Performance Analysis of Chain-Cluster based Routing Protocols in WSN

Aya Ayad Hussein¹, Rajaaalden Abd Khaled²
¹,²collage of information Engineering, Al-Nahrain University, Baghdad, Iraq

Abstract: Wireless Sensor Networks (WSNs) is one of the effective techniques that being strongly considered in scientific and engineering fields. WSN consists of small nodes with sensing, computing and communicating wireless abilities. These Sensor nodes usually have a limited lifetime duration due to the limited power supply. The main aim of WSN is to sense all the information from the environment (The environment can be an Information Technological framework, a physical world, or a biological system) based on the kind of application for which is deployed and send this information to it is Base Station (BS). Sensor nodes have to ensure that their effort to complete a task fit with their strict energy budget, this constraint makes the energy resource the most of critical importance in the WSNs. Sensor Nodes communicate with each other by different Routing Protocols, and Routing Protocols can be classified into different categories in WSNs. This paper will focus on the second category which is the Hierarchical (cluster-based) routing protocols, more precise the chain based routing protocols. As a result, we will go deeply in PEGASIS protocol backgrounds and improvements by showing each stage description, models, and analyzing their performance by simulate these protocols in Matlab simulator to show the progress of each protocol due to the number of dead nodes during time which proves increasing the rate of prolonging the network lifetime to offer a complete vision for these protocols through making a short detailed table which contain the basic performance measurements of these improved protocols.

Key Word: Wireless sensor network, Cluster-based routing protocols, LEACH, PEGASIS, chain based routing protocols

I. Introduction

Due to the advances in the WSN technology field, it becomes necessary to explore new ways or new techniques such as improving routing protocols by using many intelligent systems and optimization algorithms to keep abreast of developments that affect in WSN technologies positively. These routing protocols applied to small and cheap sensor nodes to achieve efficient communication between these sensors nodes in the whole network. The architecture of sensor nodes is showing in Figure 1. These sensor nodes are very sensitive in terms of energy that will lead to limited energy supply and in turn, will cause a short network lifetime, to recover this issue we have to use efficient routing protocols that will ensure efficient and reliable communication between these nodes. In this paper, the author presents a simple comparative study about a specific protocol, PEGASIS protocol and its improvements proposed in recent years [1-5]. The rest of the paper organized as follow. Section 2, describes simply routing protocol and it is a classification with focusing on the second type of routing protocol, hierachical (cluster-based) routing protocols. Section 3, describes the PEGASIS routing protocol with it is genesis and improvements and shows their models. Section 4 analyzes the performance of each of these protocols by simulating them to compare each protocol with the other one due to their network lifetime, and section 5 concludes the paper.
II. Routing Protocols in WSN

Routing is a mechanism that initiates a route between a source node and a destination node. The main goal of routing protocols is to discover the best route (shortest path) in terms of energy consumption, delay, quality of service and other metrics to make sure to extend the network lifetime by keeping the sensors node alive as can as possible [3-5]. Therefore, WSN routing protocols can be sorted and classified based on different metrics, based on network structure the routing protocols classified into three main categories:

- Flat routing protocols.
- Hierarchal (cluster-based) routing protocols.
- Location-based protocols.

In this paper, we will focus on the routing protocols that classified based on the network structure especially those called chain based routing protocols which are one of hierarchal routing protocols types. In a flat routing protocol, all the nodes are working together based on the same role in the same manner and because of the limited resources, this type is not preferred or not used a lot at large scale networks. In hierarchal routing protocols all the nodes working together but in different scenarios by distributed these nodes into groups called cluster heads and each grope execute its own tasks, this will provide many useful and supported features to the network such as scalability, energy efficiency, and increased network lifetime. Finally, the location-based routing protocol is that path creation between each node will depend on each node position. [5-7].
Hierarchal routing protocols

Hierarchal routing protocols also called cluster-based routing protocols. It is a concept stands for dividing the network into many regions of interconnected nodes these regions called cluster, each cluster contains a node called cluster head which has a specific feature that it owned the highest energy among all the nodes in the same region. These clusters also divided into a layered structure, it usually contains two layers. The idea of sending data or information between nodes in this hierarchal or layering structure based on sending information from the lowest energy nodes to the highest energy nodes. This means that the nodes that have the lowest energy are responsible for sensing and sending information to cluster head while the nodes have the highest energy is using for processing information and send to another cluster head or to the base station by using gateway nodes. This type of routing has many benefits and damages on an ad-hoc network, it is basically minimized on-demand route discovery traffic and routing overhead, reduce route determination delay and increase packet delivery ratio. On the other hand, it is a negative effect on the network rises when the clusters increase in size because it will increase packets overhead due to it is a routing source. Also, increasing packet size because of the operation that happened when every node of the route must be stored in the routed packet so the more route increases the more packet size will increase. That leads us to a bigger problem is the increase in transmission time because of the two previous negative points. So, it was helpful when this kind of routing protocol is classified into the different architecture of sensor node deployment based on the nature of the task or application [4-7].

III. Stages of building PEGASIS protocol and improvements

In this section, you will find a complete idea about the hierarchal routing protocols. Especially, that belongs to the chain cluster-based routing protocols. However, each protocol has its own special mechanism of communication and data transmission between their interconnected nodes but most of these protocols are depending on the first order radio model to make the process of transferring data completed perfectly. Therefore, to transmit a number of bits message with a certain distance using this radio model, the radio expends [8]:

\[
\begin{align*}
E_{tx}(k,d) &= E_{tx} - elec(k) + E_{tx} - amp(k,d) \\
E_{rx}(k) &= E_{rx} - elec(k) \\
E_{DA}(k) &= E_{DA} - elec(k)
\end{align*}
\]

Where:
- \(E_{elec}\): The energy consumed by the radio to run the transmitter or receiver circuitry, (equal to 50 nJ/bit).
- \(E_{amp}\): The required energy for transmitter amplifier, (equal to 100 pJ/b\ M2).
- \(k\): Number of bits.
- \(d\): Distance.
- \(E_{DA}\): The energy consumed by Transmitter circuitry to aggregate the data received by the child nodes, (equal to 50 nJ/bit).
Figure 4. First-order radio model (free space)

Or can be expressed such as:

\[ E_{tx}(k, d) = k \times E_{elec} + k \times E_{amp} \times d^2 \text{when } (d < d_0) \] (4)

Or

\[ E_{tx}(k, d) = k \times E_{elec} + k \times E_{mp} \times d^4 \text{when } (d \geq d_0) \] (5)

Where:
- \(d^2\): Represent power loss in free space with \(d^2\) and with \(d^4\) power loss.
- \(d^4\): Represent power loss in multipath fading
- \(d_0\): Represent threshold

Also, in addition to \((E_{elec} - E_{amp} - E_{DA})\) which are mentioned previously:
- \(E_{mp}\): The required energy for the transmitter amplifier in the multipath model.

A. LEACH protocol (Low-Energy Adaptive Clustering Hierarchy)

It is one of the most popular routing protocol. This protocol is not chain-based routing protocol but instead its block-based routing protocol and the reason behind mention it in this paper that this protocol was the first step in chain-based routing protocols revolution, Leach protocol architecture based on a distributed clustering algorithm, all sensors nodes communicating together using the first order radio model continues in negotiation between each other throughout the work period about who will take the place to be the chosen cluster head. The cluster head selection process will be divided into rounds; in each round, the process will be
done with 1/p probability for each sensor node to become the next cluster. The decision of selecting a node to be the next cluster head depends on the node number chooses between 0 and 1. If the number is less than a threshold $T(n)$, the node becomes a cluster head for the current round otherwise it will be considered as a regular node. The threshold represented below:

$$T(n) = \begin{cases} 
\frac{p}{1-p \times (r \mod \frac{1}{p})} & \text{if } n \in G \\
0 & \text{otherwise}
\end{cases} \quad (7)$$

Where:
- $P$: The desired percentage of cluster heads.
- $r$: The current round.
- $G$: is the set of nodes that have not been cluster-heads in the last rounds.

As a result, each node will be a cluster head for a once this, in turn, will save energy consumption and increase the network lifetime because of the changeable cluster head node process that will consume the node energy. Leach protocol is a self-organized protocol that chooses the cluster head randomly or based on some metrics such as the energy. The cluster head will be responsible for collecting information from the nodes, aggregate this information and forward it to the base station [9-13].

![Figure 6. LEACH protocol](image)

**B. PEGASIS protocol (Power-Efficient Gathering in Sensor Information Systems)**

A proposed protocol by [14] presents an improvement over the LEACH protocol. This protocol is a chain cluster-based routing protocol that presents a different mechanism of actions than the Leach protocol. The nodes that under the law of PEGASIS protocol are also communicating with each other depending on the first order radio model, equations (1)(2)(3) that been mentioned previously in this paper. There are some steps that should be followed when using PEGASIS protocol and can be classified into two phases:

- **Chain formation**
  The process of building a chain based on greedy algorithm between sensor nodes to guaranty the communication between each sensor node with it is the nearest neighbor, then send the data that been collected and gathered at the end of the chain to the leader node or to the node that is the closest to the base station. The construct of the chain always begins from the furthest node [13-15].
In the second phase, each node will take the data of its neighbor and merge it with its data then forward the gathered data to the next node as a cumulative process. The leader will be responsible for sending this data to the base station. In addition, in each round there is a different leader node in a different location is chosen randomly and used.

In a certain round, the leader will build a control token passing approach to get the gathered data from the further node to the nearest one and forward it to the base station. In fig. 7 the leader node is n2 so the first token will be pass from n0 through n1 and reach the leader node n2 then the second token will be pass from n4 through n3 and also reach the leader node n2. All the data from the two directors will be gathered in n2 then forward to the base station. [13-15].

C. EEPB (Energy Efficient PEGASIS based algorithm)

It is an improved protocol over the PEGASIS protocol proposed by the author in ref. [16]. The aim of EEPB is to overcome the shortcomings in the PEGASIS protocol by solving the problem of long links of chains because of the long distance between nodes that occurs on account of greedy algorithm which been used in PEGASIS protocol and in turn, will cause high energy consumption and fast node death. EEPB protocol handles this problem by initiate a distance threshold that will be applied to the average distance of formatted chain to decrease the problem of long link initialization as shown in the next equations:

$$D_{average} = \sum_{p=1}^{h} \left(\frac{D_p}{h}\right)$$

Where:
- $D_{average}$: is the average distance in the formed chain.
- $h$: is the hop number of the formed chain.
- $D_p$: is the distance of every segment in the formed chain, where ($p=1, 2, 3 \ldots h$).
- If the distance between an end node in a formed chain and a new requested node to join with that formed chain of nodes is longer than $D_{threshold}$ then the possibility of creating an LL problem will be high.

$$D_{threshold} = \alpha \times D_{average}$$

Where:
- $D_{threshold}$: is the threshold distance.
- $\alpha$: is a user-defined constant.

In addition, the leader node selected by EEPB according to two factors: the residual energy of the node with the distance between nodes and base station. Once the chain is created, the data transmission Stage begins by collecting and gathering the data from each node in an accumulative way until it reaches the leader node that will be responsible for sending these data to the base station. As a result, the power will be saved and balanced between nodes [15-17].
D. H- PEGASIS (hierarchal PEGASIS)

It is a beneficial extension, which adds to the PEGASIS protocol. This extension based on parallel processing between nodes. In other words, this technique allows transferring data simultaneously from different nodes to other different ones in a parallel way by sending the data from the lower layer to the upper layer up to the base station as a hierarchy structure. In addition, a signal coding technique will be used such as CDMA to avert any collisions of data that will result in interference. This solves the problem of delay by depending on multiple chains at the same time instead of single-chain and using a signal coding technique. Therefore, the problem of energy consumption also will be solved smoothly. This extension will be obvious in the MIEEPB protocol that will be described later in this paper. [17]

E. PDCCH (PEGASIS Double Cluster Head)

The technique of increasing the cluster head by doubled it proposed by the author of ref. [18] will affect the whole network positively by decreasing the traffic and in turn avoiding high delays. Normally PEGASIS protocol uses one CH that communicates with the BS. Nowadays, using a double cluster head is being preferred more than using a single cluster head. The technique of adding more cluster head is based on the mechanism of communicating between cluster heads of each layer in a hierarchal form [16, 18, 28, 29]. After chains formation stage is completed, the stage of cluster head selection is begun by choosing the nodes as a primary cluster head or secondary cluster head or either a regular node based on it is weights, the weights Q of nodes can be found or calculated by dividing it is residual energy with it is distance from the BS. The network decides to choose the node that has the highest weight as a primary cluster head. After choosing the primary CH, each node in the chain computes it is distance with the parent node that was assumed as p1, and compare it with the distance to sink node p2. If the distance of p2 > p1 then the node will be considered as a secondary CH otherwise it will be considered as a regular node participating in the chain.

\[ Q_i = \frac{E_i}{D_i} \]  

Where:

- \( E_i \): indicate the residual energy of sensor node i
- \( D_i \): Indicate the distance between sensor node i and sink.

This will provide many availssuch as decreasing the delay in transmission and receiving process between nodes by giving the responsibility of collecting data from the nodes in a certain layer by their cluster head so a cluster head will present in a form of main cluster head and secondary cluster head as shown in fig.10 As a result, this provides balancing the load of each node and increase network lifetime [18-21].

F. IEEPB (Improved Energy Efficient PEGASIS based algorithm)

EEPB overcomes several problems over PEGASIS but still has some deficiencies such as the long link (LL) phenomena and weakness in the mechanism of selecting the leader node, which will be mentioned with more details in the next few pages. In [22] the author proposed IEEPB as an improved protocol over the EEPB protocol. This protocol deal with the issues that faced us with EEPB by reducing the construction of long link (LL) in the chain by using a threshold. In IEEPB, the threshold process is to calculate and compares the distance between nodes double time, finds the shortest path to link the two adjacent nodes. Also, by selecting the
appropriate leader node based on the weighting algorithm. Weighting algorithm works by taking into account each node energy with the distance between each node and base station then distribute the weight coefficients on the leader node as shown in the following equations:

\[ D_{bs} = d_{To \ BS} / d_{ave} \] (11)

Where:
- \( d_{To \ BS} \): Distance between the sensor node and the BS.
- \( d_{ave} \): Average distance between sensor nodes and BS.

\[ E_p = E_{init} / E_i \] (12)

Where:
- \( E_p \): Portion energy.
- \( E_{init} \): Initial energy of node i.
- \( E_i \): residual energy of node i for round n

\[ W_i = w_1E_p + w_2D_{bs} \] (13)

\[ w_1 + w_2 = 1 \] (14)

Where:
- \( W_i \): combined weight of each node.
- \( w_1, w_2 \): Coefficient of weight factors.
- If \( w_1 > w_2 \) then means that the most affecting factor to select the leader node is the residual energy.
- If \( w_1 < w_2 \) then means that the most affecting factor to select the leader node is the distance between the node and BS.

Finally, the node that has the smallest weight will be chosen to be the leader node [22].

G. MIEEPB (Mobile sink improved energy-efficient PEGASIS-based routing protocol)

After EEPB problems been solved by the next emerged version IEEPB, IEEPB removes various deficiencies in EEPB; however, it still has some shortcomings such as major load on the single-chain leader, large delay in data delivery, sparse nodes in the network that facing instability periods. MIEEPB that is proposed in ref. [23] is an improvement over IEEPB that presents the sink mobility in a multi-chain model, therefore construct and applying smaller chains decreasing the load on the leader nodes. MIEEPB mechanism is the same as the previous protocol IEEPB by using the first order radio model to ensure efficient communication between sensor nodes. Also by using the token passing approach to transmit the data between nodes. MIEEPB and insert both extensions of a multi-chain and double cluster. DCH also selected by using the weighting algorithm to specify the primary and secondary cluster head. Finally, in MIEEPB a mobile sink is been used by specifying the sojourn time and sojourn location. Mobile sink node divides WSN area into more than one region as in the [23] divided into 4 regions and consider that in each round the sink node will complete one course around the 4 sojourn locations as follows

\[ T_s = \sum_{i \in S} (n_i) \] (15)

Subject to:

\[ x_{ij} = \begin{cases} D_{i,j} = j & \text{if } j = 4 \\ 0 & \text{otherwise} \end{cases} \] (16)

Where:
- \( T_s \): Total sojourn time of one course.
- \( x_{ij} \): The number of bits transmitted between chain leaders and the sink.
- \( i,j \): sink potential locations were \( i=1, j=4 \).
- \( D \): Total data transferred between chain leaders and the sink in sojourn time.

In other words, MIEEPB is a multi-chain model with continuous fixed path mobility of sink node and limited potential locations, this achieves proficient energy utilization of wireless sensors and to ensure maximizing of network lifetime [23-28].
IV. Comparison

Here, the paper compared the above discussed hierarchical routing protocols on different parameters of deployment, classification, number of chains, number of cluster head (CH) per group or chain, base station (BS) mobility, Sink mobility, Delay, Energy efficiency, quality of service (QoS), Sensor type, Protocol type, Algorithm approach, Chain data transmission approach, Avoidance of long links problem.

<table>
<thead>
<tr>
<th>Table 1. Protocols comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deployment</strong></td>
</tr>
<tr>
<td>Random</td>
</tr>
<tr>
<td><strong>Classification</strong></td>
</tr>
<tr>
<td><strong>No. of chains</strong></td>
</tr>
<tr>
<td><strong>No. of CH per group or chain</strong></td>
</tr>
<tr>
<td><strong>BS mobility</strong></td>
</tr>
<tr>
<td><strong>Sink mobility</strong></td>
</tr>
<tr>
<td><strong>Delay</strong></td>
</tr>
<tr>
<td><strong>Energy efficiency</strong></td>
</tr>
<tr>
<td><strong>QoS</strong></td>
</tr>
<tr>
<td><strong>Sensor type</strong></td>
</tr>
<tr>
<td><strong>Protocol type</strong></td>
</tr>
<tr>
<td><strong>Algorithm approach</strong></td>
</tr>
<tr>
<td><strong>Chain data transmission approach</strong></td>
</tr>
<tr>
<td><strong>Avoidance of long links problem</strong></td>
</tr>
</tbody>
</table>

V. Simulation and Results

In this paper, a Matlab is used as a simulator to analyze the performance of several PEGASIS protocol improvements by comparing: (LEACH, PEGASIS, EEPB, IEEPB, and MIEEPB) protocols due to their abilities to prolong the network lifetime depending on the number of alive nodes per round. The simulation concentrate to show the number of sensor nodes alive and to simplify the lifetime of network comparison to present a clear figure shows the network lifetime based on dead nodes percentage which are important indicators to measure the performance of a different algorithm.

<table>
<thead>
<tr>
<th>Table II. System parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>Network size</td>
</tr>
<tr>
<td>Number of nodes</td>
</tr>
<tr>
<td>Base station</td>
</tr>
<tr>
<td>The initial energy of nodes</td>
</tr>
<tr>
<td>$E_{elec}$</td>
</tr>
<tr>
<td>$E_{amp}$</td>
</tr>
<tr>
<td>$E_{mp}$</td>
</tr>
<tr>
<td>$E_{DA}$</td>
</tr>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$w_1$, $w_2$</td>
</tr>
<tr>
<td>Number of rounds</td>
</tr>
<tr>
<td>Data packet</td>
</tr>
</tbody>
</table>
The figures (Fig.10 & Fig.11) show the behavior of each protocol by knowing the number of nodes that is been dead during the passage of rounds. These figures indicate the remarkable progress due to dead nodes ratio decreasing for the subsequent protocols comparing with the previous ones. In addition, it offers to slow the speed of reaching the point of network death for the later protocols comparing with the previous protocols. The following table (table.3 shows the network lifetime comparison by estimate the number of rounds that have been passed by each protocol and the percentage of the dead nodes in a certain round number for each protocol.
Table III. Network Lifetime Comparison

<table>
<thead>
<tr>
<th>Nodes death-rate</th>
<th>Number of rounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEACH</td>
</tr>
<tr>
<td>1%</td>
<td>750</td>
</tr>
<tr>
<td>50%</td>
<td>1355</td>
</tr>
<tr>
<td>100%</td>
<td>1500</td>
</tr>
</tbody>
</table>

In addition, to make it more clearly, the following table (table 4) show a completed protocol comparison that presents the death of all nodes at 100% percent for each Protocol.

Table IV. Protocols comparison when all nodes died at 100% percent

<table>
<thead>
<tr>
<th>protocols</th>
<th>LEACH</th>
<th>PEGASIS</th>
<th>EEPB</th>
<th>IEEPB</th>
<th>MIEEPB</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH</td>
<td>-</td>
<td>27%</td>
<td>68%</td>
<td>87%</td>
<td>182%</td>
</tr>
<tr>
<td>PEGASIS</td>
<td>-</td>
<td>-</td>
<td>33%</td>
<td>47%</td>
<td>123%</td>
</tr>
<tr>
<td>EEPB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>11%</td>
<td>68%</td>
</tr>
<tr>
<td>IEEPB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>52%</td>
</tr>
<tr>
<td>MIEEPB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 12. Network lifetime based on dead nodes percentage

The PEGASIS protocol life span is longer than the LEACH protocol by 27%, where the EEPB protocol life span is longer than the PEGASIS protocol by 33% and so on for the rest of the protocols. On the other hand, the MIEEPB protocol life span is longer than the core protocol of this review (LEACH protocol) by 182% and this is a highly effective performance percentage.

VI. Discussion and conclusion

The main aim of this paper is to analyze the performance of several protocols which classified as chain-based routing protocols that is one of the hierarchical routing protocol types by going deeply in describing these protocols mechanism and models. In addition, examine the behavior of PEGASIS protocol with its genesis and its later improvements based on the number of dead nodes and how this can be effected on the network lifetime which is clear in the previous figure (figure 12). On the other hand, the conclusion of this paper is to show the importance of WSN technology to use in several fields that serve both human needs and network communication requirements with this Accelerated World. Routing protocols of WSN Plays a pivotal role in improvements of networks in general, as shown in this paper before due to the detailed description and the enhancement ability.
comparison table (table 1) of hierarchical routing protocols which gets all the attention in this work especially those who are specified in chain cluster-based routing protocols. This paper presents the enhancement that occurs clearly in (LEACH-PEGASIS-EEPB-IEEPB-MIEEPB) due to network lifetime using MATLAB simulator.

References
