Multi Objective Clustering Algorithm for Maximizing Lifetime in Wireless Sensor Networks

Salim EL KHEDIRI¹, Adel Dallali², Abdennaceur KACHOURI²
¹ University of Gafsa, Departement of computer sciences
Faculty of sciences of Gafsa, TUNISIA
² University of Sfax,
National Engineering School of Sfax,  TUNISIA
salim.el.khediri@gmail.com

ABSTRACT: Hierarchical architecture with fixed structure is an effective scheme to make Wireless Sensor Networks (WSNs) scalable and energy-efficient. Clustering the sensor nodes is a well-known two-layered architecture suitable for WSNs and has been extensively explored for different purposes and applications. In this paper, a novel clustering approach called the Fixed Competition-based Clustering Approach (FCBA) is proposed for WSNs. Selecting the cluster heads in the proposed FCBA is performed based upon a residual energy and the distances among the cluster heads. First, by the new proposed competition scheme, the nodes with the high residual energy which are closer to the center of the density of the nodes are elected and form an initial set of cluster head candidates. Then, the candidates collect data from its members and send it to the base station. By comparing our contribution to the recent clustering algorithms, simulation results show that our scheme is effective in saving energy and power consumption.

Keywords: Clustering, Energy Efficient, Lifetime, Cluster-head, Wireless Sensors Networks

Received: 18 June 2017, Revised 28 July 2017, Accepted 31 July 2017

© 2017 DLINE. All Rights Reserved

1. Introduction

Recently, there has been a rapid advancement in the field of wireless sensors networks (WSNs) due to their promising applications in a variety of areas, such as environmental monitoring, surveillance, acoustic and seismic detection, among others. A wireless sensor network is composed of a large number of sensor nodes connected wirelessly. All sensor nodes play the role of an event detector and the data router. Sensor nodes are deployed in the sensing area to monitor specific targets and to collect data. Then, sensor nodes send the data to a sink or a base station (BS) by using wireless transmission techniques [1–2]. The BS is connected to a switching center that acts as a gateway from the wireless network to the Internet. Fig.1 shows a direct communication protocol in the wireless sensor networks, where each sensor node directly transmits its sensing data to the BS.
The essential issue in any proposed architecture fig.1 is to maximize the lifetime nodes in the WSN. In fact, it is impossible to reload or replace the battery after their exhaustion [3]. Traditional methods of deployment nodes are not approved for use in the sensor network due to problems of complexity and high power consumption [4]. The synchronization mechanism is a phenomenon subject to many constraints which must meet several requirements. These constraints sometimes can be contradictory, such as minimizing energy consumption, reducing the associated costs and maximizing the quality and accuracy of services provided. One of the well-known solutions for minimizing the total energy consumption of WSN’s. In clustered WSN’s each sensor is dominated by elected routers and send information to its cluster-head (CH) that aggregate and transmit data to the base station directly or through other CH’s in the network.

The selection of cluster head is the key issue in the clustering algorithm, which is also a multiple criteria in decision making procedure [6]. In this paper, we propose a new technique for the selection of the sensors cluster-heads based on the amount of energy remaining after each round [7,8]. As the minimum percentage of energy for the selected leader is determined in advance and consequently limiting its performance and nonstop coordination task, the new hierarchical routing protocol is based on an energy limit value “threshold” preventing the creation of a group leader to ensure reliable performance of the whole network.

The present paper is organized as follows. Section two reviewed some clustering algorithms. Section three presents FCBA algorithm in details and argues the choice of its parameters. The simulation results were given in section 4. Finally, our conclusions were drawn in section 5.

2. Related Networks

Several studies on energy saving issues have been conducted on wireless sensor networks. With regard to cluster-based energy saving methods, analytical models have been proposed in [10]. Fig. 2 shows a cluster-based routing protocol in wireless sensor
networks, where the cluster head (CH) node is responsible for collecting information on non-CH nodes in its cluster. Then, it processes data and sends data to the BS. The non-CH node can only monitor the environment and send data to its CH node. Because a non-CH node cannot send data directly to the BS, the data transmission distance of the sensor node is shrunk. Therefore, the energy consumption is reduced for each non-CH node.

The LEACH protocol is perceived as one of the well known hierarchical routing algorithms for sensor networks [11], in homogeneous sensor networks. Basically LEACH uses a distributed clustering approach. LEACH selects randomly the nodes cluster-heads and assigns this role to different nodes according to round-robin management policy to ensure fair energy dissipation between nodes in order to reduce the amount of information transmitted to the base station, the clusterheads aggregate the data captured by the member nodes belonging to their own cluster, and then sends an aggregated packet to the base station [2]. The operation of LEACH consists of two phases, namely the setup phase and the steady state phase. In the setup phase, cluster heads are selected and clusters are organized. In the second phase, The actual data transfer to the base station is held. During the first phase, a predetermined fraction of nodes, p, elect themselves as CH’s as follows. A sensor node chooses a random number, r, between 0 and 1. If this random number is less than a threshold value, T(n), the node becomes a CH for the current round, else the node n is expected to connect to the nearest cluster head in its neighborhood by receiving the strength of the advertisement message from the CH nodes. The threshold value is determined based on an equation that incorporates the desired percentage to become a CH, the current round, and the set of nodes that have not been selected as a CH in the last $(1/p)$ rounds, denoted G [3]. It is given by:

$$T(n) = \frac{p}{1 - p(r \mod \left(\frac{1}{p}\right))}$$  

where $G$ is the set of nodes that are involved in the CH election, $r$ is the current round number (starting from round 0) and p refers to the probability of each node to become cluster head.

However, even LEACH can increase the lifetime of the network, it has some limitations. LEACH assumes that all nodes can transmit data with great power to reach the base station and each node has a computing power enabling it to withstand various MAC layers. Therefore, LEACH is not suitable for networks deployed in large areas. In addition, LEACH randomly selects a list of cluster heads and there are no restrictions neither on their distribution nor on their energy level. Thus, the cluster heads can concentrate on one place and therefore there may be isolated nodes (without cluster head) that may occur. On the other hand, in LEACH, the aggregation of data is centralized and is performed periodically. However, in some cases, the periodic transmission of data may not be necessary, which exhausts rapidly the limited energy of sensors [7].

- The authors in [4] proposed, LEACH-centralized (LEACH-C) as an improvement of LEACH that uses a centralized approach to create the clusters. In LEACH-C, the BS collects the information of the position and energy level from all sensor nodes in the networks. On the basis of this information, the BS calculates the number of CH nodes and configures the network into the clusters. LEACH-C uses a simulated annealing algorithm to create the routing structure.

- The simulation results in [4] show that the average energy consumption of LEACH-C is lower than that of the distributed clustering schemes, such as LEACH, LEACH-E and HEED.

- The authors in [14] suggest, Genetic Algorithm Based Energy Efficient Clusters (GABEEC) in Wireless Sensor Networks that is implemented for the fixed WSNs. Two phases were defined, the first one elects some nodes as Cluster-Heads randomly. In the second phase, the Base station controls the node energy and the node with the highest residual energy becomes CHs.

Based on the above discussion, we propose an efficient energy-saving routing architecture that has a uniform clustering algorithm to reduce the energy consumption and to prolong the network lifetime in different domain of wireless sensor networks. The main idea of this algorithm is to reduce the data transmission distances of the sensor nodes in wireless sensor networks by using a uniform cluster structure and multi-hop concepts. We adopt centralized and cluster-based techniques to create the cluster routing structure for the sensor nodes based on our proposed FCBA scheme. To make an ideal cluster distribution for the sensor nodes, the distances between the sensor nodes are calculated, and the residual energy of each sensor node is accounted for selecting the appropriate CH nodes and the number on nodes for each cluster-head is selected. On the basis of the uniform
cluster location, the data transmission distances between the sensor nodes can be reduced by using the adaptive multihop approach. Our proposed scheme has the ability to provide more efficient control over network condition fluctuations. In addition, our system accounts for environments that are encountered in practice. A different number of sensor nodes and sensing areas are simulated and discussed in this paper. By using the suggested scheme, the energy consumption is reduced and the lifetime of the network is extended by balancing the network load among the clusters. The main benefits of our proposed scheme are the lower energy consumption and the longer network lifetime.

2.1 Network Model
The infrastructure of the network is composed of a BS and some sensor nodes. We assume that the sensor nodes are employed and equipped to be used in defined area, and the moving situation of each user is a random movement within the sensing area. We classify all sensor nodes into non-CH nodes and CH nodes. The non-CH nodes operate in the sensing mode to monitor the environment information and transmit data to the CH node. Additionally, the sensor node becomes a CH to gather data, compress it and send it to the BS from the CH mode. All sensor nodes in the network are homogeneous and energy constrained.

We can modulate the proposition scheme by a graph \( G=(V,E) \), where \( V \) represents the set of sensor and \( E=\{(u,v) \subseteq V/D(u,v)\leq R\} \) represents the wireless connection between nodes. \( R \) is the transmission range, and \( D(u, v) \) represents the Euclidian distance between the node \( u \) and the node \( v \).

### Table 1. Performance Comparison of Hierarchical Routing Protocol

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Operation network model</th>
<th>Routing structure</th>
<th>Principle of CH selection</th>
<th>Sensor information</th>
<th>Requirement of powerful BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>Distributed</td>
<td>Cluster</td>
<td>Random</td>
<td>Local</td>
<td>No</td>
</tr>
<tr>
<td>LEACH—C</td>
<td>Centralized</td>
<td>Cluster</td>
<td>Random and residual energy</td>
<td>Global</td>
<td>Yes</td>
</tr>
<tr>
<td>GABEEC</td>
<td>Distributed</td>
<td>Cluster</td>
<td>Random and residual energy</td>
<td>Local</td>
<td>No</td>
</tr>
<tr>
<td>Proposed</td>
<td>Centralized</td>
<td>Cluster</td>
<td>Residual energy and Weight</td>
<td>Global</td>
<td>Yes</td>
</tr>
</tbody>
</table>

2.1.1 Mathematic formulation for static sink mode
According to the model where the cluster-heads are in static mode (static sink model, SSM), the base station is centrally located and remains fixed throughout the duration of operation of the network. The routing of data to the base station is done through driver’s nodes. When the data become available in a node, it is transmitted to the base station. In general, the transfer speed of a sensor node are constant and always they are always designed by \( d_i \). The problem of maximizing the lifetime of our model is formulated as follows:

The constraint (2) is the “flow conservation constraint” which states that, at a node \( I \), the sum of all incoming flows for the commodity \( d_i \). The inequality (3) is the energy constraint and it means that the total energy consumed by a node during the lifetime (T) cannot exceed the initial energy of the node. With this formulation, the routing is dynamic and allows multipath
communications. There is no assumption on fixed-path routing, such as the shortest path routing. The above optimization problem can be easily converted into a linear programming (LP) problem.

2.1.2 Energy Consumption Model
In wireless sensor networks, data communications take most of the energy consumption. The total energy consumption is the average energy dissipated by data transmission of the non-CH nodes and the CH nodes. In addition, energy consumption for data collection and the aggregation of CH nodes are considered. Fig. 3 illustrates the same radio energy dissipation model used in [8–9]. This model presents a simple model for radio hardware energy dissipation, where the transmitter dissipates energy to run the radio electronics and the power amplifier, and the receiver dissipates energy to run the radio electronics. Both the free space propagation ($d^{-2}$ power loss) and the multipath fading ($d^{-4}$ power loss) channel models are used, depending on the distance between the transmitter and the receiver. Power control can be used to invert this loss by appropriately setting the power amplifier [4]. In this model, to transmit an L-bit message between the two sensor nodes, the energy consumption can be calculated by the following:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Number of nodes</td>
</tr>
<tr>
<td>$E$</td>
<td>Initial Energy</td>
</tr>
<tr>
<td>$E_c(i)$</td>
<td>Energy consumed by the node $i$</td>
</tr>
<tr>
<td>$X_{BS}$</td>
<td>The $X$ coordinate of the BS</td>
</tr>
<tr>
<td>$Y_{BS}$</td>
<td>The $Y$ coordinate of the BS</td>
</tr>
<tr>
<td>$DisBS(i)$</td>
<td>Distance between the node $i$ and the BS</td>
</tr>
<tr>
<td>$DisBS_{Max}$</td>
<td>The maximum distance to the BS</td>
</tr>
<tr>
<td>$Deg(i)$</td>
<td>Degree of $i$</td>
</tr>
<tr>
<td>$Weight(i)$</td>
<td>Weight of $i$</td>
</tr>
<tr>
<td>$ID(i)$</td>
<td>Identity of $i$</td>
</tr>
</tbody>
</table>

Table 2. Parameters Definition

\[
\max \sum_{j \in N(i)} x_{ij} - \sum_{k \in N(k)} x_{ik} = d_{i}, \quad i, j, k \in N \tag{2}
\]

\[
\left( \sum_{j \in N(i)} C_{ij} x_{ij} + \sum_{k \in N(k)} v x_{ik} \right) - T \leq E_i \quad i, j, k \in N \tag{3}
\]

\[
x_{ij} \geq 0 \quad i, j \in N \tag{4}
\]

\[
T \geq 0 \tag{5}
\]
Based on the above equations, $d$ refers to the distance between the two sensor nodes, $E_{TX}(L, d)$ denotes the total energy consumption in the transmitting sensor node, and $E_{RX}(L)$ is the total energy consumption in the receiving sensor node. $E_{elec}$ is the electronics energy consumption per bit in the transmitting and receiving sensor nodes, which depends on factors such as digital coding, modulation, filtering, and spreading of the signal. $\mu$ amp is the amplifier energy consumption in the transmitting sensor nodes, which can be calculated by

$\varepsilon_{amp} = \begin{cases} 
\varepsilon_{fs} \cdot d^2, & \text{when } d \leq d_0 \\
\varepsilon_{mp} \cdot d^4, & \text{when } d > d_0
\end{cases}$

(8)

where $\varepsilon_{fs}$ and $\varepsilon_{mp}$ are communication energy parameters, and $d_0$ is a threshold value. If the distance $d$ is less than $d_0$, then a free-space propagation model is used. Otherwise, the multipath fading channel model is used.

The consumed energy by each sensor consists of three components: A communication (transceiver) unit, sensing unit and processing unit. Therefore, we will only consider the consumed energy by the communication unit (the consumed energy by the sensing unit and the processing unit is neglected):

$E_c(i) = E_{comm}(i)$

(9)

The consumed energy by the communication units breaks up into transmission energy $E_{TX}$ and reception energy $E_{RX}$, so:

$E_c(u) = E_{TX}(k, d) + E_{RX}(k)$

(10)

For $E_{TX}$ and $E_{RX}$, we apply the same model proposed in [11] which is as follows:

$E_{TX}(k, d) = E_{elec} \cdot k + \varepsilon_{amp} \cdot k \cdot d^2$

(11)
\[ E_{Rx}(k) = E_{elec} \times K \]

Where:

- \( E_{elec} \) = Electronic energy.
- \( K \) = Message length (bits).
- \( e_{amp} \) = Transmit amplifier.
- \( d \) = Distance between transmitting node and receiving node (m).
- \( \lambda \) = Path loss exponent.

### 2.2 Cluster Formation Phase

In this section, the CH selection phase is described. Our objective is to balance energy consumption between nodes. For each cluster, the CH selection procedure is achieved relying on three parameters: energy, degree and distance:

- **Energy**: Is a very important parameter for WSN. In fact, a node with low energy reserves is enable to transmit information especially if the distance between nodes is high and requires significant power. For this reason, many researchers such as [14–15] and [16] have used this parameter during the CHs selection phase to balance the load (energy consumption) between nodes in order to provide a stable structure and to increase the network lifetime.

- **Degree**: The node degree is the number of its neighbors (1-neighborhood). Researchers in the field such as [23, 17, 18] and [19] chose degree as CHs selection criterion. The node with the highest degree is selected as a CH. This approach provides a stability to the structure.

- **Distance**: The distance between nodes is also a very important criterion in the CHs selection phase. This parameter influences the transmission power and the number of hops to reach the destination and, consequently, the energy used to make the transmission. Many other researchers choose this setting to minimize the energy consumed during the inter cluster communication [20].

Since these parameters are very interesting to minimize energy consumption and to produce a stable structure. These will be used in the CH selection phase. The CH must have the highest residual energy to be able to manage the intra-cluster and inter-cluster communications. We choose the CH with the highest degree in order to be a neighbor of several CHs and facilitate the inter-cluster communication. To minimize the consumed energy in inter-cluster communication, the nearest node from the cluster center is selected as CH.

For each cluster, nodes calculate their combined weights, then, the node with the smallest weight becomes a CH. For each round, this procedure is repeated in order to balance the energy consumption between nodes.

The CHs selection phase is described in Algorithm 2.

Figure 4 illustrate the different steps of FCBA. The first phase is sending hello message to explore the neighborhood, then each node calculates and distributes its weight, the node with the lowest weight expresses becomes a CH and others declare themselves as Nm.

Without losing generalization all simulations are evaluated in the same conditions using MATLAB R2014a tool. We consider N nodes number varies between 50 and 200 nodes randomly distributed in a 100m x 100m field. We assume that the base station is fixed in the center of the sensing region. To compare the performance of FCBA with other protocols, we ignore the effect caused by signal collision and interference in the wireless channel. Each Sensor node is assumed to have an initial energy of 1.5 joules. The network is organized into a clustering hierarchy, and the cluster-heads execute fusion function to reduce correlated data produced by the sensor node within the clusters. The parameters of simulation are described in Table III. To avoid the frequent change of topology, we assume that the nodes are in static mode; the protocol compared with LEACH and GABEEC.
Assumptions (CH's election procedure)

**Algorithm : CHs election phase**

For each cluster head c do

For each node i do

Send ‘hello’ message

Find the degree

$E = \{(u,v) \in V/D(u,v) \leq R\}$

$\text{Deg}(u) = |E|$

Calculate the consumed energy $0.05j < E_c(i)/E_o \leq 0.5j$

Calculate the distance to the BS

$\text{DisBS}(u) = \sqrt{(x-x_{SS})^2 + (y-y_{SS})^2}$

Calculate the weight

$\text{Weight}(i) = 1/\text{Deg}(i) + 0.05j < E_c(i)/E_o \leq 0.5j$

$\text{DisBS}(i)/\text{DisBS}_{\text{Max}}$

end

if weight(i) = smallest_weight(i) then

state(i) = CH

send ‘clusterhead accepted’

else

state(i) = ON

Send ‘clusterhead-accepted’

end

end
Simulation and Results

Without losing generalization all simulations are evaluated in the same conditions using MATLAB R2014a tool. We consider N nodes number varies between 50 and 200 nodes randomly distributed in a 100m x 100m field. We assume that the base station is fixed in the center of the sensing region. To compare the performance of FCBA with other protocols, we ignore the effect caused by signal collision and interference in the wireless channel. Each Sensor node is assumed to have an initial energy of 1.5 joules. The network is organized into a clustering hierarchy, and the cluster-heads execute fusion function to reduce correlated data produced by the sensor node within the clusters. The parameters of simulation are described in Table 3. To avoid the frequent change of topology, we assume that the nodes are in static mode; the protocol compared with LEACH and GABEEC.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Area</td>
<td>100*100</td>
</tr>
<tr>
<td>Probability of a node to become cluster head</td>
<td>0.05</td>
</tr>
<tr>
<td>Initial energy</td>
<td>1.5j</td>
</tr>
<tr>
<td>Base station</td>
<td>50*50m</td>
</tr>
<tr>
<td>Transmitter/Receiver Electronics</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>100 &amp; 300</td>
</tr>
<tr>
<td>$\varepsilon_{fs}$</td>
<td>10 pJ/bit/m$^2$</td>
</tr>
<tr>
<td>$\varepsilon_{mp}$</td>
<td>0.0013 pJ/bit/m$^4$</td>
</tr>
</tbody>
</table>

Table 3. Simulation Parameters

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>50</th>
<th>100</th>
<th>150</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>0.0083</td>
<td>0.008</td>
<td>0.0081</td>
<td>0.0084</td>
</tr>
<tr>
<td>GABEEC</td>
<td>0.0059</td>
<td>0.0057</td>
<td>0.0061</td>
<td>0.0058</td>
</tr>
<tr>
<td>FCBA</td>
<td>0.0011</td>
<td>0.0031</td>
<td>0.0030</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

Table 4. Energy consumption of FCBA VS LEACH and GABEEC

Table 4 shows the energy consumption of nodes for FCBA and LEACH and GABEEC. This table clearly depicts that FCBA has a better performance than the two other schemes in terms of energy consumption. These parameters can resume all other simulations parameters like the number of rounds etc.

Fig.5 shows the average energy dissipation in sensor nodes in three types of routing protocols, namely LEACH, GABEEC and FCBA, within a range of 100 x 100 m$^2$. The energy dissipation found in LEACH is greater than that in GABEEC and FCBA as a whole. This is because LEACH adopts single-hop communications with the CH sending its data directly to the BS; GABEEC and FCBA utilize multi-hop communications that require less energy consumption from each sensor node. In addition, FCBA utilizes network topology and applies an energy consumption equation to locate the ideal CH. FCBA balances the load on each CH by considering cluster sizes and CH locations. In this way, energy spent by sensor nodes close to the BS is less than in GABEEC, so the average energy dissipation in FCBA is a lower than in GABEEC. The selection of a new CH in FCBA occurs when the energy of the current CH is below a given threshold, but the loca- tions of CHs deviate from the original ones.
Figure 5. A comparison of FCBA average energy dissipation with LEACH and GABEEC

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Topology</th>
<th>FND</th>
<th>LND</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCBA</td>
<td>Random (50)</td>
<td>2950</td>
<td>9000</td>
</tr>
<tr>
<td></td>
<td>Random (100)</td>
<td>1600</td>
<td>7000</td>
</tr>
<tr>
<td></td>
<td>Random (150)</td>
<td>500</td>
<td>3150</td>
</tr>
<tr>
<td>LEACH</td>
<td>Random (50)</td>
<td>880</td>
<td>1400</td>
</tr>
<tr>
<td></td>
<td>Random (100)</td>
<td>750</td>
<td>1070</td>
</tr>
<tr>
<td>GABEEC</td>
<td>Random (150)</td>
<td>175</td>
<td>700</td>
</tr>
<tr>
<td></td>
<td>Random (50)</td>
<td>2130</td>
<td>4000</td>
</tr>
<tr>
<td></td>
<td>Random (100)</td>
<td>1800</td>
<td>3250</td>
</tr>
<tr>
<td></td>
<td>Random (150)</td>
<td>1500</td>
<td>2600</td>
</tr>
</tbody>
</table>

Table 5. The number of nodes alive

This table shows that FCBA balance of energy consumption is better than that of LEACH and GABEEC. FCBA, [8], [6] are also tested on different WSNs with 50, 100 and 150 nodes. Table V represents a comparison between FCBA, [8] and [6] in terms of FND and LND. When considering LND, results show that FCBA is more than 80% more efficient than the optimal parameters, these results reflects the impact of energy consumption balance achieved by the proposed scheme which tends to divide the energy consumption among the remaining nodes for each round. Besides, we can resume that the position of the base station plays a crucial role in the stability and energy consumption of nodes. In fact, if the base station is misplaced, the quality and energy consumption become more defective. For example, the simulation of FCBA with 100 nodes and position of sink (50, 50) is better than 30 % of the results with the sink position (95, 95).
3. Conclusion

In this paper, a new energy-efficient clustering algorithm for fixed WSNs was designed in order to minimize energy consumption, balance energy consumption between clusters and balance energy consumption between nodes. The designed clusters are static and having unequal sizes. The clusters having the same distance to the BS have equal sizes. The CHs are selected according to their energy reserve, their degree and their distance to the clusters centers. Simulation results show that the proposed algorithm is effective in balancing energy consumption and prolong the network lifetime. Upcoming work will compute the impact of the deployment heterogeneous nodes having high energy capacity in the energy consumption and will consider more factors to select CHs.

References


