Multi-Layer Clustering Method Using Candidate Cluster Head Node in Wireless Sensor Network Environment

Youngbok Cho, Sangho Lee and Kyoungsoon Hwang*

Network Security Lab, Chungbuk National University, 410 Seongbong-ro, Heungdeok-gu, Cheongju, Chungbuk, 361-763, Republic of Korea,
*Chungbuk National University, BK21 410 Seongbong-ro, Heungdeok-gu, Cheongju, Chungbuk, 361-763, Republic of Korea
{bogicho,shlee,hks@cbnu.ac.kr}

Abstract

Wireless Sensor Network (WSN) is most important present minimizing energy consumption. Among various clustering methods, the Low-Energy Adaptive Clustering Hierarchy (LEACH) mechanism takes the hierarchical approach of segmenting multiple clusters for efficient energy management. The mechanism, however, configures new clusters in every round, so the energy consumed whenever configuring clusters shortens the useful lifetime of the entire network. For this reason, this paper generates clusters and selects candidate cluster head (CCH) in the initial round under the sensor network environment. Subsequent rounds continue, without performing re-clustering, until all candidate cluster heads selected become cluster head (CH), thereby addressing the issue of energy consumption in the setup phase for clustering. The proposed model consumes around 30% more energy than the conventional LEACH in the initial round, but its total energy consumption declines in as the round continues. The network simulation tool (NS-2) proves that its energy efficiency improves by up to 13.3% in the 1,000-node environment compared to when 100 sensor nodes are employed.

Keywords: Sensor Network, Clustering, Hierarchical, LEACH, Energy Efficient, Sensor Node

1. Introduction

In a wireless sensor network environment, the energy of sensor nodes has huge impacts on the life-cycle of the entire network, so it would be the most important resource here. The energy consumption of sensor nodes consists of various factors, but most of it is attributable to data packet transmission/reception, with individual nodes consuming enormous amounts of energy during the process of delivering packets to the base station. Nodes in close proximity to the base station consume lots of energy as they should not just deliver their own data but also relay the data from other nodes. Against this backdrop, a wide variety of methods have been proposed with an aim of efficiently managing the energy consumption of sensor nodes, with the representative of which being the clustering technique [1, 2]. The clustering technique divides the network into areas called "clusters" and incorporates all the nodes existing on the network into each cluster. The nodes included in the clusters are called member nodes, and individual member nodes communicate via cluster head without directly contacting
the base station. A cluster head puts together numerous data collected from member nodes and delivers them to the base station as the final destination. Various clustering techniques have been suggested, including the circulation of cluster head roles to make the energy consumption rate even; effective cluster head selection; and the consideration of the size of clusters and the number of member nodes within the clusters [3, 4, 5]. If the network is extended, however, limitations in the transmission distance of nodes further increases energy consumption not just for cluster heads but also for member nodes within clusters, shortening the useful lifetime of the network. Clustering techniques generate imbalanced clusters depending on cluster head selection methods or cannot guarantee that all sensor nodes consume energy in a balanced manner.

Furthermore, energy consumption deriving from frequent cluster generation may use up sensor node energy even without data delivery [6]. Therefore, the proposed method selects candidate cluster heads in the wireless sensor network to minimize energy consumption from frequent re-clustering; it also reduces distance-based communication costs taken into account for communication with the base station to extend the survival time of the network. In previous studies [3]-[7], sensor nodes consumed energy without data transmission as clusters were reconstructed and cluster heads reselected in every round. To tackle this issue, the proposed method performs clustering in the initial network environment and selects cluster heads considering the base station, neighboring nodes and residual energy. Here cluster heads are altered by comparing their critical values, instead of the conventional method of selecting candidate cluster heads in the course of cluster head selection and replacing cluster heads in each round. Communication may be possible in some clusters for several rounds without changing cluster heads, while certain nodes may change cluster heads on the basis of candidate nodes selected. The proposed method saves sensor node energy by 5.4% compared to the conventional method in [3]-[7], and the reduced energy consumption provides 13.3% greater efficiency in maintaining the entire network. The remainder of this paper is organized as follows: In Section 2, we briefly introduce the related work for the LEACH protocol where re-clustering is carried out in every round. In Section 3, we describe our proposed method of candidate cluster head based clustering. We present the simulation and analysis results for performance evaluation in Section 4. Section 5 concludes our paper.

2. Related Works

2.1 LEACH

In the LEACH protocol, nodes belong to a cluster where a single node serves as cluster head [4]. All ordinary nodes send data they detect to the cluster heads, and the cluster heads receive and incorporate the data and send them to the base station. For this reason, cluster head nodes consume more energy than ordinary nodes [10]. Figure 1 visualizes the operation process of a LEACH protocol.
In this case, the energy consumption of nodes selected as cluster head is greater than that of ordinary nodes, so their time of cluster maintenance becomes shorter. To address this problem, the proposed method ensures all sensor nodes reselect cluster heads fairly and justly in each round. A protocol focusing on the energy efficiency of the entire network, the LEACH protocol transmits data in a dispersed cluster structure that equalizes energy consumption across all nodes and thereby minimizes the network survival time. Cluster head selection is determined by the probability function described in Equation (1).

\[
T(n) = \begin{cases} 
  p & : \text{if } n \in G \\ 
  1 - p(r \mod \frac{1}{p}) & \\
  0 & : \text{otherwise}
\end{cases}
\]  

(1)

In Equation (1), \( p \) is the proportion of cluster heads selected among all nodes, \( r \) the current round and \( G \) the combination of nodes not selected as cluster head during the \( 1/p \) round. Without considering cluster forms, LEACH ensures that the nodes are selected as cluster head equally during the \( 1/p \) round. In the real network, however, nodes are bound to have different levels of residual energy due to carrying communication distances and external environmental factors. This shows that cluster head selection through probabilistic methods like LEACH imposes limitations on the efficient use of energy. Also, the clustering technique holds data that non-cluster head nodes will be sending for a certain period of time while neighboring nodes sense similar data. The transmission of similar data results in unnecessary energy consumption not only for cluster heads but also for other individual member nodes. Therefore, the problems with the conventional LEACH protocol can be summarized as follows:

- First, the energy consumption rate of cluster heads varies by the location of base station. In other words, the amounts of energy consumed by cluster heads in close proximity to the base station may differ from those by cluster heads far away from the base station, leading to network imbalance.
Second, the locations of cluster heads are uneven as cluster heads are selected by a probabilistic method. Clusters become uneven in size as a result, leading to energy imbalance owing to node communication and thereby giving rise to the imbalance of the network as a whole.

2.2 The Amounts of Energy Consumption in the Cluster Based

The amounts of energy consumed in the clustering model are defined as in Equation (2).

\[
E_{Tx}(l,d) = E_{Tx-elec}(l) + E_{Tx-amp}(l,d) = \begin{cases} 
        l \cdot E_{elec} + l + e_{fs} \cdot l \cdot d^2, & d > d_0 \\
        l \cdot E_{elec} + l + e_{amp} \cdot d^4, & d \geq d_0 
\end{cases}
\]  

(2)

In Equation (2), \( E_{Tx} \) is the entire amount of energy needed for transmission, \( l \) the number of data bits for transmission, \( E_{Tx-elec} \) the amount of energy by bit consumed by the circuit, \( E_{Tx-amp} \) the amount of energy consumed for signal amplification and \( d \) the transmission distance. The wireless sensor network field is assumed to be free space, and \( d \) is assumed as \( d^2 \). In a cluster, the energy consumption of nodes differs by distance. Therefore, \( \Delta E_d \) is summarized into Equation (3) to calculate Equation (2) as the amount of energy consumed by distance.

\[
\Delta E_d = 2E_{Tx-amp}ld\Delta d + E_{Tx-amp}\Delta d^2
\]

(3)

In Equation (3), \( \Delta d \) is the amount of change in transmission distance. In the clustering technique, the size of clusters determines the transmission distance between member nodes. For this reason, energy consumption owing to data transmission between member nodes under the WSN environment is defined as Equation (3).

3. Multi-Layer Clustering Method based on Candidate Cluster Head Nodes

This paper suggests a candidate cluster heads based multi-layer clustering method to maximize the energy efficiency of nodes under the WSN environment. The proposed model performs clustering and selects cluster heads in the initial round, and cluster heads select the candidate cluster heads to be subsequently used among nodes within the cluster. Candidate cluster head selection is done on the basis of distance between cluster heads and nodes. Once candidate cluster heads are selected, critical cluster values are calculated to determine when to replace the initially selected cluster heads. If the selected cluster heads can no longer participate in communication, new cluster heads are selected among candidate nodes chosen in the initial round. If the critical values of initial cluster heads are smaller than the critical cluster values, the cluster heads are switched to candidate cluster heads. Figure 2 illustrates the round format of the proposed model, which conducts clustering, data collection and data delivery through the set-up process in a single round.
Figure 2. Round Format of the Proposed Model

Here cluster heads and candidate cluster heads are selected and the weights of cluster heads and candidate cluster heads calculated in the set-up phase. If the conditions for critical cluster head values initially selected in the steady round phase are unmet, new cluster heads are selected among candidate cluster heads. In order to address the problem of energy consumption for re-clustering and cluster head selection in each round, the proposed technique continues the steady round without any particular settings if the values exceed cluster head and candidate cluster heads weights, and start from the beginning round for clustering purposes if the values fall below the weights.

3.1 Clustering Phase

Clustering in the proposed method is divided into the two stages of physical clustering (PhC) and logical clustering (LoC) processes. PhC segments all nodes constituting the entire network by communication radius (R); here the size of the entire network is divided with node distribution area (N×N) as member node communication radius(R). Figure 3 visualizes the proposed method’s hierarchical clustering.

Figure 3. Hierarchical Structure of Proposal Scheme

Areas are divided by PhC, and the density of distributed nodes is calculated for each cluster. Node density calculation requires information on the neighboring nodes of each node. The distribution of neighboring nodes can be calculated with the relative distance between
nodes; the SP$_i$ value received from neighboring nodes is used as relative distance here. Vector \( V \) represents the distribution of neighboring nodes and is written as Equation (4).

\[
\forall u_j \in U, j = 1,2,...,N_{sp}, \quad \forall d_i \in D, \omega = |D|, i = 1,2,...,\omega
\]

\[
f(u_j) = \left[ \frac{\omega(u_j - SP_{min})}{SP_{max} - SP_{min}} \right], \quad d_{f(u_j)} = d_{f(u_j)+1}
\]

In Equation (4), \( U \) is the combination of SP$_i$ values collected from adjacent nodes \( u \), \( u_j \) an element of \( U \) and the SP$_j$ value of adjacent node \( j \), and \( N_{sp} \) the number of signals received from adjacent nodes. \( \omega \) is the number of intervals dividing the relative distance (i.e. the number of PhC clustering). \( SP_{max} \) and \( SP_{min} \) are the minimum and maximum values of SP$_i$ that can be received within the maximum radius of clusters. \( d_i \) represents node density for the \( i \)-th interval, with \( i \)-th derived using the \( f(u_j) \) function. If certain nodes become cluster heads, each interval divided by relative distance becomes a cluster radius interval. Nodes selected as cluster heads in the PhC clustering stage send ADV$_{CH}$ messages to member nodes. The cluster head nodes receive responses to ADV$_{CH}$ messages from member nodes and determine the density of member nodes. If the \( i \)-th weight of the cluster head \( (W_{C,R}) \) is greater than cluster weight \( (TW_{k,R}) \), this cluster undergoes the LoC clustering stage, which divides \( (r/N) \times (r/N) \) segmented in PhC clustering into \( ((r/2)/N) \times ((r/2)/N) \) and selects cluster heads. Figure 4 visualizes the PhC and LoC clustering processes.

**Figure 4. Clustering of WSN**

With \( (N_r) \) in Figure 4 as the basis, four physical clusters (PhC) are generated. As \( \epsilon \) has high node density, re-clustering is carried out to produce four logical clusters (LoC1-LoC4). Figure 4(b) shows the table structure of each member node after clustering is completed. Each node has a node identifier (NID), the amount of residual energy \( (E_{rem}) \), the distance from the base station (DBS) and information on neighboring nodes within one-hop distance (HopN).

### 3.2 The Selection of Candidate Cluster Head Node

Candidate cluster head selection is carried out in the clustering stage using the voronoi diagram. In a single cluster, a cluster head is based on the voronoi minimum distance matrix to the cluster boundary point for the purpose of CCH selection. Cluster heads calculate the
distance to member nodes and generate edge nodes including cluster heads as matrices. Edge nodes may contain different voronoi cells, so each sub-cluster completes a final matrix with the minimum distance from nodes to the cluster. The minimum distance matrix from the cluster head to the boundary point is based on [Lemmas. 1].

**Lemmas 1. Minimum distance from CH to boundary point**

\[
A = \frac{(x_1 - x_2)(y_2 - y_1) - (x_2 - x_1)(y_1 - y_2)}{2}
\]

\[
\text{MinDist}(CH, p) = \text{Min}(d(CH, b_i) + d(b_i, b_j) + d(b_j, p)) \quad (0 \leq i \leq n_1, 0 \leq j \leq n_2)
\]

\[
n_1 = \text{Boundary of Voronoi Cell with}
\]

\[
n_2 = \text{Boundary number of voronoi cell}
\]

In [Lemmas 1], A can search for minimum distance among all nodes from cluster head s to the cluster boundary point, and this minimum distance translates into the minimum distance matrix. In the given minimum distance matrix, \(d_{\text{min}}\) completes the highest-proximity table among nodes that are within one-hop distance from the cluster heads. Candidate cluster heads are selected using this highest-proximity table.

![Candidate Cluster Heads Selection](image)

**Figure 5. Candidate Cluster Heads Selection.** It visualizes the connection status of internal nodes, which include multiple voronoi cells among individual voronoi cells generated on the basis of [Def. 1]. As shown in Figure 5, candidate cluster heads are selected from nodes within the voronoi boundary from the minimum distance matrix, with seconds as the basis.

A can search for minimum distance among all nodes from cluster heads to the cluster boundary point, and this minimum distance translate into the minimum distance matrix. In the given minimum distance matrix, \(d_{\text{min}}\) completes the highest-proximity table among nodes that are within one-hop distance from the cluster heads. Candidate cluster heads are selected using this highest-proximity table.
3.3 Weight-based Cluster Header Replacement

To replace cluster heads, cluster head weight \((TW_{ch}R_i)\) and cluster weight \((TW_kR_i)\) are calculated. \((TW_{ch}R_i)\), \((TW_kR_i)\) are defined on the basis of [Lemmas. 2] and [Lemmas. 3].

[Lemmas. 2] Cluster Header Weight \((W_{ch}R_i)\)

In a certain round \((r)\), if \(n_j(ch)\) is the number of member nodes belonging to cluster header CH and \(a_{1(ch)}, a_{2(ch)} \ldots a_{d(ch)}\) the weight of individual nodes applied to cluster head CH, the weight of cluster head CH \((W_{ch}R_i)\) is defined as follows:

\[
W_{(CH)}R_i = n_j(CH) \times \sum_{k=1}^{d} a_k(CH)
\]

\(d : Number\ of\ Weight, \ 1 \leq j \leq r, \ r : Number\ of\ total\ CH\) \hspace{1cm} (5)

[Lemmas. 3] Critical Cluster Values \((TW(k)R_i)\)

In a random round \((r)\), \(k\) is the combination of all cluster heads, with CHj belonging to \(k(CH_j \in k)\). When cluster head weight is set at \((TW_kR_i)\), the sum of weights for \(k\) in all clusters, \((TW_kR_i)\), is defined as below:

\[
TW_{(K)}R_i = \sum_{k=1}^{n} W_{(CH)}R_i
\]

\(n : Number\ of\ Total\ Clustering\) \hspace{1cm} (6)

As specified in [Lemmas. 3], \((TW_kR_i)\) is calculated in the PhC and LoC clustering stages. LoC is calculated, as a response to a \(ADV_{ch}\) message, considering the density of neighboring nodes. When the total number of clusters is \(m\), for instance, the \(n_{th}\) cluster can be expressed as \(C_n\). Here the range of \(n\) is \((1<n<m)\). The specific radius interval of \(C_n\), whose cluster radius is \(R_i\), is expressed as \(R_{i,n}\). i represents the number of the segmented radius interval, and the range of \(i\) is \((1<i<\alpha)\). Node density varies by communication radius, so the cluster heads of LoC clusters, where node density is not taken into account, consume more energy than the cluster heads of other clusters. For this reason, this paper undertakes initial PhC clustering, with cluster heads calculating cluster member node density based on CHADV messages from member nodes. If node density is greater than the average, PhC clusters perform LoC clustering to prevent the intensive energy consumption of cluster heads. Instead of replacing cluster heads in each round, the proposed method uses cluster weight \((TW_kR_i)\) as the criterion for replacing cluster heads with candidate cluster heads. If the cluster head weight of \(W_{ch}R_i\) is smaller than the cluster weight, cluster heads are replaced with candidate cluster heads by order of priority.
Table 1. Pseudo Code for Algorithm of Clustering

<table>
<thead>
<tr>
<th>Algorithm Clustering</th>
</tr>
</thead>
</table>
| **Input**: Node_Table, Round, Node  
**Output**: Node Table  
**Method**: WEIGHT(c), ADDTable(nt,K) |
| Begin  
Node_Table *nt;  
For each Round R in CT  
total_weight=0;  
For each Node n in K  
if (n.count >0) then  
n.weight=WEIGHT(c);  
total_weight=total_weight+n.weight;  
end if  
end For  
if (total_weight < threshold) then  
ADDTable(nt,K);  
end if  
end For  
return nt;  
End |

3.4 Total Energy Consumption: $\text{Tot}_E$

For comparative analysis of total energy consumption, the amount of energy consumed in the conventional method is defined as Equation (7).

$$LEACH_Tot_E = \sum_{R \in \text{Round}} \left( \frac{LEACH}{Round} \right) \times N_R$$

(7)

In the proposed model, all member nodes send messages to cluster heads, which then deliver data to the base station. Therefore, the total amount of energy consumed in the proposed model is defined with Equation (8).

$$Proposal_Tot_E = \sum_{R \in \text{Round}} \left( \frac{ACK}{Round} \right) \times N_R \left( \frac{N_{tot}}{N_{CCH}} \right)$$

(8)

$R$ is round, $ACK$ is Acknowledge, $NR$ is number of round, $N_{tot}$ is number of node in total network and $N_{CCH}$ is number of candidate cluster head.

4. Simulation

A testing environment is set up to prove the energy efficiency of candidate cluster head based clustering in the proposed technique. In the test, the efficiency of the entire network’s lifecycle is demonstrated for each round.
4.1 Simulation Environment

The simulation environment and energy consumption model for evaluating the energy efficiency of the proposed candidate cluster head-based clustering method are described in Table 2.

<table>
<thead>
<tr>
<th>Item</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>Linux CentOS 5.3</td>
</tr>
<tr>
<td>CPU</td>
<td>Inter Core Quad CPU Q8200 1.86GHz</td>
</tr>
<tr>
<td>Memory</td>
<td>2GB RAM</td>
</tr>
<tr>
<td>Tool</td>
<td>NS-2</td>
</tr>
<tr>
<td>Network Size</td>
<td>100x100</td>
</tr>
<tr>
<td>Node Count</td>
<td>100,500,1000</td>
</tr>
<tr>
<td>Base Station</td>
<td>(0,0), (50,50)</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>2J</td>
</tr>
</tbody>
</table>

For simulation purpose, nodes are randomly distributed in 100m x 100m coordinates—with initial energy set at 2J and time for each round at 20 seconds—and the base station is set at (50, 50) for energy efficiency.

This paper increases the total number of nodes from 100 to 500 to measure the number of nodes participating in communication. In the proposed technique, the number of clusters and candidate cluster heads serves as a very important factor in the entire network, so the proportions of candidate cluster heads are set at 5%, to 20% of all nodes under the testing environment.

Test results indicate that, in WSN communication, the energy consumption rate of cluster heads varies by the location of base station. Figure 6 visualizes test results with the base station located at (50, 50) each. Setup phase is undertaken at base station (50, 50) in Figure 6(a) and Figure 6(b) to compare the respective amounts of energy consumption with the conventional LEACH protocol. Figure 6(c) and Figure 6(d) are represent energy consumption rates in five round and throughout 25th rounds, respectively. Test results indicate that the proposed technique consumes 30.6% more energy than the conventional method in the initial setup process of round 5. Its energy consumption, however, decreases by up to 32.8% in round 25; overall energy consumption over round 25 also declines by 33% compared to what is found in round 5.

The proportion of candidate cluster heads affects the energy consumption rate. Therefore, this paper increases the candidate cluster head rate from 5% to 20% in the testing process to find the optimal rate, and the results are presented in Figure 6. In Figure 6, the base station is situated in (50, 50) and candidate cluster head rates are set at 5% to 20%. Then energy consumption rates are calculated for 100 nodes Figure 6(a), 500 nodes Figure 6(b) and 1,000 nodes Figure 6(c). Figure 6(d) shows the number of nodes that it can participate in communication when 100 nodes and 500 nodes in total are tested with the candidate cluster head rate of 5%.

As shown in Figure 6(a), Figure 6(b) and Figure 6(c), the optimal candidate cluster head rate is set at 5%. Therefore, the proposed method is proved very effective in terms of energy efficiency when the number of nodes is large (e.g. WSN environment).
Figure 6. According to Location of Base Station Energy Consumption is Different

Based on these test results, the candidate cluster head rate is set at 5% and the nodes participating in communication in the proposed method are compared to those in ILEACH, as described in Figure 7.

Figure 7. Base Station Location, Depending on the Energy Consumption in the Setup Phase

For the test, the total number of nodes is fixed at 100 and 500 each, and the number of nodes participating in communication is measured by an increase in the number of nodes under the identical settings. In the proposed method, the number of rounds for 100 nodes is
about 45, representing approximately 5.6% improvement compared to the ILEACH method. For 1,000 nodes, however, the proposed method shows 13.3% improvements with 300 rounds.

5. Conclusion

The proposed technique suggests a candidate cluster head method to minimize energy consumption in the set-up stage for cluster head selection in the course of initial clustering. While conventional methods performed re-clustering in each round, the proposed technique enables communication without changing cluster heads for several rounds. As clusters are maintained for a certain period of time, the technique reduces energy consumption in the set-up phase, which would take place in each round, and thereby extends the lifecycle of the entire network. It also undertakes multi-layer clustering given the density of nodes to minimize the energy consumption of cluster heads. Test results show that the proposed method is very suitable for wireless sensor network with a large number of nodes. Compared to the ILEACH technique, it improves the entire energy efficiency rate by 5.4% for 100 nodes, 6.0% for 500 nodes and 13.3% for 1,000 nodes. Though it provides a longer survival time for the entire sensor network than ILEACH, the proposed method involves considerable energy consumption for cluster heads in the initial clustering process, but the overall energy consumption rate declines in subsequent rounds. Further research will be needed on ways to minimize energy consumption in the initial set-up process.

References


Authors

Youngbok Cho received the M.S. degree in the department of computer science from Chungbuk National University, Korea, in August 2003. She is currently in the process of a Ph.D. Her research interests include mobile ad hoc network and wireless sensor networks.

E-mail: bogicho@cbnu.ac.kr

Sang-Ho Lee received the M.S. degree in the department of computer science, Soongsil National University in February 1981. He received the P.H.D. degree in the department of computer science, Soongsil National University in February 1989. He is currently working professor in the department of electrical and electronics engineering, Chungbuk National University. His research interests also include Protocol Engineering, Network Security, Network Management and Network Architecture.

E-mail: shlee@chungbuk.ac.kr

Kyoungsoo Hwang received M.S. and Ph.D. from Dept. of Computer Science, Chungbuk National University, in 2003 and 2010 respectively. Her research interest includes machine learning, Bioinformatics, medical protocol description, and wireless sensor networks.

E-mail: hks@cbnu.ac.kr