EFFICIENT AND DISTRIBUTED CLUSTERING SCHEME WITH MOBILE SINK FOR HETEROGENEOUS MULTI-LEVEL WIRELESS SENSOR NETWORKS

ALAMI CHAIBRASSOU, AHMED MOUHSEN, ISMAIL LAGRAT

1 Research Laboratory in Mechanical, Industrial Management and Innovation, Faculty of Science and Technology, Settat, MOROCCO
2 National School of Applied Sciences, Khouribga, MOROCCO
E-mail: alami70@yahoo.fr, mouhsen.ahmed@gmail.com , i.lagrat@yahoo.fr

ABSTRACT
Wireless sensor nodes are typically powered by non-rechargeable batteries and deployed in an area of interest to supervise or monitor various phenomena (temperature, pressure, humidity...). So, constrained energy is the most important feature of wireless sensor networks (WSNs). In WSNs, saving energy and extending network lifetime are great challenges. Clustering is a key technique used to optimize energy consumption in WSNs. However, sensor networks with fixed sink node or Base Station (BS) often suffer from a hot spots problem since nodes near to sink have more traffic burden to forward during a multi-hop transmission process. The use of mobile sinks has been shown to be an effective technique to enhance network performance features such as latency, energy efficiency, network lifetime, etc.

In this paper an Efficient and Distributed Clustering Scheme with Mobile Sink (EDCSMS) for Heterogeneous Multi-level Wireless Sensor Networks is proposed. In EDCSMS, the cluster-heads (CHs) are elected by a probability based on the ratio between residual energy of each node and the average energy of the network, the BS moves towards each CH by a distance proportional to the weight of its probability. Furthermore, normal nodes select the optimal cluster-head based on the cost function.

Finally, Simulation using MATLAB software shows that our proposed protocol achieves longer lifetime, stability period and more effective messages to BS than LEACH, DEEC, EDCS in multi-level heterogeneous environments.

Keywords: Wireless Sensor Networks, BS Movement, Cost Function, Lifetime, Multi-Level Heterogeneous Environments.

1. INTRODUCTION
The advancement in the areas of Micro-Electro-Mechanical Systems (MEMS) and wireless communication technologies, have allowed the rapid development of wireless micro-sensors for wireless communications. The WSNs is composed by a large number of micro sensors called nodes communicating with each other through radio links independently and randomly distributed over an area of interest. nodes are powered by battery, which is impossible to get recharged after deployment. As a large part of energy is consumed when communications are established, so it is imperative to develop an energy efficient routing protocol, taking into account the constraints by these sensors (lifetime, Quality of Service etc.).

For this reason several routing protocols have been designed for wireless sensor networks to satisfy energy utilization and efficiency requirement. Efficiency, scalability and lifetime of wireless sensor network can be enhanced using hierarchical routing. Here, sensors organize themselves into clusters and each cluster has a cluster head [1]. The main role of the cluster head is to provide data communication between sensor nodes and the base station efficiently [2].

In WSN the BS can be either a mobile or a fixed node that connects the sensor network to other types of network such as Internet or satellite where reported data are accessible to the user [3].
However, the use of mobile sinks can potentially provide energy-efficient data collection with well-designed networking protocols for WSNs [4]. When using the mobile sink in practice, the sink nodes can be attached to vehicles, animals or people that can move inside the region of interest. Usually, static sink nodes are not very efficient [5], nodes located in the vicinity of the sink, depleting their energy much earlier compared to the nodes located farther away from the sink due to higher data-relaying load. Although single-hop data collection is feasible in networks deployed in small regions, the multi-hop transmission manner is more commonly used in large sensor areas [6]. Intuitively, mobile sinks gain advantages by mitigating the so-called hot spot problem, balancing energy among sensor nodes, prolonging network lifetime, reducing transmission latency and improving network performance by periodically accessing some isolated nodes into the network.

Extending network lifetime is the most important design issue in WSNs. In this paper we propose an EDCSMS for Heterogeneous Multi-level WSNs in which the use of mobile sink permits to minimize CH-to-BS transmission distances, thereby minimizing the energy consumption of the CHs nodes, which constitutes the majority of the energy consumption of the entire network. Hence the network’s lifetime is extended.

The remainder of the paper is structured as follows: Related work is introduced in Section 2. In section 3, Heterogeneous model for wireless sensor network. Section 4 describes The EDCSMS Protocol, section 5, Simulations and Results and section 6, Discussion. We draw the conclusion in Section 7.

2. RELATED WORK

There are two types of clustering schemes. The clustering algorithms applied in homogeneous networks are called homogeneous schemes, and the clustering algorithms applied in heterogeneous networks are referred to as heterogeneous clustering schemes. Many existing clustering techniques such as LEACH [7], PEGASIS [8], and HEED [9] etc., are designed for homogeneous sensor networks where all sensor nodes are equipped with the same battery energy, while DEEC [10], SDEEC [11], EEHC [12] and EDCS [13] etc. are considered as Heterogeneous schemes. In this part, the key features of the most popular and recent clustering algorithms are explained.

Low-Energy Adaptive Clustering Hierarchy (LEACH) is a distributed algorithm which makes local decisions to elect cluster-heads. It partitions the network into groups (clusters). Nodes transmit their data to representatives of groups called cluster heads (CH), which in turn send the data after processing and aggregation to the desired destination or base station. Therefore, LEACH includes randomized rotation of cluster head locations to evenly distribute the energy dissipation over the network. Its advantages can be summarized as follows: first, it can prolong the network lifetime compared to the original flat routing protocol. Second, Energy consumption is shared across all nodes, thus extending the lifetime of the network. However, LEACH protocol has some disadvantages. First, The use of probabilistic model to select CHs can generate CHs too close in an area of the network. Second, if the cluster head dies in round R, the whole cluster is unable to transfer its data to the base station until the next round.

Qing et al [10] proposed a distributed energy efficient clustering scheme for heterogeneous wireless sensor networks, which is called DEEC. In DEEC, the cluster heads are selected by a probability based on the ratio between residual energy of each node and the average energy of the network. The epochs of being cluster heads for nodes are different according to their initial and residual energy. The authors have assumed that all the nodes of the sensor network are equipped with different amount of energy, which is a source of heterogeneity. DEEC is also based on LEACH, it rotates the cluster head role among all nodes to expend energy uniformly and estimates the ideal value of network lifetime, which is used to compute the reference energy that each node should expend during a round. Simulation results show that DEEC achieves longer lifetime and more effective messages than LEACH, SEP and LEACH-E.

Zhen HONG et al. proposed an efficient and dynamic clustering Scheme (EDCS), for heterogeneous multi-level wireless sensor networks. It solves the drawback that the general routing protocols in homogeneous networks can not be directly applied to heterogeneous multi-level environments. The EDCS protocol focuses on energy heterogeneity. To guarantee the process of cluster head selection accurately, the average residual energy of network in the next round is estimated through the average energy consumption forecast in ideal state and the reference value of historical energy consumption simultaneously to determine the probability of which node will be a cluster head. Furthermore, analogous universal gravitation is introduced to make non-cluster head node join the cluster in terms of gravitation during
cluster formation. Simulation results have proved that EDCS is more suitable and efficient than LEACH, SEP, DEEC and EDFCM for multi-level heterogeneous WSNs.

3. HETEROGENEOUS MODEL FOR WSNs

There are three common types of resource heterogeneity in sensor nodes: computational heterogeneity, link heterogeneity and energy heterogeneity[14]. In this paper we just consider heterogeneous WSNs.

In multi-level heterogeneous WSNs, each node is equipped with a random initial energy over the interval of \( [E_0, E_0 (1+\lambda)] \), where \( E_0 \) is the lower bound and \( \lambda \) is a constant \((\lambda > 0)\) which determines the value of the maximal initial energy. Usually, the heterogeneous network will become homogeneous when \( \lambda = 0 \). Mathematically, we can express our network as Set \( S = \{ s_i | s_i = (x_i, y_i), s_i \in \mathbb{R}^2, i = 1, 2, \cdots , N \} \).

Each sensor node \( s_i \) is equipped with initial energy \( E_0 (1 + \lambda_i) \), which is \( \lambda \) times more energy than the lower bound \( E_0 \).

The total initial energy of the multi-level heterogeneous networks is given by:

\[
E_{\text{total}} = \sum_{i=1}^{N} E_0 (1 + \lambda_i) = E_0 \sum_{i=1}^{N} (1 + \lambda_i)
\]  

From equation (1), the total initial energy in a heterogeneous WSNs can be treated as \( \sum_{i=1}^{N} (1 + \lambda_i) \) nodes which are equipped with initial energy of \( E_0 \) in a homogeneous WSNs.

4. THE EDCSMS PROTOCOL

In this section, we present details of our EDCSMS protocol. Our proposed protocol implements the same idea of probabilities for CHs selection based on initial, remaining energy level of the nodes and average energy of network as supposed in DEEC[10].

4.1 Estimating Average Energy of Network

Let us assume the ideal scenario where all sensor nodes are uniformly distributed and will die at the same time as a result of load balancing. The average energy of \( r \) round from [10] is given by:

\[
\bar{E} = \frac{1}{N} E_{\text{total}} \left( 1 - \frac{r}{R} \right)
\]  

Where \( R \) is the total rounds of the network lifetime. and can be estimated from [10] as:

\[
R = \frac{E_{\text{total}}}{E_{\text{round}}}
\]  

where \( E_{\text{round}} \) is the energy dissipated in the network during single round. We use the same radio energy dissipation model that was proposed in [7].

The total energy dissipated \( E_{\text{round}} \) can be approximated to:

\[
E_{\text{round}} = L \left( 2NE_{\text{elec}} + NE_{\text{DA}} + k_{\text{imp}} \cdot d_{\text{BS}} \cdot N_{\text{elec}} \cdot d_{\text{toCH}}^2 \right)
\]  

Where,

\( E_{\text{elec}} \) : Energy dissipation to run the radio

\( K \) : Number of clusters.

\( E_{\text{DA}} \) : Data aggregation cost expended in CH.

\( d_{\text{BS}} \) : Average distance between CH and BS.

\( d_{\text{toCH}} \) : Average distance between cluster members and CH.

\( L \) : Packet size.

Assuming that the nodes are uniformly distributed, we can get [15]:

\[
d_{\text{BS}} = \sqrt{x^2 + y^2}, \quad d_{\text{toCH}} = \sqrt{\int (x^2 + y^2)p(x,y)dx\,dy} = \frac{M}{\sqrt{\pi}}
\]  

We can find the optimum number of cluster \( k_{\text{opt}} \) by setting the derivative of \( E_{\text{round}} \) with respect to \( k \) to zero, we can get:

\[
k_{\text{opt}} = \frac{\sqrt{\pi}}{\sqrt{2\pi}} \frac{M}{\sum_{i=1}^{N} d_{\text{toBS}}(x_i, y_i)}\frac{\Delta y}{\Delta x}
\]  

The optimal probability of a node to become a cluster head, \( p_{\text{opt}} \), can be calculated as follows:

\[
p_{\text{opt}} = \frac{k_{\text{opt}}}{N}
\]  

4.2 Cluster Head Selection

At start of each round, node \( s_i \) decides whether to become a CH or not based on threshold calculated by the following equation and as supposed in [7][10]:

\[
T(s_i) = \begin{cases} 
\frac{p_i(s_i)}{1 - p_i(s_i)} \left( \frac{1}{p_i(s_i)} \right) & \text{if } s_i \in G \\
0 & \text{otherwise,} 
\end{cases}
\]
where, G is the set of nodes eligible to become CH for round r and $p_i(s_i)$ is the desired percentage of CH in real scenarios. And $p_i(s_i)$ can be calculated as follows:

$$p_i(s_i) = \frac{N_{opt} \cdot E_{res}(i)}{(N+\sum_{k=1}^{i} E_{r}(k))} \cdot E_{r}(r)$$

(10)

where, $E_{res}(i)$ is the residual energy of node $s_i$ in $r^{th}$ round.

In each round $r$, when node $s_i$ finds whether it is eligible to be a cluster head, it will choose a random number between 0 and 1. If the number is less than the threshold $T(s_i)$, then node $s_i$ becomes a cluster head during the current round.

Let $n_i = 1/p_i$ denote the number of rounds to be a cluster head for the node $s_i$, and we refer to it as the rotating epoch.

$$n_i = \frac{(N+\sum_{k=1}^{i} E_{r}(k)) \cdot E_{r}(r)}{N_{opt} \cdot E_{res}(i)}$$

(11)

From equation (11) the rotating epoch $n_i$ of each node fluctuates around its reference epoch $I_i$ based on the residual energy $E_{res}(i)$. If $E_{res}(i) > E_{r}(r)$, we have $n_i < I_i$, and vice versa. That means the nodes with more energy will have more chances to be the cluster head than the nodes with less energy due to different initial energies and energy dissipated every round in such a heterogeneous environment. Thus the energy of network is well distributed in the evolving process.

### 4.3 BS Movement

In this section, we consider a simple example in a two-dimensional space to illustrate the BS movement strategy in our approach. In the following example, we consider four CHs and show their initial parameters (See table1) and initial BS location to be at (50,50).

**Table 1: Example Parameter.**

<table>
<thead>
<tr>
<th>Cluster Head (CH)</th>
<th>Location</th>
<th>CH Probability ($p_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH1</td>
<td>(10,10)</td>
<td>0.2</td>
</tr>
<tr>
<td>CH2</td>
<td>(20,40)</td>
<td>0.6</td>
</tr>
<tr>
<td>CH3</td>
<td>(50,70)</td>
<td>0.8</td>
</tr>
<tr>
<td>CH4</td>
<td>(80,100)</td>
<td>0.1</td>
</tr>
</tbody>
</table>

First the vectors between a BS and all the CHs must be calculated as shown in Figure 1.

$$V_1 = <10-50>,(10-50) = <-40,-40>$$
$$V_2 = <40-50>,(20-50) = <-10,-30>$$
$$V_3 = <50-50>,(70-50) = <0,20>$$
$$V_4 = <80-50>,(100-50) = <40,50>$$

Next step is to multiply each vector by the corresponding CH’s probability.

$$V_1' = <-40,-40> \times 0.2 = <-8,-8>$$
$$V_2' = <-10,-30> \times 0.6 = <-6,-18>$$
$$V_3' = <0,20> \times 0.8 = <0,16>$$
$$V_4' = <40,50> \times 0.1 = <4,5>$$

Next we add the resulting vectors to produce a net vector ($V_n$) as shown below:

$$V_n = V_1' + V_2' + V_3' + V_4'$$
$$= <-8,-8> + <-6,-18> + <0,16> + <4,5>$$
$$= <-10,-5>.$$}

Finally we can add the net movement vector to the current BS location to obtain the final BS location as shown in Figure 2.

$$<X_{BS}, Y_{BS}> = <50,50> + <-10,-5> = <40,45>$$

**Figure 1: Vectors between CHs and BS.**

**Figure 2: The final BS location.**
In general, the BS movement is given by:

\[
\begin{align*}
&< X_{BS}, Y_{BS} > = < X_{BS} + \sum_{i=1}^{K} (X_{CHi} - X_{BS}) \ast p_i(CH_i), Y_{BS} + \sum_{i=1}^{K} (Y_{CHi} - Y_{BS}) \ast p_i(CH_i) > \\
&\text{(12)}
\end{align*}
\]

Where,

- \(X_{BS}, Y_{BS}\): Are the initial coordinate of BS at the start of round.
- \(X'_{BS}, Y'_{BS}\): Are the final coordinate of BS in the end of round.
- \(K\): Is the number of cluster heads in the current round.

4.4 Cluster Formation

Once the cluster heads nodes have selected, they must let all non-cluster head nodes in the network know through broadcasting message. Meanwhile, they wait for other non-cluster head nodes to join in. Each non-cluster head node receives invitation message packets from multiple cluster heads for this round in its communication range and computes the distance \(d(i, j)\) between the sender and the receiver based on the received signal strength indicator (RSSI). In order to make normal nodes choose optimal cluster head, we introduce the cost function. This can be shown as:

\[
\text{cost}(i, j, r) = \frac{E_{res}(r)}{d^2(i)}
\]

Where, \(\text{cost}(i, j, r)\) is cost value between non-cluster head \(s_i\) and cluster head \(s_j\) in the \(r\)th round, and \(E_{res}(r)\) is the residual energy of the cluster head \(s_j\) in the \(r\)th round.

From Equation (10), in every round, each normal node selects from their inviting cluster heads one that allows to have a maximum value of the cost function. Which leads to a load balancing in the whole network.

After the cluster head receives all the Join messages, in order to avoid collisions during messages transmission among sensors, a TDMA (time division multiple access), schedule is made up and transmitted to the sensor nodes in its cluster. Which allows the sensor nodes to be turned off if not on duty. This can effectively reduce the energy consumption for sensor nodes and prolong the lifetime of the network. The huge amount of data gathered in the cluster head ought to be fused into a single data message and transmitted to the BS, the process of this round is complete. Afterwards, a new process starts as before round by round until the energy is depleted by each node, which means the network lifetime is completely ended as illustrated in Figure 3.

5. SIMULATIONS AND RESULTS

5.1 Simulation Environment

We use the MATLAB simulator to evaluate the performance of our proposed EDCSMS algorithm. Simulation parameters are listed in Table 2, where the number of nodes \(N=100\) sensor nodes are distributed randomly in a square region of \(100 \times 100\) m\(^2\). For simplicity, we consider that all nodes are either fixed or micro-mobile and the initial position of BS at every round is placed in the center of the network field. With the same network parameters setup, the EDCSMS protocol is compared with three different clustering algorithms, namely LEACH, EDCS, DEEC in multi-level heterogeneous WSNs.

The performance metrics used for evaluation of our approach are stability period (the first node dies.), lifetime (the last node dies.) and data packets which are successfully sent to the BS.

Table 2: Simulation Parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes number (N)</td>
<td>100</td>
</tr>
<tr>
<td>Round</td>
<td>5 TDMA</td>
</tr>
<tr>
<td>Node initial energy (E0)</td>
<td>0.5 joule</td>
</tr>
<tr>
<td>the threshold distance (d0)</td>
<td>87.7 m</td>
</tr>
<tr>
<td>Packet size (L)</td>
<td>4000 bits</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>2.4</td>
</tr>
<tr>
<td>Network size</td>
<td>100*100m(^2)</td>
</tr>
<tr>
<td>(E_{elec})</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>(E_{DA})</td>
<td>5nJ/bit</td>
</tr>
<tr>
<td>(s_{fs})</td>
<td>10 pJ/bit/m(^2)</td>
</tr>
<tr>
<td>(s_{mp})</td>
<td>0.0013 pJ/bit/m(^4)</td>
</tr>
</tbody>
</table>
5.2 Simulation Results

In this section, the performance of our protocol is evaluated by considering two simulation scenarios:

5.2.1 Scenario 1

In this scenario, \( N=100 \), Rounds Number =5000 and \( \lambda =2 \) which means the initial energy of nodes is randomly distributed in interval \([E_0, 3E_0]\), to prevent the affection of random factors, the network is equipped with the same amount of initial energy.

From Figure 4, we can see that the LEACH performances are the poorest as its stability period and lifetime both are very short. This is because it considers that all nodes have the same rotating epoch. DEEC has longer stability period and lifetime than LEACH just because it considers the rotating epoch according to their initial and residual energy. However EDCS and EDCSMBS have both better performances because, in addition to the advantages of DEEC, they present a load balancing in clusters formation. But the BS movement favors our protocol EDCSMBS to EDCS and allows it to have a very large stability period and lifetime.

Figure 5 shows the number of packets messages received by the sink. As illustrated in the figure above, our protocol EDCSMBS has a higher number of data received by the sink compared to LEACH, DEEC and EDCS. In the first 350 rounds, the four algorithms have nearly the same packets delivery number. However, after 2292 rounds LEACH stops sending packets, while DEEC, EDCS and EDCSMBS continue respectively delivering and forwarding data at 5000 rounds a large amount of packets sent to the sink is scored for EDCSMBS. The summary of Simulation scenario 1 is given in Table 3.

Table 3: Scenario 1: FND, LND and PR.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>First Node Dies (FND)</th>
<th>Last Node Dies (LND)</th>
<th>Packets Received (PR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>350</td>
<td>2292</td>
<td>85452000</td>
</tr>
<tr>
<td>DEEC</td>
<td>1198</td>
<td>2663</td>
<td>424176000</td>
</tr>
<tr>
<td>EDCS</td>
<td>1857</td>
<td>2829</td>
<td>545084000</td>
</tr>
<tr>
<td>EDCSMBS</td>
<td>2062</td>
<td>3145</td>
<td>643096000</td>
</tr>
</tbody>
</table>

5.2.2 Scenario 2

In this scenario, \( N=100 \), Rounds Number =10000 and \( \lambda =4 \) and each node is randomly equipped with an initial energy in the interval \([E_0, 5E_0]\).

As seen in Figure 6 and Figure 7, the performances of the four algorithms keep almost the same order of ranking obtained in scenario 1, with a remarkable enlargement of measurement parameters values due to the increasing of initial energy band. The summary of simulation scenario 2 is given in Table 4.
6. DISCUSSION

From the simulation results, our proposed algorithm is seen as an energy efficient routing protocol. In fact, if we consider Table 4, the stability period is increased at nearly 76%, 17% and 12%, while the network lifetime is increased by 19%, 11% and 13% compared with LEACH, DEEC and EDCS respectively. These EDCSMBS performance improvements reflect the impact of the mobile sink. Hence, the use of fixed sink for LEACH, DEEC, and EDCS protocols, causes hot spots problems around the sink due to the higher data relaying load. Thus, the nodes close to the sink deplete their batteries more quickly than further nodes. However, the use of mobile sink for EDCSMBS permits to minimize the CH-to-BS transmission distances, which allows load balancing and consequently increase the lifetime of the whole network.

Simulation results have proved that EDCSMBS is an energy efficient routing protocol compared to LEACH, DEEC and EDCS in multi-level heterogeneous wireless sensor networks.

7. CONCLUSIONS

In this paper, EDCSMBS an Efficient and Distributed Clustering Scheme with Mobile Sink for Heterogeneous Multi-level Wireless Sensor Networks is proposed in which cluster heads selection like that of DEEC protocol and clusters formation like the EDCS protocol. What is new in our algorithm is the use of mobile sink; at the beginning of every round, the default sink location is in the center of the field, afterward it adjusts its position according to the CH’s probabilities. This strategy minimizes CH-to-BS transmission distances, which allows load balancing and consequently increase the lifetime of the whole network.

REFERENCES:


