Evaluation and Enhancement of Data Gathering Algorithm based Load Balance (E²DGLB) for Wireless Sensor Networks

Taha M. Rajeh¹, Ahmed I. Saleh² and Lbib M. Lbib³.
¹,²,³Dept. of Computer Eng. & Systems, Mansoura University, Egypt

Abstract—Wireless sensor network is a special kind of ad-hoc network and have many difficulties that raised a number of research issues such as routing, sensing, data gathering and decision making. However, the real critical issue in these networks is how to save energy of sensor nodes during the operation of the network to serving requirements of applications or users. In the literature there are many energy balancing algorithms proposed to prolong the WSN lifetime. In this paper we propose a new load balance data gathering algorithm for energy saving by clustering the nodes into different layers based on the hop counts between sensor nodes and the sink named E²DGLB Evaluation and Enhancement of Data Gathering Algorithm based Load Balance. The aim of our solution is to improve lifetime of WSN, by calculate the residual energy ratio of forwarding node and consider it in the calculation of the probability of each route, to ensuring a fair energy consumption during routing process. We have implemented our algorithm and compared it with DGLB, simulation results show that our algorithm outperforms DGLB by enhancing the network lifetime.

Keywords—Wireless sensor network, Network lifetime, Energy consumption, Routing protocol, Hop count, Data aggregation, energy balance.

I. INTRODUCTION

Ad-hoc is less infrastructure network, has a number of nodes connected by wireless links. Each node in the network can generate its own traffic and can be worked as a routing device between the source and the destination node.

WSNs are built by deploying the sensor nodes randomly or manually to collect data in an environment, this environment maybe has access limitations like in harsh environments [1, 2]. The collected data are delivered to the central sink node, this node intended to manipulate collected data. To reach the sink, data can be sent directly if the sink within its transmission range or through intermediate nodes. The main goal for a WSNs which consists of mainly battery-powered nodes, is to prolong the lifetime of the network as it senses information and delivers it to the sink.

The algorithm for establishing paths for transmission data from source node to any destination node is called a routing process. The network layer is the main responsible of the routing establishment. In WSNs the tasks for finding and maintaining routes are not easy tasks because energy restriction and the topological changes are unpredictable. The multi hop method is necessary in the applications of WSNs, since these applications require a large number of sensors to be deployed in widely geographical areas. In multi hop networks, nodes are not always in the transmission domain of each other. So, the intermediate nodes act as routers to receive and send the routing data packets from and to the nodes in their transmission. Routing in wireless sensor networks a very challenging mission because WSNs have an unique nature which distinguish it from other wireless networks like mobile ad hoc network (MANETs) and cellular network etc[4-10], these challenges includes Addressing, Constrained resources, Movement etc.

In this paper we propose a new routing algorithm named Evaluation and Enhancement of Data Gathering Algorithm based Load Balance (E²DGLB) where each sensor node forwards data to sink by passing data to its neighbor node. Our propose algorithm is based on two main ideas. The first idea is that network is divide to several layer and nodes of the network calculate routes costs with different probabilities use. This is the idea of the load balance data gathering for energy saving in
wireless sensor networks (DGLB)[11] protocol. The second idea that comes to improve the first is the calculation of residual energy ratio of forwarding nodes and use it in the calculation of the probability of each route. The rest of this paper is organized as follows, related work is presented in section 2. Section 3, detailed study of the DGLB algorithm . Section 4, describes the proposed algorithm. Simulation scenario and results are presented in section 5. Finally section 6, presented the conclusion. 

II. RELATED WORK

Sensor nodes are usually battery-operated with limited resources like energy supply, thus, issues such as energy aware routing[12] and sensor coordination [13] have the attention of many research groups. For the clustering, almost the published approaches for selection of a cluster-head depends on various factors like degree of connectivity [14, 15], highest energy, randomized [16] or cluster ID [17]. When the topology constantly changing or the load has to be distributed over all the nodes, the cluster heads must be selected frequently. Use the traditional clustering approaches will lead to elect same node as cluster-head every time, thus, this sensor drain its energy very fast.

Heinzelman, et. al. [18] proposed a clustering based algorithm for sensor networks, called Low Energy Adaptive Clustering Hierarchy (LEACH). In LEACH there are a few sensor nodes randomly selected as cluster heads (CHs), and this role is rotated among nodes to distributed the energy load between the nodes in the network. Each node within cluster sends data to the respective CH, CH compress data and send an aggregated packet to the sink, thus, the amount of the information transmitted to the sink is reduced . Younis, et. al. [19] proposed a hybrid, energy-efficient distributed clustering. where the cluster head periodically selected depending on a hybrid of the residual energy in node and a secondary parameter like node degree or proximity to neighbors.

Another clustering based algorithm proposed by Guan.Xin, et. al.[11] to prolong network lifetime and balancing the load among nodes, called a load balance data gathering for energy saving in wireless sensor networks( DGLB). The main idea of DGLB is clustering the nodes into different layers based on the hop counts between sensor nodes and the sink to achieve better load balance in the inter-cluster, also the cluster head would not be switched frequently and this can save the energy consumption on cluster head selection. This protocol will be detailed further in the next section, since our improvement is based primarily on it.

III. DGLB ALGORITHM

Guan. Xin and L. Guan propose DGLB a load balance data gathering for energy saving in wireless sensor networks. They take the view that, Frequent using the lowest energy or the minimum hop count path may not be proper from the point view of network lifetime. Due that, the nodes in this path will consume more energy and deplete their energy early which leads to divided the network into isolated parts and shorten network lifetime. Thus, they have proposed a hierarchical clustering algorithm to increase the network lifetime by balance the energy load among all sensor nodes and minimizing unnecessary energy consumption because of transmitting redundant sensed data. The main idea is cluster the network based on the hop counts between sensor nodes and the sink to achieve better load balance in the inter-cluster, also the cluster head would not be switched frequently and this can save the energy consumption on cluster head selection .

DGLB contains three phases. (1) After Sink broadcasts packet for divides sense area into several layers according to the hop count between the sink and sensor nodes, this lead to that each node having the same hop count will organized in the same layer. (2) Routing tree generation and data gathering , sink broadcast the pool packet to the whole network for generate the routing tree for forwarding data. (3) Data aggregation mechanism, data aggregation is an effective technique to reduce the network traffic.
3.1 Illustrative example of DGLB

After the presentation of the functioning of DGLB, we present here an example how the DGLB works, and to illustrate a case where DGLB cannot make a good decision for load balancing energy during routing process. In the example shown in Fig.(1), there are five nodes A, B, C, D and S assume that there is one source node S sends data to the sink has residual energy $E_r=0.5$, the residual energy of the other nodes A, B, C and D are 0.4, 0.3, 0.4 and 0.3 prospectively. Nodes of this example use the DGLB protocol for routing.

In the first phase, sink broadcasts an information packet. Each node receive this packet directly will assign one to its hop count and organized in layer one. Sensor node in layer one will rebroadcast the information packet to their neighbor each nodes receives this packet will organized in layer two and their hop count is two and so on.

Second phase, sink broadcast a pool packet. Each sensor node receive the pool packet puts their residual energy in the pool packet then resend it to their neighbor nodes to refresh their neighbor nodes list. nodes also calculate the distance between itself and its neighbor according to the received signal. Now each node calculate the link cost $C_{ij}$ to its forwarding nodes according to the following formula:

$$Cost_{ij} = \delta * \frac{L_{ij}}{E_{residual,i}}$$  \hspace{1cm} (1)

Where $\delta$ is a variable with value interval is (0, 1), $L_{ij}$ is the distance between node i and node j and $E_{residual,i}$ is the residual energy in node i. Once cost is calculated, nodes assign probabilities to each forwarding node according to the following formula:

$$P_{ik} = \frac{C_{ik}^{-1}}{\sum_{n=1}^{\left|FT_i\right|} \frac{1}{C_{in}}}$$  \hspace{1cm} (2)

Where, $P_{ik}$ denotes the forwarding probability from node i to node k, $\left|FT_i\right|$ is the number of nodes in the forwarding table of node i, $C_{ik}$ denotes the link cost between node i and node k.

Fig.(1), present an illustrate case where DGLB cannot make a good decision for load balancing energy during routing process. Assume that, the sender node S has hop count equal 3 and residual energy $E_r=0.4$ and in the routing table of S there are two forwarding nodes A and B. Nodes A and B having a hop count equal 2 and residual energy $E_r=0.4$ and $E_r=1$ consequently. Table (1), shows the routing tables in nodes S.

![Fig. 1 Example With DGLB](image-url)
As we see the distance from S to the A is 10m, and from S to B is 8m. Now node S has data to send. Assuming the value of $\delta$ is 1, depending on Eq.(1,2) the cost from node S to node A and from node S to node B are calculated as follows:

$$
\text{Cost}_{SA} = \frac{\delta \cdot L_{SA}}{E_{\text{residual}_S}} \gg \frac{10}{0.5} = 20
$$

$$
\text{Cost}_{SB} = \frac{\delta \cdot L_{SB}}{E_{\text{residual}_S}} \gg \frac{8}{0.5} = 16
$$

The probability $P_{SA}$ and $P_{SB}$ are:

$$
P_{SA} = \frac{C_{SA}^{-1}}{\sum_{n=1}^{\mid FT \mid} \frac{1}{e_{Sn}}} \gg \frac{0.05}{0.1125} = 0.444
$$

$$
P_{SB} = \frac{C_{SB}^{-1}}{\sum_{n=1}^{\mid FT \mid} \frac{1}{e_{Sn}}} \gg \frac{0.0625}{0.1125} = 0.555
$$

According to the above results, node B has a high probability, so the node S will select node B as a forwarding node rather than node A. Because the sender S chooses the closest neighbor without taking into account the remaining energy in neighbor. But, as we see, residual energy in node B is less than residual energy in node A, thus, node B depletes its energy and dies faster than node A. In the proposed algorithm will try to solve this problem to prolong network lifetime.

### IV. PROPOSED ALGORITHM E$^2$DGLB DESIGN

In this section, we present our new routing algorithm $E^2$DGLB. The aim of our algorithm is to improve lifetime of WSNs by calculating the residual energy ratio of forwarding node and consider it in the calculation of the probability of each route, to ensure a fair energy consumption during routing process.

For the our proposal algorithm we assume the following properties:

- There are N node randomly deployed in the sense area.
- The network is static.
- There is one sink gathers sensed data.
- The energy of sensor node can't be recharged.
- Sensor node are location-unaware.
- The radio power can be controlled.

The proposed algorithm is based on two main ideas. The first idea is that network is divided to several layers and nodes of the network calculate routes costs with different probabilities use. This idea is that of the DGLB protocol. The second idea is an improvement the first by calculating the residual energy ratio of forwarding nodes and use it in the calculation of the probability of each route.

### Table 2 Routing table in node S

<table>
<thead>
<tr>
<th>Node</th>
<th>Energy</th>
<th>Neighbor</th>
<th>Hop count</th>
<th>Distance</th>
<th>Energy</th>
<th>Cost</th>
<th>Probability (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.5</td>
<td>A</td>
<td>2</td>
<td>10</td>
<td>0.4</td>
<td>20</td>
<td>0.444444</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>2</td>
<td>8</td>
<td>0.3</td>
<td>16</td>
<td>0.555556</td>
</tr>
</tbody>
</table>

The cost from node S to node A and from node S to node B are calculated as follows:

$$
\text{Cost}_{SA} = \frac{\delta \cdot L_{SA}}{E_{\text{residual}_S}} \gg \frac{10}{0.5} = 20
$$

$$
\text{Cost}_{SB} = \frac{\delta \cdot L_{SB}}{E_{\text{residual}_S}} \gg \frac{8}{0.5} = 16
$$

The probability $P_{SA}$ and $P_{SB}$ are:

$$
P_{SA} = \frac{C_{SA}^{-1}}{\sum_{n=1}^{\mid FT \mid} \frac{1}{e_{Sn}}} \gg \frac{0.05}{0.1125} = 0.444
$$

$$
P_{SB} = \frac{C_{SB}^{-1}}{\sum_{n=1}^{\mid FT \mid} \frac{1}{e_{Sn}}} \gg \frac{0.0625}{0.1125} = 0.555
$$

According to the above results, node B has a high probability, so the node S will select node B as a forwarding node rather than node A. Because the sender S chooses the closest neighbor without taking into account the remaining energy in neighbor. But, as we see, residual energy in node B is less than residual energy in node A, thus, node B depletes its energy and dies faster than node A. In the proposed algorithm will try to solve this problem to prolong network lifetime.
The proposed algorithm has three phases. As shown in Fig(2-a) in the first phase, the sink broadcasts an information packet to the sense area through a specific power level. Nodes C and D receive the information packet directly assign one to their hop count and organized itself as nodes within a layer one. The sensor nodes belong to the layer one, will send the information packet to their neighbors A and B then each neighbor receives this packet will organized within layer two and their hop count is two.

The rules of formation layers and routing table can be summarized as follow, initially the sink hop counts will set as \( H_s = 0 \), other sensors hop counts to sink node is \( H = \infty \). After sink node broadcasts the information packet with \( H_s = 0 \). Each neighbor node N received this packet modify the hop counts \( H_n = H_s + 1 = 1 \). After that node N broadcasts the information packet with \( H_n = 1 \). Then, Each node z receives the packet which has \( H_n=1 \), it will work according the following rules:

1. \( H_z < H_n + 1 \), discard packet and do nothing.
2. \( H_z = H_n + 1 \), node z adds node n to its relay list.
3. \( H_z > H_n + 1 \), node z cleans up its relay list and adds node n to its relay list and refreshes \( H_z = H_n+1 \), broadcasts the packet \( H_z \).

According to the above rules, nodes will organized according to their hop count Fig.(2.b) shows the hop count of each node after first phase.

\[
\text{Fig. 2 (a) nodes hop counts before first phase. (b) nodes hop counts after first phase.}
\]

Second phase, sink broadcast the pool packet to the whole network. Each sensor nodes receive the pool packet must report their residual energy to their neighbor nodes. Due that, sensor node will receive poll packet can use the residual energy to refresh their neighbor nodes list. Nodes also need to calculate the distance between itself and its neighbor sensor node according to the received signal.

When Each node in the network know the residual energy of its neighbor and the distance between its neighbors and themselves, the routing tree generation process will begin.

There is a necessary condition when the sensor node try to select it forwarding node, they must select the one whose hop count is less than their own.

Now each node before sends data, will calculate the link cost to its forwarding nodes. Assume that the node \( N_i \) need to calculate the link cost \( \text{Cost}_{ij} \) to the forwarding node \( N_j \). \( \text{Cost}_{ij} \) is directly proportional to the distance between \( N_i \) and \( N_j \) and inversely proportional to the residual energy of \( N_i \).

\[
\text{Cost}_{ij} = \frac{\delta \times L_{ij}}{E_{\text{residual,}i}} \quad (3)
\]
Where $\delta$ is a variable with value interval is $(0, 1)$. The value can take randomly when the energy is equal for each sensor node, and it can be used to increase the influence on energy factor by decreasing the value of $\delta$. $L_{ij}$ is the distance between $N_i$ and $N_j$ and $E_{\text{residual},i}$ is the residual energy in $N_i$.

After calculating the cost for all forwarding nodes, node $N_i$ would calculate the relay probability ($P$) of each next hop neighbor node.

$$P_{ik} = \frac{1}{\sum_{n=1}^{\left|FT_i\right|} \frac{1}{\text{Cost}_n} \times E_{\text{residual},k}} \times E_{\text{residual},i}$$

(4)

Where, $P_{ik}$ denotes the forwarding probability from node $N_i$ to node $k$, $|FT_i|$ is the number of nodes in the forwarding table of node $N_i$, $E_{\text{residual},k}$ is the residual energy in the forwarding node. Data aggregation phase, is an effective technique to reduce the network traffic.

### 4.1 Illustrative example of $E^2$DGLB

We present here the same example as for DGLB to illustrate the functioning of $E^2$DGLB and its improvement compared to DGLB. Let’s see how our algorithm will solve the problem invoked in the critical study of DGLB and will achieve greater fairness in energy consumption.

In Fig.(3), Assume that, the sender node S has hop count equal 3 and residual energy $E_n=0.4$ and in the routing table of S there are two forwarding nodes A and B. Nodes A and B having a hop count equal 2 and residual energy $E_n=0.4$ and $E_n=1$ consequently. Table(2), shows the routing tables in nodes S.

![Fig. 3 Example With E2DGLB.](image)

<table>
<thead>
<tr>
<th>node</th>
<th>energy</th>
<th>neighbor</th>
<th>hop count</th>
<th>dist</th>
<th>energy</th>
<th>cost</th>
<th>Probability(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.5</td>
<td>B</td>
<td>2</td>
<td>10</td>
<td>0.4</td>
<td>20</td>
<td>0.253968</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>2</td>
<td>8</td>
<td>0.3</td>
<td>16</td>
<td>0.238095</td>
</tr>
</tbody>
</table>

The distance from S to the A is 10m, and from S to is 8m. Now node S has data to send. Assuming the value of $\delta$ is 1, depending on Eq.(3,4) the cost from node S to node A and from node S to node B are calculated as follows:
\[
\text{Cost}_{SA} = \frac{\delta \times L_{SA}}{E_{\text{residual}}_S} \gg 10/0.5 = 20 \\
\text{Cost}_{SB} = \frac{\delta \times L_{SB}}{E_{\text{residual}}_S} \gg 8/0.5 = 16
\]

The probability \( P_{SA} \) and \( P_{SB} \) are:

\[
P_{SA} = \frac{\frac{1}{\text{Cost}_{SA}}}{\sum_{n=1}^{\mid FT_s \mid} \frac{1}{\text{Cost}_{sn}} \times \sum_{n \in FT_s} E_{\text{residual}}_n} \times E_{\text{residual}}_A \gg 0.02/0.07875 = 0.2539
\]

\[
P_{SB} = \frac{\frac{1}{\text{Cost}_{SB}}}{\sum_{n=1}^{\mid FT_s \mid} \frac{1}{\text{Cost}_{sn}} \times \sum_{n \in FT_s} E_{\text{residual}}_n} \times E_{\text{residual}}_B \gg 0.01875/0.07875 = 0.2380
\]

According to the above results, node A has a high probability, so the node S will select node A as a forwarding node rather than node B unlike the example in 2.1. Because the residual energy of each forwarding node includes within probability equation. Through this example, it is clear that nodes A and B will be consume probably the same amount of energy through routing process.

Comparing between DGLB and E2DGLB through this example, we see that E2DGLB achieves more fairness in energy consumption between sensor nodes better than DGLB. In DGLB, node B is used more than A, so it will consume its energy quickly and dies early, so the network also dies. This example shows that the E2DGLB improves and increases the lifetime of sensor networks.

V. SIMULATION AND ANALYSIS

We evaluate the E2DGLB performance by simulation using Matlab software (R2013a) [20]. We compare the performance of the proposed algorithm E2DGLB with DGLB under the network topology contains 25 sensor deployed randomly in 25*25 m area, there is one sink placed randomly either inside or outside the sensing zone. Each sensor in the network can be sense event and send data to sink. There are four metrics will be calculated: network lifetime, total network lifetime, stability period and residual Energy. Network lifetime and stability period, is time span from the deployment to the instant when the first node dies. The total network lifetime, represents the total network time until the last node die. The residual Energy, is the amount of energy consumption during simulation time. A simulation parameters that used in this simulation are listed in Table.(2).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sensor node</td>
<td>25</td>
</tr>
<tr>
<td>Size of sense zone</td>
<td>25*25m</td>
</tr>
<tr>
<td>Transmission range</td>
<td>7 m</td>
</tr>
<tr>
<td>Initial energy</td>
<td>2J</td>
</tr>
<tr>
<td>( E_{\text{elec}} )</td>
<td>50nj/bit</td>
</tr>
<tr>
<td>( E_{\text{efs}} )</td>
<td>10pj/bit/m²</td>
</tr>
<tr>
<td>( E_{\text{amp}} )</td>
<td>0.0013pj/bit/m⁴</td>
</tr>
<tr>
<td>( E_{Da} )</td>
<td>5nj/bit/signal</td>
</tr>
<tr>
<td>Size of Data packet</td>
<td>500bits</td>
</tr>
<tr>
<td>Sink location</td>
<td>Inside or out side</td>
</tr>
</tbody>
</table>

5.1 Simulation results and analysis
- **Network lifetime**
  The network lifetime as mentioned before is the time till the first node runs out of energy. Fig.(4) shows the time of death of first node for the E$^2$DGLB and DGLB algorithms. The time of death of first node in E$^2$DGLB is greater than the time for DGLB algorithm. This means that the network lifetime is improvement in E$^2$DGLB by 10% more than DGLB.

- **Total Network lifetime**
  Fig.(5) represents the nodes died through the simulation time for the four protocols. From Fig.(5),the number of died nodes at round =5000 are 3 and 5 for E$^2$DGLB and DGLB respectively. The total lifetime of the network for E$^2$DGLB algorithm is increase From Fig.(5). The proposed algorithm E$^2$DGLB has longer stability period. The E$^2$DGLB network stability has achieved stability of 10% more than DGLB stability. These results show that the stability of E$^2$DGLB is enhanced and the network is more stable.
Residual Energy

Fig. (6), presents the residual energy of the proposed algorithm $E^2$DGLB and DGLB. Selection of appropriate forwarder node when sending data leads to save energy. to transfer packets to sink, our algorithm $E^2$DGLB uses a different forwarder node in each sending time, this decrease the load on particular node. Fig(6) shows that the residual energy of $E^2$DGLB is 52% of initial energy. In DGLB the residual energy is 40% of initial energy.

![Fig. 6 analyses of remaining energy](image)

VI. CONCLUSION

The lifetime of wireless sensor networks is a critical property in the routing protocols designed. There are many routing algorithm proposed to prolong the network lifetime by minimizing the consumption of energy of each node but the balance energy consumption among nodes do not consider in algorithm design. Thus, the aim of this paper is to present a new routing protocol for WSNs to prolong the network lifetime. The main idea of this algorithm is cluster the nodes based on the hop counts between sensor nodes and the sink to achieve better load balance in the inter-cluster, also the cluster head would not be switched frequently and this can save the energy consumption on cluster head selection. This helps to prevent depleting the energy of some nodes more than others, thus prolong the network lifetime. The simulation results show that our solution outperforms those of DGLB in terms of balance the energy consumption and network lifetime.

REFERENCES


