Load-balanced routing algorithm based on cluster heads optimization for wireless sensor networks

Yong-wen Du, Zhang-min Wang, Gang Cai and Jun-hui Gong

Lanzhou Jiaotong University, The School of Electronics & Information Engineering, Lanzhou 730070, China

Abstract: In order to solve the problem of unbalanced load consumption of nodes for wireless sensor networks (WSNs), this paper proposes a load-balanced routing algorithm based on cluster heads optimization for wireless sensor network. The proposed algorithm first applies first-order wireless transmission model to calculate the optimal number of clusters, then calculate nodes competitiveness rating by fuzzy algorithm considering the residual energy of node and distance from the node to base station, cluster head selection uses unequal clustering algorithm according to the competitiveness of nodes. By node competitiveness and energy management mechanism which cooperate with each other to select the best cluster heads. Use connected optimization between clusters to search multi-hop paths base station for reducing energy consumption of node, and consider transmission energy consumption, residual energy, transmission distance and other factors. The experimental results show that the proposed algorithm compared with LEACH and UCDP algorithm, can balance loading and effectively extend the life cycle of wireless sensor network.

1 Introduction

There are a large number of nodes in wireless sensor networks, which are difficult to replenish with limited energy consumption. If the energy consumption is too fast, the unbalanced load will reduce the node lifetime and affect the network performance. Therefore, it is an important issue to study how to reduce the energy consumption of sensor nodes and improve the energy utilization rate of nodes, so as to prolong the network lifetime. Clustering routing algorithm is a clustering to optimize the data fusion, thereby reducing the node direct communication with the base station algorithm. It can reduce the node energy consumption at the same time, improving network scalability. Many clustering protocols have been proposed for WSN. Clustering algorithm includes the election of cluster heads (CHs) and the communication between the ordinary node and CH sand between CHs and base station (BS). Clustering is a useful technique through which we can affect the network lifetime, scalability and load balancing[1]. LEACH[2] algorithm is a kind of the earliest clustering routing algorithm, using single jump directly communicate with CHs. With the deepening of research, the clustering algorithm begins to adopt the multi-hop mode to communicate, so as to achieve the purpose of saving energy. Compared with the single communication, the CHs that near the base station energy consumption will increase, but the overall energy consumption of the network is reduced. Because CHs must communicate within the cluster and other CHs, the energy consumption of CHs are more than the ordinary nodes. If the CHs energy is depleted, it will lead to cluster failure in the current round of communication, and the formation of the energy hole, thus reducing the network lifetime.

In order to solve the unbalanced load among clusters, literature [3] proposed a CHs election mechanism based on the adaptive inertia weight adjustment strategy based on the particle swarm optimization considering the load balancing and energy consumption. Aiming at the problem that the randomly selected cluster heads in the clustering routing protocol easily deviate from the optimal value, literature [4] considers the distribution of cluster load and the residual energy of the nodes and improves the threshold of the CHs to ensure that each CHs' number is within the expected range. literature [5] proposed a hybrid integer linear programming model to solve the problem of integrated topology and clustering routing, in order to determine the optimal CHs position. In order to solve the hot zone problem caused by the uneven load of nodes, literature [6] proposed a clustering routing protocol based on dynamic partition load balancing. Aiming at the problem of excessive energy consumption in long-distance communication between multiple CHs and base stations in LEACH, literature [7] considered the energy and location of nodes in clustering to optimize the cluster structure.

Although the above clustering routing algorithms are effective under certain conditions, they do not completely solve the problem of node energy

[a] Corresponding author: duyongwen@mail.lzjtu.cn

imbalance and increase the energy consumption during each cluster head election. At the same time, the path factors are not considered comprehensively, and the path information is only updated locally, which leads to the unbalanced path load easily. The closer the CHs are to the base station, the more data tasks need to be forwarded, the more energy is consumed, which may easily lead to node premature death.

In this paper, we propose a load-balanced routing algorithm (ECPF) based on cluster heads optimization in Wireless Sensor Network. The Algorithm is divided into two stages: unequal clustering and establishing paths between clusters. The stage of unequal clustering calculates the optimal clustering number, then calculation node competitiveness through fuzzy algorithm according to the residual energy and distance factor. Using unequal clustering algorithm for CH selection, thus selecting the best CH node; Stage of establishing paths between the CHs connected multiple hops routing algorithm to BS according to influence the factors such as transmission of energy consumption,residual energy and transmission distance. At the same time with the energy management mechanism of low energy limit and high energy strengthen, to implement the local equilibrium which could balance energy consumption locally.

2 Network Model and Energy Model

2.1 Network model

The network model used in this paper has the following properties:

1) All sensor nodes are randomly distributed;
2) BS is constant and abundant energy;
3) Sensor nodes have limited energy and can sense their own location information;
4) Sensor nodes have the power control capability.

2.2 Energy model

The nodes have a variety of operating modes, corresponding to different levels of power consumption. Based on the state of the communication, the power of nodes can be divided into four states: transmission state, receiving state,idle listening state, sleep state[8]. The working time and power consumption of each state are different. In this paper, we used energy consumption model, ignoring the node in the process of energy consumption, only considered the energy consumption of communication.

$E_{rf}(t)$: Communication energy consumption for the time t. $E_{send}(t)$: transmitting energy. $E_{rec}(t)$: receiving energy. $E_{idle}(t)$: Idle listening energy. $E_{sleep}(t)$: Sleep energy. So, Energy dissipation of CH node is calculated by:

$$E_{rf}(t)=E_{send}(t)+E_{rec}(t) \tag{1}$$

Every time transfer n bit in the process of information through the *d* distance. Transmitting energy of node is calculated by:

$$\begin{aligned} E_{send}(t) &= (n\times E_{send}+n\times\varepsilon\times d^2)\times\sum_{i=1}^{n}Ni_{send} \\ &= (n\times P_{send}\times\sum_{i=0}^{n}Ti_{send}+n\times\varepsilon\times d^2)\times\sum_{i=1}^{n}Ni_{send} \end{aligned} \tag{2}$$

P_{send} is transmission power. ε is energy amplification of distance unit to send information. *d* is the distance between the sending and receiving nodes. Ti_{send} is the time takes to send. Ni_{send} is the number of *i*-th transmission. Similarly, the energy receiving formula is:

$$E_{rec}(t)=n\times P_{rec}\times\sum_{i=0}^{n}Ti_{rec}\times\sum_{i=1}^{n}Ni_{rec} \tag{3}$$

2.3 Data fusion model

Algorithm use data fusion technology to reduce network transmission of data, save the network energy consumption and prolong the network life cycle. The cluster data fusion model[6] is: CHs receive each ordinary node sent data *n* bit, the CHs fusion data compression for *n* bits. Data fusion for energy consumption $E_{DA} = 5*l*10^{-12}$ J/bit.

3 ECPF Algorithm

In this study, first-order wireless transmission model is used to calculate the optimal clustering number, and it use fuzzy competition mechanism for cluster head node selection balance load and energy consumption on the whole; it uses energy management mechanism of low energy limit and high energy strengthen, to implement the local equilibrium which could balance energy consumption locally.

3.1 Optimal cluster distribution

CH nodes are responsible for most of the activities, such as intra-cluster communication, data fusion, need to consume more energy. The clustering algorithm mostly concern large messages and computational complexity of the algorithm. It is difficult to ensure the appropriate number of clusters, load balancing within the cluster and inter-cluster communications communication. The number of CH nodes excessive or little can cause performance degradation, and nodes are not easy to manage, can lead to increase in the number of communicate directly by nodes and base stations, increase telecommunications energy. If the number of clusters is small, the number of the CH nodes will excessive. So the management difficulty of the CH nodes will increase. Unbalanced load will cause the local node energy consumption too large and premature failure. It will affect the overall performance of the network. In order to save more energy, the protocol needs to minimize the number of active CHs during each round. This can be achieved by minimizing the following function. The number of clusters in the ECPF algorithm is constant, it is based on a rule: first-order model. And analysis of sub-optimal deployment cluster numbers through a random network.

In first-order model for the wireless transmission of energy model[3], *N* nodes deployed randomly in side length *r*, optimal clustering number is:

$$k = r \times \sqrt{N\varepsilon / (2(\varepsilon d^2 - E_{rf}))} \tag{4}$$

d is the average distance from CH node to BS. ε is the information transmission unit distance when the power amplification factor. $E_{rf}(t)$ is the energy consumed by the communication process.
Prove as follows:

Suppose the nodes of n are evenly divided k clusters in the monitoring area, each cluster is $\frac{n}{k-1}$, According to the apparent consumption model , E_{CH} includes receiving energy of receiving data within CH and send energy of transmitting data to BS.

The nodes in the monitored area are a uniform distribution of random. *S* represents the area of the simulation field, and the area of each is $\frac{r^2}{k}$. Distribution density function for the node region.

$$\rho(x, y) = \begin{cases} \frac{k}{r^2} & (x, y) \in S \\ 0 & (x, y) \notin S \end{cases} \tag{5}$$

With the square of the distance of cluster head nodes mathematical expectation

$$E\left[\mathrm{d}^2_{\mathrm{MEMtoCH}}\right] = E\left[x^2 + y^2\right] = \frac{r^2}{2\pi k} \tag{6}$$

Total energy consumption of network

$$\begin{aligned} E_{\mathrm{rf}} &= \mathrm{k} \times E_{cluster} = k \times E_{CH} + k \times (N / k - 1) \times E_{NEN} \\ &= k \times ((N / k - 1) \times lE_{elec} + lE_{elec} + l\varepsilon \mathrm{d}^2_{\mathrm{toBS}}) \\ &\quad + (N - \mathrm{k}) \times (lE_{elec} + l\varepsilon d_{toCH}) \\ &= l(2NE_{elec} - kE_{elec} + k\varepsilon \mathrm{d}^2_{\mathrm{toBS}} + (N-k) \times \varepsilon \mathrm{d}^2_{\mathrm{toCH}}) \end{aligned} \tag{7}$$

$$\begin{aligned} E_{total} &= l(2NE_{elec} - kE_{elec} + k\varepsilon \mathrm{d}^2_{\mathrm{toBS}} \\ &\quad + (N-k) \times \mathrm{r}^2 / 2k) \end{aligned} \tag{8}$$

Strives the first order derivative for formula.

$$k = r \times \sqrt{N\varepsilon / (2(\varepsilon d^2 - E_{rf}))} \tag{9}$$

The members of the energy consumption of nodes E_{NEN} including data-aware energy and the energy of send packets to CH. The optimal number of cluster head within a radius of 150m monitoring area in Table 1 below.

Target monitoring area of the square, BS located at the side, as shown in figure 1. According to the distance of the sensor node to the CH is divided into far, mid and near three parts

$$\begin{aligned} n_{\mathrm{near}} &= \alpha \times r \times \sqrt{n\varepsilon / (2(\varepsilon d^2 - E_{rf}))} \\ n_{\mathrm{mid}} &= \beta \times r \times \sqrt{n\varepsilon / (2(\varepsilon d^2 - E_{rf}))} \\ n_{\mathrm{far}} &= \gamma \times r \times \sqrt{n\varepsilon / (2(\varepsilon d^2 - E_{rf}))} \end{aligned} \tag{10}$$

α, β and γ is the probability of the weighting factor that divide the number of cluster head, and α+β+γ=1. Due to the more close to BS, the CH need to more undertake the task of data transmission, To avoid this undesirable feature, algorithm makes increase the number of the CHs that close to the base station, increase the weight of α appropriately.

Table 1. The number of CHs of different nodes

The number of nodes	Clustering number	The average distance of the CH to BS
100	10.08	59.30
200	14.38	58.80
300	17.75	58.39
400	20.76	57.64
500	23.38	57.20

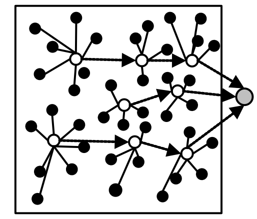

Figure 1. Cluster diagram of WSNs

3.2 Fuzzy competition mechanism

The factors that affected the selection of cluster head in wireless sensor networks are diverse, and fuzzy agreement select the cluster head node, according to these parameters reduce energy consumption that excellent performance and reasonable position[9].

According to the fuzzy reasoning method, input fuzzy sets that node residual energy, to the base station distance, elected as cluster head frequency. The fuzzy sets obtained through competitive fuzzy calculation. The larger the greater competitiveness node, become the ultimate possibility of the cluster head.

$$y_{energy-low} = \begin{cases} 1 & 0 < x < 0.3 \\ 4 - 10x & 0.3 < x < 0.35 \end{cases} \tag{11}$$

$$y_{energy-mid} = \begin{cases} 10x - 3 & 0.35 < x < 0.4 \\ 1 & 0.4 < x < 0.6 \\ 7 - 10x & 0.6 < x < 0.65 \end{cases} \tag{12}$$

$$y_{energy-high} = \begin{cases} 10x - 6 & 0.65 < x < 0.7 \\ 1 & 0.7 < x < 1 \end{cases} \tag{13}$$

Using Equation (11),(12),(13) were analyzed for residual energy, the remaining energy of the node quantized to three fuzzy subset {high, mid, low}, the greater the residual energy of sensor nodes, cluster head competition the better.

Similarly, the distance between the node to the base station is divided into three fuzzy subset {far, mid, near}. Node to the base station $D=\frac{D_{to\sin k}}{D_{\max}}$, D_{tosink} is distance from node to sink; $D_{\max}$ is nodes in the network to the maximum distance from the base station. The closer from the sink, the more pro-competitive for choose cluster head. The number of elected CH node times is divided into three fuzzy subset {high, mid, low}. The proportion of elected cluster head node $D=\frac{D_{\sin k}}{D_{all}}$, D_{Sink} is the number of elected cluster head node; D_{all} is the number of elected CH node times. Node cluster head elected fewer times, more favorable competitive take for cluster head.

The competitiveness of nodes is shown in the equation:

$$y=c_1 y_{\text{energy}}+c_2 y_{\text{distance}}+c_3 y_{num} \quad (14)$$

Among them, C1, C2 and C3 are the probability factors of remaining energy, compound distance, weight of the elected cluster heads when there are nodes in the cluster, and C1 + C2 + C3 = 1.

Nodes compute pro-competitive according to the fuzzy rules established. Then broadcast Send_Message (include node ID, location information and pro-competitive), and receive Send_Message from other nodes with their ability to compete for comparison. According to the distribution of CHs, if the competitiveness ranking of the node is higher than the number of CH, the cluster head node broadcasts H_Message (includes node ID, location information and competitiveness); if no information received, the node will broadcast H_Message as the cluster head. Node that receives broadcast, according to information from the CH node and choose from the recent CH, became member of the cluster.

3.3 Inter-cluster routing algorithm

CH nodes have Inter-cluster routing to the BS. Take a single-hop communication between members of the cluster nodes and cluster head, which makes a limited range of clusters. When the cluster head communicates with the sink node, the cluster head node that is farther away from the sink node consumes more energy. So the energy consumption of the cluster head node is too fast, which does not make use of the formation of a large-scale wireless sensor network.

To avoid the remote nodes directly communicate with BS by a lot of energy, the Inter-cluster routing algorithm should be use. Inter-cluster routing algorithm uses the algorithm to select the next hop node, a cluster head always selects the neighbor cluster head with the largest weight as its next hop relay node. We give a CH forwarding diagram shown in figure 2.

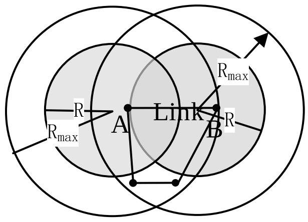

Figure 2. LMST link

LMST (Local minimum spanning tree) protocol ensures network performance. LMST algorithm is based on the neighboring graph, power control and minimum spanning tree structure topology control algorithm[10]. It builds a minimum spanning tree based on the competitiveness of the nodes.

The process of building a minimal spanning tree is as follows:

(1) Information exchange phase. The CH node sends H_Message message with maximum power (including node ID, distance information and competitiveness);

(2) Topology construction phase. The CH node obtains the local minimum spanning tree according to the received H_Message packet;

(3) Transmission phase. The CH node determines the next hop based on the spanning tree structure;

(4) When the CH node is less than a certain distance from the sink node, the cluster head node can directly send information to the BS.

3.4 Energy Management System

3.4.1 Low energy restriction

When the wireless sensor networks working for some time, some CH nodes that closer the BS, inter-cluster communication energy consumed, the more to die in wireless sensor networks.

Set the thresholds T_a and T_b as the low energy threshold of the next hop of inter-cluster communication, where $T_a > T_b$. If the node is selected as cluster head node, then take the appropriate interval forwarding mechanism based on the remaining energy value lies.

Specific procedures are as follows: Analyzing the interval where the residual energy of nodes; When the remaining energy is greater than T_a, the CH node will participate in the competition.

When the remaining energy is less than T_a and the remaining energy is greater than T_b, the cluster head node involvement competition. But only transfers the data to the other cluster head node, rejected as a next hop inter-cluster communication; If the residual energy is less than T_b, the cluster head node is not involved in the competition.

3.4.2 High energy enhancement

When the wireless sensor networks working for some time, some nodes high residual energy rarely elected cluster head in wireless sensor networks. These nodes far away to the sink, the probability of elected cluster head becomes small. Increase the probability of high-energy cluster head node election, can achieve energy balance run, extending the life cycle of sensor networks. Design of High T_c energy threshold for enhancing opportunities for high-energy sensors elected cluster head node. If the residual energy is greater than T_c, increase the competitiveness of the node.

The pseudo code of routing algorithm as shown in Table 2.

Table 2. Pseudo code of routing algorithm

ECPF Algorithm

ClusterHead Number Selection Algorithm
for every node vj:
compute all ClusterHead Number of vj according Eq(4)
compute ClusterHead Number of every Area according Eq(10)
ClusterHead Selection Algorithm
for every node vi
compute P of vi according Eq(7), Eq(8),Eq(9)
for every node vi in Area aj:
if Ta <yenergy<Tb
give up ClusterHead retransmission communication
end if
if yenergy<Ta
give up ClusterHead competition
end if
if P(vi) is large than each of its neigbor node's P
beClusterHead←TRUE
for every node vk in Area aj:
if beClusterHead Number > ClusterHead Number then
give up ClusterHead competition
end if
end if

4 Simulation

In this paper, we use the original simulation charts and parameters of LEACH and UCDP to demonstrate the correctness of our protocol. The input parameters are the initial energy of each sensor nodes, energy consumption rate, transmit, receive, and sense process costs, and send/receive buffer size. All three protocols are performed in parallel and have the same values for the parameters. Therefore, the simulation results will have a high degree of confidence.

4.1 Simulation environment

we use TinyOS2 simulation, experimental environment as follow: 100 sensor nodes distribute in the 150 m *150 m square area randomly. Table 3 shows the various parameters adopted in the experiment

Table 3. Values for network parameters

Parameter names	value	parametric description
α,β,γ	0.3,0.3,0.4	probability factor
E_{DA}	5 nJ/bit	Data fusion energy consumption
n	8bit	The length of the data
ε	10 pJ/(bit·m2)	Energy consumption coefficient of power amplification circuit
C1,C2,C3	energy =low 0.5,0.3,0.2 energy =mid 0.4,0.4,0.2 energy =high 0.3,0.4,0.3	probability factor

4.2 lifecycle

The life cycle of wireless sensor network is an important measure of load balance. Experiments with one hundred sensor nodes, intuitive shows the relationship between the number of network round and invalid nodes in figure 4.

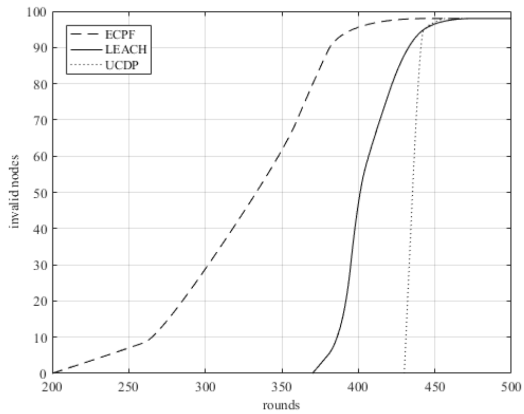

Figure 4. The curves of failure nodes number

LEACH algorithm first node death round number is 220 round, 392 rounds of all nodes were killed; UCDP algorithm from the 381th began to 430 rounds all death; ECPF algorithm from the 428th began to 448 rounds all death. LEACH algorithm by using single jump directly with the convergence of cluster head nodes communication, consumes a lot of energy to reduce the network life cycle, UCDP algorithm and ECPF algorithm used the multiple hops to communicate between clusters, reduce the energy consumption to extend network cycle. UCDP algorithm in the selection of cluster head take more than the round head of the election, increased consumption of energy, reduce the network life cycle.

4.3 energy-balancing

The formula for calculating the variance of the node death time:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n} \quad (15)$$

According to the formula, the variance of LEACH UCDP and ECPF was 1923.8 73.9 and 5.6. The greater of different of variance, the greater scattered of the node death. It can be seen from the data in this paper, ECPF algorithm makes the network energy consumption balance. This is due to the fuzzy competition mechanism and the energy management mechanism can effectively control the probability that the election of cluster head nodes, and low energy refusal to avoid node premature death and prolong the lifecycle of the node, Enhancement of high-energy node increase enough nodes utilization of resources, avoid the energy consumption distribution imbalance, realize the energy balance of wireless sensor network.

4.4 energy efficiency

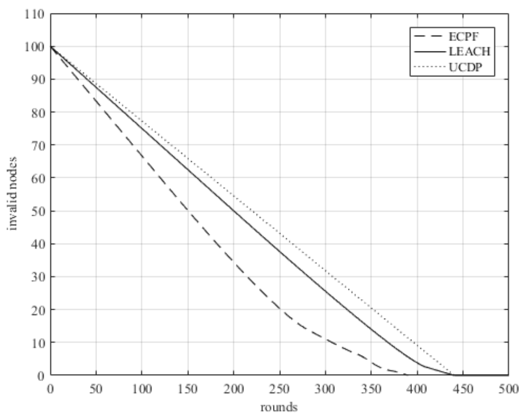

Figure 5. Remaining energy curves

Definition: when the network nodes 90% death, network stop working. Compare the curve of the three kinds of algorithm of network residual energy consumption over time, it can be seen from the figure 5, the energy consumption speed of ECPF algorithm is lower than other two algorithms. When LEACH algorithm failure in the network, the network energy of 4.3%; When UCDP algorithm failure in the network, the network energy of 1.1%; When ECPF algorithm failure in the network, the network energy of 0.4%. It is shown that ECPF algorithm has a longer survival time of network, can effectively balance energy consumption between nodes, prolong the life cycle of wireless sensor network.

5 Conclusions

Usually the feasible methods of saving energy for WSN are to improve the network routing protocols and to optimize the network topology structure. In this paper, we propose a new novel routing protocol for prolonging the network life of a wireless sensor node. Our algorithm divides the sensor field into different clusters and elects a node as the cluster head for each cluster and constructs the routing tree for data forwarding to the sink node. Each node within the cluster sends its data to the cluster head with single hop transmission and cluster heads aggregates the received data and transmits it to the base station via Routing tree. The simulation results show that the proposed method can effectively save the energy consumption of sensor nodes in a cluster.

References

1. H Taheri, P Neamatollahi, OM Younis, et al. Ad Hoc Networks. *An energy-aware distributed clustering protocol in wireless sensor networks using fuzzy logic*,**10**.7,1469-1481, (2012)
2. RK Kodali, SKA Venkata, S Bhandar, et al. International Conference on Advances in Computing Communications and Informatics IEEE. *Energy efficient m- level LEACH protocol*, 973-979, (2015)
3. SU Jin-Shu, WZ Guo, YU Chao-Long, et al. Chinese Journal of Computers. *Fault-Tolerance Clustering Algorithm with Load-Balance Aware in Wireless Sensor Network*, **37**.2, 445-456, (2014)
4. C Lai, W Jiang, Journal of Chinese Computer Systems. *Energy-balanced routing algorithm based on cluster heads expectation for wireless sensor networks*, **36**.12, 2685-2689, (2015)
5. H Lin, H Uster, IEEE/ACM Transactions on Networking. *Exact and Heuristic Algorithms for Data-Gathering Cluster-Based Wireless Sensor Network Design Problem* **22**.3,903-916, (2014)
6. YQ Sun, J Peng, T Liu, et al. Journal on Communications. *Uneven clustering routing protocol based on dynamic partition for wireless sensor network*, **35**.1,198-206, (2014)
7. HU Feng-Song, Q Xiao. Journal of Chinese Computer Systems. *Multi-hop Routing Algorithm of Energy-balancing Based on LEACH*,**35**.1,70-73, (2014)
8. C Sun, C Yang, S Fan, et al. Information & Control. *Energy Efficient Distributed Clustering Consensus Filtering Algorithm for Wireless Sensor Networks*, **44**.3, 379-384, (2015)
9. B Chen, H Yao, M Yang, et al. Chinese Journal of Sensors & Actuators. *A Inter-Cluster Multi-Hop*

Routing Protocol Improved Based on LEACH Protocol, **27**.3,373-377, (2014)
10. SEL Khediri, N Nasri, A Wei, et al. Procedia Computer Science. *A New Approach for Clustering in Wireless Sensors Networks Based on LEACH*. **32**.32,1180-1185, (2014)