http://ansinet.com/itj



ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL



Asian Network for Scientific Information 308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

EEHCA: An Energy-Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks

Guan Xin, Wu Hua Yang and Bi DeGang School of Information Science and Technology, HeiLong Jiang University, Harbin, China

Abstract: In this study, a hierarchical clustering algorithm for long-lived sensor network is proposed. EEHCA (an energy-efficient hierarchical clustering algorithm for wireless sensor networks) achieves a good performance in terms of lifetime by minimizing energy consumption for communication and balancing the energy load among all the nodes. EEHCA adopts a new method for cluster head election, which can avoid the frequent election of cluster head. In order to improve the performance of fault-tolerance, we introduce the concept of backup cluster head. Furthermore, when nodes finished the communication within their own clusters and the cluster heads have finished the data aggregation, the head clusters will transmit aggregated data to the sink node by a special multi-hop mode. Simulation results have shown that EEHCA has the better performance than LEACH (low energy adaptive clustering hierarchy) and HEED (Hybrid Energy-Efficient Distributed clustering) in terms of network lifetime.

Key words: Sensor networks, clustering algorithm, energy-efficient

INTRODUCTION

Continued advances of MEMS and wireless communication technologies have enabled the deployment of large scale wireless sensor networks (WSNs) (Estrin et al., 2001). The potential applications of WSNs are highly varied, such as environmental monitoring, target tracking and military (Akyildiz et al., 2002). Sensors in such a network are equipped with sensing, data processing and radio transmission units while the power is highly limited. Due to the sensors' limited power, innovative techniques that improve energy efficiency to prolong the network lifetime are highly required.

Data gathering is a common but critical operation in many applications of WSNs, where data aggregation and hierarchical mechanism are commonly used techniques. Data aggregation can eliminate the data redundancy and reduce the communication load (Krishnamachari *et al.*, 2002). Hierarchical (clustering) mechanisms are especially effective in increasing network scalability and reducing data latency, which have been extensively exploited. LEACH (Heinzelman *et al.*, 2002) is the first clustering protocol, proposes a two-phase mechanism based on single-hop communication. The plain node transmits the data to the corresponding cluster head and the cluster head transmits the aggregated data to the sink node. HEED (Younis and Fahmy, 2004) selects cluster heads through O(1) time iteration according to some metric and

adopts the mutil-hop communication to further reduce the energy consumption. During the execution of HEED, the worst case occurs when a node has a very low $E_{residual}$ (current residual energy in the node). This node will start executing HEED with CH_{prob} (the probability of one node to be cluster head) set to P_{min} (a certain threshold) However, CH_{prob} doubles in every step and the phrase of the protocol terminates one iteration after Ch_{prob} reaches 1. Therefore, $2^{Niter-1*}$ $p_{min} \ge 1$ and hence

$$N_{tter} \le \left\lceil \log_2 \frac{1}{P_{min}} \right\rceil + 1$$

Therefore, Niter ©(1). PEGASIS (Lindsey and Raghavendra, 2002) improves the performance of LEACH and prolongs the network lifetime greatly with a chain topology. But the delay is significant although the energy is saved. Akkaya and Younis (2005) surveys recent routing protocols for sensor networks and presents a classification for the various pursued. The three main categories explored are data-centric, hierarchical and location-based. And each category has its own advantage and shortage. There are some other related works which efficiently use energy through clustering.

In this study, we propose and evaluate an energy efficient hierarchical clustering algorithm (EEHCA) for data gathering applications in WSNs. In the cluster heads determination phase, we raise a new concept of backup cluster heads. That is to say each cluster have two cluster heads, one is primary cluster head which is working normally, the other is backup cluster head which is

prepared to be a primary cluster head when the primary cluster head is destroyed or run out of energy. The determination of cluster head is based on the distances between nodes and the centre of the clusters. We adopt the multi-hop mode to balance the load among cluster heads during the transmission between sink node and cluster heads. EEHCA is more energy efficient and the simulation results show that it prolongs the network lifetime than the LEACH and HEED.

EEHCA: A NEW HIERARCHICAL CLUSTERING ALGORITHM

Network model: In order to describe the algorithm more clearly, we make the following assumptions:

- N nodes are uniformly dispersed within a square field N.
- All nodes and sink are stationary after deployment.
- The communication in the cluster is based on the single hop.
- Communication is symmetric and a sensor can compute the approximate distance based on the received signal strength if the transmission power is given.
- All nodes are location-aware and equipped with directional antenna.
- All nodes are of equal significance.

We use a simplified model shown in Heinzelman *et al.* (2002) for the radio hardware energy dissipation as follows. To transmit an L-bit data to a distance d, the radio expands:

$$E_{\text{TX}} = \begin{cases} L * E_{\text{dec}} + L * \epsilon_{\text{fs}} * d^2 & d < d_{\text{crossover}} \\ L * E_{\text{dec}} + L * \epsilon_{\text{mp}} * d^4 & d \ge d_{\text{crossover}} \end{cases}$$
(1)

The electronics energy, E_{elec} , depends on factors such as the digital coding, modulation, filtering and spreading of the signal, whereas the amplifier energy, $\epsilon_{\text{fi}} * d^2$ or $\epsilon_{\text{mp}} * d^4$, depends on the distance to the receiver and the acceptable bit-error rate. When receiving the data, the radio expends: $E_{\text{RX}} = L * E_{\text{elec}}$. Additionally, the operation of data aggregation consumes the energy as E_{DA} .

In Table 1, we give the meaning of the notations which will have readers understand the working process of the EEHCA much easily.

Obtaining the location information of nodes: Before the network starts to work, sink node broadcast a hello message to all the nodes. According to the signal it

Table 1: Meaning of the Notations

Notations	Meaning
r_c	Optimal radius of cluster
L	Length of data
d_{loch}	Distance between common node and cluster head
A	Area of sense zone
$\mathrm{E}\left[\mathbf{d_{toch}}^{2}\right]$	Expect of distance between common nodes and cluster head
N	Amount of sensor nodes

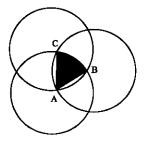


Fig. 1: Maximum overlap angle among 3 neighboring clusters

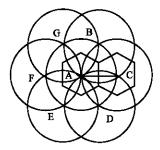


Fig. 2: Minimum area of cluster

received, nodes compute the approximate distance and obtained the angle information through the directional antenna. And then, all nodes transmitted the distance, angle and their own ID to the sink node. Because we have known the sink's coordinate, we can compute the coordinates relative to the sink according to the distance and angle information which the nodes transmitted.

The computation of optimal amount of cluster heads:

According to Choudhury and Kravets (2004) and Yang et al. (2006), it has been proved that among 3 unconnected neighbor nodes, the largest overlap angle θ is $\pi/3$ and it is shown in Fig. 1. Obviously, each cluster has the minimum zone, if in the network any 3 neighboring cluster heads all have the largest overlapping angle. In Fig. 2, the six neighbor nodes (B, C, D, E, F, G) of cluster head A constructed a hexagon, the distance between any two nodes is $r_c+\epsilon$. Here, $\epsilon\to 0$ and r_c denote radius of the cluster (Liu et al., 2007). Obviously, if the distance between any two clusters less than or equal to r_c cluster head A will be covered by the six neighbor nodes. In fact, the cluster which is denoted by cluster head A is a hexagon whose side length is $r_c/\sqrt{3}$ and the cluster area is

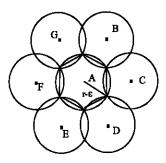


Fig. 3: Maximum area of clusters

 $\sqrt{3}$ * $r_c/2$ and then, we can conclude that the maximum cluster heads amount are $K_{max} = \lceil 2||A||/\sqrt{3}r_c^2 \rceil$.

As is shown in Fig. 3, the six neighbor nodes (B, C, D, E, F, G) of cluster head A construct a hexagon, the distance between any two nodes are $\sqrt{3}(r_c - \epsilon)$, $\epsilon \to 0$. If the distance of any two nodes is longer than $\sqrt{3}\,r_c$, sense zone will generate blind spot. Figure 3 shows that the status minimum overlapping area among neighboring clusters. Cluster head A denotes a hexagon whose side length is $\sqrt{3}r_c$ and the area of the hexagon is $3\sqrt{3}*r_c^2/2$. Therefore, the minimum amount of cluster heads is $K_{min} = \lceil 2\|A\|/\sqrt{3}r_c^{-2} \rceil$.

According to the description above, for any of a cluster C_i , the actual area of the cluster accords with the inequality: $\sqrt{3r_c^2}/2 \le |C_i|| \le 3\sqrt{3r_c^2}/2$. The maximum cluster area and the minimum cluster area are in direct proportion to r_c^2 , so we can conclude that the actual area is also in direct proportion to r_c^2 . Through the value of maximum and minimum cluster heads, we can confirm that the actual amount of cluster heads (K_{exp}) in the whole network is $K_{exp} = \lceil (4||A||)/(3\sqrt{3}r_c^2)^2 \rceil$.

Confirmation of optimal cluster radius: In the sense zone, the energy consumption model of common nodes and cluster heads are different. So we will discuss the two sorts of nodes separately.

To the common nodes, at their own TDMA slots, they always need to transmit the sensed data to their own cluster heads. Here, we assume the data length is a constant L. Then based on the wireless communication model, we can know the common node's energy consumption on transmitting the data.

$$E_{\text{member}} = L * E_{\text{elec}} + L * \epsilon_{\text{fs}} d_{\text{toch}}^{2}$$
 (2)

We assume that the actual cluster is a circularity whose area is

$$[(4||A||)/(3\sqrt{3}r_c^2)].$$

Then we have the equation:

$$E\left[d_{toch}^{2}\right] = \rho \int_{0}^{2\pi} \int_{0}^{\frac{\sqrt{2}\sqrt{2}}{2\sqrt{\pi}}} r_{c}^{3} d_{rc} d\theta \tag{3}$$

To a circularity whose area is

$$\left\lceil (4||A||)/(3\sqrt{3}r_c^2)\right\rceil,$$

the nodes in this circularity are distributed randomly. To the cluster head, the location of nodes is random too. So we can only evaluate the distance between the common nodes and the cluster heads. Since nodes are distributed randomly, we can obtain

$$\rho = 4/3\sqrt{3}re^2$$

and then we have the Eq. 4.

$$E\left[d_{loch}^{2}\right] = \frac{9}{8\sqrt{3}\pi}r_{c}^{2} \tag{4}$$

To the cluster heads, energy consumption is composed with four parts: (1) energy consumption of receiving and aggregating the data sensed by common nodes, (2) energy consumption of receiving the data sent by remote cluster heads, (3) energy consumption of relaying the data of remote cluster heads and (4) energy consumption of transmitting their own data.

From the description above, the distance between each circle's centre and its nearest neighbor circle's centre is less than 2r_c. In order to ensure the connection of the network, we assume the communication radius is $2r_c$. Since the amount of cluster heads K $_{exp}$ and nodes N are determined, we can conclude that the amount of the nodes in each cluster is N/Kexp. We just assume that the cluster head's energy consumption of the data of aggregating each mode is E_{DA} and then, the energy consumption of aggregating all the nodes in the cluster is L^* $(N/K_{exp}^{-1})^*E_{DA}$. At the same time, the energy consumption of receiving the data of N/ K_{exp}^{-1} is L* E_{elec}^{*} (N/K_{exp}⁻¹). According to above description, we concluded that the equation of the energy consumption of receiving and aggregating the sensed data of common nodes is given by:

$$E_{\text{T&A}} = L * \left(\frac{N}{K_{\text{exp}}} - 1 \right) * E_{\text{DA}} + L * E_{\text{elec}} * \left(\frac{N}{K_{\text{exp}}} - 1 \right)$$
 (5)

Assuming the data length which needs to be relayed is still L after cluster head aggregating the data of the common nodes, then the equation of the energy consumption of this cluster sending the aggregated data to the relaying cluster head is given by:

$$E_{TL} = L * E_{elec} + L * \varepsilon_{fs} * (2r_c)^2$$
 (6)

Since the sense zone is a square and the amount of cluster heads is K_{exp} , in some relaying path, the amount of cluster heads approximately equal $\sqrt{K_{exp}}$, the relaying amount of the furthest cluster head (relative to sink) is 0, whereas the nearest is $\sqrt{K_{exp}} - 2$. The relaying amount of all the cluster heads is the span $[0, \sqrt{K_{exp}} - 2]$. So the mean of relaying amount is $(\sqrt{K_{exp}} - 2)/2$ Then the energy consumption of relaying remote cluster heads' data is given by:

$$E_{TR} = \left(L * E_{\text{elec}} + L * \epsilon_{\text{fs}} * (2r_{\text{c}})^{2}\right) * \left(\frac{\sqrt{K_{\text{exp}}} - 2}{2}\right)$$
(7)

The energy consumption of receiving the remote cluster heads' data is given by:

$$E_{RX} = (L * E_{elec}) * \left(\frac{\sqrt{K_{exp}} - 2}{2} \right)$$
 (8)

The summation of (5), (6), (7), (8) is the energy consumption of one cluster head finishing data gathering in one round.

$$\begin{bmatrix} L * \left(\frac{N}{K_{exp}} - 1\right) * E_{DA} + L * E_{elec} * \left(\frac{N}{K_{exp}} - 1\right) + \\ L * E_{elec} + L * \varepsilon_{fs} * \left(2r_{c}\right)^{2} + \left(L * E_{elec} + L * \varepsilon_{fs} * \left(2r_{c}\right)^{2}\right) * \\ E_{CH} = K_{exp} * \left(\frac{\sqrt{K_{exp}} - 2}{2}\right) + \left(L * E_{elec}\right) * \left(\frac{\sqrt{K_{exp}} - 2}{2}\right) + \\ \left(\frac{N}{K_{exp}} - 1\right) * \left(L * E_{elec} + L * \varepsilon_{fs} * \frac{9}{8\sqrt{3\pi}} r_{c}^{2}\right) \end{bmatrix}$$

$$(9)$$

As is mentioned earlier, the energy consumption of one common node is given by:

$$E_{\text{member}} = L * E_{\text{elec}} + L * \epsilon_{\text{fs}} * \frac{9}{8\sqrt{3}} r_{\text{c}}^{2}$$
 (10)

So we can obtain the energy consumption of one cluster in one data gathering round, as the following equation:

$$E_{\text{chuster}} = E_{\text{CH}} + \left(\frac{N}{K_{\text{exp}}} - 1\right) * E_{\text{member}}$$
 (11)

The energy consumption of the whole network is given by:

$$E_{\text{total}} = K_{\text{extr.}} * E_{\text{cluster}}$$
 (12)

According to the above equation, we conclude that $E_{\mbox{\tiny total}}$ is shown as the following.

$$E_{\text{total}} = \frac{4\|A\|}{3\sqrt{3}r_{c}^{2}} * \left(\frac{N}{K_{\text{exp}}} - 1\right) * E_{\text{elec}} + L * E_{\text{elec}} * \left(\frac{N}{K_{\text{exp}}} - 1\right) + L * E_{\text{elec}} + L * \varepsilon_{\text{fs}} * \left(2r_{c}\right)^{2}\right) * \left(\frac{\sqrt{K_{\text{exp}}} - 2}{2}\right) + \left(L * E_{\text{elec}}\right) * \left(\frac{\sqrt{K_{\text{exp}}} - 2}{2}\right) + \left(\frac{N}{K_{\text{exp}}} - 1\right) * \left(L * E_{\text{elec}}\right) * \left(\frac{\sqrt{K_{\text{exp}}} - 2}{2}\right) + \left(\frac{N}{K_{\text{exp}}} - 1\right) * \left(L * E_{\text{elec}}\right) * \left(\frac{N}{2}\right) + \left(\frac{N}{2}\right) * \left($$

Take the cluster radius r_{c} as the variable, we differentiated the Eq. 13. Then we obtained the optimal cluster radius r_{c} .

Cluster heads determination and cluster working mode:

After the determination of cluster radius, sink node work on forming cluster according to the cluster radius. During the process of forming cluster, sink node can ensure the coordinate of each circle's centre. All the circles' centre set up a set. After getting the coordinate of all nodes, compute the distance between each node and all the circles' centre. After computing the distance, the first and second nearest nodes can be ensured. The determination of the first and second nearest nodes is to take the first nearest node as the initial cluster head and the second nearest as the backup one. On the sink node's aspect, all the ID of the nodes which are first nearest to the circle center form up a set and the ID of the second nearest one form up another set. Sink node broadcast the two sets to the whole network. After all the nodes receiving message from sink node, they begin to compare with their own ID. If they discover their own ID in the set, then the primary cluster head and the backup cluster head form up.

After the determination of the primary cluster head and backup one, the clusters are determined. Sink node appoints a unique ID to each cluster in order to have the intra-cluster send message easily.

After the determination of nodes, by adjusting the transmitting power the cluster heads broadcast the inviting message to the common nodes within the known

radius. When the message is received by the common nodes, they will send the joining message to the cluster heads. It is possible that some node receives two or more inviting message, then, it ensures that the nearest cluster head join in. The operation finished, the primary cluster heads would ready to work and the backup cluster heads would get into sleeping state.

If the following two conditions happen, the backup cluster heads will upgrade to the primary cluster heads:

- When the primary cluster heads are destroyed accidentally and the common nodes can not send their data to the primary cluster heads. The backup cluster heads will be waked up, upgrade to be the primary cluster heads.
- After working for several times, the primary cluster heads will consume much energy than the common nodes. Therefore, the declination of the primary cluster heads is very fast. When their current energy is less than 30% of their initial energy, they will send a message to their common nodes announcing that they degrade to be a common node and never act as a cluster head. At the same time, the backup cluster heads are waked up, upgrade to primary cluster heads.

Intra-cluster communication and order of transmission within clusters: After the stabilization of the clusters, the primary and backup cluster heads are determined. The primary cluster heads broadcast the TDMA slots to all their own common nodes in order to get known the exact time of every common node transmitting the data. The LEACH protocol is definitely works on by round. So at the beginning of each round, the TDMA slots should be determined. But EEHCA protocol is different. That is to say, on the occasion of not changing cluster heads, each common node would transmit to the head cluster for more times. In order to solve the problem that a common node would transmit to the cluster head for more times, we set a threshold T. Here, T denotes the summation of the time of all nodes in the cluster transmitting data in one TDMA round. When the time T passes, the cluster heads will send new TDMA slots to all their own common nodes.

Mode of inter-clusters transmission: At present, there are two types of inter-transmission mode, single hop and multi hop (Mao *et al.*, 2005). Here, we adopt the multi hop mode to achieve the inter-cluster transmission. It means that the cluster heads transmitted their aggregated data to the sink node by passing several other cluster heads. When the furthest cluster heads want to transmit their own aggregated data to the sink node, they would think over the following questions:

- Which angle is the smallest between two lines which are separately from relaying cluster head to sink node and the furthest cluster head to sink node?
- Which relaying cluster head is the nearest to the transmitted cluster head?
- The distance between relaying cluster head and the sink node must shorter than the distance between transmitted cluster head and the sink node.

After the above operation, the network can work regularly. Then we will make simulation to prove that the EEHCA has better performance than other protocols.

SIMULATION RESULTS

Here, we evaluate the performance of EEHCA protocol implemented with MATLAB. For simplicity, we assume the probability of signal collision and interference in the wireless channel is ignorable. And we adapt the same MAC protocols in EEHCA as in LEACH and HEED. MAC protocols have been developed to assist each node to decide when and how to access the channel. This problem is also known as channel allocation or multiple access problems. MAC protocols have been extensively studied in traditional areas of wireless voice and data communications. Time division multiple access (TDMA), frequency division multiple access (FDMA) and code division multiple access (CDMA) are MAC protocols that are widely used in modern cellular communication systems. Their basic idea is to avoid interference by scheduling nodes onto different sub-channels that are divided either by time, frequency or orthogonal codes. Since these sub-channels do not interfere with each other, MAC protocols in this group are largely collision-free.

In order to explain the relations between the network scale and the parameters in EEHCA, we run each kind of simulation in two different scenes, which are normal scale scene and large scene respectively. The parameters of simulations are shown in Table 2. Unless otherwise

Table 2: Pa	arameters of s	imulations			
Parameters Scene 1		Scene 2	Scene 3	Scene 4	Scene 5
Area	100*100	200*200	300*300	400*400	500*500
Location	(50,100)	(100,200)	(150,300)	(200,400)	(250,500)
N	400 or 1000	400 or 1000	400 or 1000	400 or 1000	400 or 1000
of sink					
Initial	0.5J or 1.0J	0.5 Jor 1.0J	0.5J or 1.0J	0.5J or 1.0J	0.5J or 1.0J
energy					
$\mathrm{E}_{\mathrm{elec}}$			50 nJ/bit		
ϵ_{fs}			10 pJ/bit/m ²		
$\varepsilon_{ m amp}$			0.0013 pJ/bit/m ⁴		
d _{crossover}			87 m		
E_{DA}			5 nJ/bit/signal		
Packet size			4000 bits		

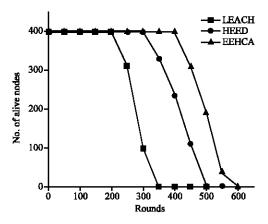


Fig. 4: Lifetime of 400 nodes

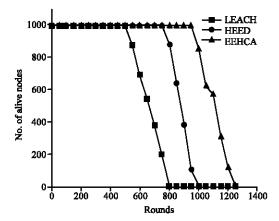


Fig. 5: Lifetime of 1000 nodes

specified, every simulation result shown below is the average of 100 independent experiments where each experiment uses a different randomly-generated uniform topology of sensor nodes. After we finish these independent experiments, we obtain the maximum value and the minimum value of network's lifetime and their difference are all less than 40 rounds. We use the mean values which are obtained by the independent experiments as the basic data to draw the Fig. 4-7.

Lifetime is the criterion for evaluating the performance of sensor networks. In the simulation, we measure the lifetime in terms of round when the last node dies out.

Firstly, we examine the impact on the amount of nodes on the network lifetime. We compare the data by doing experiments in different number of nodes. In the first experiment, the number of nodes is 400; in the second experiment, the number of nodes is 1000. In these two experiments, the sink node is located at the (50,100); the area of sense zone is 100*100; the initial energy of all

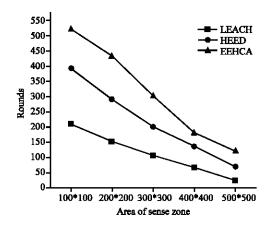


Fig. 6: Impact of area on lifetime odonlifetime

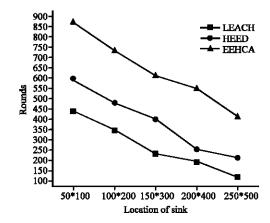


Fig. 7: Impact of sink location on lifetime

nodes is 1.0J. Others relative parameters are shown in Table 2. The simulations are shown in Fig 4 and 5.

Base on the results of the two experiments which are shown in the Fig. 4 and 5, we can observe that the EEHCA protocol has a much better performance than the LEACH and HEED protocol in the aspect of network lifetime. From the trend of the curve's variation, we can find that the basic graphic does not change with the increase of the nodes. Just because of the increase of sensor node, it prolongs the lifetime of the network. Therefore, EEHCA has a better performance no matter the nodes are dense or sparse.

In Fig. 4 and 5, LEACH takes about 150 rounds and 300 rounds, respectively to die from the first node to the last one, while HEED takes about 180 rounds and 210 rounds, EEHCA takes about 170 rounds and 280 rounds. Therefore, EEHCA has a restricted improvement in its performance comparing to the above two protocols in terms of node energy consumption load. This is also the problem we are going to solve in the future.

In the experiment shown in Fig. 6, we examine the impact of sense zone's area on network's lifetime. We have done experiments in different area of sense zone whose sense zones changing from 100*100 to 500*500. In these scenes, the sink node is located at the (100,200); the number of nodes is 400; the initial energy of all nodes are 0.5J; others parameters are shown in the Table 2. It is obvious that with the increase of sense zone the network lifetime become short, while, the performance of EEHCA are still better than the other two. With the expanding area of sense zone, the communication distances increases between the nodes. However, long distance data transmission is the main factor for energy consumption. We can see from Fig. 6, under a comparatively large sense zone, the network lifetime of EEHCA lasts longer than the above two protocols by 80%.

Finally, we compare the performance of EEHCA with LEACH and HEED with the different location of sink node. In order to finish the comparison, we have done the experiments in different location of sink node whose location of sink node is (50,100), (100,200), (150,300), (200,400), (250,500). In all the experiments, the number of nodes is 400; the area of sense zone is 100*100; the initial energy of all nodes are 1.0J; others parameters are shown in the Table 2. The simulation result is shown in the Fig. 7. The distance between sink node and sense zone determines the energy declining speed of cluster head node. EEHCA does not need to elect cluster head frequently. Therefore, in course of increasing the distances between sink node and sense zone, the energy consumption is much slower than the other two protocols. It fully demonstrates the excellent performance of EEHCA.

CONCLUSIONS

In this study, we present a novel energy efficient and load balanced clustering scheme applying for periodical data gathering. In the determination of cluster head, the EEHCA produced a new method to balance the energy consumption among all the nodes. At the same time, backup cluster head is produced to improve the performance of fault-tolerance. What's more, a novel approach has been introduced to distribute the energy consumption among the cluster heads in the inter-clusters transmission. Simulation results show that EEHCA prolongs the network lifetime than LEACH and HEED much more.

If sensor nodes will not equipped with directional antenna, then we can hardly determine the location of cluster's centre during the phrase of cluster heads determination. It will have great impact on the process of EEHCA protocol. But we can try to solve the problem according to the following method. After the determination of optimal cluster radius, we can adopt the minimum overlapping area mode during the coverage process. Because of the sense zone is square, we can define the (0,0) point of the cartesian coordinates located at the centre of the square. Furthermore we regard the point as the centre of a cluster, therefore, we can determine the coordinates of the others clusters' centre.

Our contributions here focus on the cluster setup stage and improving the fault-tolerance performance. There is much space to improve the performance of data transmission. In the large scale of sensor networks, we will improve the inter-clusters communication mode. Since those nodes that have directional antenna will cost much, we won't make an assumption that the nodes equip the directional antenna and design a location unaware protocol to adapt more application in the future work.

REFERENCES

Akkaya, K. and M. Younis, 2005. A survey of routing protocols in wireless sensor networks. Elsevier Ad Hoc Network J., 3/3: 325-340.

Akyildiz, I.F., W. Su and Y. Sankarasubramamiam, 2002. A survey on sensor networks. IEEE. Commun. Mag., pp: 102-114.

Choudhury, R.R. and R. Kravets, 2004. Location-Independent coverage in wireless sensor networks. http://www.crhc.uiuc.edu/~croy/pubs.

Estrin, D., L. Girod, G. Pottie and M. Srivastava, 2001. Instrumenting the world with wireless sensor networks. In: The International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Heinzelman, W.R., A. Chandrakasan and H. Balakrishnan, 2002. An application-specific protocol architecture for wireless microsensor networks. IEEE. Trans. Wireless Commun., 1 (4): 660-670.

Krishnamachari, B., D. Estrin and S. Wicker, 2002. The impact of data aggregation in wireless sensor networks. In: The 22nd International Conference on Distributed Computing Systems Workshops (ICDCSW'02).

Lindsey, S. and C.S. Raghavendra, 2002. PEGASIS: power efficient gathering in sensor information systems. IEEE Aerospace Conference Proceedings, 3: 1125-1130.

- Liu, M., J.N. Cao and G.H. Chen, 2007. EADEEG: An energy-aware data gathering protocol for wireless sensor networks. J. Software, 18 (5): 1092-1109.
- Mao, Ye, F.L. Cheng and W. Jie, 2005. EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks. Proceedings of the IEEE International Performance Computing and Communication Conferences New York: IEEE Press, pp: 535-540.
- Yang, S., F. Dai, M. Cardei, J. Wu and F. Patterson, 2006.
 On connected multiple point coverage in wireless sensor networks. J. Wirelsss Inform. Networks, 13 (4): 289-301.
- Younis, O. and S. Fahmy, 2004. HEED: A hybrid, energy efficient, distributed clustering approach for Ad Hoc sensor networks. IEEE. Trans. Mobile Comput., 3 (4): 660-669.