

An Energy-Balancing Unequal Clustering Protocol for Wireless Sensor Networks

^{1,2}J. Yang and ¹D. Zhang

¹School of Electronics and Information Engineering,
Xi'an Jiaotong University, Xi'an 710049, China

²School of Mathematics and Computer Science, Ningxia University,
Yinchuan 750021, China

Abstract: In this study, a novel energy-balancing unequal clustering protocol (EB-UCP) for wireless sensor networks is presented. EB-UCP achieves a good performance in terms of lifetime by unequal clustering and balancing the energy load among all the nodes. An unequal clustering algorithm from probability view is employed to form clusters. Clusters closer to the sink node have smaller sizes than those farther away from the sink node, thus cluster heads closer to the sink node can preserve more energy for the purpose of inter-cluster data forwarding. In addition, the distribution of sensor nodes is deployed according to the energy-balancing layered algorithm and therefore the energy consumption in every layer is nearly equal. Finally, an energy-efficient data transmission mechanism on basis of the above is proposed. Simulation studies show that EB-UCP leads to more uniform energy dissipation and enlarges the lifetime of networks than EEUC (Energy Efficient Unequal Clustering) and LEACH (Low-Energy Adaptive Clustering Hierarchy).

Key words: Wireless sensor networks, unequal clustering, energy-balancing, network lifetime

INTRODUCTION

Recent advances in wireless communication and electronics have enabled the development of low-cost, low-power, multi-functional sensors that are small in size and able to communicate untethered within a short distance. In order to achieve energy efficiency and increase network scalability, sensor nodes can be organized into clusters. Communication within a cluster as well as communication between different clusters can be organized as a combination of one-hop and multi-hop communication. In one-hop communication, every sensor node can directly reach the sink node, while in multi-hop communication, nodes have limited transmission range and therefore are forced to route their data over several hops until the data reach the sink node. In both models, there is an unavoidable problem of unbalanced energy dissipation among different nodes, leading to the situation where some nodes lose energy at a higher rate and die much faster than others, possibly reducing sensing coverage and leading to network partitioning. For single-hop communication, the nodes furthest away from the sink node are the most critical nodes, while in multi-hop communication, the nodes closest to the sink node are burdened with a heavy relay traffic load and die first, i.e., the hot spot problem (Perillo *et al.*, 2005).

LEACH (Heinzelman *et al.*, 2002) is a well-known hierarchical architecture routing algorithm, which uses randomization to rotate the Cluster Heads (CHs) and achieves significant improvement compared to direct approach. However, this approach allows only one-hop clusters to be formed and the hot spot problem does not be considered. Li and Mohapatra (2005) presented a mathematical model to analyze hot spot problem and characterize energy hole problem in Wireless Sensor Networks (WSNs). Based on a traffic perspective, this analytical model examines the validity of several possible schemes aiming to mitigate or solve the energy hole problem. It is observed that in a uniformly distributed sensor networks, clustering deployment and data compression have a positive effect. Olariu and Stojmenovic (2006) presented the first theoretical work to analyze avoidance of the energy hole problem and it has been proven that the unbalanced energy depletion is intrinsic and no routing strategy can avoid under some specific conditions in this study. Perillo *et al.* (2004) formulated the transmission range distribution optimization problem and derived a linear programming model to obtain the maximizing WSNs lifetime. EEUC (Li *et al.*, 2005) use the competition radius to produce the unequal clustering, but this mechanism does not consider density of sensors deployed that might cause much

workload on CHs and too much communication overhead. Soro and Heinzelman (2005) proposed an unequal clustering size mode for network organization. However, it shows somewhat an unrealistic architecture since they assume that CHs with super power are deployed or move to the best position where minimizes energy consumption. It also considers only 2 layers, which means scalability is not in concern.

In this study, a novel Energy-Balancing Unequal Clustering Protocol (EB-UCP) for WSNs is proposed. In EB-UCP, the entire network has been divided into some layers and clusters from probability view. The number of clusters closer to the sink node is bigger than those farther away from sink node. However, the number of nodes in each cluster closer to the sink node is smaller than those farther away from sink node. Thus inner CHs can reduce the energy consumption within a cluster and preserve more energy for inter-cluster data relaying. In order to balance the energy dissipation in every layer, the reasonable ratios of distribution density among layers are derived according to energy-balancing principle. Therefore, the energy consumption in every layer is nearly equal after such a deployment and the hot spot problem can be effectively mitigated by this way. Finally, an energy-efficient data transmission mechanism is presented on basis of unequal clustering algorithm and energy-balancing layered algorithm. Theoretical analysis and simulation studies show that EB-UCP outperforms in terms of balancing energy dissipation among sensor nodes and enlarging the lifetime of networks than EEUC and LEACH.

EB-UCP: AN ENERGY-BALANCING UNEQUAL CLUSTERING PROTOCOL

Network model: All the nodes are assumed to be deployed in a circular area with a radius of R . The only sink is located at the center of the area. This represents a layered network where every layer contains a particular number of clusters, as shown in Fig. 1.

All the sensor nodes are homogeneous and have an ID number and the i -th corona is denoted by layer L_i . The maximum layer number is k and the width of each layer is r . For sake of simplicity, each sensor node is assumed to generate and send h bits of data per unit time. Nodes belonging to layer $\{L_i | i \neq k\}$ will forward both the data generated by them and the data generated by nodes from layers $\{L_j | (i+1) \leq j \leq k\}$. The nodes in the outermost, L_k , need not forward any data. The model for energy dissipation is taken from LEACH, where, for our multi-hop forwarding scheme we assume a free space propagation channel model. The energy spent for transmission of a h -bit packet over distance d is:

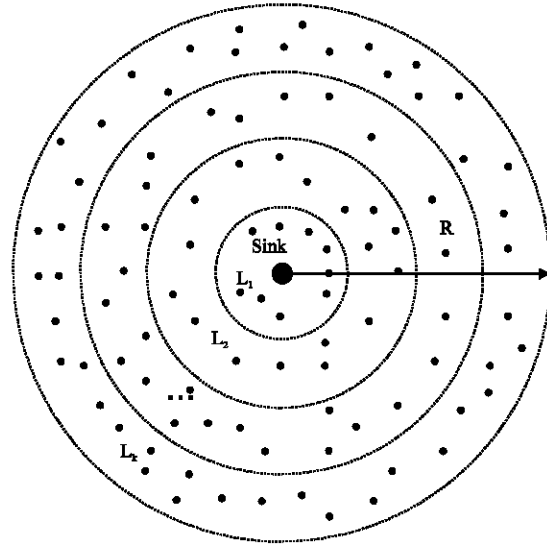


Fig. 1: A circular area consisting of coronas

$$E_{tx}(h, d) = h(e_1 + e_2 d^2) \quad (1)$$

And the energy spent on receiving a h -bit packet is:

$$E_{rx}(h, d) = h e_1 \quad (2)$$

The electronics energy, e_1 , depends on factors such as the digital coding, modulation, filtering and spreading of the signal. The amplifier energy, e_2 , depends on the distance to the receiver and the acceptable bit-error rate.

Unequal clustering algorithm: At network deployment stage, sink node broadcasts a hello message to all sensor nodes. Through the Received Signal Strength Indication (RSSI), each node can compute the approximate distance to the sink and the layer number of node belonging according to algorithm which is proposed by De *et al.* (2006). After the computation, each sensor sends its corresponding level number to the sink node. The sink knows the maximum layer number k and broadcasts it to all sensor nodes.

The key idea of clustering algorithm is to utilize clusters of unequal sizes to mitigate the hot spot problem. Layers are assigned different probabilities according to the distance from sink node and the layers closer to the sink node have greater probabilities than those farther away from sink node. Low probability indicates small number of CHs, which also means that each CH will occupy a huge cluster area with probably many non-CH nodes. Therefore, it is fair that closer sensor nodes get higher probability to be chosen as CHs in the perspective of balanced energy dissipation through the whole

network. In order to balance energy depletion among nodes, the CH selection in every layer is primarily based on the residual energy of tentative cluster heads. The node i 's probability of becoming a CH p_i will be computed as follows:

$$p_i = p_{\min} + \frac{k-j}{k-1} \times (p_{\max} - p_{\min}) \quad (3)$$

where, p_{\max} is the node's probability of becoming a CH in first layer, p_{\min} is the node's probability in the k th layer, j represents node i 's layer number.

In our algorithm, let $p_{\max} = 1$, That is, nodes in first layer all become CHs as they have direct transmission to sink node with small amount of energy and fully devote to heavy relaying traffic. The pseudocode for an arbitrary node i at the cluster head selecting stage is given in Fig. 2.

Each node in every layer become a tentative CH with the same probability p_i which is defined by Eq. 3. Other nodes keep sleeping until the CH selection stage ends. Each tentative CH maintains a set S_{CH} of its adjacent tentative CHs. Tentative head q is an adjacent node i if q is in i 's intra-cluster transmission range and two nodes are in the same layer. In lines 4-5 of Fig. 2, each tentative CH broadcast a CompeteCH_msg which contains its id, layer_id and residual energy. After the construction of S_{CH} has been finished in lines 7-8, each tentative CH checks its S_{CH} and makes a decision whether it can become a CH in lines 9-19. Before deciding what its role is going to be, i needs to know what each node q in its S_{CH} such that $q.E_{\text{current}} > i.E_{\text{current}}$ has decided for itself. In lines 10-12, once node i finds that its residual energy is more than all the nodes in its S_{CH} , it will win the competition and broadcast

```

1.   Compute  $P_i$ ;
2.   if (Rand (0,1) <  $P_i$ )
3.       TentativeHead = TRUE;
4.   if (TentativeHead == TRUE)
5.       CompeteCH_msg (id,  $E_{\text{current}}$ , layer_id);
6.   Rec_msg (On receiving from node q);
7.   if (i.layer_id == q.layer_id)
8.       Add node q to  $i.S_{CH}$ ;
9.   while (TentativeHead == TRUE){
10.      for ( $q \in i.S_{CH}$ )
11.          if ( $i.E_{\text{current}} > q.E_{\text{current}}$ )
12.              FinalCH_msg (id, layer_id);
13.   Rec_msg (FinalCH_msg) from q;
14.   if ( $q \in i.S_{CH}$ )
15.       QuitEle_msg (id, layer_id);
16.   Rec_msg (QuitEle_msg) from q;
17.   if ( $q \in i.S_{CH}$ )
18.       Remove q from  $i.S_{CH}$ ;
19.   }
```

Fig. 2: Cluster head competitive algorithm

a FinalCH_msg to inform its adjacent tentative CHs. In lines 13-15, if q is in i 's S_{CH} and i receives a FinalCH_msg from q , i will give up the competition immediately and inform all nodes in its S_{CH} by broadcast a QuitEle_msg. In lines 16-19, if i receives a QuitEle_msg from q and q belongs to $i.S_{CH}$, i will remove q from its S_{CH} .

After CHs have been selected, each CH broadcasts a CH_ADV_msg across the network field. Each ordinary node chooses its closest CH with the largest received signal strength and then informs the CH by sending a JOIN_CLUSTER_msg.

Energy-balancing layered algorithm: The energy balancing in every layer is considered in this subsection. In order to measure the data aggregation by CH within a cluster, all CHs are assumed to equally compress the data, where this efficiency is expressed by the aggregation coefficient α and energy dissipation for unit data aggregation is e_3 . If let N_i and ρ_i denote the number and distribution density of layer L_i , respectively. The total energy consumption in each layer can be calculated. According to the above assumptions and network model, the outermost layer L_k only needs to forward data generated by themselves. Thus energy dissipation per unit time of each CH in outermost layer L_k is:

$$E_{\text{ch}_k} = (1/p_{\min} - 1)h e_1 + 1/p_{\min} h e_3 + 1/p_{\min} \alpha h (e_1 + e_2 r^2) \quad (4)$$

Equation 4 shows that the smaller α is, the higher efficiency of data aggregation by CH is. The energy dissipation per unit time of each non-CH in outermost layer L_k is:

$$E_{\text{nonch}_k} = h(e_1 + e_2 r^2) \quad (5)$$

Thus, the energy consumption per unit time of all nodes in L_k is:

$$E_k = p_{\min} N_k E_{\text{ch}_k} + (1 - p_{\min}) N_k E_{\text{nonch}_k} \\ = N_k h \{ [2(1 - p_{\min}) + \alpha] e_1 + (1 - p_{\min} + \alpha) r^2 e_2 + e_3 \} \quad (6)$$

All the CHs in other layers have to transmit data generated by them as well as data originated from outer layers. Therefore, the energy consumption per unit time of all nodes in L_i ($1 \leq i \leq k$) is:

$$E_i = N_i h \{ [2(1 - p_i) + \alpha] e_1 + (1 - p_i + \alpha) r^2 e_2 + e_3 \} \\ + \sum_{l=i+1}^k \alpha h N_l (2e_1 + e_2 r^2) \quad (7)$$

Ideally, when the energy consumption of all the layers is equal, the network lifetime and the energy efficiency are maximized, That is,

$$E_1 = E_2 = E_3 \dots = E_k \quad (8)$$

If let

$$E_k = E_{k-1} \quad (9)$$

Then we can obtain:

$$\frac{N_{k-1}}{N_k} = \frac{[2(1-p_{\min})-\alpha]e_1 + (1-p_{\min})r^2e_2 + e_3}{[2(1-p_{k-1})+\alpha]e_1 + (1-p_{k-1}+\alpha)r^2e_2 + e_3} \quad (10)$$

Let S_i stand for the area of layer L_i . Then the node density is given by

$$\rho_i = \frac{N_i}{S_i} = \frac{N_i}{\pi(2i-1)r} \quad (11)$$

By Eq. 10 and 11, the ratio between the node densities of L_i and L_{i-1} is obtained as:

$$\frac{\rho_{k-1}}{\rho_k} = \frac{2k-1}{2k-3} \cdot \frac{[2(1-p_{\min})-\alpha]e_1 + (1-p_{\min})r^2e_2 + e_3}{[2(1-p_{k-1})+\alpha]e_1 + (1-p_{k-1}+\alpha)r^2e_2 + e_3} \quad (12)$$

Thus, if distribution densities of nodes between L_k and L_{k-1} is deployed by Eq. 12, energy consumption of two layers is balanced. If let

$$E_k = E_{k-2} \quad (13)$$

$$\frac{N_{k-1}}{N_k} = m \quad (14)$$

Then,

$$\frac{N_{k-2}}{N_k} = \frac{[2(1-p_{\min})-\alpha-2\alpha m]e_1 + (1-p_{\min}-\alpha m)r^2e_2 + e_3}{[2(1-p_{k-2})+\alpha]e_1 + (1-p_{k-2}+\alpha)r^2e_2 + e_3} \quad (15)$$

$$\frac{\rho_{k-2}}{\rho_k} = \frac{2k-1}{2k-5} \times \frac{[2(1-p_{\min})-\alpha-2\alpha m]e_1 + (1-p_{\min}-\alpha m)r^2e_2 + e_3}{[2(1-p_{k-2})+\alpha]e_1 + (1-p_{k-2}+\alpha)r^2e_2 + e_3} \quad (16)$$

After iterative calculating, we can obtain the ratios of $\rho_k, \rho_{k-1}, \dots, \rho_i \dots, \rho_2, \rho_1$. The nodes' distribution densities of whole network under energy-balancing is achieved when the ρ_k is determined. Therefore, the more uniform energy dissipation and the longer lifetime of networks are obtained by this energy-balancing layered algorithm.

```

1.   FinalCH_msg (id,layer_id);
2.   Rec_msg (FinalCH_msg) from q;
3.   if (q.layer_id == i.layer_id+1)
4.       Add node q to i.Parent;
5.   else if (q.layer_id == i.layer_id-1)
6.       Add node q to i.Children;
7.   if (Rec_msg (InqResEnergy) from Parent)
8.       Send_msg (id,i.ResEnergy,layer_id);
9.   if (Rec_msg (AckInqResEnergy)from Children)
10.      Update (i.Children.ResEnergy);
11.  while (Data_arrival == TRUE){
12.      if (Rec_msg (Forwarddata)from Parent){
13.          u = MaxResEnergy (i.Children);
14.          Send_data (u, Forwardingdata);
15.      }else{
16.          u = MaxResEnergy (i.Children);
17.          Send_data (u, i.data);
18.      }
19.  }
    
```

Fig. 3: Inter-cluster data transmission algorithm

Data transmission mechanism: The data transmission mechanism of EB-UCP is based on unequal clustering algorithm and energy-balancing layered algorithm. Firstly, sensor nodes can be deployed according to the distribution density based on the results derived in the previous section. Then, nodes use the unequal clustering algorithm to compete CH. When CHs deliver their data to the sink node, each CH first aggregates the data from its cluster members and then sends the packets to the sink node via a multihop path through other intermediate CHs. Thus the data transmission of whole network includes inter-cluster and intra-cluster data transmission. The organization of intra-cluster data transmission is similar to LEACH after clusters have been set up, so it was omitted in this section. The pseudocode of the inter-cluster algorithm for CHi is presented in Fig. 3. In order to gain balanced energy depletion among the CHs, the CH i selects one relay CH with maximum energy resource. When selecting the relay CH with maximum residual energy, the CHi has to exchange energy information with all candidate relay CHs in lines 7-10 of Fig. 3. If there is more than one candidate with the same maximum residual energy, choose one of them randomly. After the CH i selects the relay CH, it forwards data of its own as well as those from the upper layer CHs in lines 11-19. And the selected relay CH will repeat this process until the data arrive at sink node.

Analysis of EB-UCP complexity: The EB-UCP is composed of CH competitive algorithm, energy-balancing layered algorithm and inter-cluster data transmission mechanism. Since energy-balancing layered algorithm can be calculated by sink node, we only consider the

clustering algorithm and inter-cluster data transmission algorithm. According to previous sections, both algorithm are message-driven, thus we discuss its message complexity. For CH competitive algorithm, $N_i \times p_i$ tentative CHs are produced in every layer and each of them broadcasts a CompeteCH_msg. Then each of them makes a decision by broadcasting a FinalCH_msg to act as a final CH, or a QuitEle_msg to act as an ordinary node. Suppose s CHs are selected, they send out s CH_ADV_msg and then $(N_i - s)$ ordinary nodes transmit $(N_i - s)$ JOIN_CLUSTER_msg. Thus the messages add up to $2N_i \times p_i + s + N_i - s = (2p_i + 1)N_i$, i.e., $O(N_i)$. The inter-cluster data transmission mechanism mainly includes energy inquiring message and data forwarding message. The message complexity of both is $O(N_i^2)$. Therefore, the complexity of EB-UCP is $O(N_i^2)$.

SIMULATION RESULTS

Here, the performance of the EB-UCP protocol with simulations is evaluated. Lifetime is the criterion for evaluating the performance of WSNs. In the simulation, we measure the lifetime in terms of the round when the first node dies (Chang and Tassiulas, 2004). For simplicity, an ideal MAC layer and error-free communication links are assumed. The simulation parameters for our proposed protocol are given in Table 1. The experiments described in this part have been conducted with the aid of a well known simulation tool, Ns2. The parameters and characteristics of our architecture are first observed and then the performance of proposed idea is examined comparing with the EEUC and LEACH.

The value of k verification: The suitable number of layers is important to the energy consumption. Too big or too small value of k will decrease the lifetime of network. From Fig. 4, the value of k is 7 which is suitable for the range of 6-8.

Impact of node number of the first layer: In order to investigate the impact of node number of the first layer, its range is varied from 8 to 52 and the aggregation coefficient ($\alpha = 0.1, 0.5, 1$) is adopted as shown in Fig. 5. Figure 5 shows that the rounds of lifetime increase almost proportionally for EB-UCP as the number of nodes in the first layer increases. With more CHs closer to the sink node, the lifetime of the network improves significantly. In addition, when CH forward 10% of the cluster load ($\alpha = 0.1$), the maximum rounds of lifetime is increased by 1-2 times than $\alpha = 0.5$.

Table 1: Parameters of simulations

Parameter	Value
Network size	R=200 m
Sink location	(0,0)
Data packet size	500 (bytes)
Initial energy	2 (J)
Number of nodes	500-2000
e_1	50 (nJ b ⁻¹)
e_2	10 (pJb/m ²)
e_3	5 (nJ/b/signal)
k	7
p_{max}	1
P_{min}	0.05

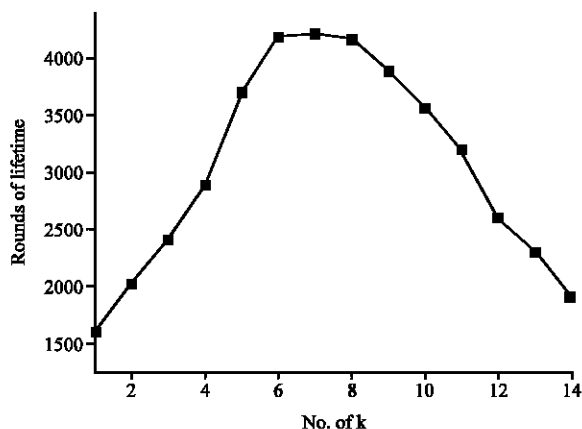


Fig. 4: The value of k verification

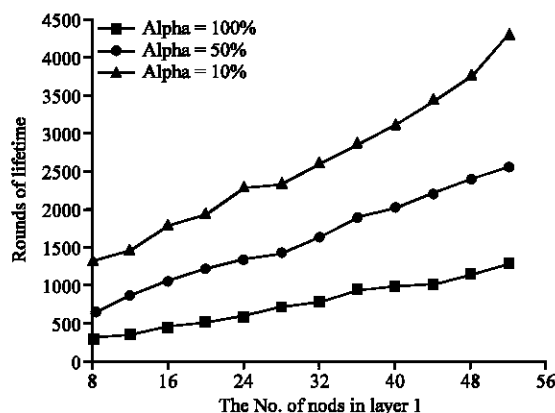


Fig. 5: Impact of node number of the first layer

Energy efficiency: In this subsection, the performance of EB-UCP, EEUC and LEACH in terms of network lifetime are compared. From Fig. 6a-d, the ratio of alive nodes of EB-UCP is much bigger than LEACH and EEUC when CHs forward different ratio of the cluster load ($\alpha = 0.1, 0.2, 0.7, 1.0$). Compared with the EB-UCP and EEUC, CHs in LEACH sends their packets to the sink node via single hop and the energy consumption is much higher. In EB-UCP and EEUC, CHs transmit their data to the sink node via multihop, thus a considerable amount of energy is

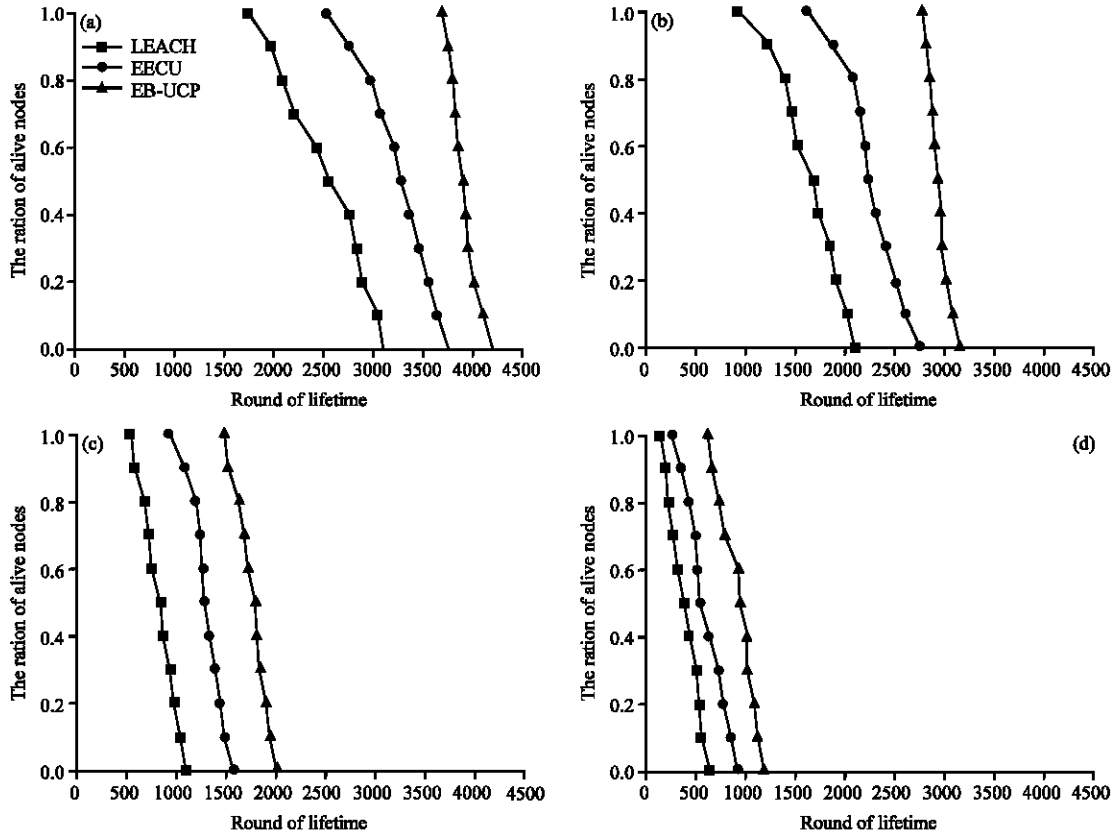


Fig. 6: Performance comparison, (a) $\alpha = 0.1$, (b) $\alpha = 0.3$, (c) $\alpha = 0.7$ and (d) $\alpha = 1$

saved. By node distribution of energy balancing in each layer, the energy dissipation of EB-UCP is more uniform than EECU and LEACH. On average, EB-UCP achieves a 33-54% improvement in network lifetime against EECU and LEACH depending on the aggregation efficiency.

CONCLUSION

The hot spot problem appears when employing the multihop routing in a clustering approach. To address the problem, a novel energy-balancing unequal clustering protocol for WSNs has been introduced. An unequal clustering algorithm from probability view is presented and clusters closer to the sink node have smaller sizes than those farther away from the sink node. Thus CHs closer to the sink node can preserve some energy for the purpose of inter-cluster data forwarding. Then, the balance of energy consumption among layers is analyzed and an energy-balancing layered algorithm is proposed. Based on the above, the data transmission mechanism is discussed. By unequal clustering and node distribution of energy balancing, EB-UCP ensures more balanced energy consumption as well as network lifetime than previous

researches. Simulation results show that EB-UCP significantly improves the network lifetime over LEACH and EECU.

To ease the analysis of EB-UCP, several simplifying assumptions that have been made in this study will be addressed in further research. For example, the effect of errors and collisions and event-based networks where the data generation rate at each node is a function of phenomena in the environment rather than constant at each node will be considered in EB-UCP.

ACKNOWLEDGMENTS

This research was financially supported by National Natural Science Fund of China with grant No. 60472074 and China's Next Generation Internet Demonstration Project with grant No. CNGI-04-2-1D.

REFERENCES

Chang, J. and L. Tassiulas, 2004. Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Trans. Network.*, 12: 609-619.

- De, S., A. Caruso, T. Chaira and S. Chessa, 2006. Bounds on hop distance in greedy routing approach in wireless ad hoc networks. *Inter. J. Wirel. Mobile Comput.*, 1: 131-140.
- Heinzelman, W.B., A.P. Chandrakasan and H. Balakrishnan, 2002. An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans. Wireless Commun.*, 1: 660-670.
- Li, J. and P. Mohapatra, 2005. An analytical model for the energy hole problem in many-to-one sensor networks. *IEEE 62nd Semiannual Vehicular Technology Conference*, September 25-28, IEEE Press, Dallas, USA., pp: 2721-2725.
- Li, C., M. Ye, G. Chen and J. Wu, 2005. An energy-efficient unequal clustering mechanism for wireless sensor networks. *IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, November 7-10, IEEE Press, Washington, USA., pp: 1-8.
- Olariu, S. and I. Stojmenovic, 2006. Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. *Proceedings on 25th IEEE International Conference on Computer Communications*, April 23-29, IEEE Communications Society, Barcelona, Spain, pp: 1-12.
- Perillo, M., C. Zhao and W. Heinzelman, 2004. On the problem of unbalanced load distribution in wireless sensor networks. *IEEE Global Telecommunications Conference Workshops 2004*, 29 November-December 3, IEEE Press, Dallas, USA., pp: 74-79.
- Perillo, M., Z. Cheng and W. Heinzelman, 2005. An analysis of strategies for mitigating the sensor network hot spot problem. *Proceedings of the 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05)*, July 17-21, IEEE Computer Society, San Diego, USA., pp: 474-478.
- Soro, S. and W.B. Heinzelman, 2005. Prolonging the lifetime of wireless sensor networks via unequal clustering. *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05)*, April 4-8, IEEE Computer Society Press, San Francisco, USA., pp: 236-240.