An energy-efficient non-uniform clustering routing protocol based on improved shuffled frog leaping algorithm for wireless sensor networks

Qi Huamei1 | Lin Chubin1 | Gao Yijiahe1 | Xiong Wangping2 | Jiao Ying1

1 School of Information Science and Engineering, Central South University, Changsha, People’s Republic of China
2 School of Computer, Jiangxi University of Traditional Chinese Medicine, Nanchang, People’s Republic of China

Correspondence
Qi Huamei, School of Information Science and Engineering, Central South University, Changsha, People’s Republic of China.
Email: sans_hua@163.com

Abstract
Wireless sensor networks have been used widely in environmental tracking and monitoring. Nevertheless, WSNs are facing a number of challenges, such as unreasonable cluster head selection and energy-hole problems while nodes have unbalanced energy consumption. Thus, the paper aims to propose an energy-efficient non-uniform clustering routing protocol to enhance nodes energy efficiency and balance the energy consumption in WSNs. In addition, a non-uniform clustering network partition is introduce to reduce the probability of energy hole occurrences and optimize cluster heads dynamical selection method to suggest an improved shuffled frog leaping algorithm. Ultimately, the simulation experiment has demonstrated a 20% gain of energy efficiency to extend network lifetime by the proposed protocol and enhanced algorithm.

1 INTRODUCTION

WSNs stands for wireless sensor networks, which include many low-cost and low power-sensing tools [1]. A large number of sensor nodes self-organize to form a large-scale network, and sensor nodes can monitor data in the physical environment. Due to the weak communication capability and energy of sensor nodes, the energy problem seriously limits the application and development of WSNs. Because of this, in the study of wireless multi-sensor networks, how to extend the survival time of the network, balance the energy consumption of the network, reduce or eliminate the energy hole, and control the hotspot problem of the network has drawn researchers’ attention.

Many scholars have done a lot of research on energy-efficient routing algorithms [2]. The cluster-based hierarchical routing algorithm embodies better energy conservation and adaptability, compared with the flat routing algorithm [3]. The clustering routing protocol (CRP) [4] is divided into two categories: uniform clustering and non-uniform clustering. First, the monitoring area is divided into multiple regions. Each region is a cluster. Every cluster consists of a head node and member nodes, and the head node of a low-level cluster can be treated as a member node of a high-level cluster. In this way, the network is hierarchically divided, and the head of the top layer cluster takes responsibility to communicate with the base station or the sink node. For large-scale networks, when same-sized clusters are formed, the head node of the closest cluster to base station often acts as a relay node and bears a much larger communication load than other nodes. Therefore, the relay node in the cluster drains out energy rapidly, which causes the energy hole. To improve this problem, non-uniform clustering routing has been proposed. For a cluster in the WSNs, the closer it is to the base station, the smaller the cluster size is, which can reduce...
the internal communication and the energy consumption in the nearest cluster to the base station.

In the design of non-uniform CRP, cluster head selection plays a key role. How to form and divide high-quality clusters in WSNs, and to select the best head node in a cluster for collected data movement, is the challenge. To reduce and balance node energy consumption and minimize the occurrences of energy holes, the shuffled frog leaping algorithm (SFLA) [5–6] is proposed, which is a sub-heuristic algorithm that simulates information exchange during frog foraging. It is a combination of memetic algorithm and particle swarm optimization (PSO) to simulate frogs. Above is a collaborative search method for group foraging, which has the advantages of simple structure, few parameters, and easy implementation, and it is suitable for designing energy-saving routing protocols in WSNs.

However, there are some drawbacks of those extant frog leaping algorithms, which are the complex and the weak search breadth and depth [7–9], so that it can potentially drain the cluster head node’s energy when the node do tasks of data transmitting and data merging.

Therefore, this paper aims to tackle the network "hot zone" problem in non-uniform clustering protocols, using the improved hybrid frog leaping algorithm in the cluster to find the best cluster head, and propose an energy-efficient non-uniform clustering routing (E2NUCR) protocol based on improved shuffled frog leaping algorithm (ISFLA). It is used to equalize the energy of network nodes, reduce energy holes, and delay network lifetime.

The rest of the paper is organized as follows: Section II conducts a literature review. Section III elaborates the details of E2NUCR. The experiment and simulation are discussed in Section IV. Section V concludes this contribution.

## RELATED WORK

Recently, quite a few solutions on clustering protocol and SFLA have been proposed by researchers. In this section, typical clustering protocols and SFLA are discussed.

### 2.1 Clustering routing protocol

Gao et al. [10] proposed a uniform clustering method considering average residual energy and distance from the device to base station. It can acquire optimal cluster numbers, minimize the consumption for each round, and set a scheme to elect the energy-saving cluster head.

Darabkh et al. [11] presented an improved threshold based cluster head replacement protocol considering the major shortcomings of the T-low-energy adaptive clustering hierarchy (LEACH) protocol. They also proposed a EA-DB-CRP [12] for WSNs based on energy perception and density. The basic goal of the protocol is to balance the energy consumption in the network by distributing the load to the available sensor nodes.

Bhushan et al. [13] designed a hybrid method that combines genetic algorithm (GA) with k-means algorithm in the area of EEUC.

A novel idea had been put forward by Sadek [14], which may overcome the differences between the physical WSN environment and the real heterogeneous network IoT environment. They provided an efficient hybrid energy-aware cluster communication protocol for green IoT network computing.

Toor and Jain [15] proposed a novel mobile sensor node conception-multi-hop (MEACBM) routing protocol based on a layered heterogeneous WSN.

Wang et al. [16] presented an improved energy-saving wireless non-uniform CRP. Compared with the conventional LEACH algorithm and energy-efficient unequal clustering (EEUC) algorithm, the simulation results show that the improved routing protocol is more stable to reduce the whole energy consumption with more balanced communication loads and the long networks’ lifetime.

Gupta and Shekokar [17] invented a novel k-means L-layer algorithm for non-uniform clustering in the field of WSN, it can largely prolong the network lifespan via clustering in WSNs.

Yang et al. [18] demonstrated a low-energy non-uniform clustering topology control algorithm. The algorithm adopts a three-step mechanism during the election of cluster head. Comparing with the traditional algorithm, it can effectively optimize the topology structure of clustering and balance the energy consumption of nodes. As a result, it is able to prolong the lifespan of the network.

Bozorgi and Bidgoli [19] came up with a hybrid unequal-energy efficient clustering algorithm. A new mechanism called clustering strategy was used to improve the previous approach and extend the life of the network.

Wang et al. [20] proposed an algorithm based on double cluster head clustering and a data fusion mechanism based on information entropy. In the cluster head selection algorithm, two cluster heads perform different duties. This can better share the energy consumption and extend the network lifespan.

Gupta and Jha [21] proposed an improved cuckoo search-based energy balanced node clustering protocol which uses a novel objective function for uniform distribution of cluster heads.

Gajjar et al. [22] proposed the MAC (Multiple Access Channel), routing, and unequal clustering cross-layer protocol for WSNs-based on fuzzy and ant colony optimization. To avoid hot issues, FAMACROW used an unequal clustering mechanism.

Baradaran and Navi [23] proposed a method (HQCA) for generating high-quality clusters, which uses a criterion for measuring the cluster quality and can reduce the error rate during clustering.

Sirdeshpande and Udupi [24] proposed an algorithm based on energy efficiency optimization for creating the energy-efficient routing path.
2.2 Shuffled frog leaping algorithm

A direction-changeable shuffled-frog-leaping algorithm is suggested by Kong et al. [25]. In this method, we will set a direction flag for the frog before it moves. As long as it jumps well in a certain direction, there is a larger probability that it may move further on this well. This method could effectively solve the test function. Durgadevi and Srinivasan [26] proposed an optimal hybrid algorithm combined SFLA with cuckoo search. In the algorithm, SFLA acquires a few outstanding advantages including converging fast, implementing easily, and acquiring the capability of global optimization. Li et al. [27] put forward a hybrid optimal method based on SFLA and bacterial foraging, which has solved the problem of the reliability of the system and redundancy assignment. Edla et al. [28] made good use of SFLA to treat the problem of load balance in the WSN gateway. To improve SFLA, they modified the period of frog population generation and offspring generation and introduced shift phases in SFLA. Fanian and Rafsanjani [29] made good use of grouping and choose appropriate cluster heads to handle fast sensor energy release issues. As a result, they put forward a fuzzy hybrid leapfrog algorithm. It utilizes memetic SFLA to optimize the standard base table relying on Mamdani fuzz rule. Jingjing and Dai [30] proposed an adaptive hybrid mutation SFLA. However, a few disadvantages, including the high randomness in previous partial search, weakness in global search, and the slow convergence, still exist in the method. In the later round, individual form monotony may be a great problem and the algorithm may fall into the trouble of local optimization.

3 DESIGN OF E2NUCR PROTOCOL

3.1 Energy-efficient non-uniform clustering mechanism

The traditional SFLA has relatively large randomness when performing population initialization, and it fails to satisfactorily solve the problems of the energy balance of nodes in the network and early death of cluster heads which are around sink node. In response to the above problems, we propose non-uniform clustering based on prior knowledge.

The nodes in the network are mainly composed of a sink node and a sensor node. The sensor nodes form the cluster layer in the network initialization. As shown in Figure 1, when clustering, the clusters, closer to the sink node, are smaller, and the number of sensor nodes is smaller. The clusters, farther away from the sink node, are larger, and the number of sensor nodes is also larger. In this way, the energy consumption of the cluster head node closer to the sink is greatly reduced, and the energy consumption in the network is more uniform.

First, statistically calculate the prior probability of the data distribution at the previous moment in the network to form a prior knowledge $P_i$. The $P_i$ value of a node is mainly determined by the degree of similarity between the data packet and the packets generated by the surrounding nodes and the distance of the node from the sink node

$$P_i = t \cdot DS_{ij} + (1 - t) / (DS_{ij} + 1).$$

In formula (1), $DS_{ij}$ represents the similarity of the data between the node $i$ and the node within one hop range. $Di_j$ refers to the distance between node $i$ and the sink node.

Then, according to the $P_i$ value of the node, the nodes with similar $P_i$ values are divided into the same cluster, which means the nodes in the network are clustered considering both the location and the content similarity. Weight $\alpha$ is introduced in adjusting the proportion of $DS_{ij}$ and $Di_j$ in $P_i$. The greater the weight is, the greater the similarity of the packet between node $i$ and nodes within the hop range is. In classical algorithms such as LEACH, the size of each cluster is approximately the same when clustering, and the cluster head node uses multiple hops to transmit data. This is prone to hot zone problems, which means cluster head nodes, close to the sink node, consume a lot of energy due to data fusion and data transmission. To solve this problem, clustering adopts a non-uniform algorithm. The clusters near the sink node have fewer cluster members, and those away from the sink node have relatively more cluster members, so that the cluster head closer to the sink node can save energy for data transmission and achieve the purpose of balancing the energy consumption of the cluster head.

3.2 Improved shuffled frog leaping algorithm

Based on the traditional frog leaping algorithm, an ISFLA was proposed to improve the shortcomings in the search mechanism and the breadth and depth of the algorithm. The algorithm mainly includes the following subsections:

- Fitness function design;
- Update local search strategy;
3.2.1 Fitness function design

After initializing the population with non-uniform clustering proposed above based on prior knowledge, the initial clustering ends and the node network is divided into several cluster regions. The next step is to select the cluster head node in every region. PSO is needed in the search process of this paper. PSO abstracts each individual of the population into a particle with negligible mass and volume and velocity and position, and each particle in the search area may be the optimal solution, and iteration is used to find the optimal one (i.e. the optimal candidate node selected as the cluster head). Therefore, it is necessary to design a fitness function to measure the quality of the solution, through which the optimal solution can be selected from multiple particles. The node with the highest fitness function value is the optimal candidate node selected as the cluster head. The situation is indicated in the following contents. $M$ nodes are randomly distributed in the network area. $K$ clusters are divided according to the non-uniform clustering based on prior knowledge and the sink node picks out $N$ nodes as the candidate cluster head nodes whose residual energy is greater than the threshold value. The cluster head node will be selected from candidates and there will be $C_N^K$ order to judge the quality of this group of candidate cluster head nodes, the proposed fitness function can be used to calculate and sort the candidate cluster head nodes. The candidate node with the best fitness will be selected as the cluster head node of this round.

After comprehensive consideration, the fitness function designed in this paper mainly examines the following elements.

Residual energy factors
Considering the large energy cost of the cluster head node, candidate nodes with more residual energy should be selected as priority ones when selecting the cluster head node. According to the relationship between the residual energy of nodes and the energy of all nodes in the network, the residual energy factor is introduced, as shown in formula (2), where $E(p_i)$ means node residual energy, $E(c_k)$ means selected cluster head residual energy

$$
 f_1 = \sqrt{\frac{\sum_{i=1}^{M} E^2(p_i)}{\sum_{k=1}^{K} E^2(c_k)}}.
$$

Location factor intra-cluster
Considering that the cluster-head node needs to have good transmission efficiency with the nodes in the cluster, the candidate nodes with better location should have priority to be chosen as the cluster-head. According to the location of the candidate node, by integrating its position in the whole cluster and its distance from other nodes, the position factor in the cluster is introduced into the fitness function, as shown in formula (3),

$$
 f_2 = \max_{i=1,2,\ldots,M} \left( \sum_{y \in c_k} d_{p_i,c_k} / N_i \right). (3)
$$

Location factors between clusters
Considering that the cluster head node needs to transmit information with the sink node frequently, distance between them is one of the important influencing factors as well. According to the distance between sink node and cluster head nodes, the inter-cluster location factor is introduced into the fitness function $f_3$, where $d_{p_i,Sink}$ means the distance between all nodes and the sink node, $d_{c_k,Sink}$ represents the distance between selected cluster head and sink node

$$
 f_3 = \max_{i=1,2,\ldots,M} d_{p_i,Sink}/d_{c_k,Sink}. (4)
$$

In conclusion, the formula of the fitness function is shown in formula (5), where $\alpha, \beta, $ and $\gamma$ represent the weight coefficient of residual energy factor, location factor within clusters, and location factor between clusters in the fitness function, respectively:

$$
 f(p) = \alpha f_1 + \beta f_2 + \gamma f_3. (5)
$$

3.2.2 Update local search strategy

Improvement of the update strategy of the worst individuals in the subgroup
Due to the problem that the traditional SFLA only updates the worst individuals in groups, when the algorithm goes to the late process, if some high fitness individuals remain unchanged, local search will be a linear search process all the time, which will cause the entire population to search only in the direction of high fitness individuals. The generated new individuals always go along a straight line to the global optimal individuals or to the ones in subgroups, which abates the search breadth and gives rise to decreased species diversity in later evolution; at the same time, with the continuous progress of evolution, because the worst individual in the memes set itself carries excellent information, the algorithm is prone to fall into local optimization, which affects the solution accuracy of the algorithm and reduces the search depth.

Considering that the local optimal individuals and the global optimal individuals play a role in guiding the update, changing them can not only get a better global solution but also guide the update for other nodes. Therefore, the purpose of this paper is to deep search the network space near the superior individuals to obtain better individuals. The updates of each individual in the subgroup affect the next part of the evolution, and the average fitness of all individuals in the subgroup reflects the overall level of the subgroup to a certain extent, while the
optimal individuals of the subgroup with the highest fitness in the subgroup affect the evolution of the rest. Therefore, taking the leading role of the optimal individuals in the subgroup, the moving step size of the worse individual was adjusted by the best and the worst in the group and the average of all individuals in a group and the random difference of the worst individual, to improve the breadth and the depth of local search. Based on this, the updating strategy of the worst individual in the subgroup is modified in formula (6), where \( \text{rand} \) is random number between 0 and 1, \( P_{\text{avg}} \) is the average of all the individuals in the subgroup, \( P_w \) is the worst individual in the subgroup, \( P_b \) is the best individual in the subgroup, and \( \mu \) is the proposed variation factor, whose value changes with the number of iterations of the algorithm, and specific definitions will be given below. The purpose of defining \( \text{rand} \) is to increase the randomness of the search and improve the probability of obtaining the optimal solution:

\[
P_w' = P_b + \mu \cdot (\text{rand} \cdot (P_b - P_w) + (1 - \text{rand}) \cdot (P_{\text{avg}} - P_w)).
\]

(6)

If the fitness of the newly obtained individual \( P_w' \) is better than the worst individual \( P_w \) in the subgroup, \( P_w' \) is used to update \( P_w \):

\[
P_w' = P_b + \mu \cdot (\text{rand} \cdot (P_b - P_w) + (1 - \text{rand}) \cdot (P_{\text{avg}} - P_w)).
\]

(7)

Then updated judgment is made again. If the fitness of the new individual \( P_w' \) obtained is better than that of the worst individual \( P_w \) in the subgroup, \( P_w' \) is used to update \( P_w \). Otherwise, randomly generate a new individual to replace \( P_w \). The improvement of the worst individual update strategy in the subgroup not only enables the worst individual to learn from the best individual but also takes the influence of the average fitness of all individuals in the subgroup into account. Compared with the traditional worst individual update strategy, it has a higher search breadth and reduces the possibility of the algorithm falling into local optimization.

**Improvement of subgroup optimal individual update strategy**

Combined with the strategy of elite learning, this paper makes a detailed search of the superior individuals in the population into a single elite group in a small field to realize self-update. The adaptive evolution of the better solution found in the search process is used to make the superior individual more approximate to the optimal individual, to improve the direction of the next evolution and guide the evolution of the population toward the global optimal direction, which avoids falling into the local optimal and improves the performance of global optimization.

Algorithm based on the mechanism of elite learning is to select elite individuals in the population. After finishing the local search for each round, a global search first starts, including updating all the fitness of individuals in real-time and selecting \( M \) individuals with the highest fitness to form an elite group. At the same time, individuals in the elite group search the nearby small areas of space in the search. If a more optimal solution is discovered, the algorithm will update individuals and their fitness, as shown in formula (8), where \( P_e \) is the selected elite individual, \( P_e' \) is the updated elite individual, and \( \mu \) is the proposed variation factor, whose value changes with the number of iterations of the algorithm, and specific definitions will be given as:

\[
P_e' = (1 + \mu \cdot \text{rand}) \cdot P_e.
\]

(8)

If the new elite individual \( P_e' \) is better than the original elite \( P_e \), then \( P_e' \) is used to update \( P_e \). This updating strategy can not only ensure that the original elite individuals of the population will cause individual degradation due to random movement and improve the fitness of the better individuals in the population based on the evolution of the original elite individuals but also maintains a good learning relationship between the poor individuals and the better individuals. In addition, the effective search efficiency can guide the evolution of the population toward the global optimal direction, and to a certain extent, accelerate the convergence speed of the algorithm.

### 3.2.3 Variation factor

At the early stages of the traditional SFLA, the randomness of local search is large, the global search ability is poor, and the convergence speed is slow. However, later in the algorithm, the individual form will appear drab, falling into the predicament of the local optimum. Therefore, when updating the optimal individuals in the subgroup, it is required to increase the randomness of new individuals in the early stage so that the population can have better search ability and greater search breadth and speed up the convergence. In the later stage of the algorithm, new individuals can be positioned more accurately to make the population have deeper search depth and increase the diversity of the population. Because Gaussian distribution function is relatively narrow wing, two wings wide Cauchy distribution, so formula (6)–(8) propose a variable factor \( \mu \). Its value varies with the number of iterations of the algorithm.

In the early stage of the algorithm, namely, the number of iterations is lower than the threshold, with the method of Gaussian mutation to define the variation factor, that is \( \mu \sim \mathcal{N}(0, \delta^2) \), it is a Gaussian distribution with the mathematical expectation of 0 and standard deviation of \( \delta^2 \), and the value is determined by the example.

In the later stage of the algorithm, that is, when the number of iterations is higher than the threshold value, the variation factor is defined by Cauchy variation, that is \( \mu \sim C(\eta, 0) \), where \( C(\eta, 0) \) is a distribution peak position of 0, and the scale parameter of half the width at half of the maximum value is Cauchy distribution of \( \eta \), and the value of \( \eta \) is still determined by the example.
Data transmission in clusters

3.2 Selection of cluster head

The selection process of cluster head is as follows: after the non-uniform clustering based on prior knowledge is used for initialization, the network partition is completed, and the node network is divided into several clusters with different sizes. The next step is to select the cluster head node of the subregion.

The fitness of each node is calculated by using the designed fitness function and sorted in descending order. It determines whether the distribution of network nodes meets the search conditions, such as the number of iterations reaching threshold or the distribution of network nodes reaching the global optimum. Then, use formula (6) to update the position of the member with the worst fitness in the cluster. If the new position obtained is better than the original position, replace it; otherwise, keep the original position unchanged. And then use formula (7) to update the node with the worst fitness in the cluster again. After updating the worst member nodes in the cluster, the next step is to update the elite nodes with high fitness in the cluster. First, form the elite node group with $h$ nodes with the highest fitness in the cluster, and then use formula (8) to update the position of the via elite nodes in the group. So repeat the step above, until this time the end of the iteration of the algorithm.

After each iteration, we must carry on the judging of the local search terms again, if not satisfied, immediately jump out of the cycle, then transfer to the condition of global searching, if global search conditions are not satisfied either, namely the network node distribution has amounted to global optimal in theory. The local optimal solution of each cluster area will be taken as the cluster head node of the current area for data transmission with sink. The rest nodes in the cluster automatically become member nodes, and the clustering will be completed.

Figure 2 shows the flow required for clustering after each iteration. The detailed description of ISFLA is displayed as Algorithm 1.

### Algorithm 1 Improved shuffled frog leaping algorithm

Initialization: the network nodes whose total number is $N$ are divided into $K$ clusters.

Input: The upper limit of global iteration accuracy $\omega_{global}$, the upper limit of local iteration accuracy $\omega_{local}$.

While iteration accuracy $\omega < \omega_{local}$ do

For $i = 1$ to $M$ do

For $j = i$ to $M$ do

If $P_{j} > \omega$ then

$k = j$

End if

End for

End if

If $k \neq i$ then

$P_{i} = P_{k}$

End if

End if

If iteration accuracy $\omega < \omega_{local}$ then

Update the worst individual of the subgroup by (10);

If $P_{w}^{e} > \omega$ then

$P_{w} = \omega$

End if

End if

Else then

Update the worst individual of the subgroup by (11);

If $P_{w}^{e} > \omega$ then

$P_{w} = \omega$

End if

Else then

Update $P_{w}$ by generating new individuals randomly;

End else

End else

Update better individuals of subgroup by (12);

If $P_{w}^{e} > \omega$ then

$P_{w} = \omega$

End if

End while

3.3 Design of data routing

When adaptive clustering finished using the above ISFLA, data transmission starts. Data transmission can be divided into data transmission in clusters and between clusters, as shown in Figure 3.

#### 3.3.1 Data transmission in clusters

Data transmission in clusters means that member nodes collect data at regular intervals and send it to the cluster head, and then the cluster head collects data and merges the data. Data transmission in a cluster uses the single-hop method, that is, member nodes in a cluster collect data and send it directly to the cluster head. Cluster head node collects data and merges the data, and finally sends the data packet to sink node through the established multi-hop routing algorithm.

However, in this cluster, the cluster head node needs to receive the information transmitted by other member nodes in this cluster frequently, which is likely to result in the transmission delay and the channel occupation time caused by node competing channels. To ensure the stability and timeliness of transmission in the cluster and avoid the conflict of data transmission in the cluster, TDMA (Time Division Multiple Access) scheduling is also adopted in data transmission in the cluster, that is, the discrete-time multiple access frame is constructed to eliminate the interference in the cluster by means of the average allocation of channel occupation time.

As shown in Figure 4, TDMA is used to communicate with nodes in a cluster. Considering that TDMA frames may cause a waste of resources in the light load network, a better method is to use channel contention. In this case, a multi-path can effectively increase the capacity of the channel. In the data transmission in clusters, to reduce data conflicts, the cluster head node creates a TDMA scheme for its member nodes and broadcasts it between the member nodes in the cluster. Each node sends its data to the cluster head node based on a specified TDMA schedule.
3.3.2 Data transmission between clusters

Data transmission between clusters refers to data fusion of the collected data after the cluster head node receives the information transmitted by member nodes in the cluster, and data integration is sent to the sink node through the established hybrid multi-hop routing algorithm. The hybrid multi-hop algorithm between clusters proposed here mainly contains the following steps: the head node of each cluster receives the data forwarded from the sink node and collects the information. Meanwhile, the distance between the cluster head node and sink node, the residual energy of the cluster head, and the minimum hop number with the sink node are recorded. Then, the weight function value $\text{Weight}(i, j)$ was obtained by using the cost function $\text{Cost}(i, j)$, as shown in formulas (9) and (10), where $d_{ij}$ represents the distance between node $i$ and node $j$, and $d_{i, \text{sink}}$, $d_{j, \text{sink}}$ represents the distance from node $i$ and $j$ to sink node, respectively. Among them, weight $\beta$ is introduced into the calculation formula of $\text{Weight}(i, j)$, where $\beta$ plays the role of adjusting the proportion of $E_{ij}$ and $P_i / \text{Cost}(i, j)$ in $\text{Weight}(i, j)$. The larger $\text{Weight}(i, j)$ is, the greater the proportion of energy consumed by information transmission from node $i$ to node $j$ is in the calculation of $\text{Weight}(i, j)$. However, $\beta$ should not be too large, otherwise, the cost function will lose its set significance.

According to the weight function $\text{Weight}(i, j)$, the node with the largest function value becomes the cluster head node of the next hop. Repeat the above steps until the next hop node of the current cluster head node is itself, then transmit the data directly to the sink node. The data transmission route between the
clusters is established. CDMA technology is introduced into transmission between clusters, as shown in Figure 5. CDMA (Code Division Multiple Access) technology is used to reduce the overhead of data transmission and improve the flexibility of data transmission between clusters.

\[
\text{Cost}(i,j) = \begin{cases} \\
\sqrt{d_{ij}^2 + d_{i\text{sink}}^2} & d_{i\text{sink}} \leq d_{i\text{sink}}^* \\sqrt{d_{ij}^2 + d_{j\text{sink}}^2} & d_{i\text{sink}} > d_{i\text{sink}}^* \end{cases} 
\]

\[
\text{Weight}(i,j) = \beta \cdot E_{ij} + \beta \cdot P_i / \text{Cost}(i,j)
\]

However, data transmission will be carried out between clusters until the residual energy of a certain cluster head node in the current route is lower than the set threshold. When that happens, the sub-optimal solution will be replaced as the new cluster head node in its cluster area to continue data transmission until the residual energy of the new cluster head node is lower than the set threshold, and the next optimal solution will be selected. These steps are repeated until the number of remaining nodes in the network falls below a certain threshold, at which time the network life is terminated.

### 4 EXPERIMENT RESULTS

The performance of the protocol is verified and simulated by omnet++. As an open-source simulation software platform, omnet++ can provide effectiveness analysis and modular simulation for large wireless multi-sensor networks.

To ensure the generality of the simulation experiment, 400 sensor nodes were randomly distributed in a network area of 200 m × 200 m. We compare the network lifetime, power consumption, and throughput of LEACH protocol, DEBUC protocol, and EEUUC protocol. To reduce errors, a total of 10 sets of data were used to observe the simulation effects of the four protocols. The specific parameters set in this simulation are shown in Table 1.

**TABLE 1 Simulation parameters**

| Parameter                          | Value |
|------------------------------------|-------|
| Numbers of nodes                   | 400   |
| Monitoring area                    | 200 m 200 m |
| Base station coordinates           | (100, 250) m |
| Frequency of data transmission     | 10 Hz |
| Data transmission rate             | 250 kb/s |
| Initial energy of node             | 0.5 J |
| Energy consumption of transmission | 5 nJ/bit |
| Energy consumption of sending      | 5 nJ/bit |
| Amplification coefficient ($d > d_0$) | 0.0013 pJ/bit/m^4 |
| Amplification coefficient ($d < d_0$) | 10 pJ/bit/m^2 |
| Coefficient $\alpha + \beta + \gamma$ | 1 |
| Coefficient $t$                    | 0.5 |

**FIGURE 5** CDMA adopted for inter-cluster communication

**FIGURE 6** Comparison of the number of remaining nodes in the network

#### 4.1 Network lifetime

Network lifetime is an important index that affects network performance. We take the death time of the last node of the network as the network lifetime. The network lifetime of the four protocols is compared, and the simulation results are shown in Figure 6. LEACH, EEUUC, DEBUC, and E2NUCR network lifetime are approximately 1063, 1158, 1554, and 1961 rounds, respectively. According to Figure 6, the comparison of the death time of the last node of LEACH, EEUUC, DEBUC, and E2NUCR can be seen more intuitively. LEACH is the shortest and E2NUCR is the longest. LEACH protocol adopts single-hop communication, the residual energy of nodes is not taken into account when clustering, so the network lifetime is the shortest. The other three protocols all adopt the uneven clustering method that improves the energy balance, so the network lifetime is longer than LEACH. The EEUUC protocol mainly considers the distance from the node to sink node and the remaining energy to construct different clusters and select
cluster heads. But the energy consumption in the process of cluster head selection is large. The DEBUC protocol improves the competition mechanism of cluster head selection, but only refers to the location information when clustering, and the size of clusters far from the base station node is too large, resulting in faster energy consumption. There is still much room for improvement in the inter-cluster communication strategies of the two algorithms. Based on the uneven clustering method, the E2NUCR protocol adds the ISFLA to make the cluster head node selection more reasonable. The results show that E2NUCR has the longest network lifetime.

The time of the first dead node can also reflect the network lifetime to some extent. The comparison of the first node time can be seen from the histogram in Figure 7, both LEACH and EEUC protocols had the first node death time before 500 rounds, while DEBUC protocol extended the first node death time to about 1000 rounds. E2NUCR protocol has 1260 rounds of first-node death, which is significantly better than LEACH protocol and EEUC protocol and also has advantages over DEBUC. This is mainly due to the mutation factor used in the early clustering to accelerate the convergence speed of the algorithm and uneven clustering method to achieve the purpose of energy balance. Therefore, the first node death time of the E2NUCR protocol is later than the other three protocols.

4.2 Throughput

The throughput of the four protocols is compared in Figure 8, which is the total number of data packets received by the base station. Among them, LEACH only received about 260,000 packets in the end. EEUC and DEBUC receive about 334,000 and 500,000 packets, respectively, while E2NUCR has the maximum throughput of about 596,000 packets. In the initial stage, LEACH received packets quickly, but due to the large number of redundant data containing LEACH unable to efficiently handle, later it gradually weakens. The remaining three all adopt uneven clustering method, and the cluster distribution is more reasonable, thus improving the throughput. The length of the network lifetime also affects the throughput, so E2NUCR still has advantages in terms of throughput.

4.3 Rate of energy consumption

The energy consumption rate of the protocol greatly affects the network lifetime. Figure 9 is the comparison of energy consumption rates of the four protocols. LEACH's energy consumption rate remains the highest because of a lack of consideration for residual energy in the clustering mechanism. The EEUC protocol adopts the uneven clustering mechanism to improve the energy consumption of LEACH to some extent. The DEBUC protocol adopts an improved competition mechanism for cluster head selection to make cluster distribution more reasonable than EEUC. On the basis of non-uniform clustering, E2NUCR adopts an ISFLA to select cluster head nodes and makes the cluster distribution more reasonable.

From the above simulation results, it can be seen that compared with LEACH, EEUC, and DEBUC, E2NUCR is a more efficient routing protocol with a longer network lifetime and more balanced energy.

5 Conclusion

For non-uniform clustering network, the cluster head nodes for data transmission and data merging of dual tasks at the same
time will be a large number of its energy consumption, the traditional algorithm of hybrid leapfrog algorithm is complex, weak search breadth and depth, easier to fall into a local optimal solution of the problem; this paper proposes a re-shuffled leapfrog algorithm based on improved efficient non-uniform CRP. The non-uniform method is adopted to divide the sensor nodes into clusters, and the improved hybrid leapfrog algorithm is used in the clustering to find the optimal cluster head, to improve the energy efficiency of the nodes, balance the energy consumption in WSNSs, and minimize the probability of energy hole occurrence. The protocol USES fitness function and sub-group elite individual update strategy to iterate the node performance and select the optimal cluster head. The simulation results show that this method improves the energy efficiency and the service life of the network. Our next work will continue to optimize the hybrid frog-hopping algorithm used in the E2NUCR protocol to reduce the complexity and increase the convergence rate of the algorithm. Inter-cluster transfers are considered to introduce other factors such as transmission direction to make this routing protocol more suitable for practice.

ACKNOWLEDGEMENT
This research has been sponsored by Natural Science Foundation of China (project number: 61803387, 61762051). The work is also supported by Central South University of College students’ free exploration project (project number: 201910533333) and Research-Based Learning and Innovative Experiment Projects (project number: YC20180624).

REFERENCES
1. Amer, O., et al.: Enhanced LEACH protocol for increasing a lifetime of WSNSs. Pers. Ubiquitous Comput. 23(5–6), 901–907 (2019)
2. Dabagi, F., et al.: A survey on green routing protocols using sleep-scheduling in wired networks. J. Netw. Syst. Manage. 27, 106–122 (2017)
3. Singh, S.K., et al.: A survey on successors of LEACH protocol. IEEE Access 5(2), 4298–4328 (2017)
4. Tao, W., et al.: Non-uniform clustering routing protocol for energy-balanced wireless sensor networks. J. Software 23(5), 1222–1232 (2012)
5. Eusuff, M.M., Lansey, K.E.: Optimization of water distribution network design using the shuffled frog leaping algorithm. J. Water Resour. Plann. Manage. 129(3), 210–225 (2003)
6. Shi, Q.: An improved hybrid frog leaping algorithm. J. Chongqing Univ. Technol. 5, 78–81 (2013)
7. Dai, Y., et al.: Cooperative evolutionary frog leaping algorithm. Comput. Eng. Sci. 01, 139–147 (2018)
8. Jiang, Z., et al.: A hybrid frog leaping algorithm based on elite strategy improvement. J. Shanghai Dianji Univ. 21(01), 14–19 (2018)
9. Jiang, Z., et al.: A hybrid frog leaping algorithm based on elite strategy improvement. J. Shanghai Dianji Univ. 21(01), 14–19 (2018)
10. Gao, H., Xu, X., Han, S.: Homogeneous clustering algorithm based on average residual energy for energy-efficient MTC Networks. Paper presented at the 24th Asia-Pacific Conference on Communications, Ningbo, Zhejiang Province, pp. 28–33 (2018)
11. Darabkh, K.A., et al.: MT-CHR: A modified threshold-based cluster head replacement protocol for wireless sensor networks. Comput. Electr. Eng. 72, 926–938 (2018)
12. Darabkh, K.A., et al.: Energy-aware and density-based clustering and relaying protocol (EA-DB-CP) for gathering data in wireless sensor networks. Appl. Soft Comput. 80, 154–166 (2019)
13. Bhushan, S., et al.: Energy efficient clustering protocol for heterogeneous wireless sensor network: A hybrid approach using GA and K-means. Paper presented at the IEEE Second International Conference on Data Stream Mining and Processing, Lviv State, Ukraine, pp. 381–385 (2018)
14. Sadek, R.A.: Hybrid energy aware clustered protocol for IoT heterogeneous network. Future Comput. Inf. Sci. J. 3(2), 166–177 (2018)
15. Toor, A.S., Jain, A.K.: Energy aware cluster based multi-hop energy efficient routing protocol using multiple mobile nodes (MEACBM) in wireless sensor networks. AEU - Int. J. Electron. Commun. 16(4), 41–53 (2019)
16. Wang, X., et al.: An improved unequal cluster-based routing protocol for energy efficient wireless sensor networks. Paper presented at the International Conference on Intelligent Transportation, Big Data and Smart City (ICTBS), Ha Long Bay, Vietnam, pp. 165–169 (2019)
17. Gupta, A., Shekhar, N.: A novel K-means L-layer algorithm for uneven clustering in WSN. Paper presented at the International Conference on Computer Communication and Signal Processing, Chennai, pp. 1–6 (2017)
18. Yang, X., et al.: Research on clustering routing algorithm based on K-means++ for WSN]. Paper presented at the 6th International Conference on Computer Science and Network Technology, Dalian, pp. 330–333 (2017)
19. Bozorgi, S.M., Bidgoli, A.M.: HEEC: A hybrid unequal energy efficient clustering for wireless sensor networks. Wireless Networks 25(8), 4751–4772 (2019)
20. Wang, H., et al.: Research on LEACH algorithm based on double cluster head cluster clustering and data fusion. Paper presented at the IEEE International Conference on Mechatronics and Automation, Beijing, pp. 342–346 (2017)
21. Gupta, G.P., Jha, S.: Integrated clustering and routing protocol for wireless sensor networks using cuckoo and harmony search based metaheuristic techniques. Eng. Appl. Artif. Intell. 68, 101–109 (2018)
22. Gajjar, S., et al.: FAMACROW: Fuzzy and ant colony optimization based combined mac, routing, and unequal clustering cross-layer protocol for wireless sensor networks. Appl. Soft Comput. 43, 235–247 (2016)
23. Baradaran, A.A., Navi, K.: HQCA-WSN: High-quality clustering algorithm and optimal cluster head selection using fuzzy logic in wireless sensor networks. Fuzzy Sets Syst. 389, 114–144 (2020)
24. Sridheshpande, N., Udapi, V.: Fractional ion optimization for cluster head-based routing protocol in wireless sensor network. J. Franklin Inst. 354(11), 4457–4480 (2017)
25. Kong, L., et al.: Directional shuffled frog leaping algorithm. Paper presented at the International Conference on Smart Vehicular Technology, Transportation, Communication and Applications, pp. 257–264. Springer, Berlin (2017)
26. Dungadevi, P., Srinivasan, S.: Resource allocation in cloud computing using SFLA and cuckoo search hybridization. Int. J. Parallel Program. 48, 549–565 (2018)
27. Li, Y., Yan, Z.: Improved shuffled frog leaping algorithm on system reliability analysis. Brain Inf. 6(1), 1 (2019)
28. Edla, D.R., et al.: An efficient load balancing of gateways using improved shuffled frog leaping algorithm and novel fitness function for WSNs. IEEE Sens. J. 17(20), 6724–6733 (2017)
29. Fanian, F., Rafsanjani, M.K.: Memetic fuzzy clustering protocol for wireless sensor networks using cuckoo search optimization. Int. J. Intell. Comput. Neurosci. 17(1), 1–3 (2018)
30. Li, J., Dai, Y.: Adaptive hybrid variation frog leaping algorithm. Comput. Eng. Appl. 49(10), 58–71 (2013)

How to cite this article: Huamei Q, Chubin L, Yijiahe G, Wangping X, Ying J. An energy-efficient non-uniform clustering routing protocol based on improved shuffled frog leaping algorithm for wireless sensor networks. IET Commun. 2021;15:374–383. https://doi.org/10.1049/ietcom.2012.12067