A Novel Cluster-header Selection Method in Wireless Sensor Networks

SUNGWON CHUNG, BYUNGHWA LEE, JILONG LI, ICKSOO LEE, JINSUK PAK, NAMKOO HA, KIJUN HAN*
Department of Computer Engineering
Kyungpook National University,
1370 Sankyuk-dong, Book-gu, Daegu, KOREA

Abstract: Clustering has been proved to be an effective approach to organizing the wireless sensor network into a connected hierarchy for routing of sensing data. For efficient constructing in the actual circumstance, the distributed algorithm is more practical than the centralized algorithm. However, the distributed algorithm can not guarantee even distribution of cluster heads over the network, and it can cause wasteful energy consumption due to a lot of control messages needed to be exchanged among nodes. In this paper, we present a probabilistic scheme for selecting cluster heads based on only local information at each node without exchanging control packets among its neighbors. Computer simulation results show that our scheme offers a longer network life time than other existing schemes.

Key-Words: - Energy Efficient, Routing, Wireless Sensor Networks, Clustering, HEED

1 Introduction

Wireless sensor networks can be deployed to provide continuous surveillance over an area of interest, referred to as a sensor field. Wireless sensor nodes perform collaborative sensing via wireless communication channels to retrieve information about targets that appear in the sensor field. Higher-level decision making can then be carried out based on the information received from the sensor nodes. Thus, these networks can be deployed in inhospitable terrain or in hostile environments to provide continuous monitoring and information processing for a wide variety of applications.

Many wireless sensor nodes are limited in power, computation capabilities and memory. So they have low cost and small form factors; therefore, they can be deployed in large numbers with high redundancy. In this way, the network can be made fault-tolerant. But, since nodes are deployed in a redundant fashion, power is the most important resource for wireless sensor networks.

In this paper, we propose a clustering scheme which selects cluster heads in a distributed manner based on local information such as the number of neighbors at each sensor node and distribution status of cluster heads around itself. Our scheme can minimize the wasteful energy consumption by reducing unnecessary control packets, and can increase network lifetime as they are compared with other clustering schemes.

Section 2 describes the related works. Section 3 describes the proposed scheme. Section 4 shows the simulation results. Finally, Section 5 gives concluding remarks and directions for future work.

2 Literature review

The clustering scheme reduces the energy consumption by aggregating redundant data at the cluster head and reducing the number of data transmissions. The clustering schemes can be divided into two types: centralized and distributed approach.

The first one is a centralized approach whose typical examples are LEACH-C [6], BCDCP [7], and PEGASIS [8]. In this approach, a base station (sink) which has the whole information of the network designates cluster heads as shown in Fig. 1(a). Banerjee et al.[5] proposed a centralized technique that does not require knowledge of node locations. Their technique is based on constructing a spanning tree of clusters that is rooted at an observer and forcing a bound on the maximum and minimum cluster size. The strength of the centralized approach is that the optimal number of cluster heads is determined and cluster heads is optimally chosen in the sensing field. But, it is not suitable in large networks since the enormous information is needed for the base station to construct clusters and correcting the information causes high energy and time consumption. So, the distributed approach is more suitable in large scale networks.

In distributed approach, a node decides to join a cluster or become a cluster head based on the information obtained solely from its one-hop neighbors as illustrated in Fig. 1(b). Several distributed clustering techniques have been proposed in the literature. These techniques are either iterative or probabilistic.

In iterative clustering techniques, a node waits for a specific event to occur or certain nodes to decide their role (Cluster head or a member of Cluster) before making a decision. There are Baker et al.[9], GAF [10], SPAN [11], DCA [12], MAX-MIN D-Cluster [13], ACE [14].

The probabilistic approach for node clustering ensures rapid convergence while achieving some favorable properties, such as balanced cluster size. It enables all nodes to independently decide on its role while keeping the message overhead low. LEACH [15] is the first research to select cluster heads probabilistically. It assumes that all nodes can reach to sink in a one-hop and that load distribution is uniform among every node as depicted in Fig. 1(c). However, it cannot guarantee an even distribution of the cluster heads over the sensing field, and thus it does not offer a sufficiently long network lifetime. Kuhn et al.[16] presented a probabilistic scheme for selection of the cluster heads based on the degree of nodes. The HEED [17] protocol considers a multi-hop network and assumes that all nodes are equally important. Fig. 1(d) shows the concept of multi-hop routing. A node uses its residual energy as the primary parameter to elect itself to become a cluster head. It results in the uniformly distribution of the elected set of cluster heads across the network. And each node executes a constant number of iterations to become a cluster head. However, it needs too many iterations and broadcasting of control packets until it finally decides to become a cluster head.

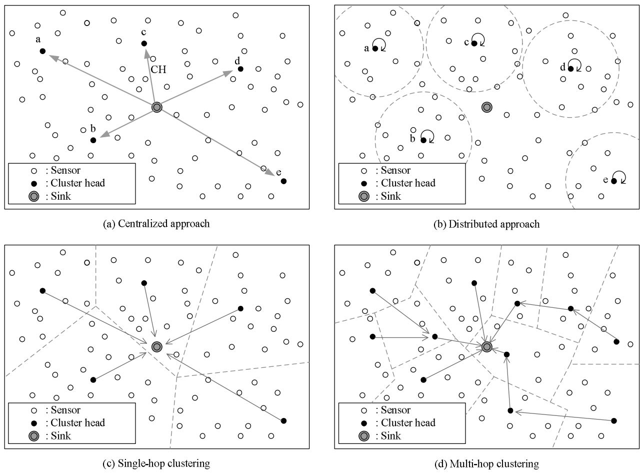

Fig. 1. Clustering techniques and routing types

3 Proposed schemes

In this section, we describe our scheme. First, we'll define some essential parameters needed to construct clusters. Next, we'll show the operation procedure of this protocol and algorithm. Third, we'll present the method to communicate between cluster heads.

We consider the following points to prolong the network lifetime.

1) Wasteful energy consumption in wireless sensor networks is mainly due to generating/handling control packets, idle listening, retransmitting owing to collisions, and overhearing.
2) The cluster heads should be created as few as possible to prolong the network lifetime, since the cluster head spends much energy for gathering and reporting data.
3) The node which has more energy should be more probable to be selected as a cluster head.

All subsequent discussions are based on the following assumptions.

- The ad-hoc sensor network is deployed randomly with sufficient nodes such that the network is connected. All sensor nodes are equipped with the same hardware, software, and initial energy level, i.e., the network consists of homogeneous sensor nodes.
- Each sensor node is left unattended after deployment. Therefore, battery recharge is not possible. It has the same energy E_{init} initially. After k round, its residual energy will be E_r.

- Each sensor node can control its transmission power flexibly. It has the same communication range for clustering r_c, and for routing between cluster heads r_t as depicted in Fig. 2.
- Each sensor node independently tries to become a cluster head.

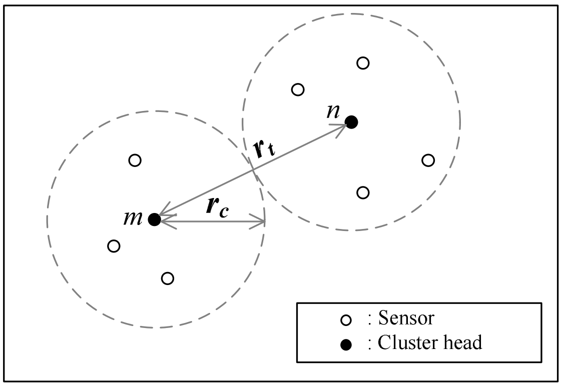

Fig. 2. Communication range (r_c) routing range (r_t)

3.1 Clustering parameters

Each round consists of two phases: Cluster Configuration Phase and Data Delivery Phase. Let the time of each phase be T_{CCP} and T_{DDP}, respectively. During T_{CCP}, each node decides its state (a cluster head or a member), and during T_{DDP}, the sensed data is delivered to the sink.

In our scheme, each node decides with a probability of p_{CH} whether or not it can become cluster head based on its local information. Local information includes the number of neighbors around itself ($\|N_i\|$), and how many times it has become cluster head during some interval (F_i) as shown in Fig. 3.

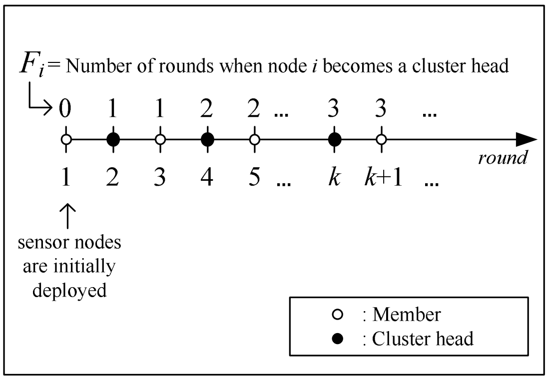

Fig. 3. The number of rounds when a node has become cluster head (F_i)

In the k-th round, the vaule of p_{CH} is given by

$$p_{CH} = \frac{1}{\|N_i\|} \cdot e^{-\frac{F_i}{k}} \qquad (1)$$

The first term of Eq. 1, $1/\|N_i\|$, means the necessary number of cluster heads within a cluster range (r_c) of node i. In other words, if there are n neighbors in the cluster range of node i, at least one cluster head would be needed to process the data in the cluster. So, the p_{CH} at node i will be $1/n$.

The second term, $e^{-\frac{F_i}{k}}$, dynamically adjusts the p_{CH}. It measures how many times each node has become a cluster head since the network is initially configured. A node which has had less opportunity will be given a higher F_i. Fig. 5 shows p_{CH} decreases exponentially as $\frac{F_i}{k}$ is increased. This gives every node almost the same qualification to become the cluster head in the long run.

To compute p_{CH}, node i makes a set of its neighbors, denoted by N_i, through a neighbor discovery procedure. And then it tries to become a cluster head with a probability of p_{CH}. After the attempt to become a cluster head, node i will be in one of the following three cases.

Case 1: The case when node i becomes a cluster head. In this case, it broadcasts *Advertisement message* to call for its members for construction of a cluster.

Case 2: The case when node i fails to become a cluster head. In this case, it hears *Advertisement messages* from cluster heads located within one-hop from it. At this time, it should decide which cluster it has to join. For this, we introduce a concept of intra-cluster communication cost (CC_i) which means the sum of squared distance between the cluster head and its neighbors. The energy consumption for transmitting data is proportional to square of distance between two nodes. Therefore, the energy consumption for transmitting data between cluster members in a cluster would be proportional to the total of squared distance between the cluster head and its members. In this aspect, we define an intra-cluster communication cost (CC_i) as follows

$$CC_i = \sum_{k=1}^{n} distance(i,k)^2 \qquad (2)$$

The value of CC_i is computed at each node when every round begins. For example, when node i has 7 neighbors (a,b,…,g) as shown in Fig. 4, it has CC_i = $distance(i,a)^2 + distance(i,b)^2 + \ldots + distance(i,g)^2$. If a node becomes a cluster head, its value of CC_i will be broadcast included in *Advertisement message*.

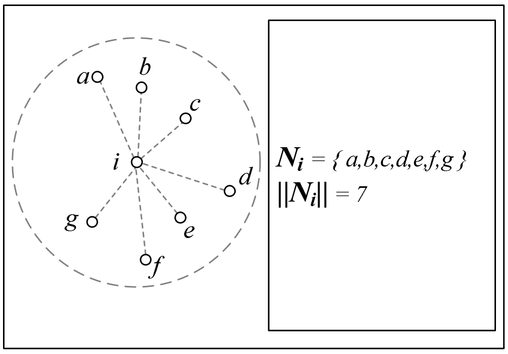

Fig. 4. A neighbor set (N_i) and its size

Let S_{CH_i} be a set of cluster heads within one-hop from node *i*. Then, on hearing *Advertisement messages*, it selects the one with the smallest CC_i in S_{CH_i} as its cluster head. This can contribute to reduction in total energy consumption.
Case 3: The case when node *i* as well as its neighbors fail to become a cluster head. In this case, S_{CH_i} is empty. At that time, it doubles the value of p_{CH}, and then repeats the same procedure until it fixes its state. If p_{CH} becomes greater than 1, the algorithm would be finished.

3.2 Inter-cluster communication method

After clustering, all data sensed by sensor nodes have to be routed to the sink. There are many researches how to deliver the sensed data efficiently to sink. It is generally assumed that the transmission range is fixed, and sensors deliver the data by transmitting the fixed range. However, in this paper, since the transmission range and clustering range are flexible, we have to fix the transmission range with respect to the clustering range. So, the transmission range depends on the clustering range. Because the routing in this paper means data delivery between cluster heads, the average distance between cluster heads is decided by the clustering range. According to the following lemma, a CDS (Connected Dominating Set) can be made if R_t is double greater than R_c.

Lemma 1. If R_t is greater than double of R_c, every cluster head could be connected.

Proof. We assume that sensing field is dense enough to construct clusters anywhere and every node is well distributed in the network. We divide the network into cells as shown in Fig. 5. In Fig. 5, if node A becomes a cluster head, its neighbors would become its members, and the others would try to become a cluster head. If node *B′* constructs its cluster far from the cluster of *A*, a blank area would be created between the clusters of *A* and *B′*. The nodes in this blank will form a cluster head through our algorithm. Then, the distance between two cluster heads (*A* and *B′*) is shorter than $2{\cdot}R_c$. So, the maximum distance between two cluster heads is $2{\cdot}R_c$. Therefore, if R_t is greater than $2{\cdot}R_c$, all cluster heads would be connected each other.

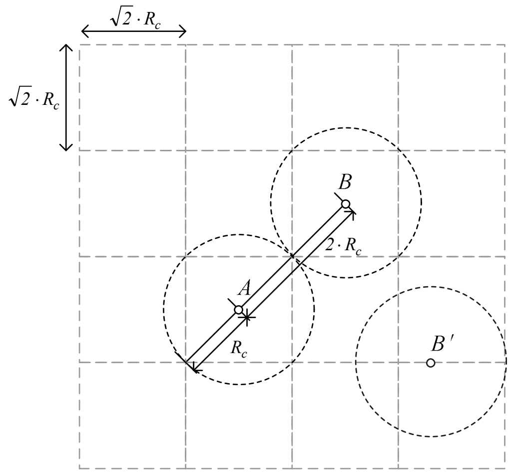

Fig. 5. Condition of transmission range to support network connectivity

4 Simulation results

We validate the performance of our scheme via computer simulation. Table 1 shows the simulation environment.

Table 1. Simulatioin parameters

Type	Parameter	Value
Network	Network size Sink position Initial energy	100 x 100 At (50, 50) 1 *J/battery*
Application	Cluster radius Data packet size Broadcast packet size	10 *m* 100 *bytes* 25 *bytes*
Radio model	E_{elec} e_{fs} e_{mp} e_{fusion}	50 *nJ/bit* 10 $pJ/bit/m^2$ 0.0013 $pJ/bit/m^4$ 5 *nJ/bit/signal*

We assume that 100 nodes are randomly dispersed into a field with dimensions 100m x 100m. For simplifying the problem, every node senses a phenomenon, reports the sensed data to its cluster head.

And then every cluster head aggregates the reported data, delivers the data to sink. A node is considered "dead" when it has lost 99.9% of its initial energy. We set the minimum probability for becoming a cluster head (p_{MIN}) to 0.0005 (which is reasonable for nodes with batteries of energy < 10 *Joule* and the same value of HEED). We use a CSMA/CA protocol for MAC. Every result shown is the average of 100 experiments. Each experiment uses a different randomly-generated topology.

Fig. 6 shows the network lifetime until every node dies. This figure indicates that our scheme offers a longer network lifetime than HEED since the energy is less consumed in our scheme mainly due to the reduced number of iterations and broadcastings of *Advertisement messages*. The number of *Advertisement messages* that should be broadcast in HEED will be at least 3 (greater than 2 in the main processing and 1 in the finalize processing) when p_{CH} < 1. When p_{CH} = 1, HEED requires at least 2 times of broadcastings. On the other hand, our scheme requires only one time of broadcasting of advertisement message at all times.

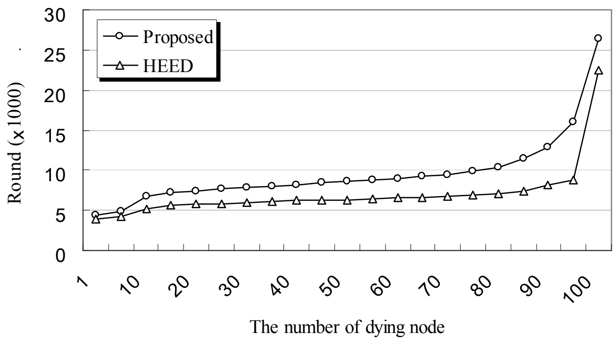

Fig. 6. The network lifetime until every node dies

5 Conclusions

In this paper, we presented a distributed, localized clustering approach for wireless sensor networks. In our scheme, the cluster heads are selected in a probabilistic manner without generating any unnecessary control packets. The probability of qualification for being cluster head at each node is based on the number of its neighbors and how many times it has been selected as cluster head in the past. We could reduce the wasteful energy consumption by eliminating necessity of transmissions of control packets and by giving fair opportunity to every node to become a cluster head.

References:

[1] K. Sohrabi et al., "Protocols for Self-Organization of a Wireless Sensor Network," *IEEE Pers. Commun.,* Oct. 2000, pp. 16–27.

[2] E. Shih et al., "Physical Layer Driven Protocol and Algorithm Design for Energy-Efficient Wireless Sensor Networks," *Proc. ACM MobiCom '01,* Rome, Italy, July 2001, pp. 272–86.

[3] A. Woo, and D. Culler, "A Transmission Control Scheme for Media Access in Sensor Networks," *Proc. ACM MobiCom '01,* Rome, Italy, July 2001, pp.221–35.

[4] Cerpa, A., Estrin, D., "ASCENT: Adaptive Self-Configuring sEnsor Networks Topologies", *Mobile Computing, IEEE Trans.,* vol.3, Issue 3, July-Aug. 2004, pp. 272 - 85

[5] S. Banerjee and s. Khuller, "A Clustering Scheme for Hierarchical Control in Multihop Wireless Networks," *Proc. IEEE INFOCOM,* Apr.2001, pp. 1028-37

[6] Heinzelman, W.B., Chandrakasan, A.P., Balakrishnan, H., "An Application-specific protocol architecture for wireless microsensor networks" *Wireless Communications, IEEE Trans.,* vol. 1, Issue 4, Oct. 2002 , pp. 660 - 70

[7] Muruganathan, S.D., Ma, D.C.F., Bhasin, R.I., Fapojuwo, A.O., "A centralized energy-efficient routing protocol for wireless sensor networks", *Communications Magazine*, *IEEE*, vol. 43, Issue 3, Mar. 2005, pp. 8 - 13

[8] Lindsey, S., Raghavendra, C., Sivalingam, K.M., "Data gathering algorithms in sensor networks using energy metrics", *Parallel and Distributed Systems, IEEE Trans*., vol. 13, Issue 9, Sept. 2002, pp. 924 - 35

[9] D. J. Baker and A. Ephremides, "The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm," *IEEE Trans. Commun.,* vol. 29, no. 11, 1981, pp. 1694–701.

[10] Y. Xu, J. Heidemann, and D. Estrin, "Geography-Informed Energy Conservation for Ad Hoc Routing," *Proc. ACM MOBICOM,* Rome, Italy, July 2001, pp. 70–84.

[11] B. Chen et al., "SPAN: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks," *ACM Wireless Networks,* vol. 8, no. 5, Sept. 2002, pp. 481–94.

[12] S. Basagni, "Distributed Clustering Algorithm for Ad Hoc Networks," *Proc. Int'l. Symp. Parallel Architectures, Algorithms, and Networks,* 1999, pp. 310–15.

[13] A. D. Amis et al., “Max-Min D-Cluster Formation in Wireless Ad Hoc Networks,” *Proc. IEEE INFOCOM,* Mar. 2000, pp. 32–41.
[14] H. Chan and A. Perrig, “ACE: An Emergent Algorithm for Highly Uniform Cluster Formation,” *Proc. 1st Euro. Wksp. Sensor Networks,* Jan. 2004, pp. 154–71.
[15] Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H., “Energy-efficient communication protocol for wireless microsensor networks,” *System Sciences, 2000. Proc. of the 33rd Annual Hawaii International Conference on,* 10 pp. vol.2, Jan. 2000
[16] F. Kuhn, T. Moscibroda, and R. Wattenhofer, “Initializing Newly Deployed Ad Hoc and Sensor Networks,” *Proc. ACM MOBICOM,* Sept. 2004, pp. 260–74.
[17] 0. Younis and S. Fahmy, “Distributed Clustering in Ad Hoc Sensor Networks: A Hybrid, Energy-Efficient Approach,” *Proc. IEEE INFOCOM,* Hong Kong, Mar. 2004; an extended version appeared in *IEEE Trans. Mobile Comp.,* vol. 3, no. 4, Oct.–Dec. 2004, pp. 366–79.