsilon_{fs}d^2, d < d_0 \\ kE_{elec} + k\varepsilon_{mp}d^4, d \ge d_0 \end{cases} \tag{1}$$

$$E_{Rx}(k) = k \times E_{elec} \tag{2}$$

$$E_c = (M+1) \times k \times E_{DA} \tag{3}$$

$$d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mp}} \tag{4}$$

Where $k$ is the size of the transmitted data, $d$ is the distance between sender and receiver, $d_0$ is threshold of node communication distance, $E_{Tx}(k,d)$ is the energy consumption of the sending end, $E_{Rx}(k,d)$ isthe energy consumption of the receiving end, $E_c$ is the energy consumption of data fusion, $E_{elec}$ is the energy consumption of 1-bit data during transmission or reception, $\varepsilon_{fs}$ and $\varepsilon_{mp}$ are all constant distributions that indicates amplification factors of circuit signal amplifiers, $E_{DA}$ is the energy consumption in the process of 1-bit data fusion, $M$ is the number of nodes in the cluster.

## 4. Optimal deployment of heterogeneous nodes based on WPA

### 4.1. An optimal deployment of heterogeneous nodes

The energy consumption in WSN is mainly related to the path length during data transmission, so the problem of optimal deployment of heterogeneous energy nodes can be transformed into the problem of solving the minimum value of the sum of distances from all common nodes to sink nodes [32]. The objective function of the optimization problem can be expressed by Eq. (5):

$$f = \min \sum_{i=1}^{N} d_i \tag{5}$$

Where $N$ is the number of ordinary nodes, $d_i$ is distance from common node to sink node via multiple hops.

In a randomly distributed heterogeneous sensor network, the optimal deployment of heterogeneous nodes represented by Eq. (5) is actually a NP-hard problem. In order to establish an optimal deployment model for heterogeneous nodes, the following conditions are given:

First: The number of heterogeneous nodes added is $\alpha$ and can only be selected from the positions of ordinary nodes;
Second: $v_0$ represents sink node, its position coordinate is $(x_0,y_0)$; $L$ is the total set of all nodes in the network, its location set is $(X,Y)$, then:

$$L = \{v_i, 0 \le i \le N\}$$

$$(X,Y) = \{(x_i,y_i), 0 \le i \le N\}$$

Third: The sink node is regarded as a heterogeneous node which position coordinate is $(x_0,y_0)$, Write $h_0=(u_0,v_0)$; the set of heterogeneous nodes is $H$ and the set of locations is $(U,W)$, then:

$$H = \{h_t, 0 \le t \le \alpha\}$$

$$(U,V) = \{(u_t,v_t), 0 \le t \le \alpha\}$$

Fourth: Where $d_{i,j}$ is the distance from the common node to the sink node when the common node $v_i$ forwards data to the sink node through the heterogeneous node $h=(u,w)$ at $v_j$, and because $h$ shares location information with $v_j$, $u=x_j$, $w=y_j$, then:

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, 1 \le i \le N, 0 \le j \le N$$

Based on the above conditions, the following mathematical model is established to describe the optimal deployment of heterogeneous nodes:

$$f = \min \sum_{i=1}^{N} \sum_{j=0}^{N} d_{i,j} z_{i,j} \tag{6}$$

S.T.

$$\begin{cases} ① \sum_{j=0}^{N} p_j = \alpha + 1 \\ ② \sum_{j=0}^{N} z_{i,j} = 1, 1 \le i \le N \\ ③ z_{i,j} \le p_j, 1 \le i \le N, 0 \le j \le N \\ ④ p_0 = 1, p_j \in \{0,1\}, 1 \le j \le N \\ ⑤ z_{i,j} \in \{0,1\}, 1 \le i \le N, 0 \le j \le N \end{cases}$$

Constraint ① indicates that the total number of heterogeneous nodes in the network is $\alpha+1$; Constraint ②indicates that any common node can only send data to sink node through a heterogeneous node; Constraint ③indicates that if heterogeneous nodes are not deployed at $v_j$, it is impossible for any common node to directly send data from $v_j$ to sink node; Constraint ④ indicates that $p_j$ is a 0–1 variable, and when $p_j=1$, heterogeneous nodes are deployed at node $v_j$, otherwise $p_j$ is 0; Constraint⑤indicates that $z_{i,j}$ is a variable of 0–1. when $z_{i,j}=1$, the data of common node $v_i$ is forwarded to sink node through heterogeneous nodes at $v_j$; otherwise $z_{i,j}$ is 0; The objective function $f$ represents the minimum value of the sum of distances from all common nodes to sink nodes under the above constraints. In this paper, WPA is used to solve the optimal value of the mixed integer programming problem, and then heterogeneous nodes are deployed in the network.

### 4.2. Wolf pack algorithm

WPA abstracts three kinds of intelligent behaviors (Running, Summoning and Besieging) as well as the "Winner takes all" generation rules and the "Survival of the strong" wolves update mechanism by simulating the predation behavior and prey distribution mode of wolves (including head wolves, fierce wolves and scout wolves), so as to ensure the reproduction and development of wolves. WPA's specific process is as follows:

(1) Initialization: set the basic parameters of the algorithm. $S_n$ is the total number of wolves, $T_{\max}$ is the maximum iteration number, $D$ is the dimension of the space to be searched, G is the number of wolf pack search directions, $W_{\max}$ is the maximum number of wolves to be searched, *stepa* is the step of searching, *stepb* is the step of running, *stepc* is the step of besieging, the initial position of the $i$-wolf in the $D$-dimensional decision space in the $N$-dimensional is:

$$X_{i,j} = X_{\min} + rand(X_{\max} - X_{\min}) \tag{7}$$

Where $1 \le i \le S_n$ and $1 \le j \le D$; *rand* is representing a random number uniformly distributed in the interval [0,1]; $X_{\min}$ and $X_{\max}$ are the upper and lower limits of $X_{i,j}$ respectively. Calculate the objective function value and fitness value of all wolves' positions, and select the wolf with the largest fitness value as the head wolf. The head wolf only guides the wolves' wandering, calling and besieging behaviors and does not directly participate until it is replaced by other wolves with better objective function value.

(2) Running behavior: Wolves must hunt in nature for survive. In order to improve hunting efficiency, the wolf scouts must run in all directions where they are located. In addition to the head wolf, $q$-wolves are selected to perform wandering behavior, the wolves run in $g$-directions around it, and the fitness value of the current position of the wolves is recorded, then the direction with larger fitness value is selected to continue wandering, and the current position of the wolves is updated, and the position of the $i$-wolve after advancing to the $e$ direction $(1 \le e \le g)$ is as follows:

$$X_{i,j}(e) = X_{i,j} + stepa \times Rand \tag{8}$$

Where *Rand* is a random number evenly distributed on the interval [−1,1]; *Stepa* is the searching step length. Repeat the above behavior until the prey whose fitness value is greater than that of the head wolf is found or the maximum number of wanderings $W_{\max}$ is reached, ending the running behavior and turning on the summoning behavior.

(3) Summoning behavior: After the wolf scouts have found the prey, the head wolf calls the nearby fierce wolf by howling. After the fierce wolf hears the call of the head wolf, the fierce wolf approaches the head wolf quickly with a large running step *stepb* and exchanges information. Then, the position equation of fierce wolf $i$ after the $T+1$ iteration:

$$X_{i,j}^{T+1} = X_{i,j}^{T} + Rand\left(X_{c,j}^{T} - X_{i,j}^{T}\right)stpeb \tag{9}$$

Where the *stepb* is the running step length, and $X_{i,j}^{T}$ represents the position of the wolf after the $T$th iteration. The equation indicates that the fierce wolf runs to the position of the head wolf under the guidance of the head wolf and starts the besieging.

(4) Besieging behavior: If the fierce wolf finds prey whose fitness is greater than the head wolf in the process of running, the fierce wolf will turn into the head wolf and direct the action of the wolf pack; otherwise, when the distance between the fierce wolf and the head wolf is less than the judgment distance, the head wolf will inform the fierce wolf to besiege through howling. Fierce wolves besieged with besieging step *stepc*, and the besieging behavior equation is as follows:

$$X_{i,j}^{T+1} = X_{c,j}^{T} + Rand \times stpec \tag{10}$$

The determination distance here can be calculated from Eq. (11):

$$d_{\lim it} = \frac{1}{D \times \omega} \times \sum_{j=1}^{D} \left|\max_j - \min_j\right| \tag{11}$$

Where the $[\max_j, \min_j]$ is the value range of the $j$th variable where the wolf is located; $\omega$ is the distance control factor. If the fitness value of the target is greater than the current value after the wolf scout attacks, it will replace the current position. Otherwise, it will remain unchanged.

(5) Population update: The wolf pack is updated by distributing food, the wolf with the worst function value is eliminated, and the same number of wolves are randomly generated to replace the wolf pack according to the initialized wolf pack position Eq. (7) until the iteration number reaches the maximum value $T_{\max}$, and the fitness function value and the position information of the optimal deployment of heterogeneous nodes are output, otherwise, (2) running behavior is executed.

### 4.3. Algorithm improvement

WPA has various search strategies and strong global search capability, but there are still some deficiencies in solving optimization problems. For example, in the summoning and besieging behaviors, the running step size and besieging step size are fixed values, which reduces the global searching ability of the algorithm, makes the algorithm fall into local optimization prematurely, and causes the algorithm to be premature. In view of the above problems, the running step length and the besiege step length in WPA are improved.

#### 4.3.1. Improvement of running step

In the basic WPA algorithm, the summoning behavior uses a fixed running step size, which reduces the algorithm's local search capability and diversity of search strategies and makes the algorithm easily fall into local optimization prematurely. Therefore, Logistic Function is introduced in this paper. Logistic function is bounded, continuous, derivable and strictly monotonic: when $x$ approaches negative infinity, $y$ approaches zero; When $x$ approaches positive infinity, y approaches 1; When $x=0$, $y=0.5$. The calculation equation is shown in Eq. (12):

$$y = \frac{1}{1+e^{-x}} \tag{12}$$

At the beginning of the summoning behavior, the fierce wolf needs to stride towards the position of the head wolf and gradually reduce the running length as the distance decreases, moving slowly towards the prey. Through Logistic function, the running step can be transformed into variable and mapped into interval (0,1), so that the running step decreases in interval (0,1) and the optimal solution can be searched more accurately. The improved running step equation is shown in Eq. (13):

$$stepb = \frac{1}{1+e^{2\ln 100 \times \frac{T}{T_{\max}} - \ln 100}} \times \frac{T_{\max} - T + 1}{T_{\max}} \tag{13}$$

#### 4.3.2. Improvement of besieging step

Various studies have shown that the flight behavior of many animals and insects in the process of predation shows Levy flight with power law, that is, long-time short-distance back-and-forth search trajectory and short-time long-distance search trajectory are interspersed with each other, which is the main characteristic of Levy flight. Scholars have also confirmed that many birds in nature also follow Levy flight, especially when searching for targets in a large space and with limited searchers, Levy flight is one of the most effective search strategies.

In the basic besiege behavior of the wolf pack algorithm, the fixed besieging step size reduces the search capability of the algorithm. Levy flight is applied to the besiege. The large step in the early stage of the besiege is used to find the target, which expands the search scope to avoid the algorithm falling into local optimization. The small step size in the later stage of the besiege is used for precise search, making wolves search for the global optimal solution in a small range. The besiege behavior equation of the improved wolf pack algorithm is shown in Eq. (14):

$$X_i(T+1) = X_{best}(T) + rand \otimes s \otimes |X_{best}(T) - X_1(T)| \tag{14}$$

In the equation, $X_i(T)$ is the position of the $i$th besiege wolf in the $t$th generation; $X_{best}(T)$ is the optimal solution for the current wolf pack. $s$ is the random step size of Levy flight, which can be obtained from Eq. (15):

$$s = \frac{\mu}{\upsilon^{\frac{1}{\beta}}} \tag{15}$$

Where the parameters $\mu$, $\upsilon$ conform to normal distribution: $\mu$~$N$ (0, $\sigma^2$), $\upsilon$~ (0, 1), $\sigma$ can be obtained from Eq. (16):

$$\sigma = \left\{ \frac{\Gamma(1+\beta)\sin\left(\frac{\pi\beta}{2}\right)}{\beta\Gamma\left(\frac{1+\beta}{2}\right)2^{\frac{\beta-1}{2}}} \right\}^{\frac{1}{\beta}}, \ (0 \le \beta \le 2) \tag{16}$$

Eq. (13) to (15) show that the length of the besiege behavior is random and has no fixed size and direction. In the process of searching iteration, the new solution near the local optimal solution makes the WPA easily jump out of the local optimal solution, improving the quality of the optimal solution and enhancing the searching ability of the algorithm. The flowchart of WPA with Logistic function and levy flight (LWPA) is as follows. (Fig. 1).

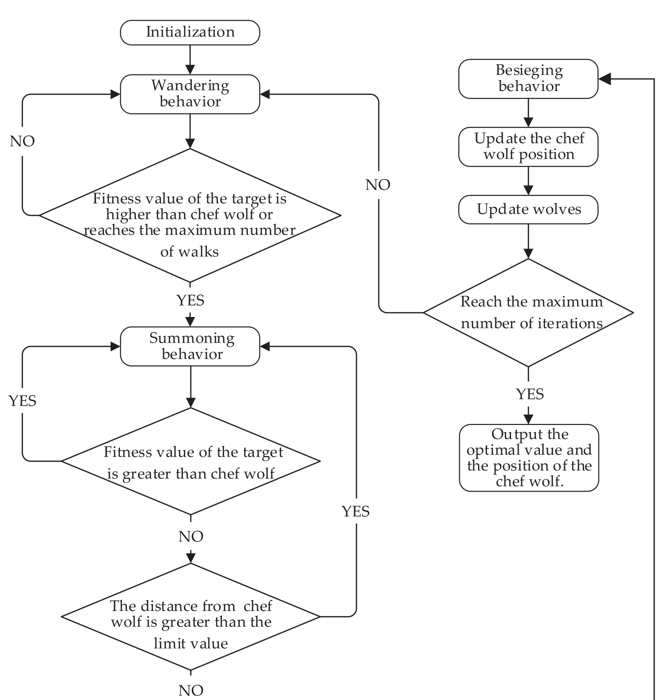


**Fig. 1.** The process of LWPA.

## 5. Cluster routing algorithm for heterogeneous sensor networks

The energy heterogeneous sensor network based on LWPA algorithm which is proposed in Section 4 is essentially equivalent to a static clustering network with heterogeneous nodes as cluster heads and corresponding cluster division. In LWPA heterogeneous sensor networks, heterogeneous nodes becoming cluster heads will not lead to premature node death due to energy consumption. However, for large-scale sensor networks, static clustering networks have problems such as fast death of non-heterogeneous nodes, short network life cycle and stability cycle, large redundancy of transmission data, high network energy consumption, etc.

In order to overcome the shortcomings of static clustering networks, this section proposes an energy heterogeneous network routing algorithm based on dynamic clustering on the basis of LWPA routing algorithm (CLWPA).The core idea of the algorithm is: Firstly, the number of heterogeneous nodes with $\phi$ is preset, and the heterogeneous nodes are deployed according to the optimal deployment algorithm of heterogeneous nodes in Section 4; Then common nodes in the network are clustered according to the improved DEEC algorithm and cluster heads are selected. After clustering, member nodes in the cluster send data to the cluster head of the cluster where they are located; Finally, the cluster head node performs corresponding processing on its own data and the received data, and sends the processed data to the sink node or the corresponding heterogeneous node through multiple hops of the cluster head layer node.

### 5.1. The election of cluster head

Dynamic clustering of wireless sensor networks can reduce the communication distance between nodes, reduce the redundancy of transmission data, and reduce the energy consumption of the network. Compared with LEACH and SEP algorithms, the classical DEEC algorithm takes into account the initial energy of nodes and the current residual energy of nodes in the process of cluster head election, which increases the probability of high initial energy nodes and high residual energy nodes becoming cluster heads, balances the network load and prolongs the network life cycle. However, DEEC algorithm does not consider the location of nodes in the cluster head election process. If a large number of cluster heads are distributed at the edge of the monitoring area and the cluster heads are far away from heterogeneous nodes or sink nodes, excessive energy will be consumed in the data transmission process, resulting in rapid death of cluster heads. To overcome this shortcoming, this section introduces the concept of edge degree to improve the threshold function $O(v_i)$ of cluster head election. On this basis, DEEC clustering routing algorithm based on edge degree (IDEEC) is proposed. The calculation is shown in Eq. (17):

$$O(v_i) = \begin{cases} \frac{Kp_i}{1-p_i\left(r \bmod \frac{1}{p_i}\right)}, & v_i \in G \\ 0, & v_i \notin G \end{cases} \tag{17}$$

Where $G$ indicates all common nodes of the $1/p_\mathrm{i}$ round that do not serve as cluster heads; $p_i$ is the probability of common node $v_i$ becoming cluster head; $r$ is the current number of rounds; $K$ is edge degree, The calculation equation is shown in Eq. (18):

$$K = A\frac{E_i(r)}{\bar{E}(r)} + B\left[1 - \exp\left(-\frac{d_{toBS}}{d_{i,j}}\right)\right] \tag{18}$$

Where $A$ and $B$ are the control factor, The value range is between intervals (0,1) and the sum of the two is 1;$E_i(r)$ is the remaining energy of the node $v_i$ of the $r$th wheel; $\bar{E}(r)$ is the average energy of all common nodes in the $r$th round; $d_{avg}$ is the average distance from all cluster heads to sink nodes;$d(i)$ is the distance from the cluster head to the sink node. Since the energy consumption of heterogeneous nodes does not need to be considered in the energy heterogeneous sensor network, cluster head nodes are selected from ordinary nodes, and the probability that each ordinary node becomes a cluster head is as shown in Eq. (19):

$$p_i = \frac{1}{1+k\lambda} \times \frac{E_i(r)}{\bar{E}(r)} \times p_{opt} \tag{19}$$

Where $P_{opt}$ is the ratio of the set number of cluster heads to the total number of nodes; $\lambda$ is the ratio of the number of heterogeneous nodes to the total number of nodes; $k$ is the ratio of the total initial energy of heterogeneous nodes to common nodes. According to Eq. (17) and (19), the calculation Eq. (20) of the optimal value $n_{CH}$ of the cluster head number of each round can be obtained:

$$n_{CH} = \frac{\sum_{i=1}^{n} p_i}{p_{opt}} \tag{20}$$

Since the nodes are uniformly distributed in the monitoring area, the expectation of the square of the distance between the nodes in the cluster and the cluster head can be calculated by Eq. (21):

$$E\left[d_{toCH}^2\right] = \frac{R^2}{2\pi n_{CH}} \tag{21}$$

In order to calculate the average distance from the cluster head to the sink node, assuming that the coordinates $(x_0,y_0)=(\lambda_R,\psi_R)$, $0 \le \lambda$, $\psi \le 1$, any cluster head coordinates are $(x,y)$ and uniformly distributed in the monitoring area, and the probability density distribution $\rho(x,y)=1/R^2$, the Expectation of the square of the cluster head to sink node is:

$$\begin{aligned} E\left[d_{toBS}^2\right] &= \tfrac{1}{R^2}\int\int_{R\times R}\left[(x-\lambda R)^2+(y-\psi R)^2 dxdy\right] \\ &= \left(\tfrac{2}{3}-\lambda-\psi+\lambda^2+\psi^2\right)R^2 \end{aligned} \tag{22}$$

*5.2. The process of cluster head election*

This section proposes an energy heterogeneous clustering routing algorithm (CLWPA) based on LWCA and IDEEC. Assuming that the number of heterogeneous nodes in the sensor network is $\varphi$, and the number of cluster heads of common nodes is calculated according to Fomula (20), Then the process chart is as follows (Fig. 2).

Step1: According to the heterogeneous node optimal deployment algorithm in section, $\varphi$ heterogeneous nodes are optimally deployed;

Step2: According to IDEEC algorithm, $n_{CH}$ cluster heads are selected from common nodes and clustered.;

Step3: The member node $\gamma$ in the cluster transmits the collected data to the cluster head of the cluster in a given time slot (TDMA), and the cluster head processes the data;

Step4: The cluster head calculates the distance from itself to the heterogeneous node: if the distance from $v_{ch}$ to the nearest heterogeneous node is within one hop, the data is directly sent to the heterogeneous node; Otherwise, $v_{ch}$ is sent to the nearest heterogeneous node through multiple hops of cluster head layer nodes;

Step5: When the energy of the cluster head node is exhausted, the cluster head election and clustering are carried out again according to IDEEC algorithm, and the transmission modes of the nodes in the cluster and the cluster head are unchanged.

The data transmission mode in CLWPA cluster heterogeneous network is (Fig. 3):

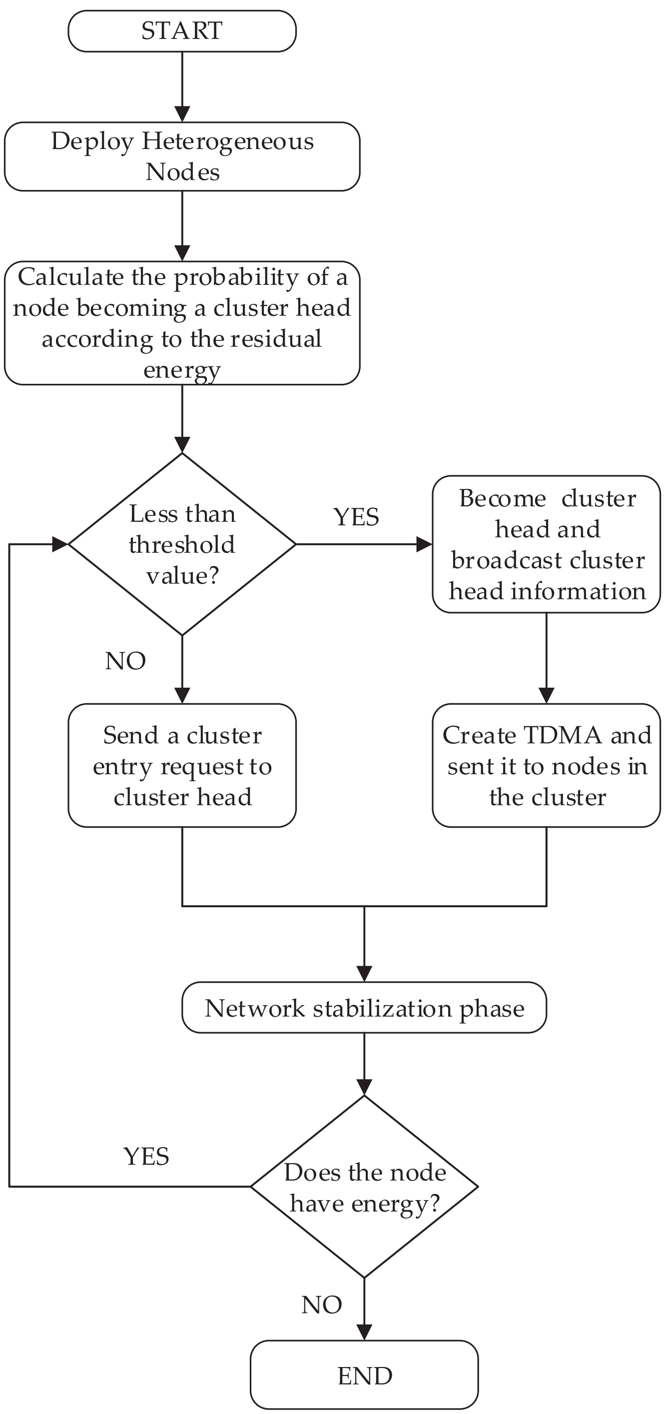


**Fig. 2.** The process of cluster head election.

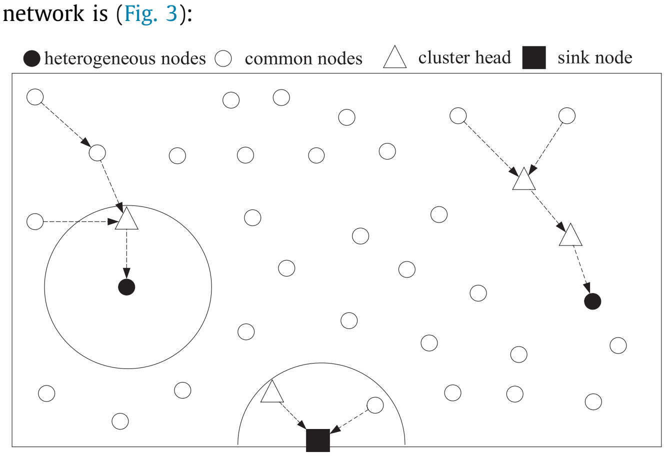


**Fig. 3.** Data transmission mode.

**Table 1**
Transmission parameters value.

| Parameters | Value |
|---|---|
| Simulation area /$m^2$ | $200 \times 200$ |
| The initial node energy /J | 0.5 |
| Number of nodes | 200 |
| $E_{elec}$/nJ•bit$^{-1}$ | 50 |
| $E_{fs}$/pJ•bit$^{-1}$•m$^{-2}$ | 10 |
| $E_{mp}$/pJ•bit$^{-1}$•m$^{-4}$ | 0.0013 |
| $E_{DA}$/nJ•bit$^{-1}$• | 5 |
| broadcast packet /bit | 200 |
| Data packet /bit | 4000 |
| $d_0$/m | 87.7 |
| $p_{opt}$ | 0.1 |

## 6. Simulation and analysis

In order to verify the validity of CLWPA, Matlab2018-b is used to simulate and test the algorithm. Assuming that the communication channel is ideal and the influence of random factors such as signal collision is ignored, the network model and energy consumption model of the monitoring area are shown in sections 2. Nodes are randomly distributed in the simulation area by manpower, and sink nodes are located outside the monitoring area. The experimental parameters in the simulation are shown in Table 1.

Assuming that there are 4 energy heterogeneous nodes in the monitoring network, EDDEEC and KCA respectively deploy the locations of the heterogeneous nodes according to their algorithm ideas. In the CLWPA algorithm, heterogeneous nodes are deployed according to LWPA optimization algorithm, and then the optimal number of cluster heads is calculated according to Eq. (15) and common nodes are clustered. Since the sensor nodes are randomly distributed in the monitoring area, 50 experiments are conducted and the average value is taken as the experimental result.

### *6.1. Networks stable period analysis and lifetime*

In the simulation experiment, the number of rounds experienced by the sensor network from the start of operation to the death of the first node is called the stable period of the network. Fig. 4 shows the stable periodic variation curves of four heterogeneous network routing algorithms in scenarios with different network area sizes. The experimental results show that the stable period of CLWPA algorithm is the longest, followed by HRMIP and KCA algorithms, and EDDEEC algorithm is the shortest. The stable period of CLWPA is increased by about 33% on average compared with HRMIP, while the stability period of HRMIP is increased by about 22% compared with KCA. The EDDEEC directly arranges heterogeneous nodes in the cluster head, increasing the number of nodes forwarding data, thus resulting in too fast energy consumption and short stability period.

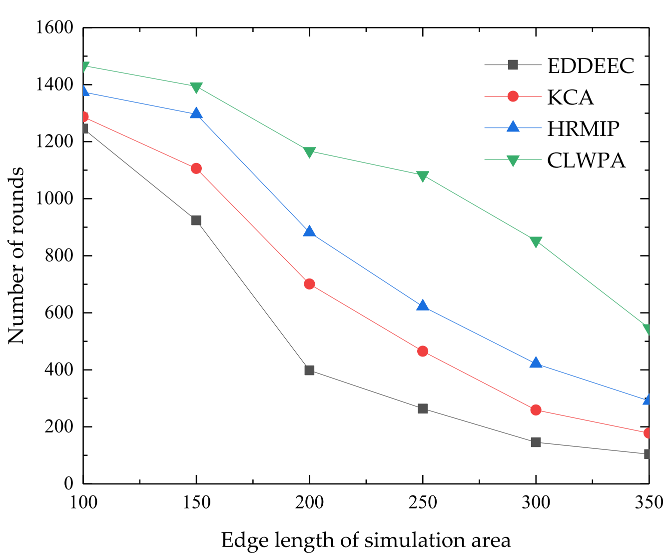


**Fig. 4.** Comparison of stable period.

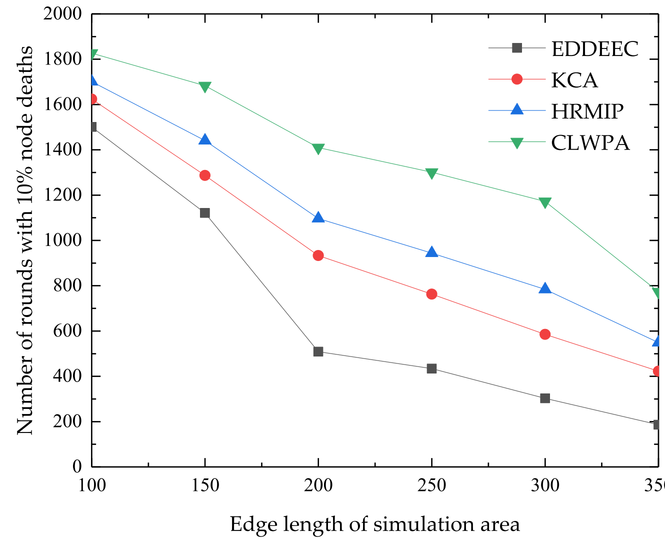


**Fig. 5.** Comparison of lifetime.

Fig. 5 shows the lifetime variation curves of four heterogeneous network routing algorithms under different network area sizes. In the simulation, when 10% of the nodes in the sensor network died, the sensor network was deemed to be invalid. The simulation results show that the lifetime of sensor network is similar to the stable period. CLWPA adopts dynamic clustering method for data transmission on the basis of optimal deployment of heterogeneous nodes, so the energy consumption of common nodes is less and more balanced, so the lifetime of sensor networks is significantly increased. Compared with the HRMIP, the lifetime is increased by 25.3% on average, while the HRMIP algorithm optimizes the deployment of heterogeneous nodes, so the lifetime of the HRMIP is increased by 16% on average compared with the KCA, while EDDEEC does not optimize the deployment of heterogeneous nodes, so the lifetime is the shortest.

### *6.2. Number of surviving nodes*

Fig. 6 shows the relationship between the number of surviving nodes and the number of rounds under the same network area size. In the simulation, a node is considered dead when it consumes 99% of the initial energy. The simulation results show that the first node death of CLWPA occurs in 1167 rounds, the first node death of HRMIP occurs in 882 rounds and KCA and EDDEEC occur in 701 rounds and 398 rounds respectively. As can be seen from Fig. 6, CLWPA ensures the survival of all nodes in a longer time. As the energy consumption of the network is relatively balanced and the residual energy of the nodes is very small, the death rate of the nodes will suddenly accelerate in the later stage of the simulation.

### *6.3. Energy consumption*

Fig. 7 shows the relationship between the standard deviation of the node's remaining energy and the number of rounds when the algorithm is running under the same simulation environment.

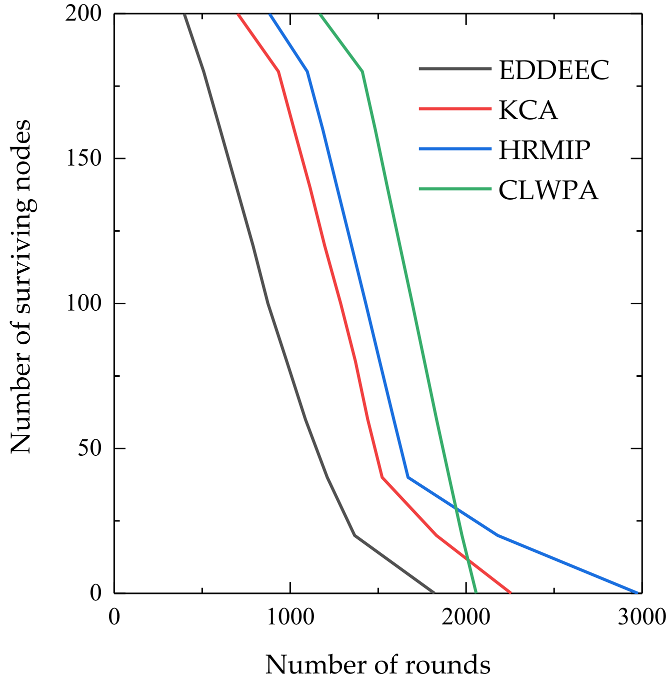


**Fig 6.** Comparison of surviving nodes.

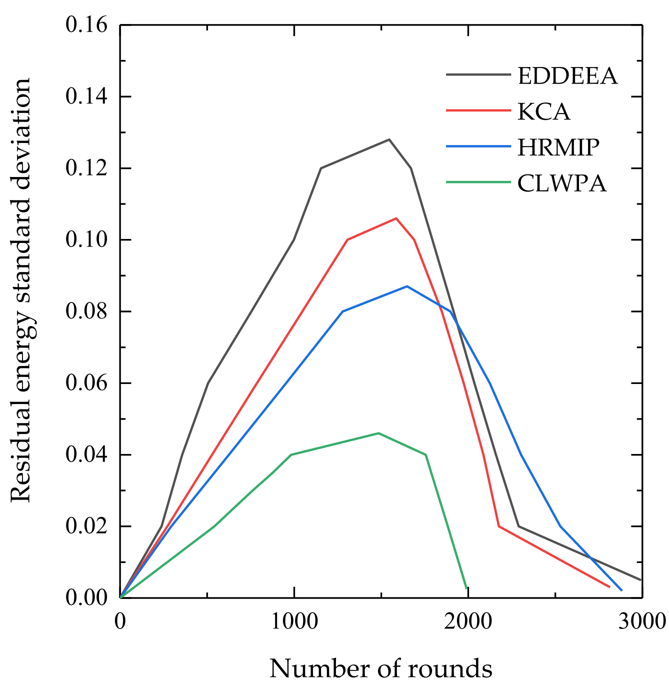


**Fig. 7.** Comparison of residual energy standard deviation.

It can be seen from the figure that in any environment, the energy consumption difference of each node in EDDEEC is the largest, because heterogeneous nodes are not optimized and nodes near sink nodes repeatedly become cluster heads, resulting in unbalanced residual energy of nodes. For KCA and HRMIP, the optimal deployment of heterogeneous nodes reduces the average path length of nodes, but the fixed data transmission structure of the network leads to unbalanced energy consumption among nodes. In CLWPA algorithm, distant nodes forward data through the cluster head,

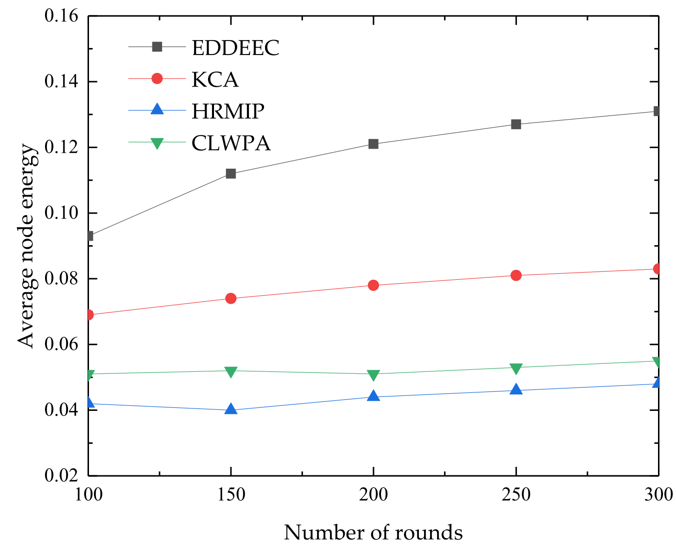


**Fig. 8.** Comparison of energy consumption.

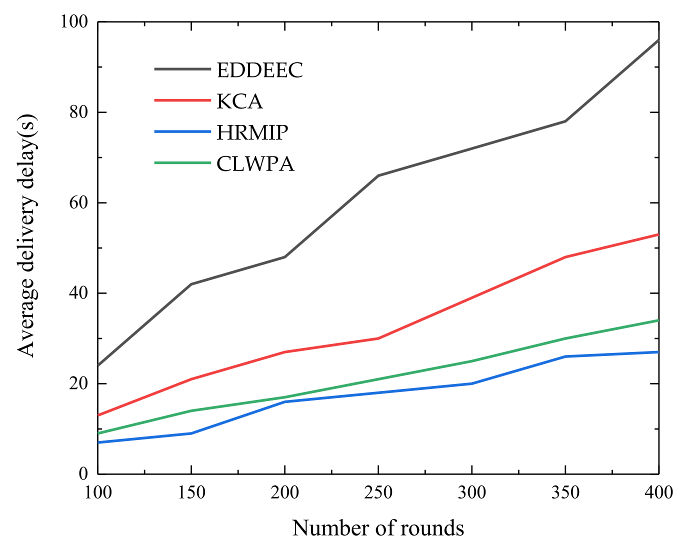


**Fig. 9.** Comparison of data transmission delay.

and the cluster head node is in a state of dynamic update, so the energy consumption is more uniform.

Fig. 8 shows the energy consumption of the four algorithms under different node numbers. As the simulation area becomes denser, more data transmission occurs in the network, which makes the average energy consumption of EDDEEC algorithm and KCA algorithm larger than that of the other two algorithms. Because the CLWPA needs to complete dynamic clustering to balance the energy consumption of nodes, and may not choose the shortest path for data transmission, the average energy consumption of the CLWPA is slightly higher than that of the HRMIP. Simulation results show that the average energy consumption of HRMIP is 8.3% lower than that of CLWPA, and the average energy consumption of CLWPA algorithm is about 31.9% lower than that of KCA.

### 6.4. Data transmission delay

Data transmission delay is one of the important standards to measure the performance of routing algorithms. The smaller the average delay, the better the stability of data transmission. Fig. 9 shows the relationship between the average transmission delay of the four algorithms and the number of rounds. As can be seen from the figure, when HRMIP and CLWPA transmit data through the shortest path, the data transmission delay is smaller than KCA and EDDEC algorithms. However, the latency of CLWPA is slightly larger than that of HRMIP because CLWPA needs to elect the cluster head and uses dynamic clustering to transmit data. The simulation results show that the average delay of HRMIP is 18% lower than that of CLWPA, and the average delay of CLWPA is 35.1% lower than that of KCA.

## 7. Conclusions

In order to effectively prolong the stable period and lifetime of the network, a clustering heterogeneous network routing algorithm CLWPA is proposed. Firstly, the improved WPA is used to optimize the deployment of heterogeneous nodes. Secondly, DEEC and heterogeneous network routing algorithm are combined to form CLWPA. Finally, the performance of the algorithm is compared with three typical routing algorithms through simulation experiments. The simulation results show that the CLWPA makes the energy consumption of nodes more uniform, ensures that all nodes remain alive for a longer period of time, effectively avoids the phenomenon of premature death of cluster heads, and makes the death time of nodes more concentrated. therefore, the CLWPA effectively prolongs the stability cycle and life cycle of the network, and increases the proportion of the stability cycle in the life cycle. CLWPA, with its excellent performance, can be widely used in a series of complex and harsh monitoring environments such as radioactive monitoring of uranium tailings ponds, ground pressure disaster monitoring of underground mines, environment monitoring of smart grid and equipment monitoring. In the future work, the optimal deployment of multi-dimensional wireless sensor nodes and the optimal deployment of mobile sink nodes are worth studying.

## Declaration of Competing Interest

None.

## Acknowledgments

This work was in part supported by the National Natural Science Foundation of China (No.11875164); Ministry of Emergency Management of the People's Republic of China Key Technologies for Safety Production and Major Accidents Prevention and Control (hunan-0001-2018AQ); Key Research and Development Projects of Hunan Province(2018SK2055).

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.comnet.2019.106994.

## References

[1] D. Zhang, L. Quan, C. Lin, et al., Multi-layer based multi-path routing algorithm for maximizing spectrum availability, Wirel. Netw. (2018) 1–13.
[2] H.D. Al-Ariki, M.N. Swamy, A survey and analysis of multipath routing protocols in wireless multimedia sensor networks, Wirel. Netw. 23 (6) (2017) 1823–1835.
[3] S. Jothimuneeswari, S. Ganapathy, A Kannan, Intelligent data gathering and energy efficient routing algorithm for mobile wireless sensor networks, Asian J. Inf. Technol. 15 (2016) 921–927.
[4] M. Selvi, P. Velvizhy, S. Ganapathy, et al., A rule based delay constrained energy efficient routing technique for wireless sensor networks, Cluster Comput. (2017).
[5] Y. Bahuguna, D. Punetha, P. Verma, An analytic study of the key factors influencing the design and routing techniques of a wireless sensor network, Int. J. Interact. Multimedia. Artif. Intell. 4 (2017) 11–15.
[6] M. Nighot, A. Ghatol, V. Thakare, Self-Organized hybrid wireless sensor network for finding randomly moving target in unknown environment, Int. J. Interact. Multimed. Artif. Intell. 5 (1) (2017) 16–28.
[7] M. Selvi, K. Thangaramya, S. Ganapathy, et al., An energy aware trust based secure routing algorithm for effective communication in wireless sensor networks, Wirel. Pers. Commun. 105 (4) (2019) 1475–1490.
[8] A.S. Rostami, M. Badkoobe, F. Mohanna, et al., Survey on clustering in heterogeneous and homogeneous wireless sensor networks, J. Supercomput. (13) (2017) 1–47.
[9] P. Kong, G. Fang, C. He, et al., Topology optimization of port wireless sensor network based on small-world network, 2017 International Conference on Circuits, System and Simulation (ICCSS), IEEE, 2017.
[10] J. An, L. Qi, X. Gui, et al., Joint design of hierarchical topology control and routing design for heterogeneous wireless sensor networks, Comput. Stand. Interfaces 51 (2017) 63–70.
[11] Z. Wadud, N. Javaid, M.A. Khan, et al., Lifetime maximization via hole alleviation in iot enabling heterogeneous wireless sensor networks, Sensors 17 (7) (2017) 1677.
[12] H. Liu, R. Sun, Q. Liu, The tactics of ship collision avoidance based on quantum-behaved wolf pack algorithm, Concurrency Comput (1) (2019).
[13] W. Liang, J. He, S. Wang, et al., Improved cluster collaboration algorithm based on wolf pack behavior, Cluster Comput. (1) (2018) 1–16.
[14] R. Logambigai, S. Ganapathy, A. Kannan, Energy–efficient grid–based routing algorithm using intelligent fuzzy rules for wireless sensor networks, Comput. Electr. Eng. 68 (2018) 62–75.
[15] M. Selvi, S.J. Muneeswari, S. Ganapathy, et al., Virtual force-based intelligent clustering for energy-efficient routing in mobile wireless sensor networks, Turkish J. Electr. Eng. Comput. Sci. 26 (3) (2018) 1444–1452.
[16] M. Selvi, R. Logambigai, S. Ganapathy, H. Khanna Nehemiah, K. Arputharaj, An intelligent agent and fso based efficient routing algorithm for wireless sensor network, in: 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM), IEEE, 2017, pp. 100–105.
[17] K. Thangaramya, K. Kulothungan, R. Logambigai, et al., Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT, Comput. Netw. 151 (2019) 211–223.
[18] J. Anzola, J. Pascual, G. Tarazona, R. González Crespo, Clustering WSN routing protocol based on kd tree algorithm, Sensors 18 (9) (2019) 2899.
[19] V. Geetha, P.V. Kallapur, S. Tellajeera, Clustering in wireless sensor networks: performance comparison of LEACH & Leach-C protocols using NS2, Procedia Technol. 4 (2012) 163–170.
[20] A. Somani, P.P Bhattacharya, Analyzing the network lifetime of heterogeneous LEACH and TEEN in three-dimensional wireless sensor networks, IEEE International Conference on Power Electronics, 2017.
[21] P.G. Naranjo, M. Shojafar, H. Mostafaei, et al., P-SEP: a prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks, J. Supercomput. 73 (2) (2017) 1–23.
[22] T. Tiwari, N.R. Roy, Modified DEEC: a varying power level-based clustering technique for WSNs, in: International Conference on Computer & Computational Sciences, IEEE, 2015, pp. 170–176.
[23] Z. Wadud, N. Javaid, M.A. Khan, et al., Lifetime maximization via hole alleviation in iot enabling heterogeneous wireless sensor networks, Sensors 17 (7) (2017) 1677.
[24] M. Shaji, S. Ajith, Distributed energy efficient heterogeneous clustering in wireless sensor network, Fifth International Conference on Advances in Computing & Communications, IEEE, 2016.
[25] N. Chaudhry, F. Saleem, Comparative analysis of clustering algorithms comprising GESC, UDCA and k-mean method for wireless sensor networks, Ursi Radio Sci. Bull. 84 (4) (2017) 12–18.
[26] C. Li, J. Bai, J. Gu, et al., Clustering routing based on mixed integer programming for heterogeneous wireless sensor networks, Ad Hoc Netw. 72 (2018) 81–90.
[27] L. Zhang, L. Zhang, S. Liu, et al., Three-Dimensional underwater path planning based on modified wolf pack algorithm, IEEE Access 5 (2017) 22783–22795.
[28] W.U. Hu-Sheng, F.M. Zhang, L.I. Hao, et al., Discrete wolf pack algorithm for traveling salesman problem, Control Decis. (2015) 0-0.
[29] Y.B. Chen, Y.S. Mei, J.Q. Yu, et al., Three-dimensional unmanned aerial vehicle path planning using modified wolf pack search algorithm, Neurocomputing (2017) S0925231217309220.
[30] W.U. Hu-Sheng, F.M. Zhang, R.J. Zhan, et al., Improved binary wolf pack algorithm for solving multidimensional knapsack problem, Syst. Eng. Electron. 37 (5) (2015) 1084–1091.
[31] D. Wang, X. Qian, K. Liu, et al., An adaptive distributed size wolf pack optimization algorithm using strategy of jumping for raid (September 2018), IEEE Access 6 (2018) 65260–65274.
[32] M.A. Abd-Elmagid, T. Elbatt, K.G. Seddik, Optimization of energy-constrained wireless powered communication networks with heterogeneous nodes, Wirel. Netw. (9) (2018) 1–18.

**Yu Xiuwu** 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 an Associate Professor in Department of Environment and Safety Engineering at University of South China, Hengyang China. His research interest includes data processing in wireless sensor networks and Big Data Analysis.

**Yu Hao** is a Ph.D. Student in safety engineering at the University of South China, Hengyang China. His research interest is wireless sensor networks.

**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.

**Xiao Renrong** is a Ph.D. Student in safety engineering at the University of South China, Hengyang China. His research interest is wireless sensor networks.