Designing a Hybrid Clustering Routing Algorithm based on Cellular Learning Automata for Optimizing Lifetime of Wireless Sensor Networks

Reza Dovlatabadi¹, Mojtaba Mohammadpoor²

¹Department of Electrical Engineering, Islamic Azad University, Gonabad Branch, Gonabad, Iran
²Department of Electrical Engineering, University of Gonabad, Gonabad, Iran
E-mail address: reza.dovlatabadi@gmail.com, m.mohammadpoor@gmail.com

Abstract
One of the most important factors in wireless sensor networks is energy consumption, hence the lifetime of these networks are strongly depending on remaining energy in the nodes. According to sensors placement farness and wireless communication between them, it is necessary to optimally consume the energy in these networks. In this study a hybrid approach is proposed by mixing two existing protocols, namely flat multi-hop routing and hierarchical multi-hop routing. Also by using Cellular Learning Automata (CLA) as clustering technique, the energy in the network will be managed and finally the lifetime of nodes will be increased. Mathematical simulation and analysis show a good performance of clustered hybrid model for energy saving that in compare with multi hop routing algorithm and hierarchical routing in non-clustered and clustered conditions, the lifetime increasing are 10.39%, 27.36% and 5.57%, 23.83% respectively.

Keywords: Lifetime, Wireless Sensor Networks, Flat Routing Protocol, Hierarchical Routing Protocol, Hole problem.

1. Introduction
Wireless sensor networks are composed of sensors, small nodes and wireless communication. [1-3] the overall structure of a wireless sensor network is composed of sensor nodes and sink nodes. The application scope of these networks is in military affairs, health care, environmental monitoring and many other topics. Due to the extension of these networks and its diverse topics, many researchers have been interested in research in this area. Wireless sensor networks consist of a large number of cheap and small wireless sensor nodes with high ability that are installed and aims to explore the events in the monitored environment. Each sensor node is equipped with a sensor, transmitter, processor, memory and battery. Wireless sensor nodes are often deployed randomly in the environment. And this issue helps to the rapid network development in inaccessible areas and high-risk environments. The random deployment of sensor networks require network self-organizing capabilities [4]. After deployment, the network collect information from the desired region and send information to rotors and the network coordinator, finally it will reach to the end users. This ability of wireless sensor networks in simultaneous communication helps to manage many unlimited sensors nodes to work as a network. Moreover, it can accept new sensor nodes without any complex reconfiguring in the network, in contrast to other common wireless networks. Different network structures are developed according to the applications and assessment process characteristic. The main structure of sensor networks can be classified in three categories:

- A flat structure in which each node plays a role of flat node and it is used for detecting and receiving the environmental information.
- Rank structure or hierarchical protocol which is composed of a sink node, rotor and an adjustment node and based on this characteristics, it performs different process.
- A structure that is based on location and the place of sensor nodes are found based on data in paths and in networks.

In some practical studies, Cellular Learning Automata (CLA) has been utilized in modeling of wireless sensor networks. Cunha et al proposed a CLA based algorithm for sleep-wake process of making a timetable of wireless sensor networks [5]. This technique also have been used to design systems to preserve the security of wireless sensor networks in other studies [6, 7]. Torbey is proposed a cellular learning automaton based local problem-solving procedure to enlarge the coverage of a network that is limited to a bounded region in [8]. However, the algorithm did not consider any random initial deployment and the connectivity protection criterion of their sensor network. Athanassopoulos and et al [9] proposed cellular automata...
for topology control in wireless sensor networks. They had used cellular automata to model a randomized WSN topology control algorithm.

One of the major problems of sensor networks is amount of energy consumed or saved in the system, because the network feeds using rechargeable batteries. The nodes close to the sink are more prone to be deactivated due to their more activities. Sometimes repetition in power discharging of batteries is so high that leads to decrease its lifetime. In most cases, because of the extensive of the assessment area, replacement of batteries is not applicable and it speed the reduction of sensor networks lifetime. The goal is to increase the lifetime of wireless sensor networks and facing with the Hole problem and its solution that effect on the network lifetime. Hole problem is caused by disconnecting a part of network. Since the data communicate among sensors are guided to reach to the sink through the close sensors to the sink, this nodes do more work and hence their lifetime is less than other nodes, and if dead it causes the separation of sink from the rest of the network and the Hole problem will occur as a result [10].

2. Available Protocols

In some proposed protocols, the energy consumption is high due to load unbalancing process. Some methods are available for optimizing energy. Most of them are relied on routing algorithms. They can be categorized as non-ranking and ranked protocols. Non-ranking protocols accelerate the energy compensation through an efficient solution in routing and data transmission. Some of techniques fall in this group are Gossiping [13], Flooding [11], SMECN [12], Data Centric [11], Directed Diffusion [14], and SPIN [15]. In these methods, there are direct path between nodes and the base station, hence the major problem of them is the high consumption of energy because the nodes should send data through this long distance. Another parts of the contracts are ranked protocols or hierarchical protocols including static clustering, such as LEACH [16], PEGASIS [17], HIT [18], LPT [19], DBS [20]. In many cases of these methods, duplicate data is received from two nodes and will be waste the energy, means sensor nodes consume energy to send the same information twice.

Routing in is different with the wireless sensor networks, in contrast to fixed networks that do not require any infrastructure, are unreliable and they need protocols to minimize their energy consumption [21]. Clustering techniques usually consist of two phases in wireless sensor network; in first step, clusters must be initialized and second step it should be involved in a re-clustering phase [22-24]. In first phase each node of the sensor network can be either selected as a head or a member. The members periodically gather data from the environment, process it and then send it out to the heads. Therefore, head nodes consume lots of energy. All of their members within clusters send data and then aggregation of data is happened. Transmission range of intra-cluster communications is higher than inter-cluster communications. As a result we encounter with energy exhaustion in heads and so to prevent this event, re-clustering process is used to exchange head or member nodes.

The main objective of this project is to increase the WSN lifetime by collecting the power pieces and using a hybrid algorithm from the energy storage protocols and then by local re-clustering the nodes in dynamic clusters as heads or members. In the next part, a brief structure of WSN will be provided. In the introduction section of the proposed algorithm, Energy Aware Routing and its different steps will be examined. In the section 4, the CLA will be introduced. Also in section 5, the proposed model which is a hybrid multi-hop protocol is presented. Finally we show the experimental results and discussion in sections 6 and 7 respectively.

3. WSN Structure

In the wireless sensor networks, the nodes can be categorized as sensor nodes and sink nodes. The sensor nodes are used for sensing and due to the different applications that are available for sensor nodes in the desired network, in general, data collection is the responsibilities of sink nodes [25]. In Fig. 1, a structure of wireless sensor networks with sink nodes and sensor nodes is displayed.

3.1. Multi-hop Routing Algorithm

Routing algorithm is looking for the best path among the available paths based on the specific criteria. The best
criterion which is the interest of researchers is to maximize the lifetime of wireless sensor networks. And its function is that a path should be chosen so that its energy consumption divide fairly among all nodes. This protocols are known as energy awareness protocols. One of these kind of protocols is Multi-hop Routing Algorithm where its schematic is shown in Fig. 2 where that sensor nodes send packets by gathering the environmental information to sink nodes in various paths.

The energy awareness protocols have three different phases:

A) **The launching phase:** in this phase, each node sends a message to its neighbors to find its neighbor flooding and make a routing table. Costs in communication channel are calculated by using (1), (2) and (3).

\[
\text{link cost}(i, j) = e_r(i) + e_r(j) \quad (1)
\]
\[
e_r(i) = \varepsilon_1 d_{ij}^2 + \varepsilon_2 \quad (2)
\]
\[
e_r(j) = \varepsilon_3 \quad (3)
\]

where \( \text{link cost}(i, j) \) is defined as the consumed energy for sending a unit data from node \( i \) to node \( j \), \( e_r(i) \) is the consumed energy to send a data unit to the receiver node \( j \), \( \varepsilon_1, \varepsilon_2, \varepsilon_3 \) are constants depend on the send and receive characteristics of sensor nodes. Other algorithms [26] have been proposed in this way and it distributes the energy consumption between nodes uniformly according to equation (4).

\[
\text{link cost}(i, j)_{\text{new}} = \frac{\text{link cost}(i, j)}{E_r^i} \quad (4)
\]

Due to the remaining energy of transmitter node in the denominator, the chance of being selected the node as rely node is depend on the remaining energy. Because the energy of specific nodes does not end quickly, so the energy consumption among nodes is distributed uniformly, while the energy consumption of the whole network will be minimal.

B) **Data transfer phase:** in this phase, each node sends the received packets based on the routing table, randomly to one of the paths. It is possible to select different paths for this purpose.

C) **Path maintenance phase:** it is necessary to check paths regularly and keep them alive.

![Figure 2. The schematic of Multi-hop Routing Algorithm](image)

3.2. Hierarchical Routing Protocols

In these methods, network nodes are converted to a series of clusters. In each cluster, there is a cluster head that gather data from other group nodes and send it to the sink. Scalability is one of the most important factors in designing the routing protocols. The main objectives of these protocols are to reducing the consumption of network nodes and this work is done in two ways. First, we can use the multi-hop approach, in other words in each cluster, data will be sent in a multi-hop manner and it will cause less energy consumption than direct mode. Second, we can use Data Fusion and Data Aggregation that reduce the number of sent packets and data will be sent with less volume [27]. One of the concrete examples that is well known in wireless network protocol and it is a hierarchical adaptive algorithm, LEACH algorithm [16, 28]. In Fig. 3 the schematic of Hierarchical Routing Protocol is shown.
A network with static hole has a hole problem even when the sink is in its desired areas mean the sensor core \(^{27}\). In Fig. 4, an example of this problem is shown. With loss of all nodes located in the hole, data aggregation from many live nodes will be impossible. Because there is no communication channel between sink node and other nodes out of its range. If the Hole is in motion, it can move around the measurement range and collecting data and this issue lead to balance in the network energy consumption. Also using the sink and mobile sensors raises the cost of network expansion. The proposed clustering technique in comparison to existing techniques presents a clustering method in which each cluster has higher number of nodes and higher unused energy for the cluster heads.

4. Cellular Learning Automata

Cellular learning automata (CLA) which is introduced in \(^{29}\) is an efficient mathematical pattern for many dispersed situation that requires a solution and environmental problems. The basic idea of Cellular learning automata is the use of learning automata to modify or adapt the state transition probability of stochastic phenomenon. Cellular automata rules can be defined as a bit string in which each bit represents the position corresponding to the number of bits. For example, Fig. 5 shows one-dimensional cellular automata with two-dimensional neighborhood that is a bit string. If the bit string convert decimal number, the name of the law will be achieved. Learning Automata (LA) is expressed by using parameters \(\{\alpha, P, \beta\text{ and } T\}\) where \(\alpha\) is a set of actions \(\{\alpha_1, \ldots, \alpha_r\}\), \(\beta\) is a set of automata entries \(\{\beta_1, \ldots, \beta_m\}\).

Also we define \(p\) the action selection probability vector \(\{p_1, \ldots, p_m\}\) and relation of \(p\) and \(T\) is defined as (6) \(^{30}\):

\[
p(n+1) = T[p(n), \beta(n), p(n)]
\]

This equation is the Learning Algorithm. If the \(\alpha_i\) action in \(n\)th level is happened, and received favorable responses from the surrounding environment, the contingency of \(p_i(n)\) will be increased while other possibilities will be decreased. However, these changes must be in such a way that the summation of \(p_i(n)\) remains constant. The favorable response of the surrounding environment is expressed as (6):

\[
p_i(n+1) = p_i(n) + A[1 - p_i(n)]
\]

And \(p_i\) is resulted from (6) and is defined based on (7):

\[
p_j(n+1) = (1-A)p_j(n) \quad \forall j \text{ that } j \neq i
\]

In similar way, we can define the undesirable response as (8) and (9):

\[
p_i(n+1) = (1-B)p_i(n)
\]

\[
p_j(n+1) = \frac{B}{r-1} + (1-B)p_j(n) \quad \forall j \text{ that } j \neq i
\]

Where \(A\) is the augmentative parameter and \(B\) is the optimal parameter. By comparing the \(A\) and \(B\), the three conditions occur based on (10):
If A = B Then the scheme is Linear Penalty \( (L_{RP}) \)
If A > B Then the scheme is Linear Reward epsilon Penalty \( (L_{RI}) \)
If A < B Then the scheme is Linear Reward Inaction \( (L_{R-\epsilon}) \)

\[ \text{max} \ E \text{max} \ E = \sum (\text{max} \ E) \]

\[ \text{If} \ A \ B \ \text{Then the} \ \text{scheme is Linear Penalty} \ (L_{RP}) \]
\[ \text{If} \ A > B \ \text{Then the} \ \text{scheme is Linear Reward epsilon Penalty} \ (L_{RI}) \]
\[ \text{If} \ A < B \ \text{Then the} \ \text{scheme is Linear Reward Inaction} \ (L_{R-\epsilon}) \]

\[ \text{CFe} = \text{w}(1-w)(\text{d}_{\text{max}1}+\text{d}_{\text{max}2}) \]
\[ \text{CFd} = \text{w}(\text{E}(i)+\text{E}(j)) \]
\[ \text{CF} = \frac{\text{CFe}}{(\text{CFe} \times \text{d}_{\text{max}1})} \]

5. Proposed Hybrid Multi-hop Routing Algorithm

In the node sets, fewer numbers are in hole areas and then
the amount of produced data by nodes is negligible in the
hole area. On the other hand in the hole, the amount of
energy consumption is high due to the data relays that are
come from outside. In order to reduce the energy in the
hole, it should:

- Reduce the data streams that are entered to the hole
area.
- The energy consumption for Relay for a data unit
that is floods from hole to the sink, should be
minimized.

In fact, by using the hierarchical multi-hop routing
algorithm outside the hole area, we can decrease the data
stream. In addition, by using flat multi-hop routing
algorithm in the hole area we can reduce the transmission
distance between nodes.

The proposed algorithm makes use of hybrid approach
idea. Before clustering with cellular learning automata, in
first step by using hybrid approach that is consists of two
models which are flat multi-hop routing and hierarchical
multi-hop routing, the topology of network is modified.
During the process, the hole problem will be solved. For
this, according to Fig. 6, in this paper the objective is to
achieve the above solutions. These two solutions will be
happen by using the hybrid multi-hop routing algorithm.
After implementation of hybrid multi-hop routing
algorithm, clustering procedure is conducted that itself is
combined two phase. First phase is initial clustering and
next phase is Re-clustering phase. In this algorithm, each
sensor node is equipped with a learning automaton (LA).

In addition, each r action \([30]\) consists of \(a_1, a_2, \ldots, a_n\). In
each r, there are number of neighboring nodes of LA and p
is probability vector of choosing of LA that is action of
choosing of i to the neighbors of the cluster-head. In Fig. 7
the proposed model of clustering by learning automata is
shown. In the proposed protocol, all sensors know the
location of base station.

The sensors calculate their distance to the base station on
the basis of this knowledge and send and receive messages.
Because the distance between the sensor and the base
station is one of the important parameters in the selection
of cluster heads and cluster size. By putting these
hypotheses in the circular loop, each node calculates its
distance from other nodes. If the calculated distance is less
than the radius, weighted function between two examined
nodes will be computed.

The distance and energy parameters are considered in the
function that calculates the amount of weight and the end a
specific weight is assigned to the node. The weight
functions are expressed in (11), (12) and (13):

\[ \text{CFe} = \text{w}(\text{E}(i)+\text{E}(j)) \]

\[ \text{CFd} = \text{w}(\text{d}_{\text{max}1}+\text{d}_{\text{max}}) \]

\[ \text{CF} = \frac{\text{CFe}}{(\text{CFe} \times \text{d}_{\text{max}1})} \]

Figure 5. This example shows dimensional cellular automata with two-dimensional neighborhood.

\[ \text{Figure 5. This example shows dimensional cellular automata with two-dimensional neighborhood} \]
It should be considered that after calculating the value of weight function, if the primary node and examined node are i and j respectively, the computed value will be stored in S(i).neighbor(j) variable and selected as primary node. In addition, if the primary node does not cover examined node, this variable will be assigned as -1. In next step, the proposed procedure selects a node as the cluster head that has chance to be candidate compared to other nodes. This means that the node that has the greatest neighborhood is selected as candidate node and it becomes Tentative. In next step, all of other nodes are checked. If it is in examined radius, and not trial node and its parameter is greater than or equal to zero, then the primary node will selected as cluster head and the end next node assigned as member of examined cluster. Also the value of machine parameter (a=0.01) will be added to the machine parameter in primary node entitled S(i).MashinPara variable and at the same time this value is subtracted from the machine parameter in next node that is member of examined cluster entitled S(j).MashinPara variable. Then by selecting a candidate node and comparing it by other nodes, if not found any cluster head, the candidate node will assigned as cluster head. Otherwise the candidate node is converted as normal node. If the normal node is not found in range of the cluster head based on radius, will considered as member of examined cluster. In proposed algorithm, we find normal nodes. Then by finding the first normal node, the algorithm will try to find first cluster head. Likewise, if algorithm finds a cluster head in range of examined radius, is selected as member of it and the end if network fails to
find either the main cluster between all of clusters head, itself will be a cluster head.
In final step of proposed clustering procedure, the cluster heads found in all iterations and as a result, the ‘a’ variable is changed to 0.01 or -0.01. The cluster head is selected as a priority that has grater machine parameter in final iteration.

6. Experimental Results
6.1. Performance Evaluation
To evaluate the effectiveness of the proposed algorithms in order to increase the WSN lifetime in dealing with the hole problem, the parameters configuration and initialization is required that are shown in Table 1. Sensor nodes are placed randomly in the central node measurement to the sink. This experiment was designed so that each sensor can produce a single package periodically and it can transfer all packages to the sink node. In this paper, we use Matlab software to simulate the proposed algorithm. We compare our results with LEACH and EEUC protocols based on mentioned condition in Table 1.
In first level, summary of the network lifetime is shown in Table 2. That the long life time is determined and it is avoiding separating sink, due to the energy analysis of all nodes in the hole. By comparing the lifetime of proposed algorithm with multi hop routing algorithm and hierarchical routing in non-clustered and clustered conditions, the lifetime increasing are 10.39%, 27.36% and 5.57%, 23.83% respectively.
In overall, the cost function of energy saving hybrid system is introduced and finally the results of simulation software will be provided.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Size</td>
<td>200×200</td>
<td>location of the base station</td>
<td>(100,250) /m</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>400</td>
<td>frame size of Control data</td>
<td>100 /Byte</td>
</tr>
<tr>
<td>$E_{elec}$</td>
<td>50 nJ/bit</td>
<td>D_MAX</td>
<td>140 /m</td>
</tr>
<tr>
<td>Data frame size</td>
<td>4000/Byte</td>
<td>Perceprion radius</td>
<td>15 /m</td>
</tr>
<tr>
<td>Distance threshold</td>
<td>86.6 (d_i)/m</td>
<td>Data fusion of energy</td>
<td>/ (pJ-bit$^{-1}$)</td>
</tr>
<tr>
<td>Maximum radius of competition</td>
<td>90 /m</td>
<td>The initial energy of nodes</td>
<td>/J</td>
</tr>
<tr>
<td>$e_0$</td>
<td>0.0012/(pJ-bit$^{-1}$,m$^{-4}$)</td>
<td>c values</td>
<td>0.6</td>
</tr>
<tr>
<td>$e_1$</td>
<td>10/(pJ-bit$^{-1}$,m$^{-3}$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The nodes distribution is hypothesized that in their density there is not much deviation. Energy consumption in the hole and network lifetime is selected as performance measures. Also, network lifetime is defined as a period that all energy capacities of nodes will be reduced. At this time, the sink node is separated from all network nodes that are out of hole.

<table>
<thead>
<tr>
<th>Routing technique</th>
<th>Non-clustered</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>5156</td>
<td>5762</td>
</tr>
<tr>
<td>Hierarchical</td>
<td>4469</td>
<td>4912</td>
</tr>
<tr>
<td>Hybrid</td>
<td>5692</td>
<td>6083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elapsed Time (Sec)</th>
</tr>
</thead>
</table>

6.2. Discussion
The hierarchical routing is highly dependent on the cluster heads. Hence, we should be careful in selecting the cluster heads. Proper distribution of cluster heads is best strategies in the design of proposed procedure that is not considered in the in the LEACH and EEUC models. The statistical comparison in distribution of cluster heads is displayed in Fig. 8 for LEACH and EEUC protocols based on the number of cluster heads and their number of nodes. Clearly, the selection of the correct cluster head according to specific situations leads to optimization of energy consumption. Thus, for each case, other protocols such as EEUC or LEACH are compared in terms of energy efficiency with proposed strategy. Even with the removal of one cluster head, the energy efficiency will increase in comparison of previous configuration. In this paper, we select 10 random cluster heads to assess the level of energy consumption as the Fig. 9. Consider that the proposed procedure consumes less energy than other two methods and thus the network lifetime increased. Efficiency at low energy consumption is best evidence of proper performance of the proposed strategy. As seen in Fig. 10, the survivability of the nodes in a simple network with specific base station location were compared in three protocols. Obviously, when the first node is dead, the interval between the half time and the last time of death node is better than two other protocols.
All aspects considered for simulation and hence the amount of energy consumed during the clustering phase is about 0.0003 of the total energy dissipated during the network lifetime.
7. Conclusion

The this study is to achieve an optimal approach for clustering in Wireless Sensor Networks to increase their lifetime based on Cellular Learning Automata (CLA) clustering and design of Hybrid clustering routing algorithm.

This procedure has less computational complexity because uses of CLA, also the hole problem can be solved. Since all operations for cluster formation are done locally a large amount of energy is saved and speed of cluster formation is increased. In addition, this proposed algorithm is extremely robust because using appropriate input for the CLA algorithm. We proposed an efficient clustering approach by combination good features of clustered hybrid model for energy saving that was combined of multi hop routing algorithm and hierarchical routing approach. The simulation results shows that this method works better than other clustering approaches such as multi hop routing algorithm.

References


