

PAPER • OPEN ACCESS

# An Energy Efficient Clustering Algorithm in Wireless Sensor Networks for Internet of Things Applications

To cite this article: Xingjie Ci *et al* 2021 *J. Phys.: Conf. Ser.* **1881** 042035

View the [article online](#) for updates and enhancements.

## You may also like

- [Data Disseminated Energy-Efficient Clustering Algorithm For Avoid Load and Bandwidth Consumption In WSN Integrated IoT Platform](#)  
Haider Ali Jasim Alshamary, Emad Hmood Salman and Yousif Allbadi
- [Secure and Efficient Fire-fly Data Routing Algorithm for Wireless Sensor Networks in IoT Monitoring Systems](#)  
R. Kowsalya and B. Rosiline Jeetha
- [Improvement Performance of Clustered Shortest Geopath Routing Protocol Using Data Aggregation in Wireless Sensor Network Environments](#)  
Rozita, W Wibisono and R Anggoro

# An Energy Efficient Clustering Algorithm in Wireless Sensor Networks for Internet of Things Applications

Xingjie Ci\*, Kang Wen, Ying Sun, Weifan Sun and Wei Deng

Chongqing University of Posts and Telecommunications, Chongqing, China

\*Corresponding author e-mail: s180131048@stu.cqupt.edu.cn

**Abstract.** Internet of things (IoT) integrates sensor, radio frequency identification, network and cloud computing technology into a large-scale monitoring area. Wireless sensor network (WSN), as the most suitable infrastructure for the Internet of Things, how to reduce its energy consumption has been a research hotspot. In hierarchical networks, cluster head nodes need to take on more energy consumption tasks. Therefore, reasonable cluster head selection scheme and normal node joining cluster mechanism become particularly important. This paper proposes a new clustering method to save energy and prolong network life. Firstly, in the cluster head selection phase, the influence of energy threshold, residual energy, node density and communication factor is considered. Secondly, in the ordinary node joining cluster phase, the residual energy and distance factor are considered to form the non-uniform cluster set of the network. Simulation results show that the algorithm has good performance in reducing and balancing energy consumption.

**Keywords:** Internet of Things (IoT); Wireless Sensor Network (WSN); Clustering Algorithm; Energy Consumption

## 1. Introduction

In recent years, the Internet is developing rapidly towards the combination of sensors and devices, known as IoT system. WSNs are the most powerful networking infrastructure that monitors an IoT system[1]. As one of the core supporting technologies in IoT system, as well as the powerful routing and network foundation of IoT, WSN plays an increasingly important role in IoT. WSN collect large-scale monitoring area data through scattered nodes, which are used in smart home, military monitoring, space exploration, environmental monitoring and other IoT applications. In addition, due to the limited energy and self-organization of sensor nodes, WSN enables IoT to meet the challenges of adaptability and energy saving.

WSN is an information transmission network consists of a large number of sensor nodes in self-organization and multi-hop mode. A single sensor node can sense, measure and collect environmental information according to the local predefined decision-making process, and then relay it to the data center for further analysis and processing through the base station (BS). The environment of WSN is usually complex and harsh, which results in many limitations, such as processing constraints, low memory, security and limited energy[2]. Network lifetime is dependent on the energy stored at the nodes[3]. Therefore, how to reduce node energy consumption and improve network lifetime has always



been the research focus and hotspot in WSN.

Researches on energy saving in WSN leads to classification in clustering techniques [4]. Clustering algorithm has always been a hot research area, because it can balance network consumption to each node. Data transmission and reception occupy the main part of energy consumption. In direct transmission, sensor nodes directly transmit the collected information to BS, which will lead to the nodes far away from BS to consume lots of energy and die too fast, thus sharply shortening the network life. Clustering technology reduces the amount and distance of data transmission by selecting the cluster head(CH). The member nodes in the cluster(CM) only need to send collected data to CH, which greatly reduces the transmission distance and saves the energy of a single node. Clustering technology divides WSN into multiple clusters, so that the entire network has a clear hierarchy and the nodes have a clear division of labor, reducing node energy consumption and balancing the network, thereby extending the entire network life cycle.

In reference[5], several works on clustering are introduced. Low energy adaptive cluster hierarchical(LEACH)[6] algorithm is one of the well-known clustering algorithms, and it is also a classic protocol in hierarchical routing protocol. However, it does not consider other factors, cluster head needs to undertake more tasks such as data receiving, fusion and transmission. Therefore, reasonable cluster head selection scheme and node joining cluster mechanism become particularly important. In this paper, an novel clustering algorithm is proposed to improve the cluster head selection and node joining cluster strategy to reach the purpose of reducing and balancing energy consumption.

The rest of this paper is: section 2 reviews the clustering work in WSN and IoT applications, section 3 describes the network model, section 4 presents the proposed algorithm, section 5 show the simulation parameters and results analysis, and conclusion is presented in section 6.

## 2. Related Work

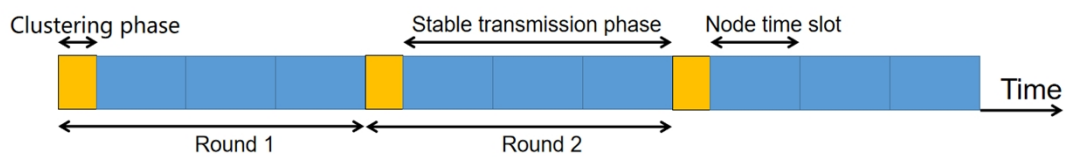
The number and position of CH are unstable due to the random cluster head selection mechanism of LEACH. Many scholars have made different improvements to it. This section will discuss the research work on clustering algorithm in WSN.

In 2011, Dasgupta et al.[7] proposed a new CH selection method, which embedded FCM algorithm according to the requirements of minimum distance and maximum residual energy. In 2018, M. Razzaq et al.[8] proposed a k-mean clustering algorithm, which determines whether a node becomes CH by two weight functions. In addition, in order to ensure flexibility, the optimal packet size is also taken into account. In 2015, Arumugam et al.[9] proposed an energy routing protocol considering the integration of data and the selection of the optimal CH. However, it only selects nodes with high residual energy and ignores any other factor. In 2017, shokrollahi et al.[10] proposed an energy-saving clustering algorithm, which combines FCM with ECAFG to optimize the selection process of CH from three aspects: residual energy, distance from node to cluster center and BS. In 2018, Elshrkawey et al.[11] proposed a CH selection algorithm. In the selection process, pay attention to the remaining energy of the node and the distance from the sink. In addition, CH only sends broadcast messages to the nearby radius to avoid a surge in long-distance energy consumption. In 2016, Murali[12] proposed a novel clustering method. Considering the relationship of communication and cooperation between nodes, the results illustrate that compared with no cooperation, energy consumption is lower when there is cooperation. However, the traffic consumption will increase as the network size increases. In 2017, K. Jung et al.[13] proposed a cluster head selection algorithm based on fuzzy logic, which took battery power, node density and data transmission frequency as consideration parameters. In 2019, Payam et al.[14] proposed a load balancing algorithm that can intelligently change to spread node energy consumption. The fuzzy logic controller forms a priority queue according to the input three parameters to allocate CH among nodes.

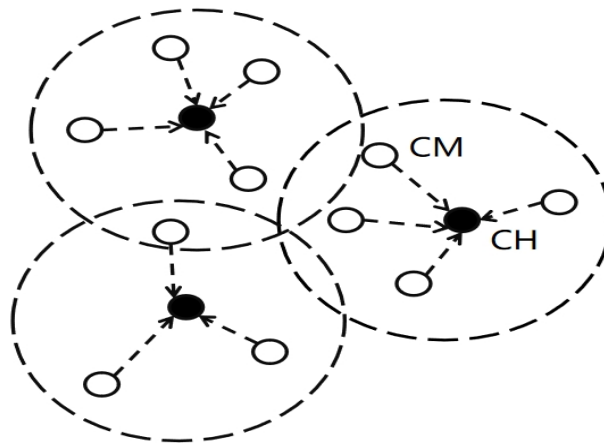
## 3. Network Model

### 3.1. Network Topology Model

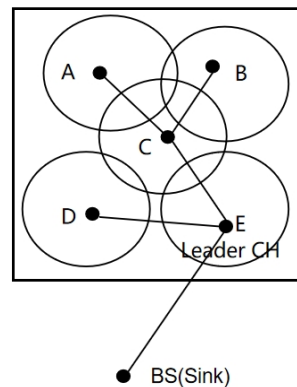
The clustering protocols generally have two phases: clustering phase and stable transmission phase, as shown in Fig. 1, after which the CH communicates with the BS. In LEACH protocol, all CH communicate directly with BS in single-hop mode. Many scholars have studied multi-hop transmission mode. Single-hop mode is straightforward and easy to execute. However, once the distance is too large, energy consumption will increase sharply, nodes will die too fast, and network cycle will be greatly shortened. Multi-hop mode divides long-distance transmission into short-distance transmission, and the energy consumption is distributed to each CH, which can reduce energy consumption and balance energy consumption, but also increase the energy consumption and delay of data forwarding, which is not suitable for real-time network. In this paper, we adopt a single-hop and multi-hop inter-cluster transmission strategy. The network topology of clustering and inter-cluster transmission is shown in Fig. 2 and Fig. 3.



**Figure 1.** Structure diagram of rounds



**Figure 2.** CH selection and CM joining cluster



**Figure 3.** Single-hop and multi-hop inter-cluster transmission mode

In Figure 3, cluster head E is the leader CH, which is responsible for communication with BS or Sink. Other CHs send data to leader CH in the form of single hop or multi hops. For example, cluster heads A and B are transmitted to leader CH through multi-hop mode, while D communicates with E directly through single hop. In this way, a routing tree with the leader CH as the root is formed to balance the network energy consumption to each CH, thus prolonging the network lifetime.

### 3.2. Network Energy Consumption Model

We adopt the energy consumption model of wireless communication in reference [6]. As shown in Fig. 4, the sensor communication module has a transmitting device and a receiving device. In this model, the transmitter sends  $l$  bits data from a distance of  $d$ , and the corresponding energy consumption of the transmitter satisfies the following formula:

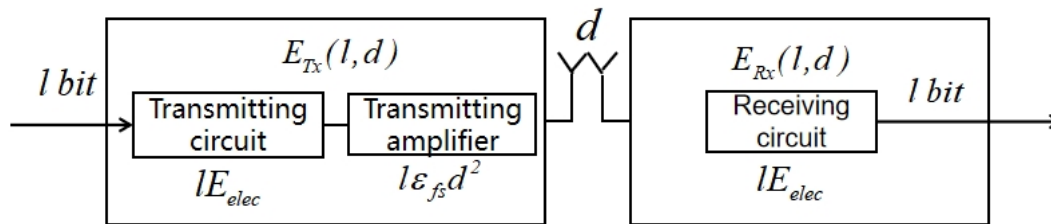
$$E_{Tx}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4, & d > d_0 \end{cases} \quad (1)$$

Where  $E_{elec}$  is the energy consumed by the transmitting circuit to transmit unit data.  $\epsilon_{fs}$  is the parameter under the free space model,  $\epsilon_{mp}$  is the parameter under the multipath model, and  $d_0$  is calculated according to the following formula:

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \quad (2)$$

The energy required by the receiver is:

$$E_{Rx}(l, d) = E_{Tx}(l) = lE_{elec} \quad (3)$$



**Figure 4.** Energy consumption model of wireless communication

### 3.3. Optimal Cluster Head Probability

The CH selection mechanism of LEACH protocol is directly related to the proportion of CHs in the network, so determining the CH ratio has a direct impact on the network power consumption. In this paper, we assume that  $N$  homogeneous nodes are randomly distributed in  $M \times M$  network, the number of CHs is  $k$ , the amount of data transmission per round is  $L$  bits, the energy consumption of data fusion per bit is  $E_{Df}$ , and the communication distance between clusters is less than  $d_0$ .  $d_{CMtoCH}$  is the distance between CM and CH,  $d_{CHtoCH}$  represents the distance between CH and CH, and  $d_{LDtoBS}$  represents distance between leader CH and BS or Sink. The steps are as follows.

Step 1. The total energy consumption of CMs is:

$$E_{CM} = (N - k)(E_{elec}L + \varepsilon_{fs}Ld_{CMtoCH}^2) \quad (4)$$

Assuming that the CH is in the center of the cluster, the node density is  $\rho(x, y) = \frac{1}{M^2/k}$ , the mean of the distance between the CM and the CH is obtained as follows:

$$E[d_{CMtoCH}^2] = \iint (x^2 + y^2)\rho(x, y)dxdy = \rho \int_0^{2\pi} \int_0^{M/\sqrt{\pi k}} r^3 dr d\theta = \frac{M^2}{2\pi k} \quad (5)$$

Incorporating formula (5) into formula (4) can obtain the total energy consumption of CMs:

$$E_{CM} = (N - k)(E_{elec} + \varepsilon_{fs} \frac{M^2}{2\pi k})L \quad (6)$$

Step 2. On average, each normal CH receives the data of one CH and  $(\frac{N}{k} - 1)$  members, and sends its own data and received data to the next CH. The energy consumption is as follows:

$$E_{CH\_NRM} = (\frac{N}{k} - 1)LE_{elec} + LE_{elec} + (\frac{N}{k} + 1)LE_{Df} + LE_{elec} + \varepsilon_{fs}Ld_{CHtoCH}^2 \quad (7)$$

The average value of  $d_{CHtoCH}$  is twice the radius of the cluster area (denoted as  $R$ ), and the average area of each cluster is  $\frac{M^2}{k}$ ,  $\frac{M^2}{k} = \pi R^2$ ,  $d_{CHtoCH} = 2R = \frac{2M}{\sqrt{\pi k}}$  can be obtained from formula  $\pi R^2 = \frac{M^2}{k}$ . Equation (7) can be changed into:

$$E_{CH\_NRM} = [(\frac{N}{k} + 1)(E_{elec} + E_{Df}) + \varepsilon_{fs} \frac{4M^2}{\pi k}]L \quad (8)$$

Equation (8) represents the minimum energy required for the normal CH to work. If it is less than this value, it will die due to insufficient energy. Therefore,  $E_{CH\_NRM}$  is the minimum standard for a node to act as an ordinary CH, that is, the energy threshold of a normal CH, which is denoted as  $E_{th}$ :

$$E_{th} = E_{CH\_NRM} = [(\frac{N}{k} + 1)(E_{elec} + E_{Df}) + \varepsilon_{fs} \frac{4M^2}{\pi k}]L \quad (9)$$

Step 3. The difference between leader CH and normal CH is that the data needs to be sent to the remote BS or Sink.

$$E_{CH\_Leader} = [(\frac{N}{k} + 1)E_{elec} + (\frac{N}{k} + 1)E_{Df} + \varepsilon_{mp}d_{LDtoBS}^4]L \quad (10)$$

Step 4. According to formula (9) and (10), the total energy consumption of all CHs is as follows:

$$\begin{aligned}
E_{CH} &= (k-1)E_{CH\_NRM} + E_{CH\_leader} \\
&= [(N+k)(E_{elec} + E_{Df}) + \varepsilon_{fs} \frac{4M^2(k-1)}{\pi k} + \varepsilon_{mp} d_{LDtoBS}^4] L
\end{aligned} \tag{11}$$

Step 5. The total energy consumption of each round can be obtained from equations (6) and (11):

$$\begin{aligned}
E_{round} &= E_{CM} + E_{CH} \\
&= [2NE_{elec} + (N+k)E_{Df} + \varepsilon_{mp} d_{LDtoBS}^4 + \varepsilon_{fs} \frac{M^2}{2\pi} (\frac{N+7k-8}{k})] L
\end{aligned} \tag{12}$$

Step 6. Calculate the partial derivative of equation (12) to obtain the optimal number of CHs:

$$k_{opt} = M \sqrt{\frac{(N-8)\varepsilon_{fs}}{2\pi E_{Df}}} \tag{13}$$

Therefore, the optimal CH ratio is:

$$p_{opt} = \frac{k_{opt}}{N} \tag{14}$$

## 4. Proposed Clustering Algorithm

### 4.1. Clustering Phase

The LEACH protocol CH selection mechanism is: in each round, The random number generated by node compares with the threshold in equation (15). If the value is less than the threshold, it becomes the CH, otherwise it is a CM.

$$T_i = \begin{cases} \frac{p}{1 - p \left( r \bmod \frac{1}{p} \right)}, & i \in G \\ 0, & i \notin G \end{cases} \tag{15}$$

Among them,  $p$  represents the proportion of CH in the network,  $r$  is the current rounds,  $i$  represents the  $i$ -th node.  $G$  is a set that nodes haven't become CH.

In this paper, the steps to improve the clustering stage are as follows:

Step 1. From equation (14), we can see that  $p_{opt}$  is the optimal CH ratio, and the normal CH energy threshold is  $E_{th}$ . Considering the  $E_{th}$  and the remaining energy of the node, the election probability of node  $i$  in each round is:

$$p_i = p_{opt} \frac{E_i - E_{th}}{E - E_{th}} \tag{16}$$

Where  $E_i$  is the remaining energy of node  $i$  and  $\bar{E} = \frac{1}{N} \sum_{i=1}^N E_i$  is the current average energy of the network. If the residual energy of node  $i$  is below the threshold, resulting in  $p_i$  less than 0 and then  $T_i$  less than 0, the node can not be selected as CH. Therefore, formula (17) can not only improve the probability of high-energy nodes being selected as CHs, but also effectively protect low-energy nodes.

Step 2. Consider the density of nodes around node  $i$ .

$$T_i = \begin{cases} \frac{p_i}{1 - p_i \left( r \bmod \frac{1}{p_i} \right)} [E_s + (1 - E_s) N_d], i \in G \\ 0, i \notin G \end{cases} \quad (17)$$

Where  $E_s = \frac{E_i}{E_0}$ ,  $N_d = n_{neighbor} * p_{opt}$ ,  $E_0$  is the initial energy of the node and  $n_{neighbor}$  is the number of nodes around node  $i$  (within the communication radius  $R$ ).

Step 3. node  $i$  randomly generates a random number between 0 and 1. If the number is below than the threshold of equation (17), the node becomes the temporary CH and immediately broadcasts the elected message to the whole network. The non CH node select the cluster with the smallest  $d_c$  to join according to formula (18).

$$d_c = \frac{d(j, ch_i) E_0}{d(ch_i, ld) E_i} \quad (18)$$

$d(j, ch_i)$  represents the distance between non CH  $j$  and temporary CH  $i$ , and  $d(ch_i, ld)$  is the distance between temporary CH  $i$  and leader CH. In the process of multi-hop communication between clusters, CH near leader CH has a large amount of tasks to converge and forward data, which leads to increased energy consumption and forms a problem of energy consumption hot spots; while the CH far away from the leader CH has a lighter task. The energy consumption is relatively small. Therefore, let non-cluster-head nodes choose the CH that is closer to themselves, have more remaining energy, and are far from the leader CH to join, so that clusters closer to the leader CH have less CM, which reduces the energy consumption of intra-cluster communication. The clusters farther from the leader CH have more CMs, forming a non-uniform network cluster, effectively disperses the energy consumption of the whole network.

Step 4. Establish the formal cluster head.  $c_i$  is defined as the  $i$ -th cluster,  $D(u, v)$  is the distance between node  $u$  and node  $v$  in cluster  $i$ , and  $m$  is the number of nodes. From the energy consumption model, we can see that energy loss is directly related to the distance  $d$ , so the definition  $\sum_{v=1}^{m-1} D^2(u, v)$  is the communication factor within the cluster. The final CH is selected by equation (19) considering the residual energy and communication factor.

$$W(c_i, u) = \frac{E_u}{\sum_{v=1}^{m-1} D^2(u, v)} \quad (19)$$

The nodes in the cluster calculate their  $W$  values, and the largest one becomes the CH of the



cluster, so that the node with more energy and the lowest communication cost becomes the final CH.

#### 4.2. Stable Transmission Phase

After the clustering phase is completed, each CM sends data to CH in its own slot according to TDMA time slot table. CH collects data packets from all sensor nodes in the cluster, and merges the collected data, then transmits them to the leader CH through a combination of single and multiple hop way. Finally, leader CH sends data to BS or Sink.

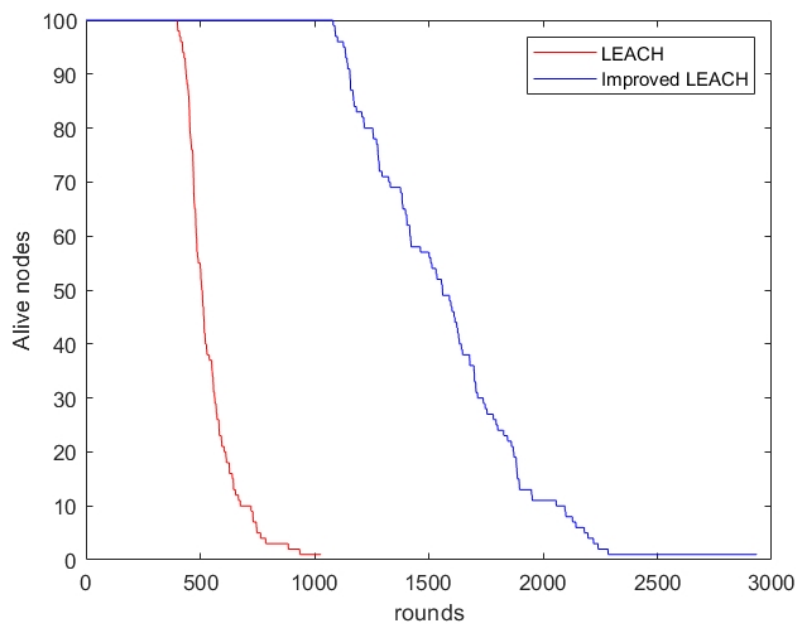
#### 4.3. Simulation Results and Analysis

In this section, we use MATLAB R2019b for simulation and verify the superiority of the proposed algorithm based on the results. The experimental parameters are shown in Fig. 5.

Parameter	Value
Network size(M x M)	100m x 100m
Number of nodes(N)	100
Base station location	(50,175)
Initial CH probability	0.05
$E_0$	0.5J
$E_{elec}$	50nJ / bit
$\epsilon_{fs}$	10pJ/bit/m <sup>2</sup>
$\epsilon_{mp}$	0.0013pJ/bit/m <sup>4</sup>
$E_{df}$	5nJ / bit

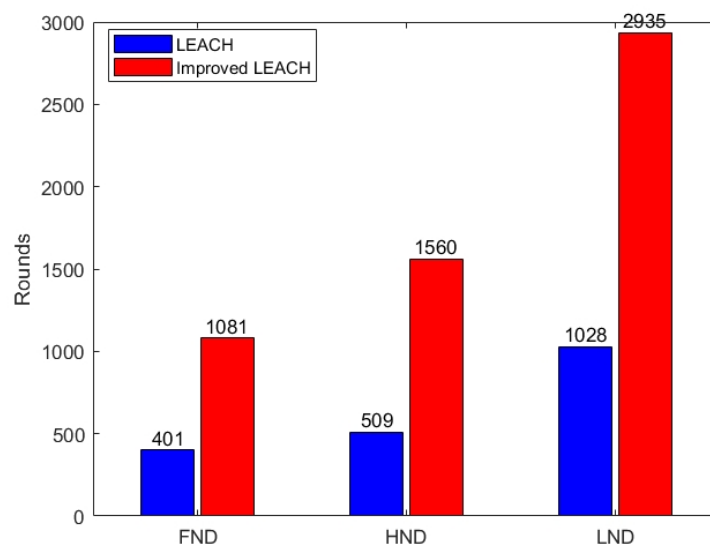
**Figure 5.** Experimental parameters

We consider four parameters to measure the performance of the protocol, namely network life (rounds of network operation), the rounds when the first node dies (FND), the rounds at half node death (HND), and the rounds when the last node dies (LND). We analyze the proposed algorithm and LEACH according to the above parameters, so as to prove the feasibility of the improved algorithm.



**Figure 6.** Alive nodes per round

Figure 6 shows the changes between the number of alive nodes and the rounds. It is obvious that the life cycle of the improved protocol is significantly longer than that of the LEACH protocol. From the beginning of the network operation to the FND, this stage is called the stable period of network operation. The stable period of the improved protocol is significantly longer than LEACH. In addition, the slope of the improved algorithm curve is significantly smaller than LEACH, which is due to the energy consumption balance of the new clustering algorithm, so that after FND, network nodes will not die sharply in short rounds. It can be seen that the improved protocol has better stability and more balanced energy consumption, which can effectively extend the network life cycle.



**Figure 7.** FND, HND and LND

Figure 7 shows the comparison between the improved protocol and LEACH's FND, HND, and LND. For LEACH, FND is 401 rounds, and no node survived after the 1028th round, while the improved protocol, FND is 1081 rounds, and the LND is 2935 rounds.

## 5. Conclusion

In this paper, we propose a new clustering algorithm to disperse energy consumption and extend network life. First, in the cluster head selection stage, consider the impact of energy threshold, residual energy, node density, and communication factors. Second, in the clustering stage of ordinary nodes, consider the remaining energy and distance factors to form a non-uniform network cluster structure to balance energy consumption. We compare and analyze the two protocol from network lifetime, FND, HND and LND aspects. The results indicate that the improved protocol is superior to LEACH and has good performance.

## Acknowledgments

National science and technology major special funded projects (2018ZX0301007-004).

## References

- [1] Xu Z, Chen L, Chen C, et al. Joint clustering and routing design for reliable and efficient data collection in large-scale wireless sensor networks[J]. IEEE Internet of Things Journal, 2015, 3(4): 520-532.
- [2] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks[J]. IEEE Transactions on wireless communications, 2002, 1(4): 660-670.
- [3] Zhou J, Chen S, Wang Y, et al. Spatial compression scheme for improving the lifetime of wireless sensor networks[C]//2017 IEEE 17th International Conference on Communication Technology (ICCT). IEEE, 2017: 922-926.
- [4] Muruganathan S D, Ma D C F, Bhasin R I, et al. A centralized energy-efficient routing protocol for wireless sensor networks[J]. IEEE Communications Magazine, 2005, 43(3): S8-13.
- [5] Fareed M S, Javaid N, Akbar M, et al. Optimal number of cluster head selection for efficient distribution of sources in WSNs[C]//2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications. IEEE, 2012: 632-637.
- [6] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks[C]//Proceedings of the 33rd annual Hawaii international conference on system sciences. IEEE, 2000: 10 pp. vol. 2.
- [7] Dasgupta S, Dutta P. An improved Leach approach for Head selection Strategy in a Fuzzy-C Means induced clustering of a Wireless Sensor Network[J]. Proceedings of the IEMCON Organised by IEM in Collaboration with IEEE, 2011: 203-208.
- [8] Razzaq M, Ningombam D D, Shin S. Energy efficient K-means clustering-based routing protocol for WSN using optimal packet size[C]//2018 International Conference on Information Networking (ICOIN). IEEE, 2018: 632-635.
- [9] Arumugam G S, Ponnuchamy T. EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN[J]. EURASIP Journal on Wireless Communications and Networking, 2015, 2015(1): 1-9.
- [10] Shokrollahi A, Mazloom-Nezhad Maybodi B. An energy-efficient clustering algorithm using fuzzy C-means and genetic fuzzy system for wireless sensor network[J]. Journal of Circuits, Systems and Computers, 2017, 26(01): 1750004.
- [11] Elshrkawey M, Elsherif S M, Wahed M E. An enhancement approach for reducing the energy consumption in wireless sensor networks[J]. Journal of King Saud University-Computer and Information Sciences, 2018, 30(2): 259-267.
- [12] Murali V, Gupta S H. Analysis of energy efficient, LEACH-based cooperative wireless sensor network[C]//Proceedings of the Second International Conference on Computer and Communication Technologies. New Delhi, Springer, 2016: 353-363.
- [13] Jung K D, Lee J Y, Jeong H Y. Improving adaptive cluster head selection of teen protocol using fuzzy logic for WMSN[J]. Multimedia Tools and Applications, 2017, 76(17): 18175-18190.

- [14] Rahimi P, Chrysostomou C. Improving the Network Lifetime and Performance of Wireless Sensor Networks for IoT Applications Based on Fuzzy Logic[C]//2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS). IEEE, 2019: 667-674.