ISSN: 2688-9323

A Novel Energy Balanced Unequal Clustering Routing Algorithm

Xiangcheng Wu
School of Computer, Central China Normal University, Wuhan 430079, China
363940934@qq.com

Abstract

With the continuous development of computer sensor technology, the Internet of things has been widely used in many fields. The terminal nodes, routing nodes and coordinator nodes in wireless sensor networks are distributed in all positions of the network. According to nodes are not rechargeable and the battery capacity is limited, some nodes in the network forward data packets. The battery energy is exhausted with the death of nodes. This will cause the network segmentation and even network paralysis, which greatly shortens the network survival time. In this paper, an improved energy balanced non-uniform clustering routing algorithm is proposed. The number of surviving nodes is considered when calculating the optimal number of cluster heads. In the final cluster head selection stage, the residual energy of nodes, the density of nodes and the location of nodes from the base station are comprehensively considered. In the construction phase of cluster, when ordinary nodes join the cluster head, they not only consider the residual energy of the cluster head and the distance from the cluster head, but also consider the distance between themselves and the base station, so as to further reduce the energy consumption of the whole network and prolong the network life cycle. At the same time, super cluster head is introduced in the data transmission stage to make the cluster far away from the base station forward by super cluster head, which reduces the communication cost of the cluster head far away from the base station and further balances the energy consumption between the cluster heads at different positions.

Keywords

Wireless Sensor Network; LEACH Protocol; Clustering; Energy Balance.

1. Introduction

With the continuous exploration of the unknown field, the acquisition, storage, transmission and utilization of information have been in-depth to all aspects of social production and life. Its wireless sensor network integrates sensor, wireless communication, distributed computing, micro electro mechanical system and other technologies, which greatly facilitates the way of information exchange between human and nature. Due to the characteristics of easy deployment, self-organization and intelligent information processing, wireless sensor networks have been widely used in military reconnaissance, medical and health, environmental monitoring, industrial control, building safety and modern agriculture. Different types of sensor nodes in wireless sensor networks are distributed in various locations in the network. Because most of the nodes cannot be charged or charging is difficult, in order to avoid premature death of nodes, energy problem has been the main limitation of wireless sensor networks. In the past decade, a lot of research work has been carried out to solve this challenge [1], [2], [3]. The base station is responsible for receiving and processing the data of nodes. In most cases, the base station (BS) is far away from the deployment area of network nodes. BS collects data periodically. [4] [5], [6], [7], [8] it is found that clustering networks with hierarchical topology provide longer lifetime than networks with direct data transmission.

ISSN: 2688-9323

2. Relation Works

For the problem that the selection probability of leach cluster heads is fixed and the proportion of fixed number of cluster heads is not optimal, the optimal number of cluster heads is calculated according to the network transmission energy consumption. Because the classical LEACH algorithm does not consider the energy consumption, the distance between nodes and base stations, and the density of nodes. Considering these factors, the algorithm can save the energy consumption of the network and maximize the network lifetime.

2.1. Network Model

WSNs are composed of two-level energy heterogeneous sensor nodes. Suppose that n sensor nodes are randomly distributed in the sensing area of L * L. A base station can supply unlimited power, and its sensor nodes assume that:

1. All nodes have the same performance except for different initial energy;
2. Each node can sense the current residual energy;
3. The node can calculate the distance between the sender and itself according to the received signal strength and adjust the transmission power in time;
4. Once the node is deployed, it will remain static and have a unique identity.

2.2. Energy Cosumption Model

The energy consumption of sensor nodes is mainly determined by communication energy consumption and data fusion energy consumption. In this paper, the classical wireless communication energy consumption model simplified by Heinzelman is mainly used.

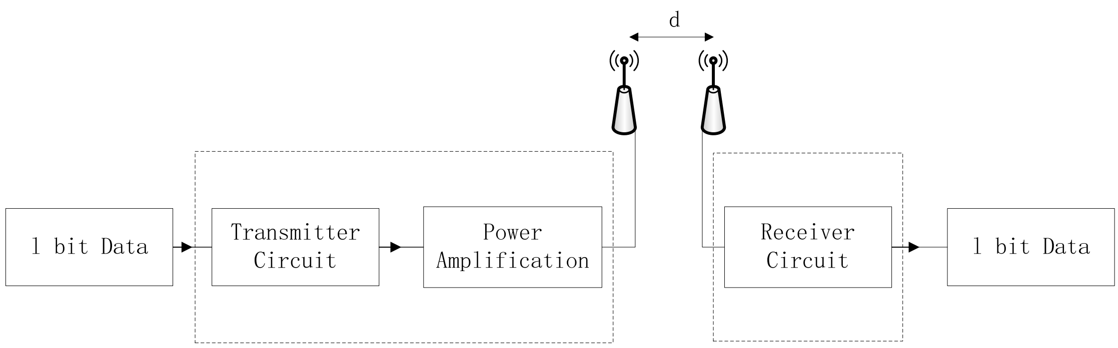

Fig. 1 Communication energy consumption model

The energy formula is

$$E_{T(l,d)} = E_{Telec}(l) = E_{Tamp}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\varepsilon_{mp}d^4, & d \geq d_0 \end{cases} \tag{1}$$

Where l is the length of the packet. E_{elec} is the energy consumption of the transmitting circuit and the receiving circuit in transmitting and receiving data per unit bit. ε_{fs} is the loss coefficient of transmission circuit in free space model, d is the distance of data transmission. It can be seen from formula (1) that the energy consumption is directly proportional to the transmission distance, and the longer the data transmission distance is, the greater the energy consumption is. ε_{mp} is the loss coefficient of the transmitting circuit in the multi-channel attenuation model, and the distance threshold is d_0 is only related to the properties of the transceiver itself, and the mathematical formula is shown in (2).

ISSN: 2688-9323

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{2}$$

2.3. Optimal Number of Cluster Heads

From the communication energy model, we can know that the clustering structure of wireless sensor network directly affects the communication energy consumption between nodes, and then determines the network generation cycle. If the number of clusters is too small, more nodes have to send data to the remote cluster. If there are too many clusters, less data fusion will reduce the advantages of network clustering, so too much or too little network clustering will lead to the increase of energy consumption. Therefore, selecting the most appropriate number of cluster heads can improve the energy consumption balance of the network and prolong the network lifetime. Theoretically, the number of optimal cluster heads can be calculated.

It is defined that N sensor nodes are randomly distributed in the L * L space, If there are K_{opt} cluster heads, then each cluster has N/K_{opt} has two nodes, including a cluster head and $\frac{N}{K_{opt}} - 1$ non cluster head nodes. The cluster head needs to receive the data of each common node and fuse it, and then send it to the base station. The energy consumed by each cluster head in each round is the sum of the energy received from common nodes, fused data and sent to base stations.

$$E_{CH} = \left(\frac{N}{K_{opt}} - 1\right). lE_{elec}^{Rx} + \frac{N}{K_{opt}}. lE_{DA} + (\, lE_{elec}^{Rx} + l\varepsilon_{mp} d_{toBS}^{n}) \tag{3}$$

Where l is the length of packets transmitted by each node, E_{DA} is the length of packets transmitted by each node represents the energy consumption of cluster head fusion data.

The members in the cluster need to send data to their cluster heads. According to the free space loss model, the energy consumption of each round is as follows:

$$E_{Node} = lE_{elec} + l\varepsilon_{fs} d_{toCH}^{2} \tag{4}$$

d_{toCH} is the distance from the member node to the cluster head, L^2/K_{opt} is the area of each cluster , density $\rho = K_{opt}/L^2$,The inner radius of the cluster is $R = L/\sqrt{\pi K_{opt}}$, $E(d_{toCH}^2)$ is the expected value of the square of the distance between the cluster members and the cluster head

$$E(d_{toCH}^2) = \iint r^2 . \rho(r,\theta) dr d\theta = \int_0^{2\pi} \int_0^{L/\sqrt{\pi K_{opt}}} r^3 dr d\theta = L^2/2\pi K_{opt} \tag{5}$$

Substituting the result of the above formula into formula (4), we can get that the energy consumption of a single member in each round of cluster is:

$$E_{Node} = lE_{elec} + l\varepsilon_{fs} L^2/2\pi K_{opt} \tag{6}$$

The sum of the energy consumed in a cluster is the sum of the energy of the cluster head and each member:

$$E_{cluster} = E_{CH} + \left(\frac{N}{K_{opt}} - 1\right). E_{Node} \tag{7}$$

Therefore, the total energy consumption of each round of network is:

$$E_{round} = K_{opt}.E_{cluster}=K_{opt}.(E_{CH} + \left(\frac{N}{K_{opt}} - 1\right).E_{Node})$$
$$=NlE_{DA} + 2NlE_{elec} + K_{opt}\left(l\varepsilon_{mp}d_{toBS}^{n} - lE_{elec}\right) + \frac{Nl\varepsilon_{fs}L^2}{2\pi K_{opt}} - \frac{l\varepsilon_{fs}L^2}{2\pi} \tag{8}$$

Find the bivalent derivative of K_{opt} as follows:

$$\frac{\partial E_{round}}{\partial K_{opt}} = l\varepsilon_{mp}d_{toBS}^{n} - lE_{elec} - \frac{Nl\varepsilon_{fs}L^2}{2\pi {K_{opt}}^2} \tag{9}$$

$$\frac{\partial^2 E_{round}}{\partial^2 K_{opt}} = \frac{Nl\varepsilon_{fs}L^2}{\pi {K_{opt}}^3} > 0 \tag{10}$$

As formula (10), when the second derivative is greater than 0 and the first derivative is equal to 0, there is an optimal solution:

$$K_{opt} = \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\varepsilon_{fs}L^2}{\varepsilon_{mp}d_{toBS}^{n}-E_{elec}}} \tag{11}$$

According to the different communication channel models of network nodes, the optimal number of cluster heads is obtained:

$$K_{opt} = \begin{cases} \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\varepsilon_{fs}L^2}{\varepsilon_{mp}d_{toBS}^{2}-E_{elec}}} & d_{toBs} \le d_0 \\ \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\varepsilon_{fs}L^2}{\varepsilon_{mp}d_{toBS}^{4}-E_{elec}}} & d_{toBs} > d_0 \end{cases} \tag{12}$$

3. Energy Balanced Clustering Algorithm

3.1. Optimal Number of Cluster Heads

To solve the problem that the ratio of cluster heads is fixed and the number of fixed cluster heads may not be optimal in LEACH protocol, the threshold formula is improved as T(n):

$$T(n) = \begin{cases} \frac{p'}{1-p'\times\left[r \bmod (1/p')\right]} & n \in G \\ 0 & otherwise \end{cases} \tag{13}$$

Among $p_{opt} = \frac{K_{opt}}{N_{alive}}$, N_{alive} represents the total number of surviving sensor nodes per round

Based on the analysis and reasoning in the previous section, K_{opt} slightly improved:

$$K_{opt} = \begin{cases} \sqrt{\frac{N_{alive}}{2\pi}}\sqrt{\frac{\varepsilon_{fs}L^2}{\varepsilon_{mp}d_{toBS}^{2} - E_{elec}}} & d_{toBs} \le d_0 \\ \sqrt{\frac{N_{alive}}{2\pi}}\sqrt{\frac{\varepsilon_{fs}L^2}{\varepsilon_{mp}d_{toBS}^{4} - E_{elec}}} & d_{toBs} > d_0 \end{cases}$$

ISSN: 2688-9323

Because of p_{opt} value is calculated according to the objective of minimizing the energy consumption. The dynamic p' value greatly optimizes the performance of the network, so the improved formula is more suitable for the dynamic changes of the network.

LEACH protocol does not consider the residual energy of nodes, based on the random number rand, the energy factor is introduced

$$rand(i) = rand.\frac{E_{ave}(r)}{E_{left}(r,i)} \tag{14}$$

$E_{left}(r, i)$ represents the residual energy of node i in round R, $E_{ave}(r)$ is the average energy of all nodes in the network in round r. The average energy is calculated as follows:

$$E_{ave}(r) = \frac{\sum_{i=1}^{N} E_{left}(r,i)}{N} \tag{15}$$

From the formula, it can be concluded that the more the residual energy of the node, the smaller the rand (i), the higher the probability of being elected as the cluster head.

Considering the uneven distribution of LEACH cluster heads, the density factor $\frac{\rho_i}{\rho_m}$ is introduced based on the threshold T (n).

$$T(n) = \begin{cases} \frac{p_{opt}}{1-p_{opt}\times\left[r \bmod (1/p_{opt})\right]} . R(r,i).\frac{\rho_i}{\rho_m} & n \in G \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

ρ_i is density of Node i,ρ_m is the max density of nodes of network:

$$\rho_i = \frac{m_i}{R^2} \tag{17}$$

$R = L/\sqrt{\pi K_{opt}} m_i$ is a circle with node i as the center and radius R. We can get the formula above, ρ_i larger is, the larger $\frac{\rho_i}{\rho_m}$ is . the higher the probability of the node being selected as the cluster head is.

From the above analysis, we can see that the improved threshold formula T(n) and node random number R(r, i). comprehensively consider the optimal number of cluster heads, node density and node residual energy factors, so that the distribution of cluster heads is more reasonable and the network energy is better balanced. The pseudo code of cluster head election is shown in the table.

Table 1. Cluster head election algorithm

Cluster Head Election Algorithm
1.Begin 2. Each node reports the energy and the number of neighbor nodes to the base station 3. Base station calculates the average network energy $E_{ave}(r)$ in this round,maximum density ρ_m, Optimal number of cluster heads Kopt 4. while (i<N) 5. According to the formula(16) Ti(n) 6. Generating random number rand(i) 7. if(Ti(n)< rand(i))
8. Node i joins set G 9. else 10.i become a normal nodei++ 11. Whileend
12. The selected cluster head broadcasts the selected messages, allocates TDMA time slots and waits for members to join
13.end

3.2. The Establishment of Cluster

According to the communication energy consumption model in wireless sensor networks, the transmission distance directly affects the energy consumption of nodes. DEEC algorithm only considers the distance between the node and the cluster head and the residual energy of the cluster head, which seems to balance the energy in the network, may reduce the energy consumption of the cluster head with less residual energy, but may increase the energy consumption of the whole network.

In order to verify the above conclusion, we calculate the energy consumption of sending data network to the base station when ordinary nodes select different cluster heads.

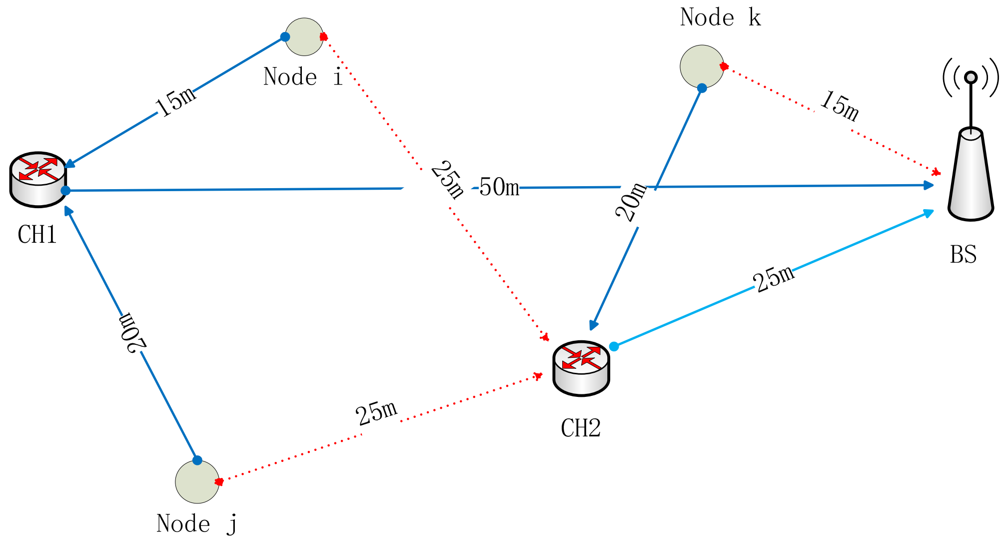

Fig 2. Schematic diagram of node joining cluster head

As shown in Fig 2 the solid line part of the figure, Node i and j add CH1 cluster heads according to the distance. The cluster head fuses the data from node i to j, and then send it to the base station. Although it may reduce the energy consumption of the cluster head with less residual energy, it may increase the energy consumption of the whole network.

When Node i selects cluster head CH1 to join, the total energy consumption of the network includes the sum of the energy consumed by nodes sending to cluster head, the energy consumed by cluster head fusion data and the energy consumed by nodes sending to base station

$$E_{CH1} = lE_{elec} + l \times 10 \times 15^2 + lE_{elec} + lE_{elec} + l \times 10 \times 25^2 = 3lE_{elec} + 27250l \quad (18)$$

If Node i chooses to join cluster head CH2, it includes the energy consumed by the node sending data, the energy consumed by the cluster head receiving data, and the energy consumed by the cluster head sending data to the base station. In the same case, if the node chooses cluster head to join, the network transmission length is the total energy consumed from data to base station

$$E_{CH2} = lE_{elec} + l \times 10 \times 25^2 + lE_{elec} + lE_{elec} + l \times 10 \times 50^2 = 3lE_{elec} + 12500l \quad (19)$$

By comparing the energy consumption of the two cases, it can be found that when a node joins a cluster reasonably, the remaining energy of the cluster head, the distance between the node and the cluster head and the base station should be considered. As shown by the dotted line in the figure, nodes should join cluster heads respectively, and transmit data from the corresponding cluster heads to the base station, which can greatly shorten the total distance of data transmission and reduce the total energy consumption of the network.

In order to further reduce the network energy consumption, when a node joins a cluster head, it first judges the distance between itself and the base station and the distance between the nearest cluster head. If the distance between the node and the base station is less than the distance between the nearest cluster heads, the Node k transmits data directly to the base station. In this way, the communication energy consumption of nodes is reduced, and the burden of cluster head near the base station is reduced. Otherwise, the node calculates the energy consumption weight value cost corresponding to different cluster heads $Cost_{CHk}$, select the maximum weight to add, and the weight calculation formula is as follows:

$$Cost_{CHk} = \alpha \frac{d_{min\_to\_CH}}{d_{i\_to\_CHk}} + \beta \frac{E_{CHk}}{E_{Left(r)}} + \gamma \frac{d_{i\_to\_BS}}{d_{max\_to\_BS}} \quad (20)$$

$d_{min\_to\_CH}$ is the shortest distance between node i and cluster head. $d_{i\_to\_CHk}$ is the distance between Node i and cluster head CHk; E_{CHk} is the residual energy of cluster head CHk; $E_{Left(r)}$ is the average residual energy of nodes in round r network; $d_{i\_to\_BS}$ is the distance between Node i and base station; $d_{max\_to\_BS}$ is the maximum distance between network node and base station; ∂, β, γ is the weight parameter and $\partial + \beta + \gamma = 1$.

Select cluster head to broadcast Head_ MSG message, ordinary nodes in the broadcast range receive the hHead_Msg message, the cluster head information is stored in the list, and the node calculates the cost corresponding to the received cluster head $Cost_{CHk}$. Node to $Cost_{CHk}$ cluster hair corresponding to chk maximum is added toJOIN_MSG message. The cluster head receives the join from the member node JOIN_MSG message, TDMA time slot is allocated for each member node. In each time slot, the member nodes transmit data to the cluster head in turn. If some nodes cannot receive the head all the timeHead_Msg message, these nodes establish contact with the nearest other nodes, and forward data through the established nodes.

Table 2. Algorithm of Cluster Establishment

Algorithm of Cluster Establishment
1. Cluster head determination 2. Cluster head broadcast information 3.if the node is closer to the base station 4. Direct communication with base station 5. else 6. Push $Cost_{CHk}$ corresponding to the received cluster head is calculated according to the formula into S_{cost}
7.if S_{cost} set is null 8. The node joins the nearest cluster head 9.else 10. Select the cluster head with the largest weight to join 11.endif
12. Node sends join message 13. Waiting for TDMA slot 14. Members send data to nodes 15. Stable transmission
16. The establishment of cluster is completed

3.3. Date Transmition

The new algorithm protocol adopts single hop data transmission within cluster and multi hop data transmission between clusters. According to the wireless transmission model, when the transmission distance exceeds the threshold, the energy consumption will increase sharply, which will lead to the rapid death of nodes far away from the base station. LEACH algorithm adopts the single hop mode, which is only suitable for small-scale networks. For a large-scale network, the lifetime of the network will be shorter if the distance between the candidate cluster head Chi and the base station is $d_{CHi} > d_0$. In this way, the formula $E = E_{elec} + l\varepsilon_{mp}d^4$, which consumes a lot of energy, leads to the premature death of nodes in the network. This paper proposes that the super cluster head is the only intermediate node for transmission. After the cluster is generated, it is necessary to select a cluster head from the cluster head set as the only super node to communicate with the BS node. In order to select the best root node, it is necessary to select the node with the largest cost function value as the super node in the cluster head set, and the cost function is:

$$p_{ch} = \alpha \frac{\overline{d_{CH\_to\_BS}}}{d_{i\_to\_BS}} + \beta \frac{E_i}{E_{avg}} + \gamma \frac{\overline{N_{men}}}{N_{i\_men}} \tag{21}$$

$\overline{d_{CH\_to\_BS}}$ is the average distance between cluster head node set and BS node,$d_{i\_to\_BS}$ is distance between Chi and BS,E_i is residual energy of CHi,E_{avg} is the average residual energy of the cluster head node set,$\overline{N_{men}}$ is average number of member nodes in cluster head node set,N_{i_men} is the number Chi members. ∂, β, γ is the weight parameter and $\partial + \beta + \gamma = 1$.

Setting ∂, β, and γ as different values can adjust the proportion of distance, energy, and the number of member nodes in the root node selection. The cluster head with the largestp_{ch} value is also the forwarding node. Although the generation of super cluster head has a certain impact on its own energy consumption, it plays a positive role in the energy balance of the whole network.

ISSN: 2688-9323

If a cluster head satisfies the following three conditions, it will have more chances to become a super cluster head: (1) the cluster head is close to the BS node; (2) the cluster head has more residual energy; (3) the number of member nodes of the cluster head is less. For the super cluster head, the communication energy consumption depends on the distance between the node and the BS node. Obviously, the closer the cluster head node is to the BS node, the more energy consumption can be reduced; at the same time, all other cluster heads have to communicate with the BS node through the root node, and the root node must have enough residual energy to undertake this task, so e-cluster is used. Finally, the less the number of member nodes of the root node, the less energy it consumes for intra cluster communication, so as to save energy for relaying messages sent by other cluster heads. Using the idea of selecting super cluster head can save energy consumption.

4. Summary

The algorithm proposed in this paper calculates the optimal number of cluster heads, selects cluster heads by integrating the residual energy, the geographical location of nodes and the density of nodes, to make the distribution of cluster heads more reasonable. In the establishment stage of cluster, considering the geographical location and energy to join the cluster, the cluster distribution in the network is non-uniform, the cluster far away from the base station is smaller, and the energy consumption is also smaller. Considering the loss model of energy transmission in data transmission, the super cluster head is proposed as the relay routing transmission, which makes the energy consumption more uniform after adding the relay.

References

[1] X. Gu, J. Yu, D. Yu, G. Wang, Y. Lv, ECDC: an energy and coverage-aware distributed clustering protocol for wireless sensor networks, Comp. Electr. Eng. 40 (2014) 384–398.

[2] V. Mhatre, C. Rosenberg, Homogeneous vs. heterogeneous clustered networks: a comparative study, Proc. IEEE, ICC 6 (2004) 3646–3651.

[3] C.Y. Chang, H.R. Chang, Energy aware node placement, topology control and MAC scheduling for wireless sensor networks, Comp. Netw. 52 (2008) 2189– 2204.

[4] S. Bandyopadhyay, E.J. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, IEEE INFOCOM, 2003, pp. 1713– 1723.

[5] S. Soro, W.B. Heinzelman, Prolonging the lifetime of wireless sensor networks via unequal clustering. Proc. of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS), USA, 2005, doi:10.1109/ IPDPS.2005.365.

[6] M. Ye, C. Li, G. Chen, J. Wu, EECS: An energy efficient clustering scheme in wireless sensor networks, IEEE International Conference on Performance, Computing, and Communications, 2005, pp. 535–540.

[7] S. Soro, W.B. Heinzelman, Cluster head election techniques for coverage preservation in wireless sensor networks, Ad Hoc Netw. 7 (2009) 955–972.

[8] W.B. Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy efficient communication protocol for wireless microsensor networks. Proc. 33rd Hawaii Intl Conference on System Sciences (HICSS'00), 2000, pp. 8020–8029.

[9] V. Mhatre, C. Rosenberg, Design guidelines for wireless sensor networks: communication, clustering and aggregation, Ad Hoc Netw. 2 (1) (2004) 45– 63.

[10] M. Perillo, Z. Cheng, W. Heinzelman, On the problem of unbalanced load distribution in wireless sensor networks. Proc. of IEEE GLOBECOM Workshops on Wireless Ad hoc and sensor networks, Dallas, TX, 2004, pp. 74–79.