Journal of Korea Multimedia Society Vol. 13, No. 6, JUNE 2010(pp. 911-923)

Density Aware Energy Efficient Clustering Protocol for Normally Distributed Sensor Networks

Xin Su[†], Dongmin Choi[††], Sangman Moh[†††], Ilyong Chung[††††]

ABSTRACT

In wireless sensor networks (WSNs), cluster based data routing protocols have the advantages of reducing energy consumption and link maintenance cost. Unfortunately, most of clustering protocols have been designed for uniformly distributed sensor networks. However, some urgent situations do not allow thousands of sensor nodes being deployed uniformly. For example, air vehicles or balloons may take the responsibility for deploying sensor nodes hence leading a normally distributed topology. In order to improve energy efficiency in such sensor networks, in this paper, we propose a new cluster formation algorithm named DAEEC (Density Aware Energy-Efficient Clustering). In this algorithm, we define two kinds of clusters: Low Density (LD) clusters and High Density (HD) clusters. They are determined by the number of nodes participated in one cluster. During the data routing period, the HD clusters help the neighbor LD clusters to forward the sensed data to the central base station. Thus, DAEEC can distribute the energy dissipation evenly among all sensor nodes by considering the deployment density to improve network lifetime and average energy savings. Moreover, because the HD clusters are densely deployed they can work in a manner of our former algorithm EEVAR (Energy Efficient Variable Area Routing Protocol) to save energy. According to the performance analysis result, DAEEC outperforms the conventional data routing schemes in terms of energy consumption and network lifetime.

Key words: Sensor Network, Clustering, Normal Distribution, LEACH

1. INTRODUCTION

WSNs consist of a large number of sensor devices cooperate together for collecting environmental information or reacting to specific events [1]. Typically, the mounted battery in a smart node has around 1 Joule of energy and is not rechargeable [2]. This is the biggest problem for sensor networks because communication module for WSNs impropriates 60% of total electric power [3]. If every sensor starts to communicate and engage in data transmission in the network, a serious network congestion and data collision are experienced. Data aggregation [4-5] and clustering [6-14] mechanisms have been proposed for dealing with this problem. In clustered networks, sensors are partitioned into small clusters and a cluster head (CH) for each cluster is selected. Sensor nodes in each cluster transmit their data to the corresponding CH, then CH aggregates data and forward

※ Corresponding Author: Sangman Moh, Address: (501-759) 375 Seosuk-dong, Dong-gu, Gwangju, Korea, TEL: +82-62-230-6032, FAX: +82-62-230-7754, E-mail: iyc@chosun.ac.kr
Receipt date: Oct. 5, 2009, Revision date: Dec. 3, 2009
Approval date: June 22, 2010
[†] Department of Computer Eng., Chosun University (E-mail: leosu8622@hotmail.com)
[††] Department of Computer Eng., Chosun University (E-mail: cdm1225@gmail.com)
[†††] Department of Computer Eng., Chosun University
[††††] Department of Computer Eng., Chosun University (E-mail: smmoh@chosun.ac.kr)
※ This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD)" (KRF-2009-L030102-해06B2005). This research was financially supported the Ministry of Education, Science Technology (MEST) and Korea Institute for the Advancement of Technology(KIAT) through the Human Resource Training Project for Regional Innovation.

them to a central base station. Clustering through creating a hierarchical WSN facilitates efficient utilization of limited energy of sensor nodes and prolongs network lifetime. In some specific applications (e.g., battle field applications), thousands of sensor nodes cannot be deployed uniformly. Air vehicles or balloons may take the responsibility for deploying sensor nodes. The random deployment of thousands of sensors may follow normal distribution. However, traditional clustering approaches such LEACH [6], TEEN [10], and RRCH [11] only considered uniform distribution in protocol design and experiments.

In this paper, firstly, we modified CH selection method in [6, 14] to ensure CH is selected with the factors of remaining energy and distance base on normal distribution function. Secondly, we proposed a new clustering protocol for normally distributed sensor network. In this protocol, each node determines whether it is located in a LD or HD cluster by broadcasting hello message within the area of a fixed signal range. Obviously, the HD nodes can receive more ACKs message than LD nodes. Then, all nodes are labeled by the information of HD/LD node, and work with different signal transmission range for the future phases. The selected HD CH should help the LD CH to send the data to central base station balancing energy consumption. Futhermore, because HD clusters are compact and have high probability to send the duplicated sensed data, we use same scheme as EEVAR for the HD clusters. The detailed explanation will be shown in section 4.

The rest of this paper is organized as follows: the related works are described in the following section by summarizing the conventional clustering schemes. In Section 3, the deployment model and energy model is explained. In Section 4, the proposed DAEEC is presented with respect to sensor nodes categories determination, cluster head selection, cluster initialization, routing and energy consumption. An additional idea TH-DAEEC is explained in Section 5, which combines the threshold value with DAEEC. The performance of the proposed DAEEC is analyzed in terms of energy consumption and network lifetime, and then compared with the conventional protocols in Section 6. Finally, conclusions are covered in Section 7.

2. RELATED WORK

2.1 LEACH

Low Energy Adaptive Clustering Hierarchy (LEACH) offers an energy balancing usage by choosing the CHs stochastically for each round. It allows for energy minimization while nodes can turn off their radio during all their scheduled time-slot. LEACH provides a good load balance between each node, a better scalability for cluster formation and better lifetime of network. Fig. 1 illustrates the initial cluster setup method of LEACH.

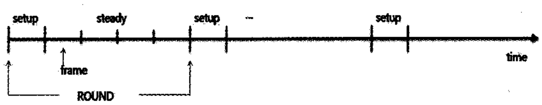


Fig. 1. Periodic round in LEACH.

2.2 TEEN

Threshold sensitive Energy Efficient sensor Network protocol (TEEN) is a reactive network, which puts two threshold values. One is Soft threshold (ST) value that triggers the node to switch on its transmitter and transmit data. The other is hard threshold (HT) value. It is the absolute value of the attribute beyond which, the node sensing this value must switch on its transmitter and report to its cluster head. Fig. 2 is the mechanism of TEEN.

Fig. 2. Periodic round in TEEN.

2.3 LELE

In the case of non-uniformly distributed networks, Leader Election with Load balancing Energy (LELE) [12] is load-balance protocol which aims to deal with the problem of sensor vacancy in the non-uniformly deployed network. In LELE, the probability of becoming CH decreased or increased with regard to the difference of the energy level of one node, as well as the number of neighbors. Regarding the distance of a node with the neighbor and the energy of the sensor, if the neighbor's energy level is lower than this node, the probability of the node becoming a CH is increased. If the neighbor energy level is high, the probability of the node's becoming CH decreases. The shorter this distance, the more the effect of this parameter, which is calculated for each neighbor and can increase or decrease according to the distance with neighbors. The amount of this parameter is calculated through the Formula. 1.

$$X=(E-E_{\ni})*(1/disti) \tag{1}$$

The initial amount of the parameter X is zero and E is the amount of energy of the node and is the amount of energy of the ith neighbor. The disti is the distance between the node and the ith neighbor. If X>0 the probability of its becoming the CH gets higher and if X<0 the probability of its becoming the leader gets lower.

2.4 EEVAR

Energy Efficient Variable Area Routing protocol (EEVAR) [13] is new approach, which can adapt to a non-uniformly distributed sensor network. It has a high performance for saving energy. During the period of initial setup time, EEVAR first sets intra clusters that operate like a node in a cluster and occurs only once at setup. Then, the next setup is to set network clusters in which the nodes are grouped as an intra cluster participates.

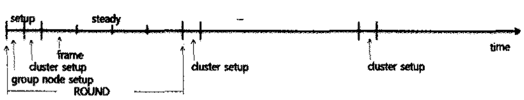


Fig. 3. Periodic round in EEVAR.

In [13], the number of intra clusters p can be calculated by:

$$S_i(t)=\frac{\lambda}{Nu} \tag{2}$$

$$p=\frac{S_t}{U_{TH}}, 1 \le p \le \lambda \tag{3}$$

The number of nodes (Nu) to be located in the 1m×1m area and the number of nodes for 5%, which is a ratio of the most effective CHs are calculated to get the number of nodes (λ) that are included in a cluster. Calculation a multiple number (UTH) of an error allowed area of data extraction assigned by a user compared to a unit area and divided it with the number of nodes in a cluster to get the value p.

In the pre-mentioned routing protocols, LEACH and TEEN were designed for uniformly distributed sensor network. CHs are selected randomly without respects of nodes topology. Therefore, electric power cannot be consumed evenly for each node, of which often leading amount of nodes dead quickly. LELE considered the problem of sensor vacancy in the non-uniformly deployed network, where the nodes vacancy area spots irregularly. EEVAR was proposed for the WSNs with high node density. For each round, only a group of node is triggered to join the routing process. Both of LELE and EEVAR can adapt to non-uniformly distributed WSNs.

3. THE DEPLOYMENT MODEL AND ENERGY MODEL

3.1 Deployment Model

In this paper, we consider the normal distribution for deployed sensor nodes because they might be dropped from air vehicles or balloons. We may acquire a probability density function (pdf) with normal distribution as:

$$f(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x/MOD-\mu)^2}{2\delta^2}}dx \tag{4}$$

where x is a pdf random variable value denoting the width of the sensing area if the flight path of an airplane defines the length of the sensing area (y). MOD is the ratio of width distance to the corresponding abscissa value, and we set it as 40. The value of function f(x) will be the percentage of all sensor nodes located in the left side of axis (x, 0), μ is the expected value and σ is the variance of random variable x. They are affected by various factors such as the height of airplane and the weather when the nodes are deployed. In this paper, we consider the standard normal distribution (SND), when μ=0 and σ=1:

$$f(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2/MOD}{2}}dx \tag{5}$$

We assume that all sensor nodes are stationary after deployment. An example of topology model is illustrated in Fig. 4. The gray sensing area is called HD area, which is a nodes compact area and has lots of sensor nodes. The white sensing area is called LD area because it has few sensor nodes.

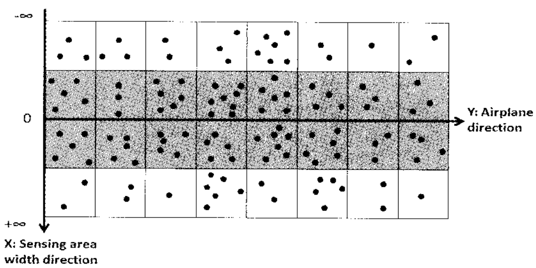


Fig. 4. An example of normally distributed topology.

3.2 Energy Model

According to the radio model of energy consumption introduced in Fig. 5, to reach an acceptable rate of Signal-to-Noise (SNR), the allocated energy to send the message with the length of l-bit

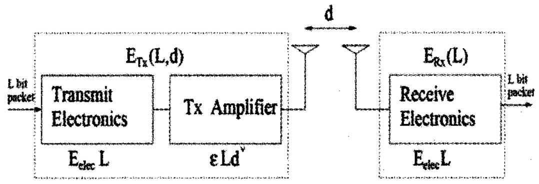


Fig. 5. The radio model of energy.

in the distance of d is calculated through the relation. The energy consumption of the proposed method can be expressed with formulas (6) and (7) used in [6].

$$E_tx(l,d)\begin{cases} lE_{elec}+lE_{mp}D^4 : d>d_0 \\ lE_{elec}+lE_{fs}d^2 \;\; : d\le d_0 \end{cases} \tag{6}$$

$$E_{rx}(l)=lE_{elec} \tag{7}$$

E_{elec} is the consumed energy for each bit in the transmitter circuit or receiver circuit, and depend on the model boosting the transmitter circuit and d is the distance between the transmitter and receiver which in case of $d=d_0$ we have $d_0=\sqrt{(E_{fs}/E_{mp})}$ to receive the l bit message the radio energy consumed is equal to relation (7).

4. THE PROPOSED SCHEME

To start of this, we make some assumptions for sensor nodes. For each node, it can use power control to vary the amount of transmit power to support different MAC protocols and perform signal processing functions. All the deployed nodes can transmit to the edge of sensing field. When it starts to form a cluster it normally decrease transmit model's power, so then, its transmission range is equal to the cluster coverage radius.

4.1 Sensor Nodes Categories Determination

Suppose all the nodes are not global location aware (without GPS system). After deployment of sensor nodes, at first, we should determine the sensor node whether it located in HD area or LD area by broadcasting hello message with a fixed signal range L. It is obvious that the number of ACKs received by a HD node is more than a LD node. Thus, we set a threshold value ε. If the number of received ACKs is less than ε, the node is located in LD area, and vice versa. Additionally, in order to avoid data collision during categories determination phase, base station send time synchronization information to all nodes that only allows some of them sending hello messages at one time with a non-persistent carrier-sense multiple access (CSMA) MAC protocol. After determination phase, each node labels itself with the information of HD/LD node and works with different signal transmission ranges for future phases. The algorithm of this step is described with pseudocode as follows.

Notations:

i: Node *i*.
ACK: Number of acknowledgement message received by *i*.
ε: Threshold value for determining the HD/LD nodes.

Algorithm 1: Determination of HD/LD nodes

```
1: if  i is the node selected to send hello message
2:   if  ACK>ε
         i label itself as HD node
3:   else
         i label itself as LD node
4: else
         Waiting for send hello message
5: end if
```

4.2 Cluster Head Selection

In [6], it is obvious that a stochastic CH selection will not automatically lead to minimum energy consumption during data transfer for a given set of nodes. A CH can be located near the edges of the network or adjacent nodes can become CH. This leads to high energy consumption since nodes have to transmit over long distance. While there are advantages for using LEACH's distributed cluster formation algorithm, this protocol offers no guarantee about the placement of CH nodes. Therefore, in order to reduce energy consumption, a good method is to choose the nodes that are closest to several specific positions as CHs at the beginning of each round. These positions are uniformly distributed in the work avoiding all CHs can be located near the edges of the network or adjacent nodes can become CHs. Once clusters are formed, each member node transmit sensor data to its CH in a schedule time slot and CHs transmit aggregated data to the base station as described in LEACH. After transmission in each frame, the node with highest energy capacity in each cluster to be the CH of next frame. This can be achieved by setting the probability of becoming a CH as a function of a node's energy level relative to the aggregate energy remaining in the network, rather than purely as a function of the number of times the node has been CH. Thus, if there are nodes in the network, each node should choose to become a CH with probability (8).

$$P_i(t) = \frac{E_i(t)}{E_{total}(t)} \tag{8}$$

where Ei(t) is the current energy of node i and

$$E_{total}(t) = \sum_{i=1}^{N} E_i(t) \tag{9}$$

Using these probabilities, the nodes with higher energy level are more likely to become cluster heads than nodes with less energy level. By modifying above CH selection, in [14], the authors use a sigmoid function to denote the possibility to be elected as a CH (versus distance),

$$g(D_i(t)) = \frac{1}{1+e^{-a_1(D_i(t)-c_1)}} \tag{10}$$

where Di(t) denotes the distance between node i to a location of ideal CH; a1 and c1 are denote the shape of function and position, respectively. Another sigmoid function denotes the possibility to be selected as a CH (versus remaining energy level),

$$f(E_i(t)) = \frac{1}{1+e^{-a_2(E_i(t)-c_2)}} \quad (11)$$

and the defined possibility of each node to be selected as a CH as

$$P_i(t) = f(E_i(t)) * g(D_i(t)) \quad (12)$$

In this paper, due to the characteristics of the nodes deployment topology, we modify the CH selection method in [6,14] by using normal distribution functions instead of sigmoid functions. The modified functions are as follows:

$$g(D_i(t)) = \frac{1}{\sqrt{2\pi}\sigma} e - \frac{(D_i(t)/\mathrm{mod}^2)}{2\sigma^2} : (D_i(t) \geq 0) \quad (13)$$

$$f(E_i(t)) = \frac{1}{\sqrt{2\pi}\sigma} e - \frac{(E_i(t)/\mathrm{mod}' - \mu)^2}{2\sigma^2} : (E_i(t) \geq 0) \quad (14)$$

Di(t) and Ei(t) are distance and energy variances similar to [14]. mod is the ratio of experiment distance and the corresponding abscissa value while mod' is the ratio of experiment energy level. Thus, we have new formula (15), which is generated by using formulas (13) and (14).

$$P_i(t) = \begin{cases} P_i(t) = f(E_i(t)) * g(D_i(t)) : C_i(t) = 1 \\ 0 \quad : C_i(t) = 0 \end{cases} \quad (15)$$

Pi (t) is the probability of a node being chosen to be the CH. If the node was recently selected as the CH, Pi(t)=0. The performance of the proposed CH selection method will be presented in the analysis section using the same simulation environments as [15].

4.3 Cluster Initialization

In [6,10-14], the node with probability Pi(t) broadcasts hello messages to the entire network in order to form a cluster. After receiving the ACK messages from network members, the node uses a more limited transmission range to send scheduling information setting up a cluster. In contrast, DAEEC does not broadcast hello message over the entire network because of the high energy consumption. We only need the selected CHs broadcast hello messages with a short distance. If we ensure that each cluster has same number of nodes (λ). Consequently, the HD CHs can use signal transmission range L to broadcast hello and scheduling messages. The LD CHs, in order to set up a cluster with λ number of nodes, expand signal transmission range to t×L to form LD clusters.

For the HD nodes, we use the same clustering method as EEVAR to set up a cluster. An intra CH is selected with the probability of Pi(t) (15), and then it participates in an HD cluster as one of the member nodes, among which an HD CH node is selected. The probability function to select an HD CH is also same to Pi(t) (15).

The intra cluster setup has a process of 3 times of transmissions and one reception, and two times of scheduling operations of an intra cluster head node. We have energy consumption in an intra cluster as:

$$E_{ic-setup-ch} = \frac{\lambda}{p} * 3(E_{tx}(l,L) + (\frac{\lambda}{p} - 1) E_{rx}(l) + 2E_{schedule}(l) \quad (16)$$

$$E_{in-setup-ch} = E_{tx}(l,L) + 3E_{rx}(l) \quad (17)$$

Combining the (16) and (17), the energy consumption of a intra CH and a participating node can provide energy consumed for intra cluster setup as shown in (18). If the number of HD clusters is denoted as high, multiplying the number of intra clusters created in the entire network (high×p) shows energy consumption for intra cluster setup like formula (19).

$$E_{ic-setup-ch} = E_{ic-setup-ch} + (\frac{\lambda}{p} - 1) E_{ic-setup-non-ch} \quad (18)$$

$$E_{ic-setup-tot} = p * high * E_{ic-setup-cluster} \quad (19)$$

The HD cluster setup of DAEEC is composed of two times of transmission one time of receipt and scheduling operation of a HD CH. We presented in a formula using (20) and (21) for comparison the formulas in [6]:

$$E_{HD-setup-ch} = (p+1)^* lE_{elec} + pE_{schedule} + 2lE_{fs}L^2 \quad (20)$$

$$E_{HD-setup-non-ch} = lE_{elce} + lE_{fs}L^2 + 2lE_{elcc} \quad (21)$$

The Enc-setup-ch is denotes energy consumption of a HD CH, while cluster source node is Enc-setup-non-ch. Then, we have energy consumption for HD clusters like (22). Multiplying the number of HD clusters (high) created in the entire in the HD area can present energy consumption like (23).

$$E_{HD-setup-cluster} = (p-1)^* E_{nc-setup-ch} + E_{nc-setup-non-ch} \quad (22)$$

$$E_{HD-setup-tot} = high^* E_{nc-setup-cluster} \quad (23)$$

The LD cluster formation is same to LEACH except LD CHs broadcasting ADVs with a limited range t×L. After receiving the Join-REQs, then the selected CH uses same expanded transmissions range t×L to send scheduling information in order to set up an LD cluster. Multiplying the number of LD clusters (low) created in the entire in the LD area can present energy consumption like (24).

$$E_{LD-setup-tot} = \lambda^* low^* (E_{tx}(l, tL) + E_{rx}(l)) \quad (24)$$

The Flow Chart of Clustering Initialization is given as follows:

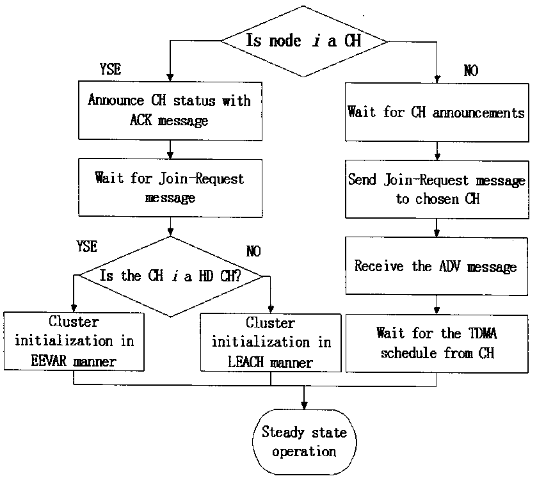


Fig. 6. The flow chart of clustering initialization.

Table 1. Notation

E_{elec}	*Electric energy*
E_{fs}	*Amplified energy of a free space model*
E_{mp}	*Amplified energy of a multiple model*
$E_{schedule}$	*Energy consumption for node scheduling*
E_{da}	*Energy consumption for data aggregation*
E_{th}	*Transmission probability (TEEN APTEEN)*
l	*Data size*
λ	*The number of member nodes in a L range of HD area/The optimal number of a cluster nodes*
t	*The coefficient for LD nodes extending transmission range*
d	*Distance for LD cluster one hop*
D	*Distance to sink*
L	*The transmission range for hello message*
low	*The number of LD clusters*
high	*The number of HD clusters*
p	*The number of intra clusters in an HD cluster*
N_u	*The number of nodes located in 1m×1m*
N_p	*The number of nodes include in an intra cluster*
U_{TH}	*Error allowed area value assigned by a user*
S_t	*Area where λ nodes locate*

4.4 Routing Scheme

In the beginning, all nodes have the same energy, but towards the end, sensor nodes may have different amounts of residual energy. We first summarize the characteristics of HD clusters and LD clusters as follows:

HD clusters: the cluster number is more than LD clusters; small coverage; densely deployed; high possibility to transmit duplicated sensed data.

LD clusters: the cluster number is less than HD clusters; large coverage; sparsely deployed; low possibility to transmit duplicated sensed data.

Thus, according to those characteristics, in the steady phase, the proposed approach uses multi-hop methods to transmit data to base station. All member nodes send sensed data to the corresponding CHs using TDMA methods. Each member works during its allotted slot, but during other times they sleep to save energy. The selected CHs then aggregate the received data, and send data to the destination using CDMA method. Each LD cluster uses two hops to get to base station: one hop to the neighboring HD CH with a transmission range d, and then one hop to base station. In this way, HD clusters can help neighboring LD clusters send data to the sink with a transmission range D, balancing the energy consumption of the network.

5. DAEEC WITH THRESHOLD (TH-DAEEC)

The suggested modified approach TH- DAEEC has the same form and structure of DAEEC except it uses threshold values. The node gets the threshold value and count value entered before it is deployed to field. The process for setting up the slope for operation is outlined below. The setting up the slope about the data collection is listed in the Fig. 7.

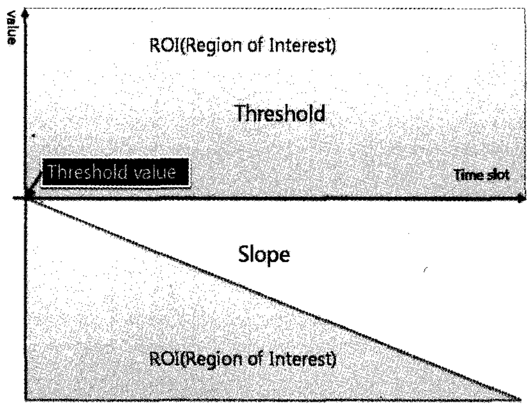


Fig. 7. Region of threshold value and slope.

5.1 Collected Data is Lower Than Threshold Value

In this case, the decision to transmit data is decided through measuring the slope of the threshold value and the collected data. The formula for slope is $S=(B_y-A_y)/(B_x-A_x)$, where Ay represents collected data at initial threshold, By represents collected data after the initial time. Ax is the time slot count assigned to node. Bx is the time slot count of By, which is collected data after Bx. The time slot number of collected data is always one larger than previously data. Thus, for S we have, S=By-Ay.

5.2 Collected Data is Higher Than Threshold Value

When the collected data is higher than threshold value, the nodes transmit the collected data into the originally assigned time slot. For TH-DAEEC, we use the same transmission probability denoted by Eth in [10].

6. PERFORMANCE ANALYSIS

In this section, we evaluate proposed algorithm performance via MATLAB. The analyzing result includes two parts. The first one is the performance comparison of proposed cluster head selection method beyond [6,14]. Second, the energy consumption of DEAAC is analyzed in terms of setup phase, steady phase, total energy consumption and the number of node alive time.

6.1 Comparison of Cluster Head Selection

We evaluate the performances of proposed method for CHs selection with the same simulation environment as [14] except setting equipped battery energy level and the deployed nodes topology with a normal distribution. Moreover, we consider the experiment ratio mod=100, and ratio mod' is chosen as 2 and 4 for two simulations with 2-5J, 1-10J energy, respectively. Last, the expected value μ for energy Ei(v) is set as 2.5 and the variance σ is 0.4.

The Fig. 8 shows the simulation results for the network lifetime using the proposed method with 100 nodes deployed following a standard normal distribution with 2J-5J of energy. The Fig. 9 exhibits the simulation result with 100 nodes deployed in a standard normal distribution with

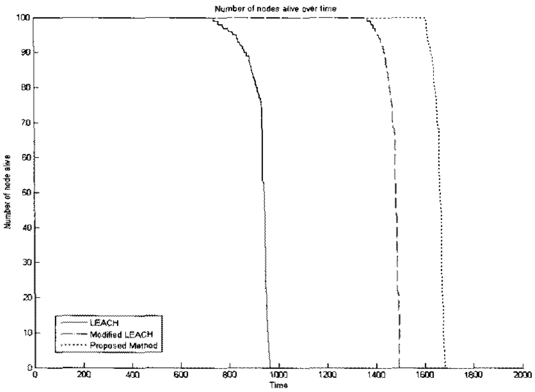


Fig. 8. Number of nodes live time with each node 2J-5J of energy.

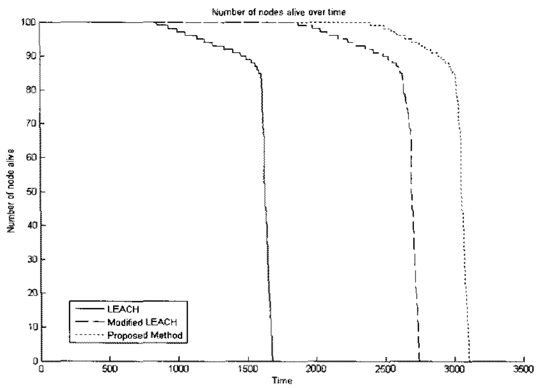


Fig. 9. Number of nodes live time with each node 1J-10J of energy.

1J-10J of energy. These plots show the number of nodes that remain alive over the number of rounds of activity for the 1000*1000m net-work scenario. It is clear that our proposed method outperforms the conventional cluster selection schemes because the node's energy level and deployed topology are normal distributed. In the first plot, the dead node of pro- posed method appears at the round of 1680. In the case of competitors, the dead node of LEACH and [15] appear at the round 750 and 1400, respectively. Moreover, with the changing of initial energy of sensor nodes, second plot shows a significant performance for proposed DAEEC. Thus, we see that if a network topology is not evenly distributed, the performance of LEACH and [14] will decrease more or less because formed network clusters may have different energy capabilities. In this case, poor load-balance will lead to some overloaded nodes dying quickly. When we consider the non-uniform distribution and generate a normal distribution function with factors of distance and remaining energy, the lifetime of the network will be extended. With the increasing of number of experiment nodes, the benefits of the proposed method will be more significant.

6.2 Performance of DEAAC with Energy Consumption

For DAEEC, we deploy N=500 hundred nodes in 200×500m sensing field. The HD area is located in the middle of sensing field, which covers 100×500m. There are two LD areas located beside the HD area. Each covers 50×500m. We set λ=20, ε=6, low=5, and high=20. If the coverage area of an LD cluster is four times as large as an HD cluster, we can calculate the following values: t=2, L=35m. Moreover, we assume that D=500m, d=100m, and the probability of transmitting a threshold value used in TEEN is 25% (Eth=0.25). To verify the superiority of our approach, energy consumption was measured using MATLAB by comparing to LEACH, TEEN, RRCH, and EEVAR. Here an experiment environment consists of a round is composed of 10 frames; Eelec=50nJ/bit, Efs=10p/bit/m2, Emp= 0.0013pJ/bit/m4, Eschedule =5nJ/bit/signal, Eda=5nJ/bit/signal, l=1000bit.

6.2.1 Energy Consumption at Setup Phase

DAEEC is designed for the scenario, in which all sensor nodes are deployed by air vehicles or balloons. With the emergence of fifth generation of smart dust nodes such as Sun SPOT [15] and Iris [16] that can have up to a maximum of 500m super-long-distance communication, our proposed scheme experiment on 200×500m field and the distance to sink is 500m. On the other hand, we aim a location-unaware approach for WSNs, which inevitably increases the cost of overhead in cluster

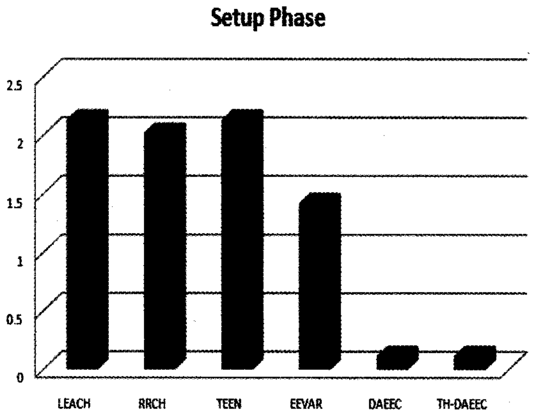


Fig. 10. Energy consumption at the setup phase.

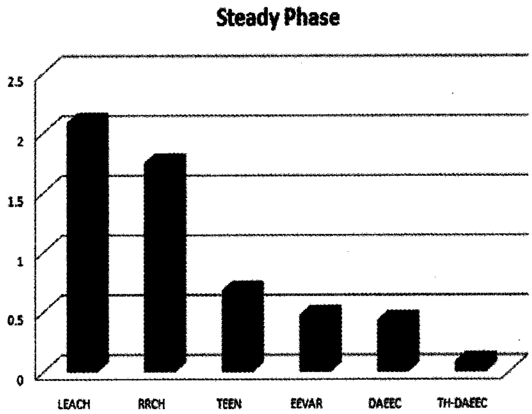


Fig. 11. Energy consumption in steady-state phase.

initialization processes. For instance, DAEEC should determine the category of node's density in the beginning of each round. Fortunately, even if the energy overhead of cluster formation is increased, DAEEC can still save plentiful energy by only broadcasting hello message with a fixed distance. Furthermore, we can only determine node's density one time. Then, each node labels itself with the information of LD node or HD node and works with different signal transmission ranges for future phases. Fig. 10 shows the energy (Joule) saving efficiency for our proposed approach DAEEC during the setup phase. The cluster head of conventional schemes wastes over than 10 times energy (2J) to form a cluster while DAEEC only spends minimum energy (0.2J) to discover cluster members. This is because the proposed DAEEC can save plentiful energy by only broadcasting hello message with a fixed distance.

6.2.2 Energy Consumption at Stead-State Phase

As shown at Fig. 11, the proposed methodcan save about three times as much energy (1.5 J) as LEACH in the steady status state. The improvement gained through DAEEC is that the HD cluster heads helps the LD heads nodes to transmit data to the base station keeping the overhead balance over each heads nodes in the system. The DAEEC using threshold values shows highest performance, because if the transmission is not triggered it will not transit any data for saving energy. For the steady phase, the energy overhead is also augmented to realize the correct data flow path. Some parameters are aggregated in the head packets of data to determine the resource data that ensures the data flows correctly from the edge of the networks to the data sink. Generally, compare to the energy used for transmission, the energy used for aggregates data is relative small which can be ignored. Plus DAEEC can offer a load-balance method to equalize the energy consumption of every cluster head. The proposed scheme also performs well in the steady phase.

6.2.3 Total Energy Consumption

Energy consumption of the proposed scheme can be shown in Fig. 12, when five rounds of 50 frames are transmitted. In Figure 11, the dissipated energy of LEACH almost reaches 120J while the DAEEC and TH-DAEEC only cost 24J, 4J, respectively. It is clear that the proposed method yields lower power consumption compared to previous works. This is because most previous works are not designed for non-uniformly distributed WSNs, and they waste energy by ignoring the allotted assignment for sensor nodes.

6.2.4 Comparison on Node Alive Time

If we assume each node has uniformly distributed 2J of energy before deployment and each

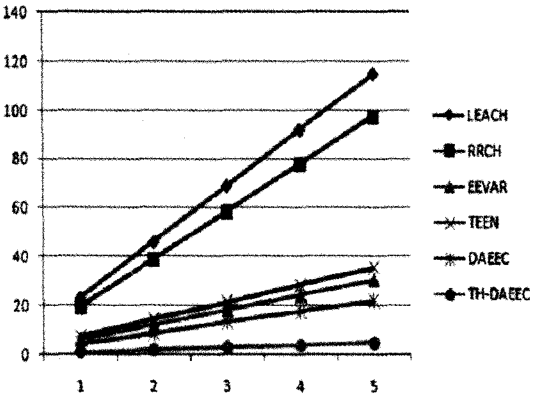


Fig. 12. Total energy consumption in 50 frames.

round costs 50s, sensor nodes lifetime will be different for each protocol. We only compare the networks lifetime for LELE, EEVAR, and DAEEC because they can adapt to the non-uniform distributed networks. The Fig. 13 shows the number of nodes that remain alive over the number of rounds of activity for network. With DAEEC all the nodes remain alive for 1500 rounds, while the corresponding number of for LELE and EEVAR are 480 and 1380, respectively. Furthermore, if system lifetime is defined as the number of rounds for which 75 percent of the nodes remain alive, DAEEC exceeds the system lifetime of LELE 880 rounds and EEVAR 180 rounds, respectively. The comparison result shows that proposed method prolongs the network lifetimes compared to the conventional schemes.

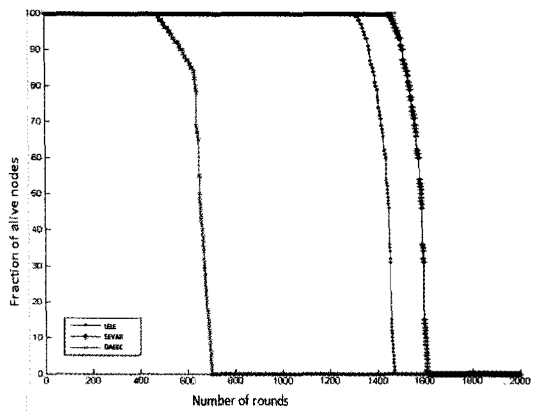


Fig. 13. Number of nodes alive time.

7. CONCLUSION

In this paper, a novel clustering algorithm called DAEEC for non-uniformly distributed sensor networks has been proposed to save energy and prolong network lifetime. We firstly modified the CH selection method in [6,14] to ensure the CH is selected with the factors of remaining energy and distance base on normal distribution function. Secondly, we proposed a new clustering protocol for normally distributed sensor network. In this protocol, each node is determined whether they are located in a LD or HD cluster by broadcasting hello message within the area of a fixed signal range. Obviously, the HD nodes can receive more ACKs message than LD nodes. Then, they are labeled by the information of HD/LD node, and work with different signal transmission ranges for the future phases. The selected HD CH should help the LD CH to send the data to the central base station balancing the energy consumption.

Futhermore, because the HD clusters are compact and have high probability to send the duplicated sensed data, we use same scheme as EEVAR for the HD clusters. Using the assumption that sensor nodes are randomly distributed according to a normal distribution, the proposed DAEEC inevitably increases the controlling overhead for determining the category of cluster in the setup phase and keep the right data flow in the steady phase. However, it can save plentiful energy by only broadcasting hello message with a fixed distance, and the DAEEC can offer a load-balance method to equalize the energy consumption of every cluster head. Therefore, the proposed DAEEC outperforms the conventional clustering algorithms in terms of energy consumption and network lifetime.

REFERENCE

[1] F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Net-

works," *IEEE Communication Magazine*, pp. 102-108, Oct. 2002.

[2] J. M. Kahn. R. H. Katz and K. S. J. Pister, "Next Century Challenges: Mobile Networking for Smart Dust," Proc. of IEEE International Conference on Mobile Computing and Networking, pp. 270-278, Aug. 1999.

[3] J. Hill, D. Culler, "A Wireless-embedded Architecture for System Level Optimization," UC Berkeley Technical Report, Mar. 2002.

[4] C. C. Intanagonwiwat etal., "Directed Diffusion for Wireless Sensor Networking," *IEEE /ACM Trans. Net.*, Vol.11, No.1, pp. 2-16, Feb. 2002.

[5] S. Lindsey, C. Raghavendra, and K.M. Silvalingam, "Data Gathering Algorithms in Sensor Networks using Energy Mertrics," *IEEE Trans. Parallel Distribute. Sys.* Vol.13, No.9, pp. 924-935, Sept. 2002.

[6] W. B. Heinzelman, A. P.Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications,* No.4, 2002.

[7] O. Younis and S. Fahmy, "Distributed Clustering in Adhoc Sensor Networks: A Hybrid, Energy-Efficient Approach," Proc. of IEEE INFOCOM, pp. 443-449, 2004.

[8] S. D. Muruganathan, et al., "A centralized energy-efficient routing protocol for wireless sensor networks," *IEEE Communications Magazine,* Vol.43, pp. s8-s13, Mar. 2005.

[9] M. Rajiullah, S. Shimamoto, "An Energy - Aware Periodical Data Gathering Protocol Using Deterministic Clustering in Wireless Sensor Networks (WSN)," Proc. of Wireless Communications and Networking Conference, pp. 3014-3018, Mar. 2007.

[10] A. Manjeshwar and D. P. Agrawal. "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks," Proc. of 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, April 2001.

[11] Do-hyun Nam, Hong-ki Min, "An Energy-Efficient Clustering Using a Round-Robin Method in a Wireless Sensor Network," Proc. of Software Engineering Research, Management & Applications 2007, pp. 54-60, Aug. 2007.

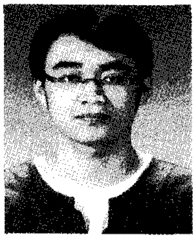
[12] Shirmohammadi. M.M, Faez. K and Chhardoli. M "LELE: Leader Election with Load Balancing Energy in Wireless Sensor Network," Proc. of Communications and Mobile Computing, 2009. CMC '09. WRI International Conference, pp. 106-110, Jan. 2009.

[13] Dongmin Choi, Sangman Moh and Ilyong Chung, "Variable Area Routing Protocol in wireless sensor networks," *Journal of Korea Multimedia Society,* pp. 1082-1092, Aug. 2008.

[14] Hsiao-Lan Hsu, Qilian Liang, "An Energy-Efficient Protocol for Wireless Sensor Networks", Proc. of Vehicular Technology Conference, 2005, VTC-2005-Fall. 2005 IEEE 62nd, pp. 2321-2325, 25-28 Sept. 2005.

[15] Sun SPOT, [Online]. Available: http://www.sunspotworld.com

[16] Crossbow, [Online]. Available: http://www.xbox.com



Xin Su

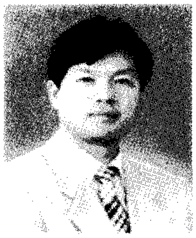
He received his B.E. degree from the Kunming University of Science and Technology, China. Now, he is a Master candidate of Dept. of Computer Science, Chosun University. His research interests are sensor network systems and mobile ad-hoc systems.

Dongmin Choi

He received his B.E. degree from the Kyunghee University in 2003 and his Masters degree in Computer Science from Chosun University. His research interests are in areas of networks and information security, sensor network systems, mobile ad-hoc systems.

Sangman Moh

He received the Ph.D. degree in computer engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea in 2002. Since late 2002, he has been a faculty member in the School of Computer Engineering at Chosun University, Gwangju, Korea. His research interests include mobile computing and networking, ad hoc networks and systems, ubiquitous sensor networks, network based computing, parallel and distributed computing, and high-performance computer systems.



Ilyong Chung

He received the B.E. degree from Hanyang University in 1983 and M.S. and ph.D. degrees in Computer Science from City University of New York in 1987 and 1991, respectively. From 1991 to 1994, he has been a faculty member at Dept. of Computer Science, Chosun University. His research interests are in of network and information security, parallel and distributed systems, computer algorithms.