
https://doi.org/10.1002/dac.2720

Copyright  John Wiley & Sons

Additional Information
Energy-Efficient Multi-Level and Distance-Aware Clustering Mechanism for WSNs

Amjad Mehmood¹, S. Khan², Bilal Shams³, Jaime Lloret⁴

¹, ², ³Institute of Information Technology, Kohat University of Science & Technology, Pakistan
⁴Integrated Management Coastal Research Institute, Universidad Politecnica de Valencia, Spain
¹amjadiitkust@gmail.com, ²skhan@kust.edu.pk, ³bilal.shams@gmail.com, ⁴jlloret@dcom.upv.es

Most sensor networks are deployed at hostile environments to sense and gather specific information. Since sensor nodes have battery constraints, therefore, research community is trying to propose energy efficient solutions for wireless sensor networks (WSNs) to prolong the lifetime of the network. In this paper, we propose an energy efficient multi level and distance aware clustering (EEMDC) mechanism for WSNs. In this mechanism the area of the network is divided into three logical layers, which depend upon the hop-count based distance from the base station. The simulation outcomes show that EEMDC is more energy-efficient than other existing conventional approaches.

Keywords: Energy-Efficient, Clustering, Multi-Level, Distance-aware, Routing, WSN

1. INTRODUCTION

WSN consists of hundreds of thousands of small-size, low-cost nodes wirelessly connected. It converts physical world into digital world. Each node has limited processing capability, low memory and battery supply [1]. Once deployed, each node uses its energy for sensing, receiving, storing and communication. As nodes are deployed at harsh environments, therefore, it is not realistic to refresh or restore the batteries. There are many limitations in the WSNs such as power-utilization, energy-utilization, robustness, sensing-coverage, organization, energy-
harvesting, and security. The WSNs are usually application specific. The most important applications of WSNs are military, environment monitoring, agricultural activities, sound, temperature, and water applications. Moreover, there are many researchers focused on sensor tracking and localization algorithms [2-4].

In WSN, network energy-preservation is a very hot issue. Due to the inherent nature of the wireless channel, nodes in WSN broadcast their sensed data to their neighbors, which in turn rebroadcast data until they reach the base station (BS). However, these broadcasting and rebroadcasting processes consume more energy. As a result, nodes’ life-time in the whole set-up is decreased. To fix the issue, researchers started to use several techniques such as virtual rings [5], collaborative MAC techniques [6], and group-based topologies [7], which group nodes into clusters [8]. Some authors have demonstrated that cluster-based techniques are more energy-efficient than regular ones [9]. Moreover, we can find in the related literature some free clustering systems [10], parameter-based clustering systems [11], and cluster-based routing protocols [12].

In clusters, one node becomes the Cluster Head (CH) and the others are the cluster members. CH congregates the information from its members, and sends the information directly to the BS in case of a single-hop. While in case of a multi-hop, it drives it to the next CH, which in turn adds its own data and sends it to the next CH. This process is repeated till data reach the BS.

A lot of cluster-based protocols have been introduced by the research community. The clustering protocols in WSN can be categorized as shown in Table 1.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering Technique</td>
<td>Distributed, Centralized, Hybrid</td>
</tr>
<tr>
<td>Network Planning</td>
<td>Single-hop, Multi-hop</td>
</tr>
<tr>
<td>Clustering intention</td>
<td>Energy efficiency, Coverage, load balancing, Fault-tolerance</td>
</tr>
</tbody>
</table>
In order to save energy, most of these routing protocols work in rounds. It has been proved by many works that it is an efficient way to save energy [13]. Each round is made-up of a Set-up phase and a Steady phase. The Set-up phase consists of cluster set-up, topology construction, CH announcement, cluster member acknowledgement to specific CH, and TDMA schedule creation. Whereas in Steady phase data transmission process is carried out from members to the CH of each cluster, from CHs to relay CHs, and finally data reaches the BS as shown in Figure 1.

In order to increase the energy-efficiency in the communication protocol, we propose a cluster-based protocol entitled as EEMDC which merges some of the characteristics of some existing works. The system divides the network area into three logical layers (obtaining a hierarchical structure, which allows to configure IPv6 address easily [14]), and takes into account the hop-count value of the nodes with regard to the BS in order to let the new node join a particular layer. Then, it calculates the average of outstanding energies divided by the hop-count values of the nodes in a cluster for the selection of the CH. It avoids clusters overlapping by taking care that each node joins only one cluster at a time. Moreover, it avoids idle lessoning due to the fact of introducing levels in the clustering, so the nodes of the particular level become active when their turn arrives and sleep otherwise. Finally, the system performs data fusion in sequence in order to save energy.

<table>
<thead>
<tr>
<th>CH Selection Process</th>
<th>Random, Deterministic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-Cluster topology</td>
<td>Fixed, Adaptive</td>
</tr>
</tbody>
</table>

Table 1: Clustering Protocols Categorization Parameters

![Figure 1: LEACH Protocol's Round Details (Adv CH: Advertisement advertise by CH, TDMA Sch: TDMA Schedule)]
The reminder of the paper is as follows. Section 2 presents the related work and our motivation to carry out this work. The protocol proposal operation is explained in Section 3. Section 4 presents some protocol simulations and contrasts our proposal with some well-known protocols. Finally, Section 5 draws the conclusion and future work.

2. RELATED WORK AND MOTIVATION

The cluster-based protocols add more robustness, adaptability, scalability, flexibility, energy-efficiency in finding routes and easy to manage [15-18] as compared to flat, tree, grid and PEGASIS based protocols in WSNs. That’s why the research community has proposed a large number of cluster based solutions for WSNs. In this paper authors are going to briefly explain the related protocols for discussion and compare them with the proposed protocol.

Mohammad Zeinali et al. worked on a novel technique called Hierarchical Routing Protocol (HRP) in-sequence to extend the lifetime of the network [19]. In this protocol, a routing tree is built on CHs for transporting aggregated data to the BS. In [20], J. Chen et al. suggested a multi-hop protocol where cluster’s sensor nodes select a coordinator. Coordinators are divided into a number of disjoint groups. One of the coordinators is chosen as a CH of this group. Hence the sensed data is routed through sensor nodes to their coordinator, from the coordinator to their head, and from the head to the BS.

LEACH (Low Energy and Adaptive Clustering Hierarchy) is a completely distributed protocol [21]. It consists of three phases. They are called advertisement, set-up, and steady phase. It gives a chance to all nodes in the cluster to become a CH. Therefore, this protocol is called as justice protocol. In this protocol, there is a threshold value to be checked for the selection of the CH. I.e.
the protocol assigns a random value of $n$ such that $0 < n < 1$, if $n < T(n)$. This means that a node is selected as CH for this particular round.

Gurprat Sign Chhabra et al. proposed a blend of cluster and tree based approach to enlarge the network lifetime by improving the FND (First Node Death) in [22]. Aimin et al. presented the enhanced version of Low Energy Adaptive Clustering Hierarchy (LEACH) by means of selecting the CH on the basis of the outstanding-energy of the nodes dynamically [23]. As in traditional LEACH, the number of CHs remains unchanging throughout the network lifetime, but in the improved version it is made variable according to the number of live nodes in the network. Consequently, simulation consequences show an upgrading for First Node Dies (FND) and Half of the Nodes Alive (HNA) over traditional LEACH.

In [24], Gou et al. introduced the Partition based LEACH (pLEACH). In this protocol, node having highest residual-energy is selected as CH in each subdivision. Each node knows its location information, which means that that the system is designed to provide location services in the network.

Ding et al. proposed a well-organized clustering protocol based on each node weight considered on its energy and its distance from its neighbors [25]. The node having more weight is selected CH. This technique also improves the network lifetime. Coly et al. [26] proposed a protocol which works in distributed and randomized manner to arrange the nodes into clusters. So, in order to save the energy, member nodes drive the sensed data to the CH. Aubin Jarry et al. [27] suggested to maximize the network lifetime and routing of data to the BS by balancing the energy utilization amongst the nodes of the network. Yi-huazhu et al. [28] proposed (ERAPL) an
energy-efficient routing mechanism to dynamically make longer the network lifetime while it is being consumed efficiently.

In [29], Hana Khamfroush et al. suggested a multi-hop protocol in which predefined nodes, having greatest outstanding energy, are elected as CH first, and then the members of every cluster are determined on the basis of their distance involving the nodes and the CH and also CH and the BS. Li et al. projected an imbalanced clustering technique to increase the network lifetime by two or three times the one accomplished in the most excellent imbalanced clustering approach [30].

Noman Aslam et al. [31] proposed a distributed protocol in which nodes take decisions based on the local information. S.Yi et al. [32] suggested PEACH (Power efficient and adaptive clustering hierarchical) protocol to form a multi-level clustering system. It improves the lifetime of the motes in the network. Pei et al. suggested DECSA stands for Distance-Energy Cluster Structure Algorithm [33]. It is based on LEACH clustering protocol. It considers distance between the nodes, the distance of the base station, and outstanding energy of nodes. It divides the network into three levels, in order to reduce the energy consumption of the CHs. It saves the energy by avoiding the direct communication between the CHs and the BS.

MTE protocol, proposed by Weng et al. [34], has exhibited less energy utilization and has enlarged the network lifetime by sending data to a path consuming minimum transmission energy. In this protocol, nodes’ remaining energy has not been considered for the selection of the CH. Therefore, it could lead to assign the cluster head role to a node which does not have enough energy. So it could deplete its energy earlier than others. Hence it decreases network lifetime. Other problem with LEACH is that every CH sends its information to the BS directly, which is again impractical in case of a large WSN. Thus LEACH doesn’t take care of outstanding-energy
and the distance of the CH to the BS. Hybrid Energy Efficient Distributed (HEED) clustering is a single-hop approach [35]. Selection of CH is based on remaining-energy and the degree of nodes in a cluster. D. Kumar et al presented Energy efficient heterogeneous clustered scheme for wireless sensor networks (EEHC) [36]. In the protocol it assumed that all the nodes are homogeneous in nature, and deployed randomly in the environment. Each node has the opportunity to become CH on the basis of the residual energy it has. Kim et al presented Tree-base Clustering Protocol (CBT) for WSN [37]. In the protocol as its name suggests, nodes form tree in the cluster, with CH as their root. The height of the tree is determined on the basis of distance of the nodes from the CH. Then CH sends data to the BS using single-hop method. J. Yu has presented EADC protocol [38], in which CH sends data to the BS by considering the threshold value. If the CH distance is less than threshold value then it sends it directly to the BS, otherwise selects the CH having more outstanding energy as a rely CH for sending data to the BS. M. Ye et al presented an Energy Efficient Clustering Scheme (EECS) [39], which is preferred for the applications of periodical data gathering. In the protocol the CH is elected on the basis of outstanding energy of the nodes, moreover it achieves well CH distribution. Therefore it improves the lifetime of the whole network.

Dynamic Clustering and Distance Aware Routing Protocol (DDAR) [40] is the one in which the energy consumption is controlled by considering remaining energy, distance and dynamic number of CHs in the network. In the setup phase all nodes forward their position and outstanding-energy to the BS. BS then finds the best possible number of CHs on the basis of available nodes in the network. Table 2 summarizes the state of the art of energy-aware cluster based routing protocols. All studied protocols are deployed for stationary nodes. The purpose of all of them has been the energy conservation. Moreover, it is considered a variable number of
clusters and the initial CH selection is random. All of them are a single-hop intra-cluster topology.

<table>
<thead>
<tr>
<th>Clustering protocol</th>
<th>Clustering methods</th>
<th>Routing Type</th>
<th>Cluster Size</th>
<th>Energy Efficiency</th>
<th>Additional features</th>
<th>Location awareness</th>
<th>Connectivity of CH to BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRP [19]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>ECHS [20]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>LEACH [21]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Very Low</td>
<td>Load Balancing</td>
<td>No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>GDP [22]</td>
<td>Centralized</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>Yes</td>
<td>Minimum spanning tree</td>
</tr>
<tr>
<td>pLEACH [24]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>Both Yes/No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>DWEHC [25]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>V. High</td>
<td>Load Balancing</td>
<td>No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>COCA [29]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>Equal size</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>PEACH [32]</td>
<td>Centralized</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>Yes</td>
<td>One hop communication</td>
</tr>
<tr>
<td>DECSA [33]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>MTS [34]</td>
<td>---</td>
<td>Flat Routing based</td>
<td>--</td>
<td>Very Low</td>
<td>----</td>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>HEED [35]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Moderate</td>
<td>Load Balancing</td>
<td>No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>EEHC [36]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>TBC [37]</td>
<td>Distributed</td>
<td>Random</td>
<td>Cluster based</td>
<td>Single-hop</td>
<td>variable</td>
<td>Good</td>
<td>Direct Link</td>
</tr>
<tr>
<td>EADC [38]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>Equal size</td>
<td>Good</td>
<td>Load Balancing</td>
<td>No</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>EECS [39]</td>
<td>Distributed</td>
<td>Cluster based</td>
<td>variable</td>
<td>V. High</td>
<td>Load balancing / periodical data communications</td>
<td>No</td>
<td>Direct Link</td>
</tr>
<tr>
<td>DDAR [40]</td>
<td>Centralized</td>
<td>Cluster based</td>
<td>variable</td>
<td>Good</td>
<td>Load Balancing</td>
<td>Yes</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>LEACH-C [41]</td>
<td>Centralized</td>
<td>Cluster based</td>
<td>variable</td>
<td>Moderate</td>
<td>Load Balancing</td>
<td>Yes</td>
<td>Direct Link</td>
</tr>
</tbody>
</table>

Table 2: Summary of the most important features of the routing protocols included in this section

Our technique is motivated by the research of Gautam et al. in [40]. They considered that clusters close to the BS consume more energy than those away from the BS. As those nodes do the job of relaying near the BS, their death chances are high [21, 33]. It leads to the Funnel Effect [35]. Therefore, our proposed EEMDC divides the network into logical layers with respect to distance from the BS.

3. EEMDC OPERATION
The nodes near the BS carry higher load of traffic than the nodes far away from the BS [42]. Therefore, the nodes near the BS drain off their energy earlier than nodes away from the BS. Hence, the lifetime of the network decreases. In order to increase the network lifetime, we propose a technique called Energy Efficient Multi level Clustering and Distance Aware Mechanism (EEMDC) for the cluster set-up, steady phase and routing process. This technique is based on considering the nodes placement with regard to the BS in the network, for cluster-setup, steady-phase, and routing process, using hop-count value. EEMDC characteristics include:

- Hybrid clustering process
- Overwhelms the hotspot problem using multi-level distance aware clustering with regard to the BS
- Random CH selection process through average of residual-energies by average hop-count values
- Minimum multi-hop Inter-cluster connectivity
- Energy-efficiency
- Load-balancing
- Scalability
- Adaptive intra-cluster topology

EEMDC protocol works in rounds like many well-known data gathering protocols such as Low Power Efficient Adaptive Clustering Hierarchical (LEACH), Centralized Low Power Efficient Adaptive Clustering Hierarchical (C-LEACH), and Dynamic Distance Aware Routing (DDAR), Hybrid Energy Efficient Distributed Clustering Approach protocols.

In order to fix the hotspot nodes problem and balance the workload of the network, EEMDC protocol divides the network into three logically levels. First Level Clusters (FLC) includes
nodes having a hop-count value of “one” and “two”, while Second Level Clusters (SLC) includes nodes having a hop-count value of “three”, “four” and “five”, and, finally, Third Level Clusters (TLC) contains nodes having a hop-count value of “six” and above. The discrepancy of nodes in each level of clusters is due to their position with respect to the BS. Hence, FLC contains more clusters and fewer members in each cluster. At the same time, SLC has lower number of clusters and more members than in SLC, and, lastly, TLC has lower number of clusters and more members as compared to SLC. In the EEMDC cluster formation process, BS broadcasts a cluster-initiative packet containing the hop-count value to all its neighbor nodes in the network. The hop count value parameter in the packet helps the node to determine its distance from the BS. So, on the basis of hop-count value, each node decides on which cluster level is going to be part of (e.g. FLS, SLC or TLC). After the beacon packet reaches, each node calculates its level. Hence the protocol starts working distributed.

BS sends the starting packet to its neighbors. The neighbors increment their hop count value only when the packet is received for its first time or when update messages are received. Then, they broadcast the new value to their neighbors. Then, their neighbors repeat the same process, and finally each node gets its position with regard to the BS, and the level of cluster it belongs to.

In the EEMDC cluster set-up phase, each node has equal energy at the beginning like LEACH [21]. Therefore, CHs are nominated randomly in the first round without considering the reminding energy of the cluster nodes, but in next rounds, CH nomination decision takes place on the basis of the reminding energies and the hop count values in the cluster. Each node sends its reminding energy and hop count value to its neighbors in its cluster. After that, each node in the cluster performs an average operation on the sum of remaining energies and hop count values. Let $E_{ij}$ be the remaining energy of node $j$ in the cluster $i$, and let $d_{ij}$ be the hop count value
of node $j$ in the cluster $i$. The average remaining energy of the cluster $i$ ($E_i$) and the average hop count value of the cluster $i$ ($d_i$) is given by equations 1 and 2 respectively.

$$E_i = \frac{\sum_j E_{ij}}{j}$$  \hspace{1cm} (1)$$

$$d_i = \frac{\sum_j d_{ij}}{j}$$  \hspace{1cm} (2)$$

When a node has its remaining energy divided by the hop count value is greater than the average reminding energies divided by the average hop count values of the neighbor’s node in a cluster, then that node is chosen as CH. This is analytically expressed as equation 3.

$$\frac{E_{ij}}{d_{ij}} > \frac{E_j}{d_j}$$  \hspace{1cm} (3)$$

In case of having two or more nodes in the same case, node with lower hop count value is selected as CH. All other nodes will be regular nodes in the cluster. Once a node gets nominated for CH, it sends a joining message to its neighbors. Neighbors decide to join a specific CH based on the hop-count value as it has been explained before. The node, which receives the joining message from two or more CHs of different levels, joins the cluster according to the level of the cluster it belongs to, but if it gets the joining message from two or more clusters from the same level, then it selects the one with highest received signal strength. Hence, this process avoids the overlapping of clusters in the same and different levels which sequentially mount-up the network lifetime.

Once the CH receives the acknowledgement of the joining packet, it assigns the TDMA schedule to its members. So, the members drive their data packet to the concerned CH without interruption. As WSN consists of large number of nodes therefore, fixing the IDs on the nodes before the deployment of the network is impractical in large WSNs [31], hence the node identification mechanism in EEMDC is as follows. CH adds one byte at the end of its ID and
assigns it to its neighbors. Similarly, its neighbors add one byte and assign it to its neighbors, and so on. By using such technique node identification process is carried out. Therefore this technique is considered more practical and also helps to keep away from redundant information to increase the nodes’ lifetime. In multi-level cluster, when the lower level of nodes is collecting information from its environment the next level of nodes are sleeping at that time [43] (see figure 2). This avoids channel idle-listening and hence saves energy to the cluster members.

We have designed our WSN without allowing regular cluster members having connections with regular cluster members from other clusters. So, nodes, which are able to see nodes of other clusters, will not establish a connection with them. When a node finds a node that is not from its cluster, it simply drops such communication and hence saves its transmission power.

Figure 2: Message Flow in a Cluster (FLC)

Figure 2: Level 1 Cluster  hpv: Hop-count value  FLC: First level Cluster
Once the set-up phase of the clusters is finalized, CH assigns the TDMA slot to its 1-hop neighbors. 1-hop neighbor adds some constant value and assigns TDMA slot to 2-hop neighbors. Similarly, TDMA schedule is distributed among the clusters according to their cluster level.

In order to avoid the hotspot problem we will try to maintain with small size the clusters close to the BS (i.e. nodes which belong to the FLC). Moreover, nodes in the hub have more chances to exhaust their energy earlier than the nodes far away from the BS belong to the SLC, and remaining nodes belong to the TLC. Figure 3 shows the cluster levels.

In the steady phase, each low level cluster gets the data from its environment and relays the information to its next level and hence data reaches the cluster head where data fusion and aggregation [44-47] is done, fused data is directed to BS through the shortest path using relay clusters.
3.1 DATA TRANSMISSION FROM CLUSTER HEAD TO THE BASE STATION

Routing data from the CH to BS, after performing the aggregation, is performed using the following technique. If CH belongs to the FLC, it sends the data to the BS directly without considering any relay clusters. But if the cluster belongs to the SLC or TLC, CH considers the minimum hop-count value of the available CHs in its neighbor set. Then, it sends the data to a specific CH taking into account the minimum hop-count value.

Similarly, the cluster receiving the data uses the same procedure and sends the data to its neighbor cluster taking into account the minimum hop-count value and so on. Hence the data reaches the BS using minimum hop of relay clusters.

3.2 FLOWCHART OF THE EEMDC OPERATION

EEMDC starts from the BS. It sends a packet to its neighbors, which contains the hop-count value. Each node calculates its hop-count value and consequently gets the cluster-level information as shown in figure 4. Once each node gets these two parameters, the protocol enters into the steady phase. In the steady phase, it checks that if the round value is 1, then CH is elected randomly because it is understood that each node has the same energy at the beginning. In case of a second round, each node shares its reminding energy and hop count value with its neighbors. Hence, each node is aware of the reminding energies and hop count values in the cluster. Therefore, each node computes equation 3 and a CH is elected based on the explained...
procedure. Once the CH is elected, it announces itself as CH to the nodes in the cluster. The nodes in the cluster acknowledge the request sent by the CH. Then, the CH sends a discovery message to discover CHs. When the node collects replies from two or more CHs from the same level, then, it joins the best CH based on the received signal strength. But, if it receives replies from two or more CHs from different levels, then it takes a decision based on the cluster level, which is saved on each node. Hence, a CH must establish connections with a CH of the same level.
BS sends a beacon Packet to nearby nodes containing hop count value and cluster level parameters

Each node calculates its hop count value and cluster level it belongs to

\[
\text{AvgRE} = \frac{\text{sum of Energies of neighbors}}{\text{Total Energy of neighbors}}
\]
\[
\text{AvgHcv} = \frac{\text{sum of hop count values of neighbors}}{\text{Total Hop count of neighbors}}
\]
\[
\text{NbrEnergy} = \frac{\text{AvgRE}}{\text{AvgHcv}}
\]
\[
\text{NdeEnergy} = \frac{\text{NdeRE}}{\text{NdeHCv}}
\]

- \(NdeEnergy > NbrEnergy\)
  - Set as Cluster Head
  - CH announces Join message according to the cluster level it belongs to
  - CM Joins its Cluster
  - CH distributes TDMA schedule among its cluster level members
  - CH Receives Data from its members
  - Cluster Head performs aggregation process on the received data
  - Cluster Head sends the data to BS using SH or MH technique
  - \(\text{Round}++\)

- \(NdeEnergy \leq NbrEnergy\)
  - \(\text{Set as non-Cluster Head}\)
  - Get Join Request from Cluster heads belong to same or different levels of clusters
  - \(\text{Req from Same Level or diff.}\)
    - \(\text{Same Level}\)
      - Join CH on Strong Signal
    - \(\text{Diff. Level}\)
      - Join CH on the basis of Hop Count value
      - CM gets TDMA Schedule from Cluster heads
      - CMs send Data to Cluster Head

<table>
<thead>
<tr>
<th>CH</th>
<th>Cluster Head</th>
<th>AvgHcv</th>
<th>Average Hop count value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM</td>
<td>Cluster Member</td>
<td>NdeEnergy</td>
<td>Node’s Energy</td>
</tr>
<tr>
<td>AvgRE</td>
<td>Average Residual Energy</td>
<td>SH, MH</td>
<td>Single hop, Multi hop</td>
</tr>
</tbody>
</table>

**Figure 4: Flowchart of EMCDR Working**
Once the CH and the regular nodes (CM), belonging to the cluster, have taken a decision, the CH sends the TDMA schedule among the nodes in the cluster in order to send their sensed data to the CH without interruption.

Once the nodes receive the TDMA schedule, they start sending their sensed data to the CH, and the CH performs the aggregation of the sensed data to fuse them. Hence, CH transmits its fused data to the neighbor CH that has less hop count value to the BS. The neighbor CH sends its fused data to its neighbor CH that has the lowest hop count value and so on. Consequently data finally reaches the BS.

Figure 5 explains how the BS sends the packet containing the hop count value to the nodes in the network. Every node checks whether it is getting the message for the first time or not. If the node has received the message before, then it simply discards the packet. Alternatively, if the node has not received the message before, then it records the hop count value, increments it, puts the new value in the packet, and broadcast it to its neighbors (except to the one from it was received). The same process is repeated by their neighbors and hence each node records its hop count value. Then, the cluster level decision is taken because it is based on the hop-count value saved at each node.

Hence, after the execution of the initialization phase, each node gets its position in terms of hop count to the BS. Finally, all nodes know at which cluster level it belongs to.
Figure 6 illustrates the aggregation performed by the CH on the sensed data received from the nodes. It checks that if the CH belongs to the FCL, then it sends the data to the BS directly because it is close to the BS. But if it belongs to SLC or TLC, then that CH has to send the data through relay clusters before reaching the BS. Therefore, CH selects the next cluster head in the network as a relay CH which has smallest hop count value because it is believed that it is close to the BS. Using the same technique of selecting the minimum number of relay CHs, data finally reaches the BS.
Figure 6 shows the code used to deploy our routing protocol. It shows how each node is included in each level (FLC, SLC and TLC), and how they take decisions to route the information according to the aforementioned metric.

Figure 7 shows the overall working of the EEMDC algorithm. The algorithm is mainly divided into two parts i.e. initialization and repetition. In the initialization phase, each node takes a decision based on the hop count value (cluster level it belongs to i.e. FCL, SCL or TLC). While in the repetition phase, the algorithm selects the CH, announces the message of the CH, the node joins the CH, and passes the TDMA schedule among the nodes. Hence, after collecting data from the cluster members, CH performs aggregation, and route the data to the BS, through the relaying CHs that have lowest hop count value.
break;
}

RepeatitionPhase{
    If(Round==1){
        Cluster Heads (C.H) are selected randomly
        for(i=0 : i<FLC.Length;i++)
            if(RandomlySelected(nodei)==true) // Randomly Takes decision of CH
                CHs[i] ← node i;
        for(i=0 : i<SLC.Length;i++)
            CHs[i] ← node i;  //Randomly Takes decision of CH
        for(i=0 : i<TLC.Length;i++)
            CHs[i] ← node i;  //Randomly Takes decision of CH
        for(i=0;i<CH.length;i++)
            CHs[i].Join_Message; ←→ announces a join message // (according to cluster of levels specified in nodeiCLI)
        CH[i]s. Schedule TDMA slots among the member in the cluster
        Round++;
    }
    else{
        AvgRE= (sum of Energies of neighbors /Total Energy of neighbors)
        AvgHCv= (sum of hop count values of neighbors /Total Hop count of neighbors)
        NbrEnergy= AvgRE/AvgHCv
        NdeEngery= NdeRE/NdeHCv
        If (NdeEnergy>NbrEnergy){
            // Nodei will be selected as C.H.
            CH[i] ← nodei;
            CH[i].Join_Message ←→ announce a join message // (according to cluster of levels specified in nodeiCLI)
        }
        CH[i]s. Schedule TDMA slots among the member in the cluster }
        //other nodes will be non-C.H node
        //Wait to lesson a Join Message send by Cluster head on same as well different cluster level
        //Store Join Messages in an Array
        If(nodei belongs to two clusters on the level)
            Decision will be made on RSSI
        else if(nodei belongs to two cluster on different levels)
            Decision will be made on HCv(if HCv=1 or HCv=2 join FLC,if HCv>= & HCv<=5 join SLC
            else join TLC)
            //non-C.H member sends the data then CH then it will performs aggregation on the data
            for(i=0;i<CH.Length;i++)
                if(CHs[i]_HCv==1 || CH[i]_HCv==2)
                    sends data directly to BS
                else if(CH[i]_HCv==3 & CH[i]_HCv<=min(Nbr_CH[i]_HCv)
                    Send data to BS  //assume that it is close to BS
                else{
                    CH[i]_HCv=min(Nbr_CH[i]_HCv)
                    Send data to new CH[i]_HCv // sends its data to new cluster head
                }
            CH[i]++;
        }if(CH[i].lastIndex)
            break;
    }
}

End

Figure 7. EEMDC’s algorithm: Process of Clustering in WSNs
4. SIMULATION RESULTS

Simulation results have been produced by considering the five conventional protocols namely MTE, LEACH, LEACH-C, DDAR, HEED against the proposed protocol (EEMDC). The parameters used for comparison are First Node Die Time (FNDT), Half Node Die Time (HNDT), Last Node Die Time (LNDT), Energy Consumption against Time, and number of nodes alive against Time. As we will see, in all aforementioned parameters, we will conclude that EEMDC exhibit better results than the other protocols.

4.1 Network and Radio Model.

All the sensor nodes have same energy at the beginning and they are densely deployed in the area. We will assume that the BS is outside the network. In the initialization phase, EEMDC protocol starts working from the BS by generating a beacon packet to help nodes to calculate its hop count value and the cluster level they belong to. After that it operates distributed and adaptively. Therefore, it avoids the congestion in the network due to distributed approach. They use the hop-count to determine the distance from the BS, and the level they belong to. They are static in nature, because they remain stationary after the deployment. Node’s battery is not chargeable once it depletes its energy. Generally, the network lifetime is defined as first node failure, average node failure or last node failure. So we developed our protocol based on these cases and it is able to deem the network lifetime in terms of them. As due to hotspots problem, the nodes near the BS deplete their energy earlier than far nodes, therefore, we divide the network into three logical levels. We also assumed that each level has approximately the same number of nodes. So, for numerical purposes, first-level contains 33%, second-level contains 33%, and third-level contains the remaining 34% of nodes. The percentage of nodes being CHs
in the first, second and third levels are 7%, 4%, and 2% respectively. Accordingly, clusters close to the BS (first level clusters) are small in size clusters (up to 4 to 5 nodes) because they have to do their own tasks and more relay jobs. Second level clusters are one-hop bigger in size than first level clusters (about 8 to 10 nodes) because they perform lower relay tasks than first-level clusters. Third level has a maximum of 16 to 20 nodes in a cluster, because they have less probability to deplete their energy. So, overall division of the network into the three layers has amplified the network’s lifetime. The connectivity of the network is such that using minimum hop count value of the CH, third-level clusters are connected with second-level clusters, second-level clusters are connected with first-level, and first-level clusters drive the data to the BS.

We adopt a typical radio and energy consumption model whose specific details can be found in [44]. Communication energy consists of two factors: electronic energy and amplification energy. The electronic energy (sending and receiving) of all nodes is the same. Consequently, the factor which affects the communication energy is the amplification energy. It depends on the space of the node from the BS. There are two types of power losses: free-space ($d^2$ power loss) and the multi-path fading ($d^4$ power loss) channel. The energy exhausted for the transmission of a p-bit packet over a distance $dist$ is given by equation 3 and 4:

$$E_t(p, \Dist) = pE_{elec} + p\epsilon_f \Dist^2 \quad \text{if } \Dist \leq \Dist_t$$  \hspace{1cm} (3)

$$E_t(p, \Dist) = pE_{elec} + p\epsilon_m \Dist^4 \quad \text{if } \Dist > \Dist_t$$  \hspace{1cm} (4)

And the radio consumes the power presented in equation 5:

$$E_R(p) = p E_{elec}$$  \hspace{1cm} (5)

The notation used is included in table 3.

<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{dist}$</td>
<td>Transmission Distance</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>( \text{dist}^2 )</td>
<td>free-space Distance</td>
</tr>
<tr>
<td>( \text{dist}^4 )</td>
<td>multi-path fading</td>
</tr>
<tr>
<td>( E_{\text{elec}} )</td>
<td>Electronic Energy</td>
</tr>
<tr>
<td>( P )</td>
<td>p-bit packet</td>
</tr>
<tr>
<td>( \text{Dist} )</td>
<td>Threshold distance</td>
</tr>
<tr>
<td>( \varepsilon_{\text{fs}} )</td>
<td>Coefficient for free-space fading</td>
</tr>
<tr>
<td>( \varepsilon_{\text{mp}} )</td>
<td>Coefficient for multi-path fading</td>
</tr>
</tbody>
</table>

**Table 3: Notation used in the equations**

### 4.2 Simulation Model

For the performance evaluation of EEMDC, we have used OMNET++ 3.2.2 Object Oriented modular discrete event based simulator having Castalia 3.2 as the model for WSN [48-50]. The network area is about 100m * 100 m. We have simulated two cases: 100 nodes and 200 nodes. All these nodes have been placed randomly in the area. The division of nodes in 100 and 200 nodes cases is such that fist, second, and third levels contain equal number of nodes respectively. The initial energy of each node is fixed to 12 Joule. The packet header size is 25 bytes, the data size is 500 bytes, and the control packet size is 100 bits. We have produced the results for both 100 and 200 nodes in same network model for checking the effects of growing the number of nodes in the simulation model. See the Table 4 for each parameter’s value.

<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Area</td>
<td>100m*100m</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>12 Joules</td>
</tr>
<tr>
<td>CH Percentage</td>
<td>0.05</td>
</tr>
<tr>
<td>Packet header Size</td>
<td>25 byte</td>
</tr>
<tr>
<td>Data Size</td>
<td>500 byte</td>
</tr>
<tr>
<td>Num of Nodes</td>
<td>100, 200 nodes</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>500 Sec</td>
</tr>
</tbody>
</table>
Control Packet Size | 100 bit  
Broadcast packet size | 25 bytes  
Round Length | 20 sec  
\( E_{\text{elec}} \) | 50 nJ/bit  
\( E_{\text{fs}} \) | 10 pJ/bit/m^4  

Table 4: Network Model Parameters

4.3 Performance Evaluation

By using the above simulation model, EEMDC is evaluated against five conventional protocols such as LEAC, LEACH-C, MTE, DDAR, and HEED. The following five performance metrics have been considered for evaluating EEMDC protocol against the conventional protocols.

- First Node Die Time (FND): Time when first node is died in the network
- Half Node Die Time (HND): Time when half nodes died in the network
- Last Node Die Time (LND): Time when last node died in the network
- Energy Consumption over Simulation Time (ECoST): Energy consumed by nodes with respect to simulation time
- Number of Alive Nodes over Simulation Time (NAoST): Represents the total number of nodes alive regarding to the simulation time.

Fig. 8 and Fig. 9 show the FND Time in case of 100 and 200 nodes. Fig. 8, shows that, for 100 nodes, the die time for the first node is 16.5152 seconds in EEMDC, whereas the die time for the first nodes are 8.35 sec., 10.96 sec., 13.25 sec., 14.77 sec., and 15.12 sec. for MTE, LEACH, LEACH-C, DDAR, and HEED respectively. When we increase the number of nodes up to 200 (see Fig. 9), for EEMDC protocol the first node died at 15.12 seconds, while the die time for the first node of the studied protocols (MTE, LEACH, LEACH-C, DDAR, and HEED) are 5.35 sec., 8.96 sec, 10.25 sec., 12.77 sec., and 14.12 sec. respectively.
Figure 8: Total 100 Nodes: First Node Die (FND) over Simulation Time [Seconds]
Fig. 10 and Fig. 11 show the HND Time in case of 100 and 200 nodes. Fig. 11 demonstrates that, for 100 nodes, 50% of the nodes died at 764.996 seconds in EEMDC, whereas half nodes died at 350.46 sec., 399.92 sec., 590.5 sec., 601.3 sec. and 799.78 sec. in case of MTE, LEACH, LEACH-C, DDAR, and HEED respectively. When we increase the number of nodes up to 200, the die time for the half nodes of the studied protocols (MTE, LEACH, LEACH-C, HEED, and DDAR) are 213.46 sec., 299.22 sec., 350.35 sec., 401.3 sec. and 523.78 sec. respectively, while half nodes died at 604.0 sec in EEMDC protocol (see Fig. 11).
Figure 10: Total 100 Nodes Half Node Die (HND) over Simulation Time [Second]

Figure 11: Total 200 Nodes Half Node Die (HND) over Simulation Time [Second]
Fig. 12 provides LND when there are 100 nodes in the WSN. It provides for MTE, LEACH, LEACH-C, HEED, and DDAR a LND time of 460.92 sec., 465.84 sec., 706.34 sec., 999.6 sec., and 1109.79 sec. respectively, while we have obtained 123.52 sec. for EEMDC.

![Bar chart showing LND over simulation time](image)

Figure 12: Total 100 Nodes: Last Node Die (LND) over Simulation Time [Second]

Similarly, when the numbers of nodes increase in the WSN the LND time drops to 305.92 sec., 365.84 sec., 586.34 sec., 645.6 sec., and 704.79 sec. in MTE, LEACH, LEACH-C, DDAR, and HEED respectively, whereas we obtained 1099.52 sec. for EEMDC protocol. This can be seen in Figure 13.

Fig. 14 shows the energy consumption of the studied protocols along the time. The simulation results shows that EEMDC consumes 69.346% less energy than MTE, 27.32% less energy than
LEACH, 17.028 % less energy than LEACH-C, 12.1 % less energy than HEED, and 4.4356 % less energy than DDAR. So, EEMDC is better than the other protocols.

Figure 13: Total 200 Nodes: First Node Die (FND) over Simulation Time [Second]
Fig. 14: Energy Consumed by Nodes over Simulation Time [Second]

Fig. 15 shows the number of nodes alive along the time. This simulation shows that the proposed protocol have 97 nodes alive after 500 seconds. But MTE, LEACH, LEACH-C, HEED, and DDAR have 6, 9, 84, 90, 94 of nodes alive respectively after 500 seconds.
5. CONCLUSION

The EEMDC is an energy efficient multi-level clustering routing protocol. It attempts to increase the network lifetime, while providing more stability and reliability to the network. The distinctive characteristics of EEMDC, which makes this protocol different from conventional protocols, are the following ones:

- It divides the network area into three logical layers. By doing the partition of the network area, the hotspot problem is fixed.
- The distance of the nodes to the CH and the CH to the BS are taken into account when considering the hop-count value of the nodes.
- CH is elected by acquiring the average left-over energy of the nodes.
• The fused-data is delivered to the BS using the minimum distance path to the BS (in terms of hop count values).

All these four features are not collectively taken into account in any traditional protocols such as LEACH, LEACH-C, MTE, DDAR and HEED.

The simulation results show that EEMDC achieves greater energy efficiency compared to conventional protocols, thus the network lifetime is increased.

In future works we are going to include security in our system by using efficient authentication processes [47] and combine our energy-efficient clustering technique with artificial neural networks in order to prevent energy exhaustion attacks [48]. Moreover, we will include data fusion techniques in our deployment.

ACKNOWLEDGEMENTS

This work has been partially carried by the "Ministerio de Ciencia e Innovación”, through the “Plan Nacional de I+D+i 2008-2011” in the “Subprograma de Proyectos de Investigación Fundamental”, project TEC2011-27516, and by the Polytechnic University of Valencia, though the PAID-15-11 multidisciplinary projects.

REFERENCES


[50] Download Leach v2 for Castalia