An Energy-Efficient Unequal Clustering Mechanism for Wireless Sensor Networks

Chengfa Li, Mao Ye, Guihai Chen
State Key Laboratory for Novel Software Technology
Nanjing University
Nanjing, 210093, P. R. China
{licf, yemao}@dislab.nju.edu.cn, gchen@nju.edu.cn

Jie Wu
Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431
jie@cse.fau.edu

***Abstract*— Clustering provides an effective way for prolonging the lifetime of a wireless sensor network. Current clustering algorithms usually utilize two techniques, selecting cluster heads with more residual energy and rotating cluster heads periodically, to distribute the energy consumption among nodes in each cluster and extend the network lifetime. However, they rarely consider the hot spots problem in multihop wireless sensor networks. When cluster heads cooperate with each other to forward their data to the base station, the cluster heads closer to the base station are burdened with heavy relay traffic and tend to die early, leaving areas of the network uncovered and causing network partition. To address the problem, we propose an Energy-Efficient Unequal Clustering (EEUC) mechanism for periodical data gathering in wireless sensor networks. It partitions the nodes into clusters of unequal size, and clusters closer to the base station have smaller sizes than those farther away from the base station. Thus cluster heads closer to the base station can preserve some energy for the inter-cluster data forwarding. We also propose an energy-aware multihop routing protocol for the inter-cluster communication. Simulation results show that our unequal clustering mechanism balances the energy consumption well among all sensor nodes and achieves an obvious improvement on the network lifetime.**

I. INTRODUCTION

Rapid technological advances in MEMS and wireless communication have enabled the deployment of large scale wireless sensor networks. The potential applications of sensor networks are highly varied, such as environmental monitoring, target tracking, and battlefield surveillance [1]. Sensors in such a network are equipped with sensing, data processing and radio transmission units. Distinguished from traditional wireless networks, sensor networks are characterized of severe power, computation, and memory constraints. Due to the strict energy constraint, energy resource of sensor networks should be managed wisely to extend the lifetime of sensors.

In order to achieve high energy efficiency and increase the network scalability, sensor nodes can be organized into clusters. The high density of the network may lead to multiple adjacent sensors generating redundant sensed data, thus data aggregation can be used to eliminate the data redundancy and reduce the communication load [2]. In periodical data gathering applications, both methods promise to efficiently organize the network since data collection and processing can be done "in place".

Among the sources of energy consumption in a sensor node, wireless data transmission is the most critical. Within a clustering organization, intra-cluster communication can be single hop or multihop, as well as inter-cluster communication. Previous research (*e.g.*, [3]) has shown that multihop communication between a data source and a data sink is usually more energy efficient than direct transmission because of the characteristics of wireless channel. However, the hot-spots problem arises when using the multihop forwarding model in inter-cluster communication. Because the cluster heads closer to the data sink are burdened with heavy relay traffic, they will die much faster than the other cluster heads, reducing sensing coverage and causing network partitioning. Although many protocols proposed in the literature reduce energy consumption on forwarding paths to increase energy efficiency, they do not necessarily extend network lifetime due to the continuous many-to-one traffic pattern.

In this paper, we propose and evaluate an Energy-Efficient Unequal Clustering (EEUC) mechanism for periodical data gathering applications in wireless sensor networks. It wisely organizes the network via unequal clustering and multihop routing. EEUC is a distributed competitive algorithm, where cluster heads are elected by localized competition, which is unlike LEACH [4], and with no iteration, which differs from HEED [5]. The node's competition range decreases as its distance to the base station decreasing. The result is that clusters closer to the base station are expected to have smaller cluster sizes, thus they will consume lower energy during the intra-cluster data processing, and can preserve some more energy for the inter-cluster relay traffic. In the proposed multihop routing protocol for inter-cluster communication, a cluster head chooses a relay node from its adjacent cluster heads according to the node's residual energy and its distance to the base station. Simulation results show that EEUC successfully balances the energy consumption over the network, and achieves a remarkable network lifetime improvement.

The rest of this paper is organized as follows: Section II covers related work in this area; Section III describes the network model and explains the unbalanced energy consumption problem; Section IV presents the unequal clustering algorithm and inter-cluster multihop routing protocol in detail; Section V analyzes some properties of the EEUC algorithm; Section

VI details our simulation efforts and the analysis of the results obtained; Section VII concludes this paper with directions for future work.

II. RELATED WORK

Many clustering algorithms have been proposed for wireless sensor networks in recent years. We review some of the most relevant papers [4]–[7], [9], [10].

In LEACH [4], each node has a certain probability of becoming a cluster head per round, and the task of being a cluster head is rotated between nodes. In the data transmission phase, each cluster head sends an aggregated packet to the base station by single hop. In PEGASIS [6], further improvement on energy-conservation is suggested by connecting the sensors into a chain. To reduce the workload of cluster heads, a two-phase clustering (TPC) scheme for delay-adaptive data gathering is proposed in [7]. Each cluster member searches for a neighbor closer than the cluster head within the cluster to set up an energy-saving and delay-adaptive data relay link. HEED [5] extends LEACH by incorporating communication range limits and intra-cluster communication cost information. The initial probability for each node to become a tentative cluster head depends on its residual energy, and final heads are selected according to the cost. In the implementation of HEED [8], multihop routing is used when cluster heads deliver the data to the data sink. All these methods require re-clustering after a period of time because of cluster heads' higher workload.

However, few work has considered the hot spots problem when multihop forwarding model is adopted during cluster heads transmitting their data to the base station. In [9], an unequal clustering model is first investigated to balance the energy consumption of cluster heads in multihop wireless sensor networks. The work focuses on a heterogeneous network where cluster heads (super nodes) are deterministically deployed at some precomputed locations, thus it's easy to control the actual sizes of clusters. Through both theoretical and experimental analyses, the authors show that unequal clustering could be beneficial, especially for heavy traffic applications. A similar problem of unbalanced energy consumption among cluster heads also exists in single hop wireless sensor networks. Cluster heads farther away from the base station have to transmit packets over longer distances than those of heads closer to the base station. As a result, they will consume more energy. In EECS [10], a distance-based cluster formation method is proposed to produce clusters of unequal size in single hop networks. A weighted function is introduced to let clusters farther away from the base station have smaller sizes, thus some energy could be preserved for long-distance data transmission to the base station.

Many energy-aware multihop routing protocols have also been proposed for wireless sensor networks. According to different application requirements, those protocols have different goals and characteristics. In [11], the directed diffusion data dissemination paradigm is proposed. It is based on data-centric routing where the data sink broadcasts the interest. When the sensor has data for the interest, it sends the data along the aggregation tree to the sink. In [12], Gradient-Based Routing (GBR) is proposed as a variant of directed diffusion. Three different data dissemination techniques (stochastic, energy-based, and stream-based schemes) are presented to obtain a uniform distribution of the traffic throughout the whole network. However, these multihop routing protocols may not be applied to applications that require continuous data delivery to the data sink.

In [13], the authors investigate an optimization problem of transmission range distribution , *i.e.*, whether nodes can vary their transmission range as a function of their distance to the data sink and optimally distribute their traffic so that network lifetime is maximized. Simulation results show that energy balance can only be achieved at the expense of using the energy resources of some nodes inefficiently. This work reveals the upper bound of the lifetime of a flat sensor network and gives some valuable guidelines for designing multihop routing protocols for wireless sensor networks.

III. PRELIMINARIES

A. System Model

Let us consider a sensor network consisting of N sensor nodes uniformly deployed over a vast field to continuously monitor the environment. We denote the i-th sensor by s_i and the corresponding sensor node set $S = \{s_1, s_2, \ldots, s_N\}$, where $|S| = N$. We make some assumptions about the sensor nodes and the underlying network model:

1) There is a base station (*i.e.*, data sink) located far away from the square sensing field. Sensors and the base station are all stationary after deployment.
2) All nodes are homogeneous and have the same capabilities. Each node is assigned a unique identifier (ID).
3) Nodes needn't to be equipped with GPS-capable unit to get precise location information.
4) Nodes can use power control to vary the amount of transmission power which depends on the distance to the receiver.
5) Links are symmetric. A node can compute the approximate distance to another node based on the received signal strength, if the transmitting power is given.

We use a simplified model shown in [4] for the radio hardware energy dissipation. Both the free space (d^2 power loss) and the multi-path fading (d^4 power loss) channel models are used in the model, depending on the distance between the transmitter and receiver. The energy spent for transmission of a l-bit packet over distance d is:

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_o \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_o. \end{cases} \tag{1}$$

and to receive this message, the radio expends energy:

$$E_{Rx}(l) = lE_{elec}. \tag{2}$$

A sensor node also consumes E_{DA} (nJ/bit/signal) amount of energy for data aggregation. It's also assumed that the sensed information is highly correlated, thus the cluster head can always aggregate the data gathered from its members into a single length-fixed packet.

B. *The Problem of Unbalanced Energy Consumption*

In this paper, cluster heads closer to the base station act as routers of heads farther away from the base station during delivering data to the base station. The reason is that multihop communication is more realistic because nodes may not be able to communicate directly with the base station due to the limited transmission range. And even if a node can use power control to send data to a farther receiver, previous research (*e.g.,* [3]) has shown that it is obviously a waste of energy. However, the hot spots problem may arise in multihop wireless sensor networks. In a clustered sensor network, each cluster head spends its energy on intra-cluster and inter-cluster processing. The energy consumed in intra-cluster processing varies proportionally to the number of nodes within the cluster. Proposed clustering algorithms usually produce clusters of even size, thus the cluster heads tend to consume even amount of energy during the intra-cluster data processing phase. However, the heads closer to the base station consume much more energy during a data gathering circle because they have a higher load of relay traffic as compared to other heads. Thus, they will die much faster than the other cluster heads (*i.e.*, the hot spots problem), possibly reducing sensing coverage and leading to network partitioning.

A fundamental problem in wireless sensor networks is to maximize the network lifetime under given energy constraints. To achieve the goal, energy consumption must be well balanced among nodes. In homogeneous networks, the role of cluster head is usually periodically rotated among nodes to balance the energy dissipation. However, the hot spots problem cannot be completely avoided. The mainly goal of rotation is to balance the energy consumption among cluster heads and member nodes, thus it could hardly balance the energy consumption among cluster heads in the inter-cluster multihop routing scenario. We also argue that using node's residual energy as the only criterion when selecting cluster heads is not sufficient to balance energy consumption across the network. Selecting cluster heads with more residual energy can only be helpful to balance energy consumption among nodes in a localized area in the long term. It is ineffective to balance loads among different cluster heads to avoid the hot spots problem, if the cluster heads are uniformly distributed over the network like that in HEED. Because nodes closer to the base station still die faster, it cannot make efficient use of all nodes' energy.

Therefore, the primary objective of this paper is trying to wisely design the clustering and multihop routing scheme to extend the network lifetime. We adopt both the rotation of cluster heads and choosing cluster heads with more residual energy. Furthermore, we introduce a novel unequal clustering mechanism which is an effective method to deal with the hot spots problem. It can prevent the premature creation of energy holes in wireless sensor networks.

IV. THE EEUC MECHANISM

In the network deployment stage, the base station broadcasts a "hello" message to all nodes at a certain power level. By this way each node can compute the approximate distance to the base station based on the received signal strength. It not only helps nodes to select the proper power level to communicate with the base station, but also helps us to produce clusters of unequal size. Detailed descriptions of the unequal clustering algorithm and intra-cluster multihop routing protocol are in the following two subsections. Figure 1 gives an overview of the EEUC mechanism, where the circles of unequal size represent our clusters of unequal size and the traffic among cluster heads illustrates our multihop forwarding method.

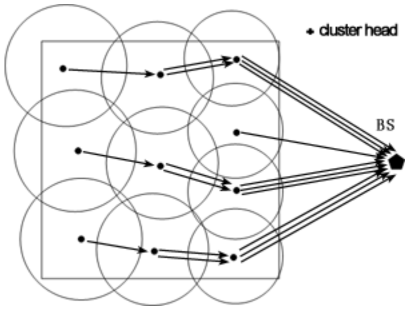

Fig. 1. An overview of the EEUC mechanism

A. *Unequal Clustering Algorithm*

Clustering a wireless sensor network means partitioning its nodes into clusters, each one with a cluster head and some ordinary nodes as its members. The task of being a cluster head is rotated among sensors in each data gathering round to distribute the energy consumption across the network. EEUC is a distributed cluster heads competitive algorithm, where cluster head selection is primarily based on the residual energy of each node. The pseudocode for an arbitrary node s_i is given in Figure 3.

First, several tentative cluster heads are selected to compete for final cluster heads. Every node become a tentative cluster head with the same probability T which is a predefined threshold. Other nodes keep sleeping until the cluster head selection stage ends. Suppose s_i becomes a tentative cluster head. s_i has a competition range R_{comp}, which is a function of its distance to the base station that we will explain later. Our goal is that if s_i becomes a cluster head at the end of the competition, there will not be another cluster head s_j within s_i's competition diameter. Figure 2 illustrates a topology of tentative cluster heads, where the circles represent different competition ranges of tentative cluster heads. In Figure 2 s_1 and s_2 can both be cluster heads, but s_3 and s_4 can not. Therefor the distribution of cluster heads can be controlled over the network. And the cluster heads closer to the base station should support smaller cluster sizes because of higher energy consumption during the inter-cluster multihop forwarding communication. Thus more clusters should be produced closer to the base station. That is to say, the node's competition radius should decrease as its distance to the base station decreases. We need to control the range of competition radius in the network. Suppose R^0_{comp} is the maximum competition radius which is predefined. We set R_{comp} of s_i as a function of its distance to the base station:

$$s_i.R_{comp} = (1 - c\frac{d_{max} - d(s_i, BS)}{d_{max} - d_{min}})R^0_{comp} \qquad (3)$$

where d_{max} and d_{min} denote the maximum and minimum distance between sensor nodes and the base station, $d(s_i, BS)$ is the distance between s_i and the base station, c is a constant coefficient between 0 and 1. According to equation 3, the competition radius varies from $(1 - c)R^0_{comp}$ to R^0_{comp}. As an example, if c is set to 1/3, $s_i.R_{comp}$ varies from $\frac{2}{3}R^0_{comp}$ to R^0_{comp} according to its distance to the base station.

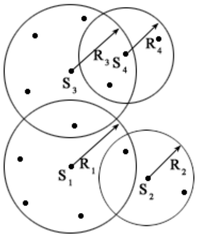

Fig. 2. The competition among tentative cluster heads

Each tentative cluster head maintains a set S_{CH} of its "adjacent" tentative cluster heads. Tentative head s_j is an "adjacent" node of s_i if s_j is in s_i's competition diameter or s_i is in s_j's competition diameter. Whether a tentative cluster head s_i will become a final cluster head depends on the nodes in $s_i.S_{CH}$ only, *i.e.*, the algorithm is distributed.

In the cluster head selecting algorithm, the broadcast radius of every control message is R^0_{comp}, thus s_i can hear all messages from node in its S_{CH}. In lines 5-6 of Figure 3, each tentative cluster head broadcasts a `COMPETE_HEAD_MSG` which contains its competition radius and residual energy. After the construction of S_{CH} has been finished in lines 10-13, each tentative cluster head checks its S_{CH} and makes a decision whether it can act as a cluster head in lines 14-26. Before deciding what its role is going to be, s_i needs to know what each node x in its S_{CH} such that $x.RE > s_i.RE$ has decided for itself. In case of a tie, the smaller node ID is chosen. In lines 15-17, once s_i finds that its residual energy is more than all the nodes in its S_{CH}, it will win the competition and broadcast a `FINAL_HEAD_MSG` to inform its adjacent tentative cluster heads. In lines 18-21, if s_j is in s_i's S_{CH} and s_i receives a `FINAL_HEAD_MSG` from s_j, s_i will give up the competition immediately, and inform all nodes in its S_{CH} by broadcasting a `QUIT_ELECTION_MSG`. In lines 22-25, if s_i receives a `QUIT_ELECTION_MSG` form s_j and s_j belongs to $s_i.S_{CH}$, s_i will remove s_j from its S_{CH}.

After cluster heads have been selected, sleeping nodes now wake up and each cluster head broadcasts a `CH_ADV_MSG` across the network area. Each ordinary node joins its closest cluster head with the largest received signal strength and then informs the cluster head by sending a `JOIN_CLUSTER_MSG`.

Algorithm 1: Cluster head Selection

```
1: $\mu \leftarrow RAND(0, 1)$
2: if $\mu < T$ then
3:     $beTentativeHead \leftarrow TRUE$
4: end if
5: if $beTentativeHead = TRUE$ then
6:     $CompeteHeadMsg(ID, R_{comp}, RE)$
7: else
8:     EXIT
9: end if
10: On receiving a COMPETE_HEAD_MSG form node $s_j$
11: if $d(s_i, s_j) < s_j.R_{comp}$ OR $d(s_i, s_j) < s_i.R_{comp}$ then
12:     Add $s_j$ to $s_i.S_{CH}$
13: end if
14: while $beTentativeHead = TRUE$ do
15:     if $s_i.RE > s_j.RE, \forall s_j \in s_i.S_{CH}$ then
16:         $FinalHeadMsg(ID)$ and then EXIT
17:     end if
18:     On receiving a FINAL_HEAD_MSG form node $s_j$
19:     if $s_j \in s_i.S_{CH}$ then
20:         $QuitElectionMsg(ID)$ and then EXIT
21:     end if
22:     On receiving a QUIT_ELECTION_MSG form node $s_j$
23:     if $s_j \in s_i.S_{CH}$ then
24:         Remove $s_j$ from $s_i.S_{CH}$
25:     end if
26: end while
```

Fig. 3. Cluster head selection pseudocode

A Voronoi diagram of sensor nodes is then constructed. Figure 4 shows an example of the clusters of unequal size, in which the base station is located at (100, 250). It is obvious that the cluster region closer to the base station is smaller than that farther from the base station.

The organization of intra-cluster data transmission is identical with LEACH after clusters have been formed, so we omit it in this paper.

B. Inter-cluster Multihop Routing

When cluster heads deliver their data to the base station, each cluster head first aggregates the data from its cluster members, and then sends the packet to the base station via multihop communication. In some proposed algorithms like PEGASIS, relay nodes can aggregate the incoming packets from other clusters together with its own packets. This assumption is unpractical because the degree of sensed data correlation between different clusters is comparatively low. In this paper, relay nodes don't aggregate the incoming packets. The routing problem here differs substantially from that of traditional ad-hoc wireless networks because of the many-to-one traffic pattern. On the other hand, neither query-driven nor event-driven routing protocols for wireless sensor networks can be applied to the cluster heads overlay. Thus we design an

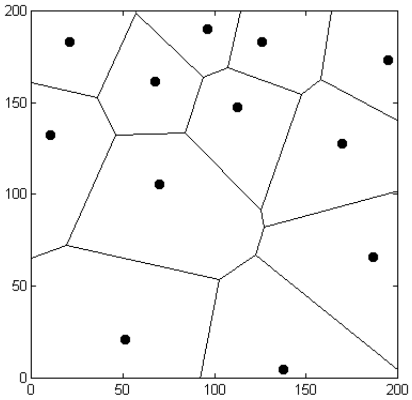

Fig. 4. Clusters formed as Voronoi cells around the selected cluster heads

energy-aware multi-hop routing protocol for the inter-cluster communication.

We introduce a threshold TD_MAX into our multihop forwarding model. If a node's distance to the base station is smaller than TD_MAX, it transmits its data to the base station directly; otherwise it should find a relay node which can forward its data to the base station. At the beginning of this process each cluster head broadcasts a message across the network at a certain power which consists of its node ID, residual energy, and distance to the base station. The concrete scheme of choosing the best relay node is explained as follows.

Cluster head s_i chooses a node to forwarding its data from its candidate set R_{CH}, which is defined as

$$s_i.R_{CH} = \{s_j | d(s_i, s_j) \leq k s_i.R_{comp}, d(s_j, BS) < (s_i, BS)\}.$$

k is the minimum integer that let $s_i.R_{CH}$ contains at least one item (if there doesn't exist such a k, define $s_i.R_{CH}$ as a null set, and s_i will send its own data together with forwarding data directly to the base station).

To reduce wireless channel interference, it's better to choose a adjacent node as the relay node. Thus we define the candidate set R_{CH} as the node's adjacent node closer to the base station. On the other hand, choosing a relay node with more residual energy helps balance the energy consumption to extend the network lifetime. However, only considering the residual energy may lead to a waste of network energy. Suppose s_i chooses s_j as its relay node. For simplicity, we assume a free space propagation channel model and s_j communicates with the base station directly. To deliver a l-length packet to the base station, the energy consumed by s_i and s_j is

$$\begin{aligned} E_{2-hop} &= E_{Tx}(l, d(s_i, s_j)) + E_{Rx}(l) + E_{Tx}(l, d(s_j, BS)) \\ &= l(E_{elec} + \epsilon_{fs} d^2(s_i, s_j)) \\ &\quad + lE_{elec} + l(E_{elec} + \epsilon_{fs} d^2(s_j, BS)) \\ &= 3lE_{elec} + l\epsilon_{fs}(d^2(s_i, s_j) + d^2(s_j, BS)) \end{aligned} \tag{4}$$

according to equation 1 and 2. Thus we define

$$d_{relay}^2 = d^2(s_i, s_j) + d^2(s_j, BS) \tag{5}$$

as the energy cost of the link. The bigger the d_{relay}^2 is, the more energy will be consumed in the relay process. Intuitively, when the node s_j is located straight along the way from s_i to the base station, it could save the network energy.

To reduce inefficiencies of energy consumption, a tradeoff should be made between the two criteria of residual energy and link cost d_{relay}^2. In our mechanism, s_i chooses s_j with more residual energy from the two smallest d_{relay}^2 nodes (if there exist) in $s_i.R_{CH}$ as its relay node.

After each cluster head has chosen a relay node or decided to transmit its data to the the base station directly, a tree rooted at the base station is constructed.

V. PROTOCOL ANALYSIS

This section presents the analysis of the unequal clustering algorithm. According to Algorithm 1, the cluster head selection process is message driven, thus we first discuss its message complexity.

Lemma 1: The message complexity of the cluster formation algorithm is $O(N)$ in the network.

Proof: At the beginning of the cluster head selection phase, $N \times T$ tentative cluster heads are produced and each of them broadcasts a `COMPETE_HEAD_MSG`. Then each of them makes a decision by broadcasting a `FINAL_HEAD_MSG` to act as a final cluster head, or a `QUIT_ELECTION_MSG` to act as an ordinary node. Suppose k cluster heads are selected, they send out k `CH_ADV_MSG`s, and then $(N-k)$ ordinary nodes transmit $(N-k)$ `JOIN_CLUSTER_MSG`s. Thus the messages add up to $2N \times T + k + N - k = (2T+1)N$ in the cluster formation stage per round, *i.e.*, $O(N)$. ■

Lemma 1 shows the message overhead of EEUC is small. In HEED, the upper-bound of message complexity is $N_{iter} \times N$ where N_{iter} is the number of iterations. Because we have avoided message iteration in the cluster head selection algorithm, the control message overhead in EEUC is much lower than that in HEED.

As described before, the threshold T determines the number of tentative cluster heads. Enough tentative cluster heads guarantee good head selection in terms of energy. On the other hand, too many tentative cluster heads cause a large message overhead. Thus proper value of T should be chosen in order to guarantee the quality of head selection and reduce the message overhead. In our previous work [10], the impact of T on the network lifetime is drawn via simulations.

Lemma 2: There is no chance that two nodes are both cluster heads if one is in the other's competition range.

Proof: Suppose s_j and s_k are both tentative cluster heads, and s_k is located within the circle of s_j's competition range. According to Algorithm 1, each node belongs to the other node's S_{CH}. If s_j first becomes a head node, then it will notice s_k its state, so s_k quits the competition and becomes an ordinary node; vice versa. ■

we simply analyze the impact of protocol parameters R_{comp}^0 and c on the network lifetime. According to equation 3, c dominates the unequal extent of the cluster sizes. The bigger c is, the bigger the range of competition radius is, and the greater difference the cluster sizes exhibit. When c is set 0, EEUC just performs as an equal clustering algorithm and cannot well balance the energy consumption among cluster heads. The number of clusters constructed in each round is determined by both R^0_{comp} and c. Intuitively, it decreases with the increase of R^0_{comp} when c is fixed, and it increase with the increase of c when R^0_{comp} is fixed. In order to balance the energy consumption well, R^0_{comp} and c should be properly set. Formulating the parameters for maximizing the network lifetime is left for future work.

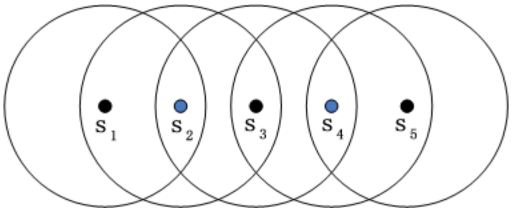

Fig. 5. A monotonic energy chain of five nodes

In order to decide whether it is going to be a cluster head or an ordinary node in Algorithm 1, each tentative node s_i waits for the decision of each node x in its S_{CH} such that $x.RE > s_i.RE$. Let's refer to Figure 5 to gain an insight into the problem of waiting time. Suppose $s_1.RE < s_2.RE < s_3.RE < s_4.RE < s_5.RE$, *i.e.*, they form an incremental energy chain. The following events will happen one after another: first s_5 claims that it is a final cluster head, so s_4 quits the competition, then s_3 announces that it wins the competition too, so s_2 decides to be a ordinary node, and at last s_1 becomes a cluster head. It takes four message steps for s_1 to make its decision in such a chain of five nodes. The example shows the waiting time depends on the the longest monotone energy chain. However, because the residual energy of tentative cluster heads is distributed randomly, the longer a monotone energy chain is, the smaller the probability is. In [14], the author analyzes a similar problem and points out that the waiting time depends on the energy topology of the network rather than on the number of nodes in the network.

VI. SIMULATION RESULTS

In this section,we evaluate the performance of the EEUC mechanism via simulations. First we study the cluster head characteristics of the unequal clustering algorithm, then we investigate how EEUC balances the energy consumption of the cluster heads and thus prolongs the network lifetime. For simplicity, an ideal MAC layer and error-free communication links are assumed. We calculate each node's energy consumption from data transmission and aggregation per round. We compare EEUC with LEACH and HEED. In our implementation of HEED, multihop routing is used during cluster heads delivering the data to the base station according to [8]. We also run extensive experiments to determine the optimal number of clusters to use in LEACH, and the optimal cluster radius to use in HEED. The simulation parameters are given in Table 1, in which the parameters of radio model are the same as those in [4]. Unless otherwise specified, we set T to 0.4, R^0_{comp} to 90m, c to 0.5 in equation 3, and TD_MAX to 150m.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Network coverage	(0,0)∼(200,200)m
Base station location	(100,250)m
N	400
Initial energy	0.5 J
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/bit/m^2
ϵ_{mp}	0.0013 pJ/bit/m^4
d_o	87 m
E_{DA}	5 nJ/bit/signal
Data packet size	4000 bits

A. Cluster Head Characteristics

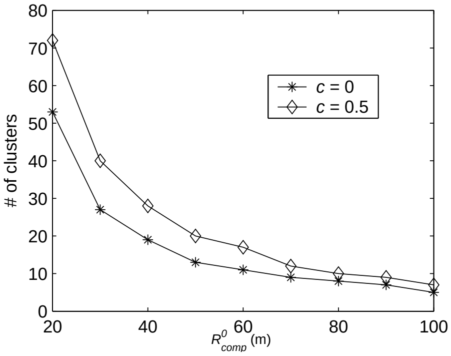

Fig. 6. The number of clusters generated by EEUC (c is fixed)

As we have explained in the previous section, the number of selected cluster heads varies according to the specified R^0_{comp} and c. Figure 6 shows the average number of cluster heads selected by EEUC. It testifies our analysis, *i.e.*, the smaller the competition radius, the larger the required number of cluster heads to cover the network. Notice that when r^0_{comp} is fixed and c increases, the competition radius decreases accordingly, thus EEUC generates more clusters when c is set to 0.5 as shown in the figure. Since each cluster head is responsible for aggregating the data from its cluster members into a single length-fixed packet, only one data packet needs to be delivered to the base station out of a cluster. Thus the more clusters are present, the more messages need to be delivered to the base station, resulting in overall energy consumption increases. In [4], the authors give an estimation of the optimum number of clusters in single hop networks. However, it cannot be applied in the unequal clustering mechanism proposed in this paper. Deriving some best values of R^0_{comp} and c for optimizing the network lifetime is left for future work.

We also examine the stability of our clustering algorithm. Figure 7 shows the distribution of the number of clusters

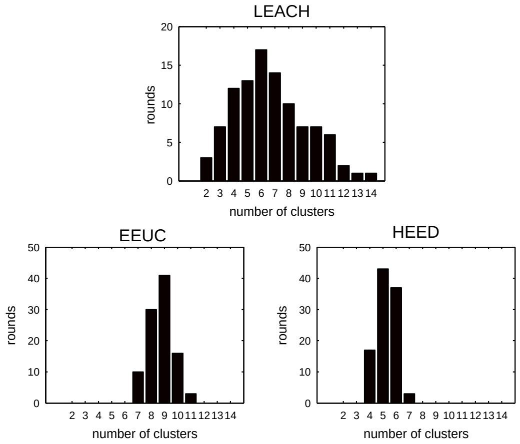

Fig. 7. Distribution of the number of clusters in each round

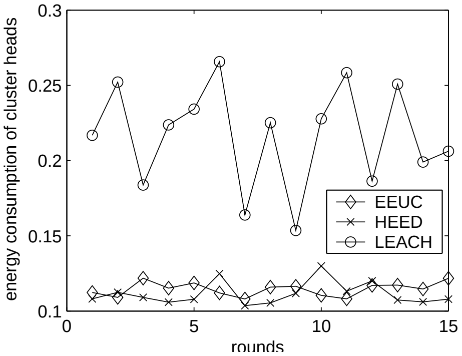

Fig. 8. The amount of energy spent by cluster heads

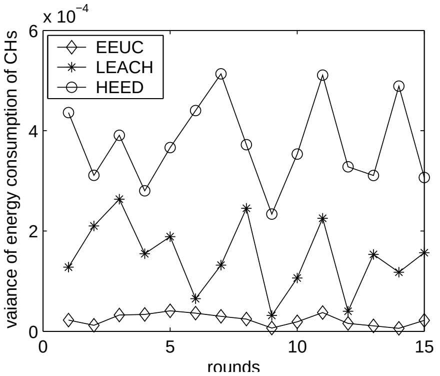

Fig. 9. The variance of amount of energy spent by cluster heads

in EEUC, HEED, and LEACH, which is calculated from randomly selected 100 rounds of the simulation. It's apparent that the number of clusters in EEUC and HEED is more steady than that in LEACH. LEACH uses a fully random approach to produce cluster heads, thus it results in a fairly variable number of clusters, although the expected number of cluster heads per round is deterministic. In EEUC, a certain proportion of nodes voluntarily join the competition of cluster heads, thus the number of selected cluster heads won't be too small. On the other hand, according to Lemma 2 the number of selected heads won't be too large. As a matter of fact, the number of clusters using EEUC depends on the competition range of tentative cluster heads. Thus EEUC achieves a steady number of clusters. HEED also uses a number of iterations to produce a steady number of clusters. It is worth mentioning that EEUC generates more clusters than LEACH and HEED because it employs extra cluster heads to afford the multihop forwarding traffic in the area closer to the base station.

B. Energy Efficiency

In this part, we investigate the energy efficiency of EEUC. First, we compare the amount of energy spent by cluster heads in three algorithms. 15 rounds of simulations are sampled and the amount of total energy spent by all cluster heads is shown in Figure 8. The energy consumed by cluster heads per round in EEUC is much lower than that in LEACH, and is about the same as that in HEED. Because cluster heads send their packets to the base station via single hop in LEACH, the energy consumption is much higher. And because the distribution of selected cluster heads is uncontrollable in LEACH, there is a dramatically variation of energy consumption of the cluster heads. In EEUC and HEED, cluster heads transmit their data to the base station via multihop, thus a considerable amount of energy is saved. Due to the stability of cluster heads topology in the two methods, the amount of energy spent by cluster heads is almost the same in each round.

Second, we study how well the energy consumption is balanced among cluster heads in three algorithms. Figure 9 gives the variance of amount of energy spent by cluster heads in 15 randomly selected rounds. It shows that EEUC balances the energy consumption among cluster heads best, and HEED performs worst. In EEUC, the unequal clustering method and the energy-aware multihop routing protocol successfully balance the energy consumption between cluster heads. And as explained before, the variance of EEUC is very steady due to the stability of the clustering mechanism. The variance of HEED is even higher than LEACH, and there are mainly two reasons. Since there exist clusters with only a single node (the cluster head), the clusters are not well balanced in HEED. What's more, HEED doesn't consider the problem of unbalanced energy consumption among cluster heads caused by the hot spots problem.

Third, we verify the unequal clustering mechanism indeed extend the network time. As we explained earlier, c determines the difference of cluster sizes. Thus we observe the relation between c and the network lifetime via varying c from 0 to 1. The result is shown in Figure 10, which justifies our unequal clustering mechanism. When c increases from 0, the effect of the unequal clustering method becomes distinct. However, the lifetime decreases when c is too big; the reason is that too many clusters will be produced closer to the base station, and each of them will deliver a data packet to the base station, thus it causes a waste of energy. Therefore, there exists an optimal value of c if other parameters are given, which is about 0.5 in this experiment as shown in Figure 10.

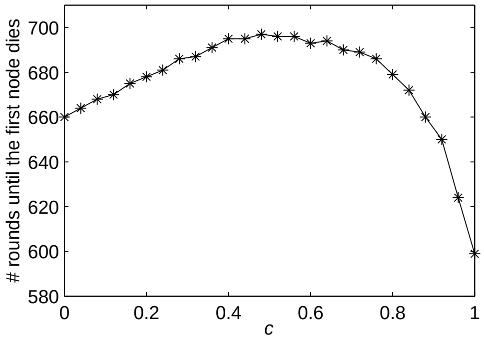

Fig. 10. The impact of c on the network lifetime in EEUC

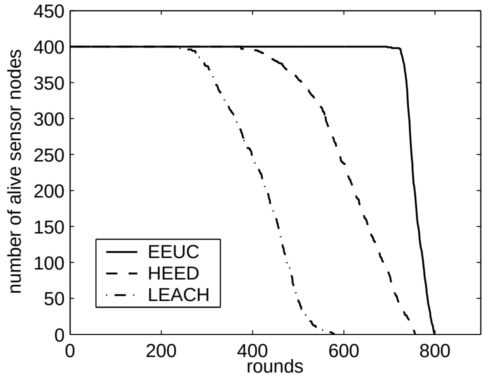

Fig. 11. The number of alive sensor nodes over time

Finally, we examine the energy efficiency of three algorithms by examining the network lifetime. Figure 11 shows the number of sensor nodes still alive over the simulation time. EEUC clearly improves the network lifetime (both the time until the first node dies and the time until the last node dies) over LEACH and HEED. In HEED, tentative cluster heads are randomly selected based on their residual energy. Therefore, sensors with low residual energy can still become cluster heads since it uses the intra-cluster communication cost to select the final cluster heads. And the energy consumption of cluster heads is not well balanced as illustrated in Figure 9. Thus some nodes die too earlier in HEED. This is avoid in EEUC because energy consumption is well balanced among nodes. The small interval between the time until the first node dies and the time until the last node dies implies that EEUC has successfully solved the hot spots problem.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a novel energy-efficient clustering mechanism for WSNs. The hot spots problem appears when employing the multihop routing in a clustering approach. We argue that both the rotation of cluster heads and the metric of residual energy are not sufficient to balance the energy consumption across the network. To address the problem, we first introduce an unequal clustering mechanism to balance the energy consumption among cluster heads. Clusters closer to the base station have smaller sizes than those farther away from the base station, thus cluster heads closer to the base station can preserve some energy for the purpose of inter-cluster data forwarding. What's more, we propose an energy-aware multihop routing protocol for the inter-cluster communication. Simulation results show that our unequal clustering mechanism clearly improves the network lifetime over LEACH and HEED.

Parameters of our mechanism, such as R^0_{comp} and c in equation 3, and TD_MAX, can be tuned to optimize energy preservation. We will try to find a solution that could determine the optimal value of these parameters according to network scale in our future work.

ACKNOWLEDGEMENTS

The work is supported by China Jiangsu Provincial NSF grant (BK2005208), China 973 project (2002CB312002) and TRAPOYT award of China Ministry of Education, and US NSF grants CCR 0329741, CNS 0422762, and CNS 0434533.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks", *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102-114, 2002.

[2] B. Krishnamachari, D. Estrin, and S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks", in *Proceedings of IEEE Int'l Conference on Distributed Computing Systems Workshops (ICDCSW)*, 2002, pp. 575-578.

[3] V. Mhatre and C. Rosenberg, "Design Guidelines for Wireless Sensor Networks: Communication, Clustering and Aggregation", *Ad Hoc Networks*, vol. 2, no. 1, pp. 45-63, 2004.

[4] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-670, 2002.

[5] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks", *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 660-669, 2004.

[6] S. Lindsey, C. Raghavendra, K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics", *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924-935, 2002.

[7] W. Choi, P. Shah, and S. K. Das, "A Framework for Energy-Saving Data Gathering Using Two-Phase Clustering in Wireless Sensor Networks", in *Proceedings of Int'l Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS)*, 2004, pp. 203-212.

[8] iHEED. http://www.cs.purdue.edu/homes/fahmy/software/iheed/

[9] S. Soro and W. Heinzelman, "Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering", in *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 2005.

[10] M. Ye, C. F. Li, G. H. Chen, and J. Wu, "EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks", in *Proceedings of IEEE Int'l Performance Computing and Communications Conference (IPCCC)*, 2005, pp. 535-540.

[11] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: a Scalable and Robust Communication Paradigm for Sensor Networks", in *Proceedings of ACM Mobile Computing and Networking (MOBICOM)*, 2000, pp. 56-67.

[12] C. Schurgers and M. B. Srivastava, "Energy Efficient Routing in Wireless Sensor Networks", in *Proceedings of IEEE Military Communications Conference (MILCOM)*, 2001, vol. 1, pp. 357-361.

[13] M. Perillo, Z. Cheng, and W. Heinzelman, "On the Problem of Unbalanced Load Distribution in Wireless Sensor Networks", in *Proceedings of the IEEE GLOBECOM Workshops on Wireless Ad Hoc and Sensor Networks*, 2004, pp. 74-79.

[14] S. Basagni, "Distributed Clustering for Ad Hoc Networks", in *Proceedings of the 4th Int'l Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN)*, 1999, pp. 310-315.